Speaker-specific generalization of pragmatic inferences based on prenominal adjectives

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

Abstract

To navigate many-to-many mappings between referents and linguistic expressions, listeners need to calibrate likelihood estimates for different referential expressions taking into account both the context and speaker-specific variation. Focusing on speaker variation, we present three experiments. Experiment 1 establishes that listeners generalize speaker-specific patterns of pre-nominal modification use across different adjective types. Experiment 2 examined a) the dimension of generalization (form-based or informativity-based); b) effects of the strength of the evidence (implicit or explicit); and c) individual differences in dimensions of generalization. Experiment 3 asked parallel questions for exposure to over-specified utterances; we predicted more conservative generalizations because in spontaneous utterances, speakers are more likely to over-modify than under-modify.

Keywords: sentence processing; adaptation; generalization; pragmatics; informativity; referential expressions

Introduction

A key feature of human language is that there are many-to-many mappings between referents and linguistic expressions. A pet dog can be referred to by many expressions (e.g., *the dog*, *Charlie*, *he*, or *my friend*) whereas the expression *the dog* can be used to refer to a real dog, a toy dog, or a contemptible person. Referential expressions can also be made arbitrarily long (e.g., *the big dog*, *the big brown dog*, *the big brown furry dog*). One long-standing issue in psycholinguistic research is how language users identify an intended referent at the rate of speed and accuracy evidenced in our real time language use (Brown-Schmidt & Hanna, 2011).

Listeners seem to deal with this variability by capitalizing on the belief that speakers behave rationally, formulating their utterances to be as economical as possible while conveying all necessary information (Grice, 1975). For example, a rational speaker is more likely to use a pre-nominal scalar adjective (e.g., *the big dog*) when there is a complement set of referents of the same semantic type (e.g., a big and one or more small dogs) in the same context. By assuming a rational model of the speaker, listeners can choose the referent that maximizes the informativeness of a linguistic element – the amount of uncertainty reduced by the element – given an array of possible referents (Frank & Goodman, 2012; 2014).

An important, but less explored, question is how the listener copes with variability in referential expressions that is not easily predictable based on the rational speaker model. In fact, a number of studies report that spontaneous utterances often contain prenominal modifiers that would be deemed unnecessary for singling out a unique referent in a given context (Belke, 2006; Deutsch & Pechmann, 1982; Engelhardt, Bailey, & Ferreira, 2006; Pechmann, 1989; Sonnenschein, 1984). For instance, 30% of speakers used superfluous adjectives in a production study in Englehardt et al., (2006) and 50% in Nadig & Sedivy (2002). Some of the variability stems from factors such as the speakers' inattentiveness or speech production difficulties, but much of it is due to differences in speaking styles, goals of the conversation, and their certainty about the likelihood of getting their intended message across.

We propose a framework in which listeners navigate the variability by adapting their referential expectations in a speaker-specific manner. In this framework, as predicted by the rational-speaker model, we expect listeners to make inferences about how a particular referent would be referred to in a given (postulated) referential domain. When the actual input deviates from the expected signal, listeners update their likelihood estimates for referential expressions, by updating their assumptions either about the context or about the speaker, so as to better predict the future input from the same speaker.

As a first step in developing this approach we focus on simple situations in which the input from one of two speakers deviates from what is expected based on the rational speaker model. Specifically, that speaker does or does not use a scalar adjective (e.g., *big*/*small*) that would be necessary for singling out a referent, or if used, would provide redundant information (*under-* and *over-modifying speakers* given the rational model). We then examine how listeners generalize the information to new stimuli, which allows us to extrapolate if and how the listener's expectations were adapted for the two speakers.

While speaker-specific adaptation has not thus far been studied extensively with respect to reference resolution (but see Grodner & Sedivy, 2011), its importance is increasingly appreciated in other domains of language processing involving rational inference under uncertainty (e.g., phonetic adaptation (see Kraljic & Samuel; 2007; Kleinschmidt & Jaeger, 2015)). One important insight coming from this literature is that rational listeners can evaluate evidence according to its reliability. In the current task, we expect that listeners should adapt their referential expectations for a particular speaker differently depending on the types of evidence.

