Speaker-specific pragmatic generalizations based on under- vs. over-informative utterances

Amanda Pogue, Chigusa Kurumada, Michael K. Tanenhaus
{apogue, ckurumada, mtan}@mail.bcs.rochester.edu
Department of Brain and Cognitive Sciences, University of Rochester

A body of work in experimental pragmatics has focused on whether speakers are optimally informative when generating referring expressions and whether listeners assume that speakers will be optimally informative [1]. Much of this work has focused on pre-nominal scalar adjectives, which are often used to select a referent from a complement set of referents of the same semantic type (e.g., *the tall cup*, when there are two cups that differ in size) [2]. Informativity of adjective use is typically evaluated based on the efficiency of singling out a target referent from an array of candidates: Under-informative use does not provide sufficient information to disambiguate referents while over-informative use provides superfluous information.

Recent research, however, demonstrates that listeners can calibrate their informativity estimates based on speaker-specific information. In particular, they down-weight a scalar adjective as a cue to a contrast in real time comprehension when that speaker over modifies and is generally unreliable [3]. An important, but less explored, question is how listeners avoid overgeneralization of the speaker-specific information, in particular, given the prevalent over-modification observed in natural discourse (30% of speakers used adjectives over-informatively in a production study in [4], and 50% in [5]). Thus a single instance of over-modification provides less reliable evidence that the speaker might be non-optimal in other domains of pragmatic language use compared to an instance of under-modification, which is less common and tends to result in referential ambiguity. If the mechanism of speaker-specific calibration of pragmatic inferences is rational, then it should take into account these prior likelihoods, generalizing conservatively from evidence of over-modification compared to evidence of under-modification.

We address this question through 3 series of experiments with an exposure phase and a generalization phase. We first demonstrate that listeners tend to infer that an under-informative speaker who fails to use of a set of scalar adjectives (big/small) is likely to omit other pairs of adjectives (e.g., skinny/wide) as well. We then examine if generalization is based on the form (i.e., use and non-use of adjectives), or informativity of the utterances. Finally, we ask whether generalization based on under- and over-informativity is symmetric.

Experiment 1 (n = 32): Generalization across different adjective types

In the Exposure Phase (20 items), participants heard two speakers giving instructions of the form "Click on the X". Each instruction was paired with a four-picture scene: a pair of items that differed in size, and two singleton items. To refer to an item in a contrast pair, one of the speakers (**the modifying speaker**) consistently used a prenominal adjective ("big" or "small") to pick out the unique referent. The other speaker (**the non-modifying speaker**) would always use a bare noun, making the utterances under-informative. In the Generalization Phase (24 items), participants were given transcribed instructions by each of the speakers, and asked to judge which of the two speakers would more likely produce a particular instruction. Crucially, half of the instructions contained bare nouns, and the other half contained modified instructions (2 "big," 2 "small," and 8 instructions with adjectives not included in the Exposure Phase (4 "wide" or "skinny," 4 "tall" or "short")). Participants were equally likely to select the modifying speaker for the modified generalization trials with both seen and unseen adjectives (83% of total responses to relevant trials). They also systematically chose the non-modifying speaker for the bare noun trials (82.6%), suggesting that speaker-specific information was generalized across seen and unseen adjective types.

Experiment 2: Form-based vs. Informativity-based generalization

Experiment 2-a (n = 33): To test whether participants generalize adjective use on the basis of speaker informativity, we replaced the bare noun instructions in the Generalization Phase with

orthogonal color adjectives (e.g., "*Click on the green car*" when both cars in the scene are green). If generalization is form-based (i.e., based solely on whether or not a speaker had used an adjective), participants should select the modifying speaker on the color-adjective trials. On the other hand, if it is based on informativity, they should select the non-modifying (previously "under-informative") speaker. We found evidence of informativity-based generalization: Participants reliably selected the non-modifying speaker for the color-modified trials (68%) and the modifying speaker for the scalar-modified generalization trails (81%). **Experiment 2-b (n = 32)**: We repeated Experiment 2-a and prefaced it with instructions to pay close attention to possible speaker differences in clarity and appropriateness. This increased the proportion of informativity-based generalizations ($p < .01$), and overall participants reliably selected the modifying speaker (93%) for the scalar-modified trials, but not for the color-modified trials (only 18%). Thus calling explicit attention to the quality of the instructions made listeners more willing to infer that the non-modifying speaker would be less than pragmatically optimal overall and therefore more likely to *use* an orthogonal, and hence under-informative, color-adjective.

Experiment 3: Under-informative vs. Over-informative utterances as evidence

Experiment 3-a (n=33): To test whether generalization based on under-, and over-, informativity is symmetric we repeated Experiment 1 with scenes that no longer contained a size contrast in the Exposure Phase, making the modified instructions (e.g., *Click on the big bike*) over-informative. Generalization Phase scenes contained contrast pairs. If participants infer that the modifying (over-informative) speaker would be pragmatically non-optimal overall, they should be more willing to select the modifying speaker for the bare noun (under-informative) generalization items, and the non-modifying speaker for modified (optimal) instructions. Participants' responses were in fact solely form-based: They preferred to select the non-modifying speaker for the non-modified generalization trials (91%) and the modifying speaker for the modified generalization trials (85%), even when we added instructions to pay attention to quality (n=33). **Experiment 3-b (n=34)** further examined whether participants would show any sign of informativity-based generalization based on over-modification. To highlight the pragmatic non-optimality of the over-informative instructions, we truncated audio stimuli for the Exposure phase, resulting in referential ambiguity for the modifying-speaker's utterances (e.g., "Click on the sma-" when there is more than one small referent) but not for the non-modifying speaker's utterances (e.g., "Click on the ca-" when a target is a "camera" and there is no onset overlap across referents). All generalization trials contained a scalar adjective, and half of the trials also contained a redundant color adjective (e.g., "Click on the small brown book"). If participants generalize based on informativity, they should choose the modifying (over-informative) speaker for these over-modified trials. However, no such informativity-based generalization was observed. Participants neither reliably selected the non-modifying speaker for the optimally-modified trials (47%), nor the modifying speaker for the over-modified trials (55%).

Discussion: Clear asymmetry emerged between the types of evidence: Listeners readily generalize under-informative uses of a scalar adjective to unseen adjectives and infer that the same speaker may make redundant (hence under-informative) use of color adjectives. On the other hand, listeners do not generalize from over-informative utterances, even when they result in referential ambiguity. We conclude that listeners generalize evidence according to 1) informativity in a given context, and 2) the prior likelihood that a particular type of utterance will provide evidence about a speaker's general pragmatic ability. We argue that appropriate (rational) inference that considers both priors and context-specific evidence helps prevent over-generalization, thereby optimizing speaker-specific pragmatic inferences.

References: [1] Frank & Goodman (2012). *Science*; [2] Sedivy et al. (1999). *Cognition;* [3] Grodner & Sedivy (2011). *The processing and acquisition of reference*; [4] Engelhardt et al. (2006). *JML*; [5] Nadig & Sedivy (2002). *Psych. Sci*.