

Exploring how speakers mark, and listeners assess, certainty
Amanda Pogue (apogue@ur.rochester.edu) & Michael K. Tanenhaus
Brain and Cognitive Sciences, University of Rochester

Successful communication would seem to require speakers to signal their degree of certainty about an utterance through lexical choices and prosodic markings (e.g., It's a dog, I think/THINK it's a dog) and listeners to successfully use these signals and adapt to variations in how different speakers signal uncertainty and whether a speaker is likely to overestimate or underestimate how certain she should be. We are beginning address a range of questions about how uncertainty is marked by speakers and inferred by listeners, and how listeners make use of this information when calibrating specifically or generally from a speaker. First, we asked whether speakers have conscious access to features of productions that can mark certainty (Pre-task). Second, we ask whether listeners have a stable preference for different lexical structures that mark uncertainty, and whether degree of uncertainty is indicated by different production cues (Experiment 1). Third, we introduce a task that asks speakers to produce labels for objects they are uncertain about, in order to see how uncertainty influences the utterances speakers choose for communicating their labels to an interlocutor (Experiment 2). Finally, we discuss a new project investigating how uncertainty expressions affect an interlocutors' choice of a name versus a description in a referential communication task.

We consider the possibility that the speaker might have two motivations for marking uncertainty in their productions: first, to appear to be producing accurate information, and second, to signal to the listener that they are a reliable speaker. These two motivations are motivated by Grice's (1975) principles of cooperation, such that if speakers want to be seen as cooperative, they should want to provide only the information they know to be true, or in this case, to do so with certainty, only when certain of the truth. Secondly, we consider the possibility that some speakers have better access to their own production choices for marking uncertainty than others. In the current tasks, and in our future research we aim to explore how speakers mark uncertainty, whether by lexical choice, prosodic contour, disfluencies, hesitations, etc, and how interlocutors use these uncertainty markings. Our goal is to extend beyond the current work in the field suggesting that listeners are sensitive to cues to speaker knowledge (see: Smith & Clark, 1993; Brennan & Williams, 1995; Swerts & Krahmer, 2005), to show that this ability allows listeners to adjust their expectations for an interlocutors' likely referential knowledge.

Pre-task: We predicted that speakers would modulate their speech when they were asked to mark uncertainty compared to just reading aloud the same sentences. Eight adult speakers of American English were recruited from the University of Rochester community, and were told that they would be making recordings for a future study. They were randomly shown individual phrases about birds that used a variety of possible hedges for uncertainty, and were asked to read each one out loud. Then they were told to imagine that they were in a task where they briefly saw the birds, but that they might be unsure that they had correctly seen the bird, either because it was displayed too briefly, or because it was partially occluded. They were shown the same phrases in random order again, and were asked to speak them out loud. We predicted that if speakers are aware of how to explicitly mark uncertainty that listeners would give lower certainty ratings when the speaker was told to produce the utterance as if she were unsure.

Experiment 1: Participants on Amazon's Mechanical Turk were randomly assigned to a read-text (n=16) or listen (n=160) task. They were told that previous workers were asked to describe pictures of birds, but that sometimes the pictures flashed quickly or were not fully visible. Their task was to rate how certain they thought the speaker was on a slider scale of 1 to 100 (1= not at

all certain, 100=completely certain). Participants either read each of the 8 target sentences plus 2 control sentences (to ensure that they were paying attention), or heard the sentences. Each participant in the listen condition heard all 8 of the target sentences; sentences were randomly ordered, and the listener heard one of 16 possible productions (8 speakers x 2 certainty levels) for that sentence. After completing these ratings, participants were asked to rank order the 8 sentences in order of certainty. **[Results]** Regardless of condition we find a stable order of the rated certainty of the utterances, and their rankings. Rank orderings, and overall ratings are provided in Table 1. We also looked at the relative differences between the ratings for each of the speakers between the read, and uncertainty emphasis productions and found that for 7/8 of the productions listeners on average rated the speakers as more uncertain when they had been instructed to mark uncertainty. We also found individual differences in the amount of certainty conveyed. Overall some speakers averaged more uncertainty between the types of productions (mean differences of 7-10 points on the certainty scale, max: 30-40 points), whereas others overall showed less difference, or even sounded more confident when they should have marked uncertainty (mean differences of 0.5-1, or an increase in certainty by 5, max: 3-11).

Phrase	Read-text Confidence	Listen (Read)	Listen (Uncertainty)	Read-text Rank	Listen Rank	Experiment 2 (mean confidence)
It could be a goose	36.994	37.706	36.283	7.125	7.063	25.163
It might be a robin	39.294	41.094	37.494	6.375	6.375	28.798
I think it's a falcon	49.918	48.918	50.919	5.688	5.644	46.458
It looks like a hummingbird	57.080	61.362	52.797	5.25	5.381	45.828
I'm pretty sure it's a woodpecker	65.476	68.110	62.842	4.063	4.319	68.577
I'm sure that it's a sparrow	84.220	87.510	80.930	2.688	2.919	80.300
It's a blackbird	86.777	88.864	84.689	2.625	2.525	91.765
It's definitely a canary	90.935	90.246	91.624	2.188	1.775	93.192

Table 1: Results from Experiment 1 and Experiment 2

Experiment 2 directly manipulated the likely degree of uncertainty. Participants (n=32) on Mechanical Turk were asked to label objects, rate their confidence in their label, and then selected one of 8 possible phrases they would use to describe that item to another person. We used line drawings from a classic perceptual recognition study (Biederman, 1987) in which parts of the images were occluded in a way that either did or did not preserve the underlying components (geons) of the image. We manipulated exposure duration, 120 or 220 ms, presented with a random dot mask (to avoid afterimage completion). Participants saw 17 images, and 3 control images (complete, simple pictures) randomly presented. **[Results]** Participants were more confident in their labels when items were presented for longer durations ($p < .02$), and when the deleted information preserved geons ($p < .001$). We also found a relationship between the confidence ratings and the kind of phrase participants used to communicate what they saw to another person (see: Table 1). Speakers' certainty in their own label, resulted in similar ordering of the phrases (7/8), as those determined by the listeners in Experiment 1.

We are now using utterances modeled on the ones we have tested in a confederate study in which a naïve participant and the confederate learn rare and common names for dogs and kitchen utensils together based on Ibarra, Runner & Tanenhaus (2017), who found that judgments of relative expertise modulated a director's use of a name versus a description following a shared learning task. The confederate will use expressions that indicate greater or lesser uncertainty. We predict that the confederate's degree of uncertainty will affect the (naïve) director's item-specific and category-based use of names in a subsequent referential communication task.

References: <http://amandapogue.github.io/docs/PogueTanenhausXPragReferences.pdf>