Likewise, to generalize meaningfully from limited instances of exposure, rational listeners integrate their prior knowledge about the statistical structure of the linguistic input with the data at hand (Xu & Tenenbaum, 2007). Given the prevalent over-modification observed in natural discourse, a single instance of a redundant adjective use provides less reliable evidence that the speaker would be non-optimal in other domains of pragmatic language use compared to an instance of under-modification. Integration of prior likelihoods would result in more conservative generalization (at the speaker-level) from evidence of over-modification compared to evidence of under-modification.

In the current paper we first establish that speakers will generalize information from a single pair of adjectives to unseen adjectives in a speaker-specific manner (Experiment 1). Focusing first on exposure to underspecified utterances, Experiment 2 examines: a) the dimension of generalization; b) effects of the strength of the evidence (implicit or explicit); and c) individual differences in dimensions of generalization. Experiment 3 then asks parallel questions for exposure to over-specified utterances, where we predict more conservative generalization.

Experiment 1

We first asked whether listeners would generalize information from observed to unobserved (new) adjectives in a speaker-specific manner. Listeners were introduced to two speakers and tasked with estimating which of four objects a speaker was likely to be talking about. The two speakers varied in their descriptions: only one speaker used adjectives to pick out a unique referent (modifying speaker). We then asked listeners to guess which speaker likely uttered transcribed instructions that were either modified (with novel, or previously used adjectives) or unmodified. If listeners generalize their assumptions about the speaker's adjective use, they should attribute both the familiar and novel modified instructions to the modifying speaker, and the unmodified instructions to the other speaker.

Methods

Thirty-two English speaking adults residing in the USA were recruited online using Amazon Mechanical Turk. Participants were naïve to the design of the experiment, and were compensated $1.00 for completing the task.

In Exposure Phase, participants listened to verbal instructions of the form "Click on the ___." and selected a referent from a 2 x 2 grid with an image in each of the four grid squares (Figure 1A). Instructions were recorded by two speakers (one male and one female). In this Exposure Phase, two of the items in the display contrasted in size. On half of the trials one of the speakers would make a request using a prenominal adjective such as, "*Click on the big/small cake.*" On the remaining trials the other (under-informative) speaker would produce unmodified instructions (e.g., "Click on the cake"). There were 20 Exposure Phase trials (10 per speaker) presented in a randomized order. Location of the target object, adjective, and which speaker modified were counterbalanced across participants. Participants were instructed to make their best guess when they thought the speaker was unclear, or if they were uncertain. Participants were not given any feedback on their responses.

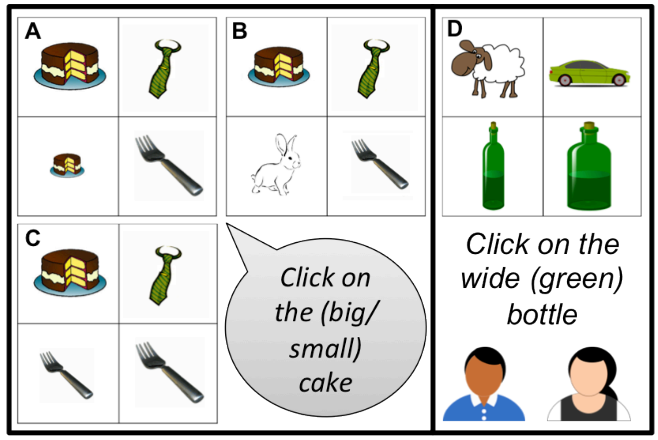

Figure 1: Four image display in the Exposure Phase of Experiment 1-2 (A); Experiment 3 Original & Explicit Instruction Conditions (B), Explicit Evidence Condition (C); Generalization Phase display (D)

In the Generalization Phase, participants read instructions that they were told had been transcribed and were asked to guess which of the two speakers most likely uttered each instruction. On half of the trials the instruction contained a modifying adjective. One third of the modified trials contained the same adjectives as in the Exposure Phase; the remaining two thirds contained new scalar adjective pairs (skinny/wide, tall/short). On the remaining trials the written instructions were unmodified. There were 24 trials presented in a randomized order. Participants clicked on an image of a male or female avatar that represented which speaker they thought uttered the written instructions. An example of the design of the Generalization Phase can be seen in Figure 1D (In this version the transcription read, "Click on the wide bottle.").

