What allows listeners to rapidly adapt to an unfamiliar talker?

Maryann Tan, Rachel Sabatello, Iva Savic, & T Florian Jaeger

Stockholm University

UNIVERSITY of ROCHESTER

Research questions

Adaptivity in response to talkers with unexpected pronunciations is central to robust speech perception. Yet, much remains unknown about:

1. The expectations that listeners hold in the earliest moments of a new talker encounter. (see poster *1pSC21*)

2. **How these expectations change as more information about the talker is revealed.** (this poster)

Insight into these questions will illuminate our understanding of how listeners maintain robust and adaptive speech perception.

Experiment (N = 122): How quickly and how much do listeners adapt their categorisation behaviour?

Increased ecological validity
- Natural and human sounding
- VOT covaries with F0 and vowel duration
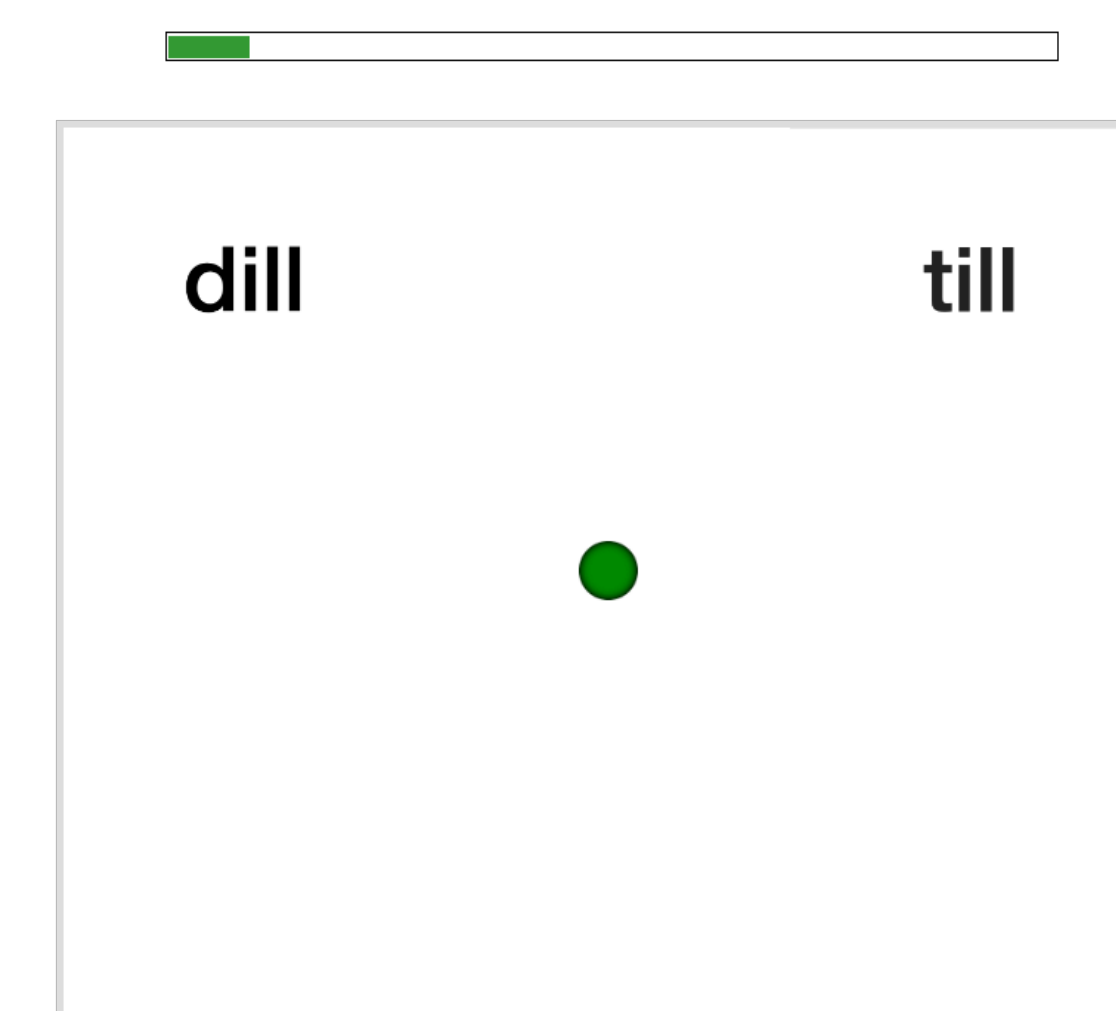- Unequal variances between categories

