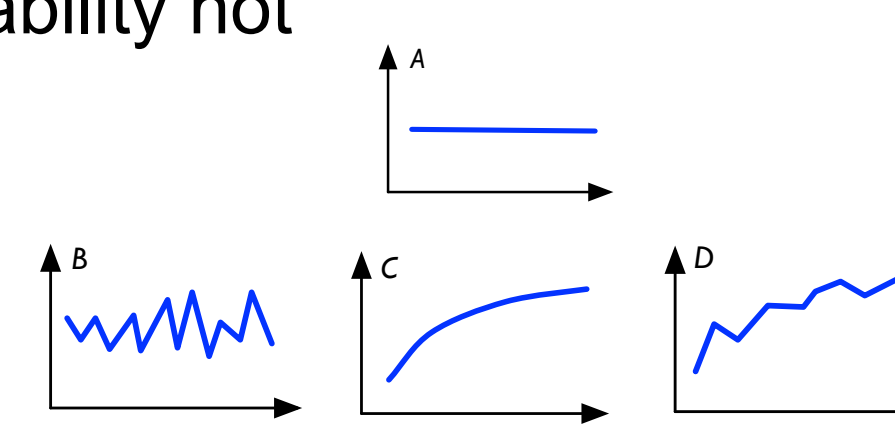



# Trajectories of phonetic variability in spontaneous speech on reality TV

Morgan Sonderegger (McGill University)

## Introduction

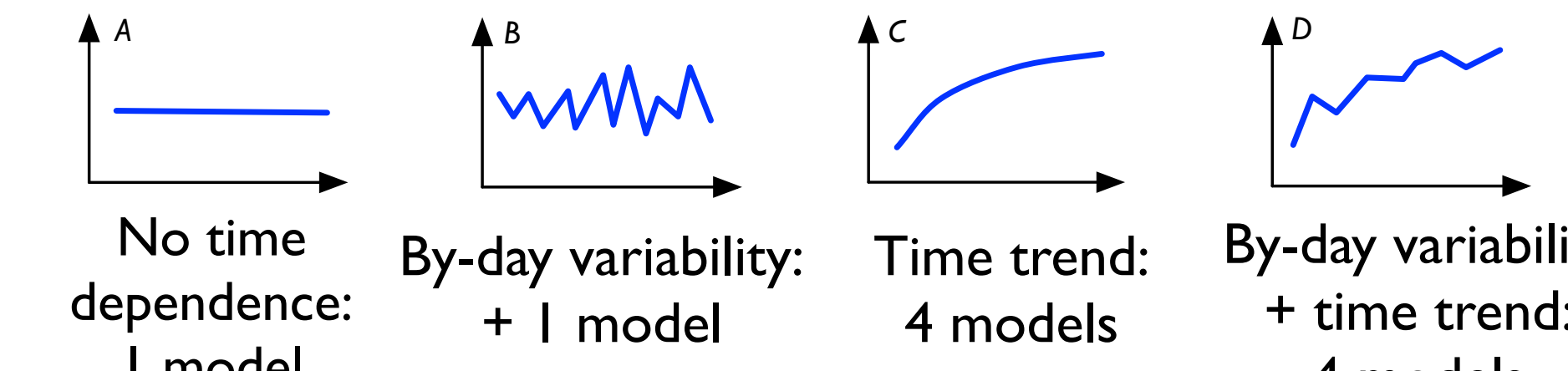
- Huge variability in spontaneous speech, many sources (coarticulation, speaking rate, social factors..)
- Less known: **variability over time**
  - Existing work: Imitation/accommodation (minutes-hours), panel studies (years) (Babel, 2011; Harrington et al., 2000; Nielsen, 2011; Prince, 1987; Sankoff & Blondeau, 2007; Shockley et al, 2004; Vallabha & Tuller, 2004)
- Our focus: **how does speech vary from day to day within individuals**, and over timescales in between (months)?
  - Very little known (Pisoni, 1980)
- Null hypothesis:** phonetic variability not time-dependent
- Alternatives:** **by-day variability** and/or **time trend**

- Motivations:
  - Test assumption that accommodation effects accumulate => could lead to sound change (Delvaux & Soquet, 2007)
  - If by-day variability exists, how much? (Nahkola & Saanilahti, 2004)
    - Relevance for panel studies

## Data

- Big Brother 9 UK**
  - Contestants spend 3 months in BB house
  - One/week voted off +sporadic additions
  - No outside contact: closed system, many opportunities for accommodation
  - Continuous surveillance
- Dataset** (builds on Sonderegger 2012)
  - 10.5 hours of semi-spontaneous speech
  - 12 speakers on show for >50 days
  - 630 clips (~1-10 min) from "diary room"
    - not conversations: give baseline beyond short-term shifts
- Phonetic variables**
  - automatic annotation ([github.com/mlml/autovot/](https://github.com/mlml/autovot/)) + manual correction
  - automatic annotation (FAVE: Rosenfelder et al, 2011) + manual correction ([github.com/mlml/plotmish/](https://github.com/mlml/plotmish/))