Results and Discussion

Choices in the Generalization Phase are plotted in Figure 2. Participants selected the modifying-speaker, who used *big/small* in the exposure phase, for the sentences that include *big/small* (83%), and the non-modifying (under-informative) speaker in the unmodified trials (80%). Choice patterns for new adjectives were almost identical to those for exposure adjectives: 84% and 84% for skinny/wide, and 83% and 84% for tall/short. We constructed a mixed-model logistical regression of the responses given for the modifying speaker in the Generalization Phase with Adjective Type (big/small, tall/short, skinny/wide), and Test Trial Type (modified or non-modified) as the fixed effects, and subject as a random effect. As predicted, Test Trial Type was the only significant predictor of whether participants would choose the modifying speaker ($\beta = 2.988$, $p < .001$), confirming that listeners track speaker-specific

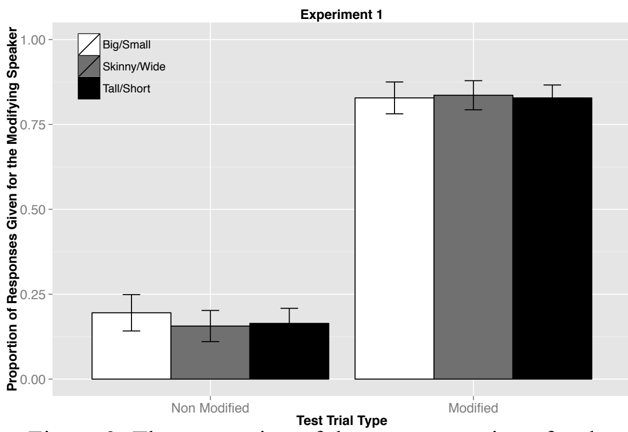

Figure 2: The proportion of the responses given for the modifying speaker on the different test trials by adjective types in Experiment 1

differences in adjective use, and generalize to unseen adjectives.

Experiment 2

We now examine the dimension of generalization. We ask whether participants generalized based on utterance length (form-based), or inferred that one of the speakers was more or less informative (informativity-based). To this end, we replaced bare noun instructions in the Generalization Phase of Experiment 1 with orthogonal color adjectives (e.g., *Click on the green car* when both cars in the scene are green). If generalizations are form-based (i.e., based solely on whether or not a speaker had used an adjective), participants should select the modifying speaker on both the color-adjective trials and the scalar adjective trials. On the other hand, if generalization is based on informativity, participants should select the previously non-modifying (under-informative) speaker. Finally, participants could be conservative in their generalizations, treating the classes of adjectives differently. Listeners do not treat all adjectives as equal cues for contrast (Sedivy, 2003). Thus participants might be hesitant to generalize to a new class of adjectives. If this were the case, participants would evenly choose between the speakers for the color trials, because they contained a new type of adjective (conservative evidence-based generalizations).

We also asked whether the strength of the evidence might affect the dimension of generalization. While explicit cues are not always necessary for listeners to make inferences about a speaker's pragmatic incompetence (Kurumada et al., 2014), Grodner & Sedivy (2011) found that explicit instructions enhanced speaker-specific pragmatic adaptation. They suggested that the explicit cue called attention to the low-level redundancy in the signal. Therefore, in a separate condition we included instructions that explicitly asked listeners to pay close attention to potential speaker differences in clarity and appropriateness. This provided an explicit top-down cue that emphasized the task of tracking differences between speakers in terms of clarity. We hypothesized that while participants might pick up on these speaker differences implicitly, the addition of an explicit cue might increase informativity-based generalizations by highlighting these low-level ambiguities in the signal. If that is the case, then adaptation is modulated both by the signal and by the listeners' task.

Methods

Sixty-five English speaking adults residing in the USA who had not previously participated in a study in this series, and were naïve about the design, were compensated $1.00 for completing the task on Mechanical Turk.

Half of the participants (n=33) read the same instructions as in Experiment 1 with no special focus on speaker differences (Original Instructions). The remaining participants (n=32) were instructed to pay attention to differences between speakers in terms of clarity and naturalness, and report any oddities at the end of the task (Explicit Instructions).