dill   till

Figure 1: Building on paradigm presented in Clayards et al. (2008). A recording is played and participants click on the word they heard.
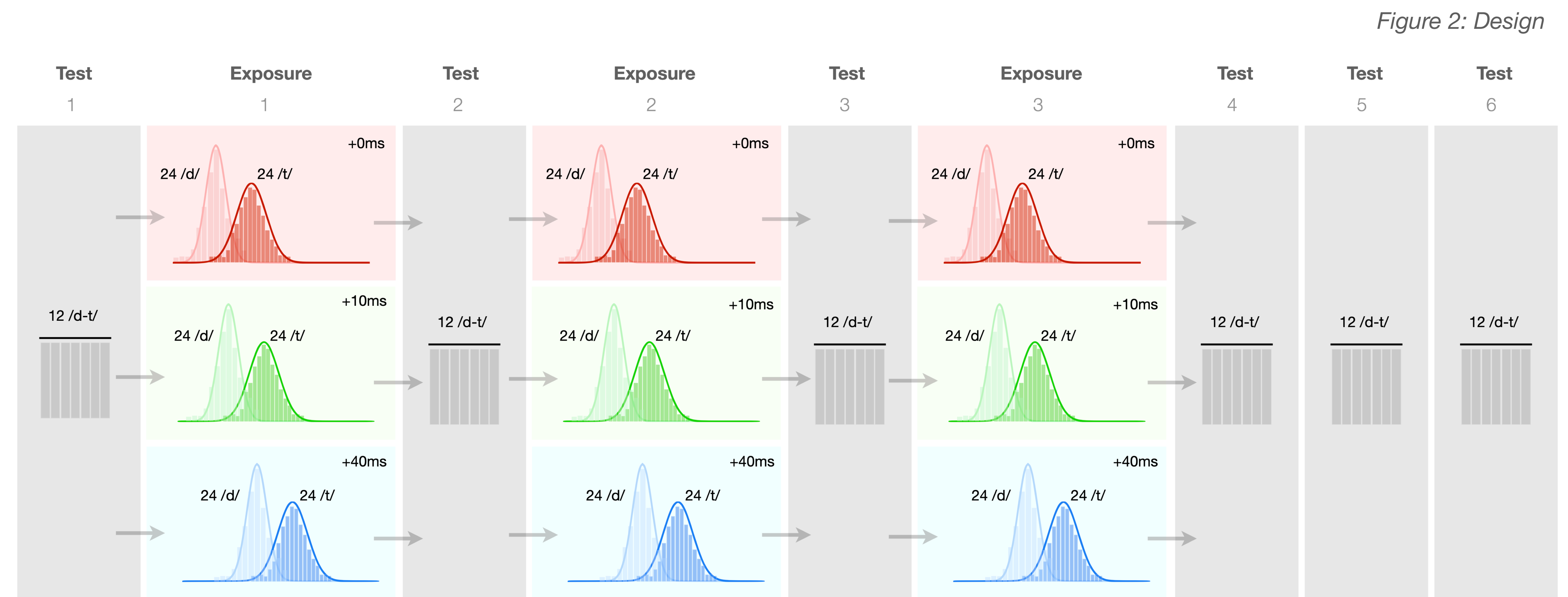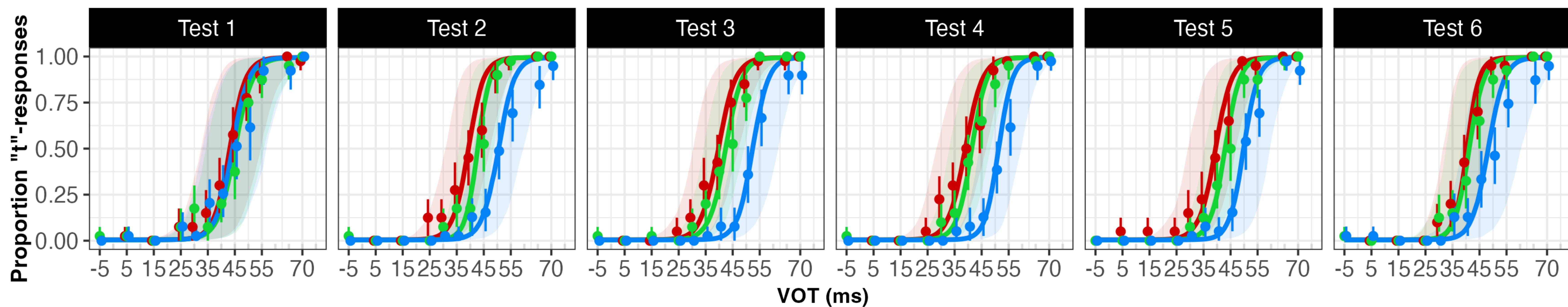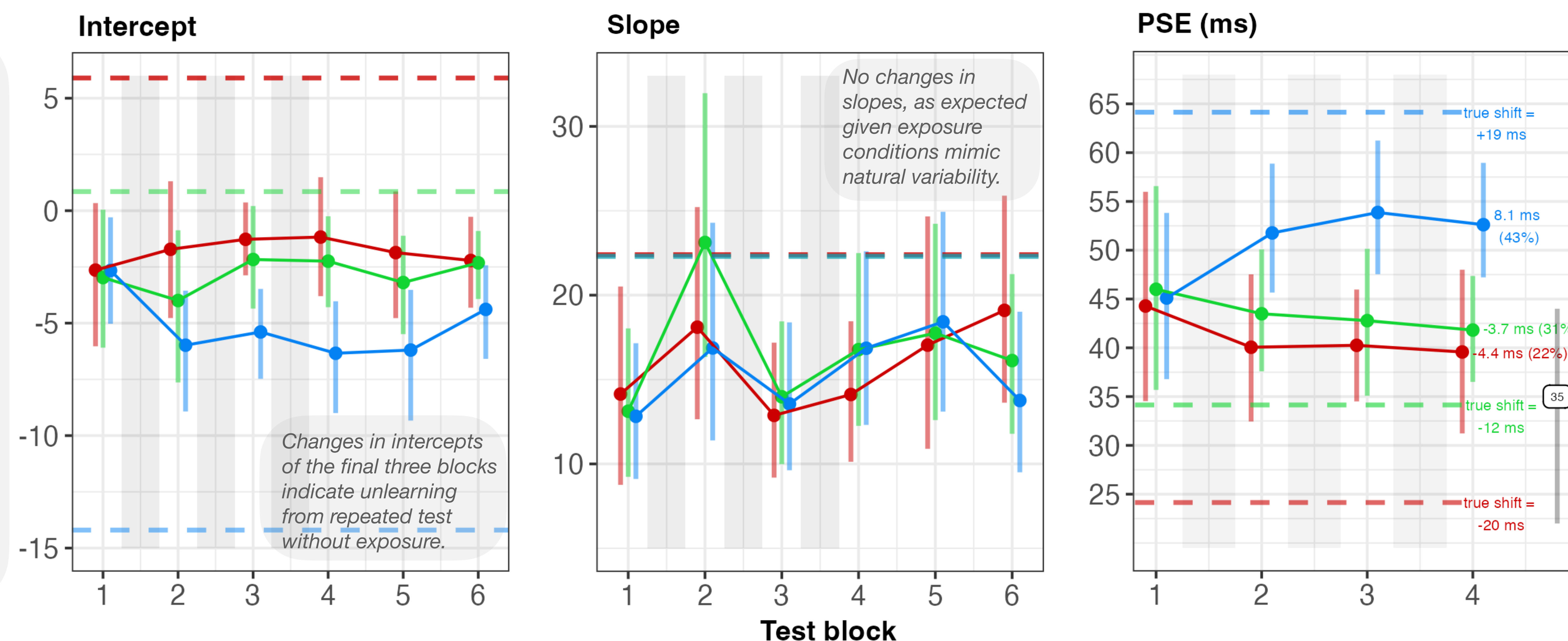

