Questions:

Theoretical:

* What is context here (more than just ‘talker’)?
  + Was this really talker-specific, or rather gender (or something broader)?
  + Other context elements (same experimental setup)?
* What does it mean to adapt? Are new models built for each new talker, or are weights of existing models adjusted in a sort of MoG approach?
* What is the nature of the ‘processing cost’ associated with exposure to multiple talkers?
  + More evidence needed, why?
  + Should cost be reflected in RTs? In exposure, test, or both?
* Does modality of evidence factor into processing cost?
  + Might AV, LG, and Text-based recalibration operate in parallel?

Pragmatic:

* Did any of these effects reach ceiling?
* Compared to comparable experiments, what did RT look like (over course of session as well, esp. when talker ID was the same throughout)?
* Did effect atrophy over test block(s)?

Potential extensions:

1. Both talkers biased in same direction: do listeners adapt twice as fast?
   1. LMM forces separation of models/predictions because of opposite biasing; but theoretically talker/gender difference alone should facilitate this separation?
2. Different talkers of same gender (or even same talker dubbed to multiple videos)
   1. How specific are these models?
   2. Are models driven by bottom up similarity of cue distributions, or top-down talker/gender/context ID?
3. Use lexical endorsement task in both exposure and test, examine RT to determine where a processing cost may emerge
   1. If storing in or generating separate models, RT should be longer during exposure (or early in block). If cost is rather in which subset of priors to apply, RT should be longer during test (or later, trading off with ‘accuracy’ or endorsement of variant as a word).
4. Interleave test!

From Rachel:

1. E1 is interleaved exposure and test, really best to test in the case of trial by trial talker variability; pilot data confirm this works
   1. Is the cost real (verify LMM21)
2. E2 is within gender talker manipulations with same acoustics for each talker; this is the only way to test whether acoustics are truly conditioned on talker
3. E3 interleaves all of exposure and test
   1. Do talker differences emerge over time?

Indexical versus phonetic processing: ‘talker-specific’ adaptation is probably NOT linked to talker identification!

If I’m exposed to one talker, and then tested on a phonetically identical but indexically different talker, do I show shifted categorization?

What is the nature of the processing cost associated with multiple representations?

VS

What is the nature of the representations, and what drives maintenance of one versus multiple?