The Exposure Phase was identical to Experiment 1. The Generalization Phase was modified such that the bare noun instructions were replaced with instructions containing color adjectives. On these trials the contrastive item pair differed in size along the same dimensions as the scalar adjectives used in the scalar modified trials, but did not differ in color. Thus, instructions such as, "*Click on the big bottle"* (in Figure 1D) would pick out a unique referent, whereas instructions such as, "*Click on the green bottle"* would not. Thus, all transcribed utterances in the Generalization Phase contained either a scalar or a color adjective.

Results and Discussion

In the Original Instructions Condition, participants' responses on scalar-modified trials were similar to those in Experiment 1. For both seen and unseen adjective types, they primarily picked the modifying speaker (81%). For the color-modified trials, they exhibited a trend towards the informativity-based generalization, preferring the non-modifying (under-informative) speaker (68%). When explicit attention was called to the clarity of the two speaker's utterances (Explicit Instructions conditions), responses showed more pronounced trends towards the informativity-based generalization. Participants selected the speaker who previously used no adjective in the exposure phase in the (under-informative) color-modified trials 88% of the time. We conducted a mixed-effects logistic regression analysis with Instruction Condition, Adjective Type, and Test Trial Type as fixed effects, and subject as a random effect. Instruction Condition (β = -1.297, p <.01), an interaction of Instruction Condition by Test Trial Type (β = 3.827, p <.001), an interaction of Adjective Type by Test Trial Type (β = 2.152, p <.05), and a three-way interaction between the fixed effects (β = -1.950, p <.05) were significant predictors of whether participants chose the modifying speaker. These results are consistent with the hypothesis that explicit instructions biased participants to generalize more on informativity. With explicit instructions

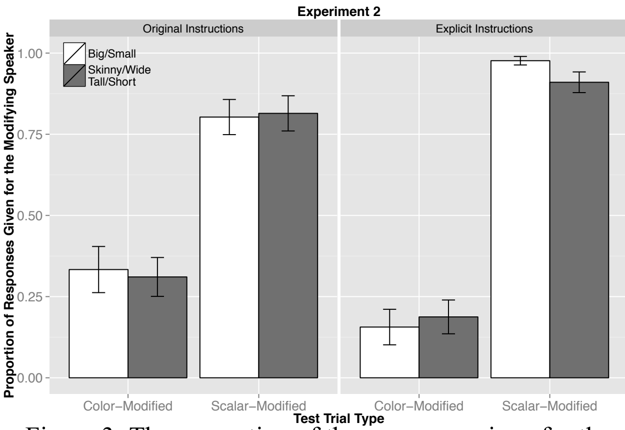

Figure 3: The proportion of the responses given for the modifying speaker on the different test trials by adjective types in Experiment 2

fewer listeners attributed the color-modified instructions to the modifying speaker, and attributed more of the scalar-modified instructions to the modifying speaker (see Figure 3).

One interesting possibility is that listeners differ in the degree to which they make form-based and informativity-based generalizations. We tested for different patterns of generalization across participants by fitting multivariate mixture models to the data. Separate models were fit for each Trial Type in each of the conditions for 1-6 components using the mixtools package (Benaglia, Chaveau, Hunter, & Young, 2009) in R, which uses expectation maximization (EM) to estimate the optimal parameter values.

In scalar-modified trials, listeners primarily attributed these instructions to the modifying speaker, and more so in the Explicit Instructions condition. In the Original Instructions condition our mixture model analysis found that the majority (73%) of the participants selected the modifying speaker for the scalar adjective trials on average 98.2% of the time, and the remaining 27% of participants selected the modifying speaker on average 35% of the time. In the Explicit Instructions condition the model found that the majority (88%) of the participants selected the modifying speaker for the scalar adjectives on average 98.2% of the time. The remaining 12% of the participants selected the modifying speaker on average 59% of the time.

