Informativity in adaptation: Supervised and unsupervised learning of linguistic cue distributions

Dave F. Kleinschmidt, Rajeev Raizada, & T. Florian Jaeger (University of Rochester)
dkleinschmidt@mail.bcs.rochester.edu

During language acquisition, infants learn the probabilistic mapping between observable cues and underlying linguistic structures. Recent work has shown that this learning continues into adulthood: comprehenders seem to rapidly adapt their linguistic expectations in response to experience (with, e.g., syntactic structures, phonetic categories, pragmatic interpretations, etc.). While there are many similarities between acquisition and adaptation in adults (cf. [1,2]), there is one major difference: acquisition is *unsupervised*, while adaptation is typically *supervised*. During adaptation, adult comprehenders have many sources of information that *label* each observed cue value, such as lexical or sentential contexts. It is well known that comprehenders take advantage of such top-down information during *processing*, integrating them with bottom-up cues to change the inferences they make (e.g. [3-6]). What is not known, however, is whether this labeling information is also taken advantage of during adaptation, or whether adaptation is purely a bottom-up learning process. To begin to address this question, we investigate supervised and unsupervised phonetic adaptation.

Exp1: Following [7], subjects heard spoken words, all members of /b/-/p/ minimal pairs (beach/peach, bees/peas, and beak/peak) synthesized with VOTs ranging from -20ms to 90ms . Between subjects, the bimodal VOT distribution (the /p/ and /b/ means) were either left normal (0ms for /b/ and 40ms for /p/, implied /b/-/p/ boundary at 20ms) or shifted up by 10ms (implied category boundary at 30ms), to assess adaptation to the shift.

On each trial, two pictures (target + distractor) were shown, and subjects were instructed to click on the picture that matched the target word (e.g., *peach*). There were two types of trials. On *unlabeled* trials, the distractor picture was minimal pair neighbor (e.g., a beach). On *labeled* trials, the distractor picture's onset was a minimal pair neighbor, but the rest of the word was unrelated (e.g., bees), thereby effectively *labeling* the acoustic percept. Subjects were randomly assigned to one of two conditions. In the *unsupervised* condition, all trials were unlabeled. In the *supervised* condition half were labeled and half unlabeled.

Learning (measured only on unlabeled trials) was nearly perfect: listeners' perceived category boundaries matched the category boundaries implied by distributions of VOTs they heard, in both the shifted and unshifted conditions. In the supervised condition, listeners consistently used the label information provided on labeled trials to guide their responses (98% consistent with label), showing integration of top-down cues during processing. However, *learning* was neither more complete nor faster in the supervised than in the unsupervised condition. This was replicated in **Exp2** and **Exp3** with 20ms & 30ms boundary shifts, resulting in poorer learning but still no benefit of supervision.

Conclusion: This suggests that the top-down label information provided by context (the distractor) is unavailable to learning processes that result in adaptation, and that these processes are primarily bottom-up. This is surprising given that labels are highly informative about the statistical structure of the environment, and should substantially reduce the difficulty of the learning problem. In ongoing work, we test whether labeling by lexical, rather than visual context can be integrated during learning. We discuss alternative explanations, including the possibility that adaptation might have been too quick to detect effects of supervision.

References: [1] Chang et al. (2006). *Psych. Rev.*; [2] McMurray et al. (2009). *Dev. Sci.*; [3] Ganong (1980). *JEP:HPP*; [4] McGurk & MacDonald (1976). *Nature*; [5] Tanenhaus et al. (1995) *Science*; [6] Trueswell et al. (1993). *JEP:LMC*; [7] Clayards et al. (2008) *Cognition*;