Figure 2: Design

Adaptation is rapid…

Figure 3: Fitted proportion of "t"-responses by psychometric lapsing mixed-effect regression.

Condition +0ms +10ms +40ms

*Figure 4: Position of category means–uncentred (**unfilled points**) and centred (**filled points**) relative to distribution of 92 L1-US English talkers' word-initial VOT-F0. **Top:** Non-parametric quantiles. **Bottom:** Bivariate Gaussian quantiles.*

… but strongly constrained

*Figure 5: Analysis of block-by-block changes in intercepts (**left**) and slopes (**middle**). Summary is of 8000 draws from the maximum a posteriori estimate of the mixed-effects logistic regression. Points are the mean of all posterior draws. Line ranges are 95% QIs. **Right:** Change in point-of-subjective equality (PSE) in the first 4 test blocks where maximal shift was observed. **Dashed lines:** predicted intercept, slope, and PSEs by the IOs of the respective conditions that have perfectly learned the exposure distributions (and are thus not guided by prior expectations).*

Changes in intercepts of the final three blocks indicate unlearning from repeated test without exposure.

No changes in slopes, as expected given exposure conditions mimic natural variability.

Predicted median PSE value from the distribution of 92 L1-US English talkers' centred VOT-F0 distributions of word-initial /d/ and /t/. Line range represents the 95% quantile interval.
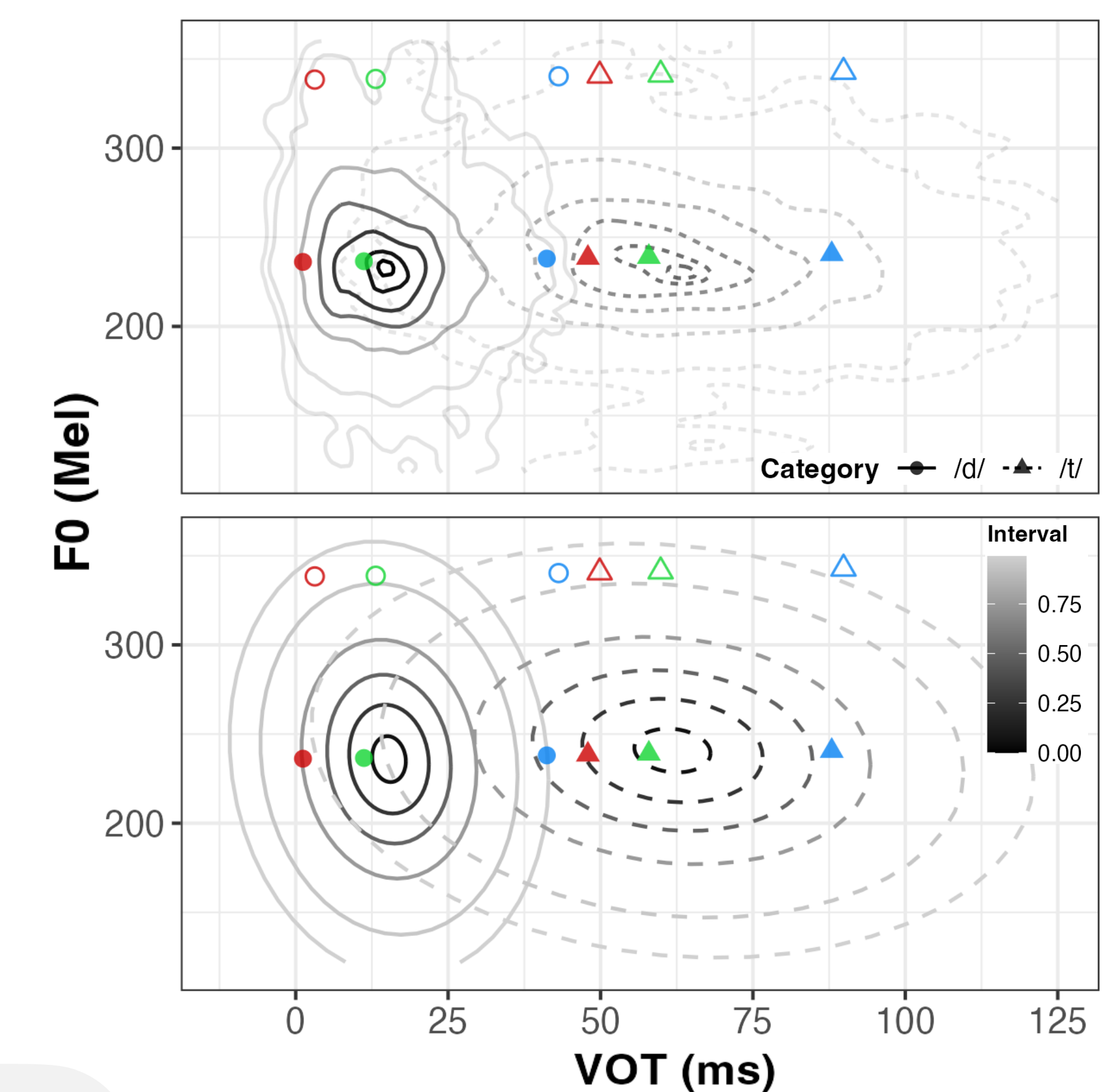
Take-home points

1. **With exposure to informative input**, listeners adapt their expectations proportionally to the statistics of the input.

2. Adaptation is rapid (< 48 trials) but highly constrained (to 43% of objective shift even after 144 trials).

3. Evidence for "shrinkage" (downward direction), unpredicted by the model but replicates that observed in Kleinschmidt & Jaeger (2016) & Kleinschmidt (2020).

4. Standard distributional learning accounts do not explain this result. Model selection/mixture accounts might.