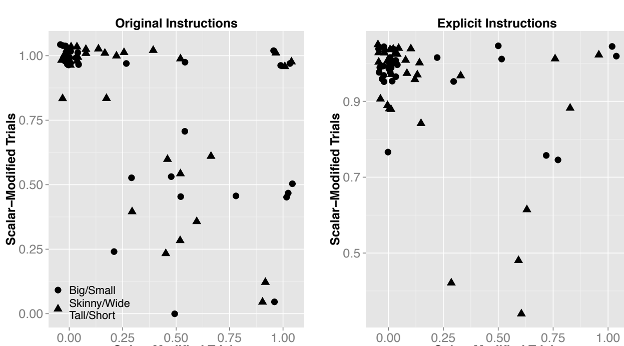

Figure 4: Proportion of responses given for the modifying speaker for color- by scalar-modified trials for each individual subject in Experiment 2

For color-modified trials a three-component model fit the data significantly better than the one-component model ($\chi^2(6) = 373.2$, $p < .001$) or the two-component model ($\chi^2(3) = 18.8$, $p < .001$) in the Original Instructions condition. Individual participants appear to be responding in three different ways, with 12% of the participants selecting the modifying speaker for these trials 98% of the time (evidence for form-based generalizations), 30% selecting the modifying speaker 57% of the time (evidence for conservative evidence-based generalizations), and the remaining 58% of the participants picking the modifying speaker only 5% of the time (evidence for informativity based generalizations). However, in the Explicit Instructions condition a two-component model fit the data significantly better than the one-component model ($\chi^2(3) = 305.1$, $p < .001$) or the three-component model ($\chi^2(3) = 0.14$, $p = 1$). This finding indicates that individual participants are responding in two different ways, with only 18.8% of the participants selecting the modifying speaker for these trials 75% of the time (evidence for form-based generalizations), and the remaining 81.2% of the participants picking the modifying speaker only 5% of the time (evidence for informativity-based generalizations).

This analysis reveals that there is more variability in responses in the Original Instructions condition, than in responses in the Explicit Instructions condition. This is evident in the tighter clustering pattern towards the top left corner for the Explicit Instructions condition in Figure 4. In this figure, if listeners generalizations are informativity-based, we expect results to cluster in the top left, whereas if they were form-based we should expect clustering in the top right corner (where the proportion of responses for the modifying speaker is near to 1 for both test trial types). In sum, calling explicit attention to the quality of the instructions made listeners more willing to infer that the non-modifying speaker would be less pragmatically optimal overall and therefore *more* likely to use an under-informative color-adjective.

Experiment 3

Speakers rarely under-modify (except in highly collaborative tasks; Brown-Schmidt & Tanenhaus (2008)). In contrast, in tasks like the ones used here, speakers frequently over-modify and listeners are less likely to penalize over-informative utterances than under-informative utterances (Engelhardt et al., 2006). Therefore a rational listener should be less likely to generalize from over-informative input than from under-informative input. We tested this prediction by repeating 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 cake*) over-informative. As in Experiment 2, we manipulated whether or not the instructions explicitly called attention to potential differences in clarity and naturalness.

Finally we also conducted a modified version of Experiment 3 (3b) with ambiguous (truncated) instructions to create a stronger manipulation highlighting the redundancy of the over-modified statements.

The Generalization Phase was analogous to that of Experiment 2. If participants infer that speakers differ with respect to informativity they should expect that the previously modifying (over-informative) speaker would be more likely to produce over-modified statements (e.g., *Click on the wide green bottle*), whereas the non-modifying speaker should be more likely to utter the concisely-modified instructions (e.g., *Click on the wide bottle.*). However, we predict that because of prior knowledge that speakers often over-modify, listeners will be conservative and therefore hesitant to make informativity-based generalizations based on exposure to over-informative utterances. Therefore they should generalize more narrowly, making only form-based generalizations.

Methods

One hundred and five English speaking adults residing in the USA were recruited online and compensated $1.00 for completing the task on Mechanical Turk. Eight participants were excluded for: 1) having previously participated in a study in this series (n=6), 2) giving no response in the Exposure Phase (n=1), and 3) technical difficulties (n=1). Remaining participants had not previously participated in a study in this series, and were naïve about the design.