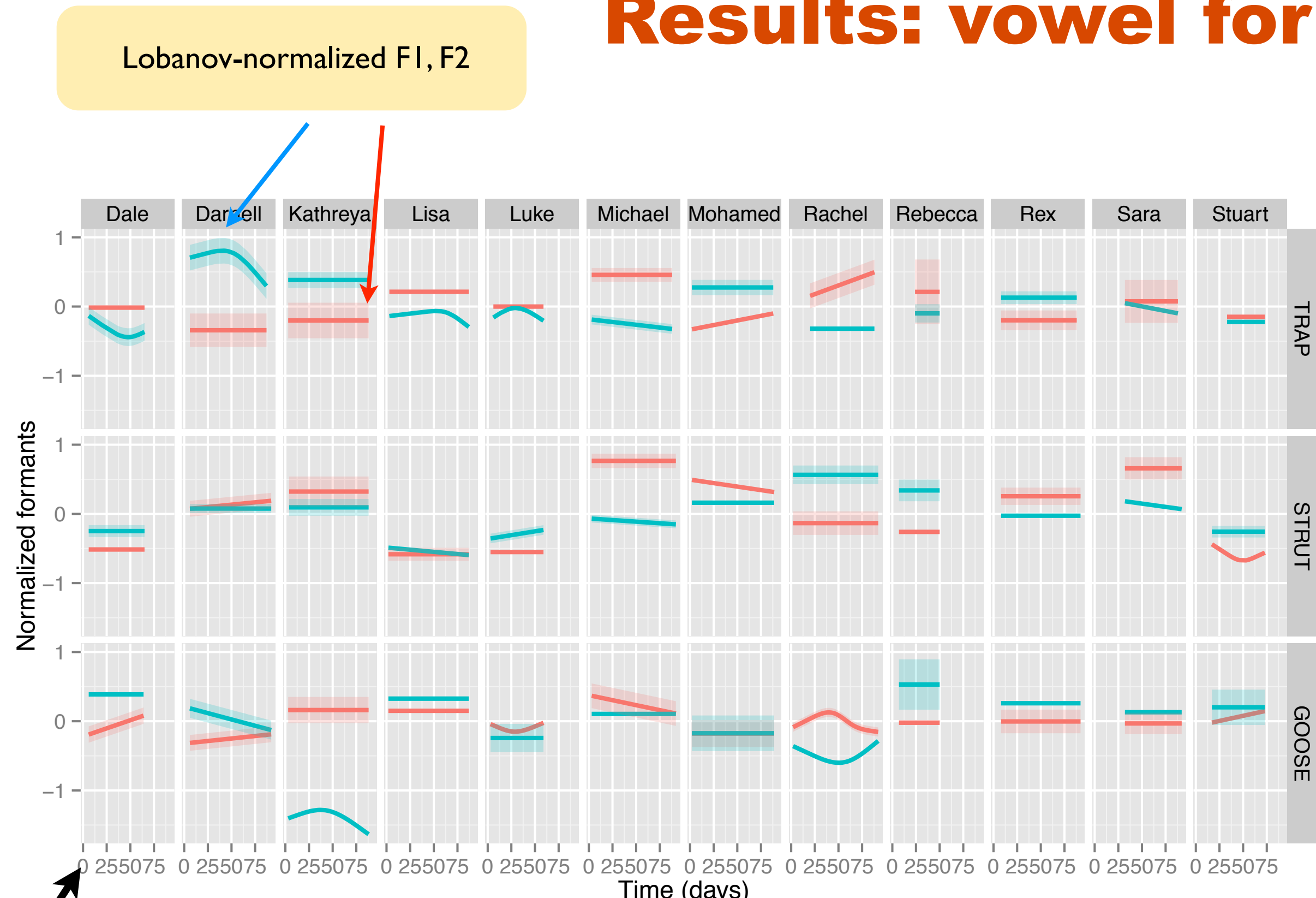
VOT (word-initial stops)	t/d deletion	F1, F2		
		STRUT	GOOSE	TRAP
voiceless: n=9.6k	n=12.8k	n=4.9k	n=2.9k	n=2.3k
voiced: n=12.8k				