Experiment 3A: In the Original Condition (n=31) the Exposure Phase contained four singleton items (see: Figure 1B). Under these circumstances, the use of a scalar modifier is over-informative. The Explicit Instructions Condition (n=33) was identical to the Original Condition, except participants were explicitly instructed to pay attention to speaker differences as in Experiment 2. Both the Original Condition and the Explicit Instructions Conditions used the same Generalization Phase trials as in Experiment 1.

Experiment 3B Explicit Evidence (n=34): the Exposure Phase contained a mix of trials with four singletons, or including a contrast pair (see: Figures 1B-C). Crucially the instructions made by both speakers never referred to an item from the contrasting pair. 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). At the end of each trial participants were told to select the referent. Generalization Phase trials contained a scalar adjective, and half of the trials also contained a redundant color adjective (e.g., *Click on the wide green bottle*).

Results and Discussion

Experiment 3A: Participants' responses to the Original and Explicit Instructions Conditions were solely form-based:

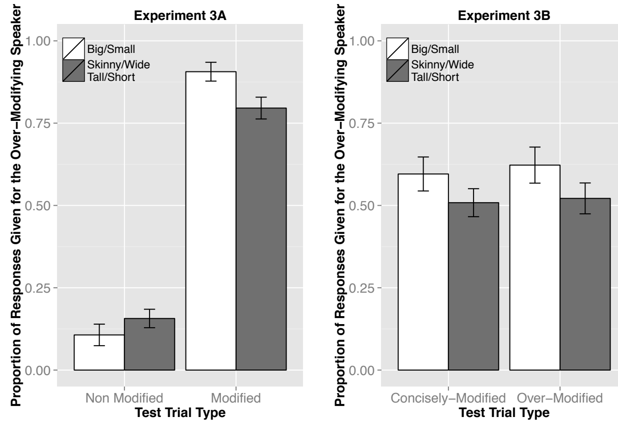

Figure 5: The proportion of the responses given for the modifying speaker on the different test trials by adjective types in Experiment 3A and 3B

participants preferred to select the non-modifying speaker for the non-modified generalization trials (86%) and the modifying speaker for the modified generalization trials (85%), even when we added instructions to pay attention to quality. A mixed regression analysis with Condition (Original, Explicit Instruction), Test Trial Type, and Scalar Dimension as random effects and subject as a fixed effect showed that only Test Trial Type was a highly significant predictor (β = 3.997, p < .001), Condition, Scalar Dimension, and the interactions were not significant predictors (*p*s >.1).

Experiment 3B: Participants neither reliably selected the non-modifying speaker for the concisely-modified trials (45%), nor the modifying speaker for the over-modified trials (57%). There were no significant predictors in the Explicit Evidence Condition (*p*s > 1)[1]. Response patterns in for both experiments are plotted in Figure 5.

In contrast to Experiments 1 and 2, participants made only form-based generalizations in Experiment 3A. In Experiment 3B listeners did not show evidence for having a speaker preference for either type of test trial. Despite explicit evidence of communicative non-optimality of over-informative utterances, participants were still unwilling to consider the possibility that one speaker would be more likely to be over-informative at test. This result stands in contrast with those from Experiment 2, suggesting that participants weigh under-informative utterances and over-informative utterances differently as evidence for informativity-based generalization.

[1] One may suspect that the truncated stimuli may be too noisy to evidence any speaker-specific information. However, post-hoc tests suggest that participants were more likely to pick the over-modifying speaker for over-modified test utterances that contained Exposure Phase adjectives (62%, p < .01), and trended towards the same pattern for the concisely modified trials (p = .06), but not for any other test trial type (ps > .1). This indicates that listeners had at least registered the over-modifying speaker's adjective uses.

General Discussion

We examined whether listeners track speaker-specific information about referential expressions and adapt their expectations according to the input. Results from the three sets of experiments suggest that listeners generalize information at the speaker level and adapt their referential expectations for unseen items. This approach can close the gap in the literature by providing a framework for how rational listeners can make effective use of the variable input that is commonly observed in spontaneous speech production.

Importantly, we predicted that listeners condition their adaptation on their prior beliefs about the statistical structure of the data to avoid under- or over-generalization. In our data, the effect of prior beliefs is evidenced by a clear asymmetry that emerged between how listeners made speaker-specific generalizations. Participants generalized from a speaker's use of under-modified expressions, which are generally less common, and hence a reliable indicator of the speaker being under-informative. On the other hand, informativity-based generalization was not observed for the more common over-modified expressions, even when they resulted in difficulties in referential resolution in the current communicative context. These patterns of generalization are broadly compatible with the assumptions of rational inference under uncertainty: listeners seem to optimize inferences by combining their prior expectations and newly observed data.

Finally, the effects of explicit instructions in Experiment 2 suggest that speaker-specific expectations for referential expressions can be further calibrated according to the listener's construal of the task and context. In future research we plan to investigate how listeners evaluate speaker-, task- and context-specific information to optimize interpretation of referential expectations.

Acknowledgments

Thanks to T. Florian Jaeger, the HLP and Kurumada-Tanenhaus Labs for helpful feedback and advice, Matthew Overlan and Sarah Bibyk for voice recordings, and Chelsea Marsh and Olga Nikolayeva for help in creating stimuli and pilot testing. Support provided by a NSERC Postgraduate Scholarship (AP) and NIH grant HD27206 (MKT).

References

Belke, E. (2006). Visual determinants of preferred adjective order. *Visual Cognition, 14*, 261–294.

Benaglia, T., Chaveau, D., Hunter, D. R., & Young, D. S. (2009). mixtools: An R Package for Analyzing Finite Mixture Models. *Journal of Statistical Software*, *32, 1-29*.

Brown-Schmidt, S., & Hanna, J. (2011). Talking in another person's shoes: Incremental perspective-taking in language processing. *Dialog and Discourse*, *2*, 11–33.

Brown-Schmidt, S., & Tanenhaus, M. K. (2008). Real-time investigation of referential domains in unscripted conversation: a targeted language game approach. *Cognitive Science, 32*, 643–684.

Deutsch, W., & Pechmann, T. (1982). Social interaction and the development of definite descriptions. *Cognition, 11*, 159–184.

Engelhardt, P. E., Bailey, K. G. D., & Ferreira, F. (2006). Do speakers and listeners observe the Gricean maxim of quantity. *Journal of Memory and Language, 54,* 554–573.

Frank, M. C., & Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science, 336*, 998.

Frank, M. C., & Goodman, N. D. (2014). Inferring word meanings by assuming that speakers are informative. *Cognitive psychology*, *75*, 80-96.

Grice, H.P. (1975). Logic and conversation. In P. Cole and J.L. Morgan (Eds.), *Syntax and semantics, Vol. 3: Speech acts* (pp. 225-242). New York: Seminar Press.

Grodner, D., & Sedivy, J. (2011). The effect of speaker-specific information on pragmatic inferences. In Gibson, E. A., & Pearlmutter, N. J. (Eds.). *Processing and Acquisition of Reference* (pp. 239-272). Cambridge, MA, USA: MIT Press.

Kleinschmidt, D. F., & Jaeger, T. F. (2015). Robust speech perception: Recognize the familiar, generalize to the similar and adapt to the novel. *Psychological Review, 122*, 148-203.

Kraljic, T., & Samuel, A. G. (2007). Perceptual adjustments to multiple speakers. *Journal of Memory and Language*, *56*, 1-15.

Kurumada, C., Brown, M., Bibyk, S., Pontillo, D. F, & Tanenhaus, M. K. (2014). Rapid adaptation in online pragmatic interpretation of contrastive prosody. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society (pp. 791-796).* Austin, TX: Cognitive Science Society.

Nadig, A. S., & Sedivy, J. C. (2002). Evidence of perspective-taking constraints in children's on-line reference resolution. *Psychological Science*, *13*, 329-336.

Pechmann, T. (1989). Incremental speech production and referential overspecification. *Linguistics, 27*, 89–110.

Sedivy, J. C., K. Tanenhaus, M., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, *71*, 109–147.

Sedivy, J. C. (2003). Pragmatic versus form-based accounts of referential contrast: evidence for effects of informativity expectations. *Journal of Psycholinguistic Research*, *32*, 3–23.

Sonnenschein, S. (1984). The effect of redundant communication on listeners: Why different types have different effects. *Journal of Psycholinguistic Research, 13*, 147–166

Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychological review*, *114*, 245.