## Methodology

- Don't know what type of time dependence to assume
- => for each phonetic variable, for each speaker, build range of regression models (mixed effects), including:
  - static factors (controls)
    - VOT:** speaking rate, place of artic., stress, following V height, ...
    - t/d deletion:** following C vs V, speaking rate morphological class...
    - F1/F2:** following segment manner, preceding segment type...
  - Time dependence
    - No time dependence: 1 model
    - By-day variability: + 1 model
    - Time trend: 4 models
    - By-day variability + time trend: 4 models
- Choose best model using AIC

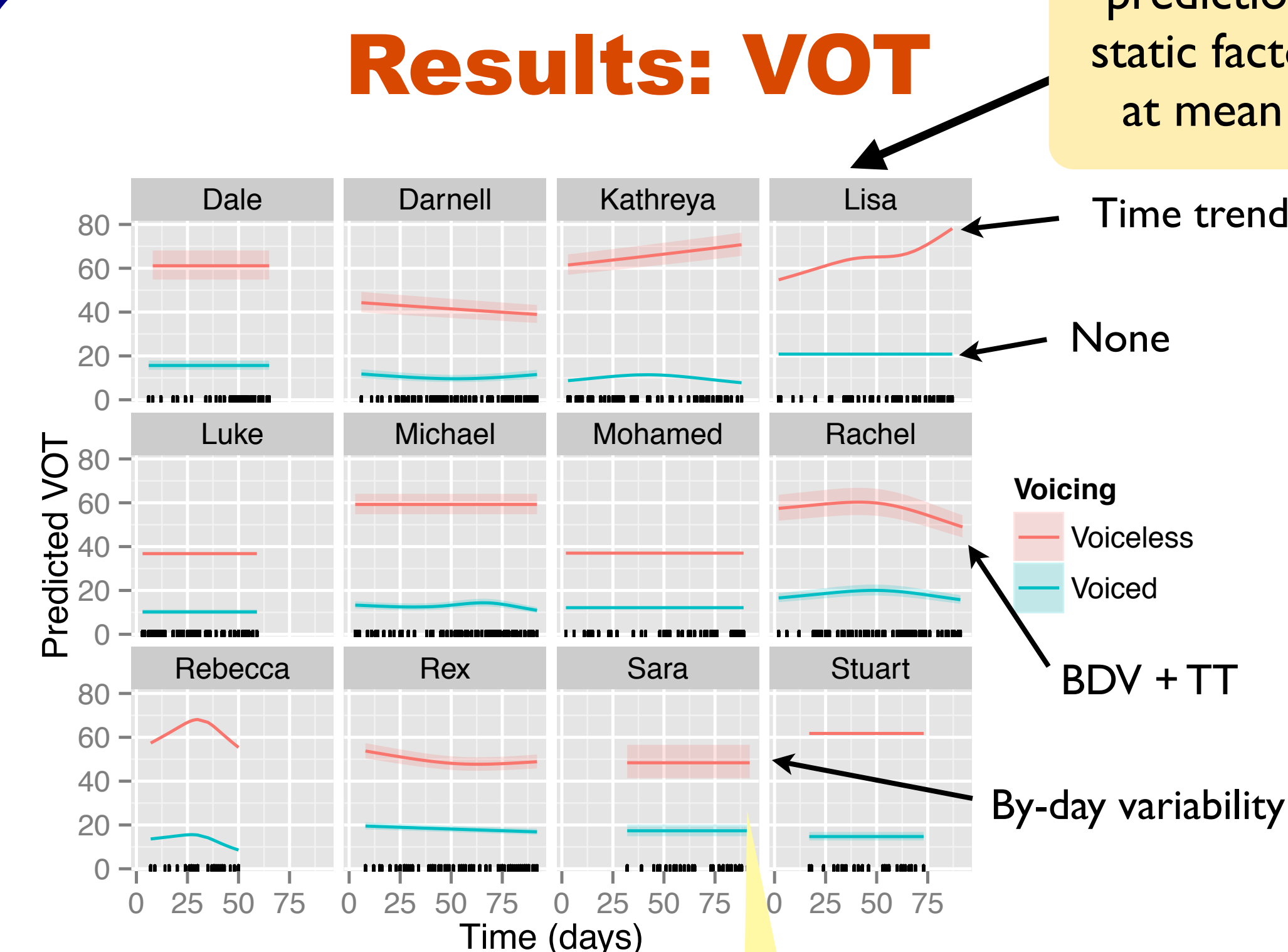
## Results: vowel formants

Lobanov-normalized F1, F2



- Variability over time: 94% of cases**
  - By-day variability: 86%
  - Time trend: 61%
- Magnitude of BDV (2σ):**
  - F1: 0.13-0.94
  - F2: 0.11-0.72
  - Comparable to other sources of variability:
    - Babel (2011) vowel imitation: most subjects <0.15
    - static factor effect sizes: F1: 0.18-0.26, F2: 0.06-1.04

## Results: VOT



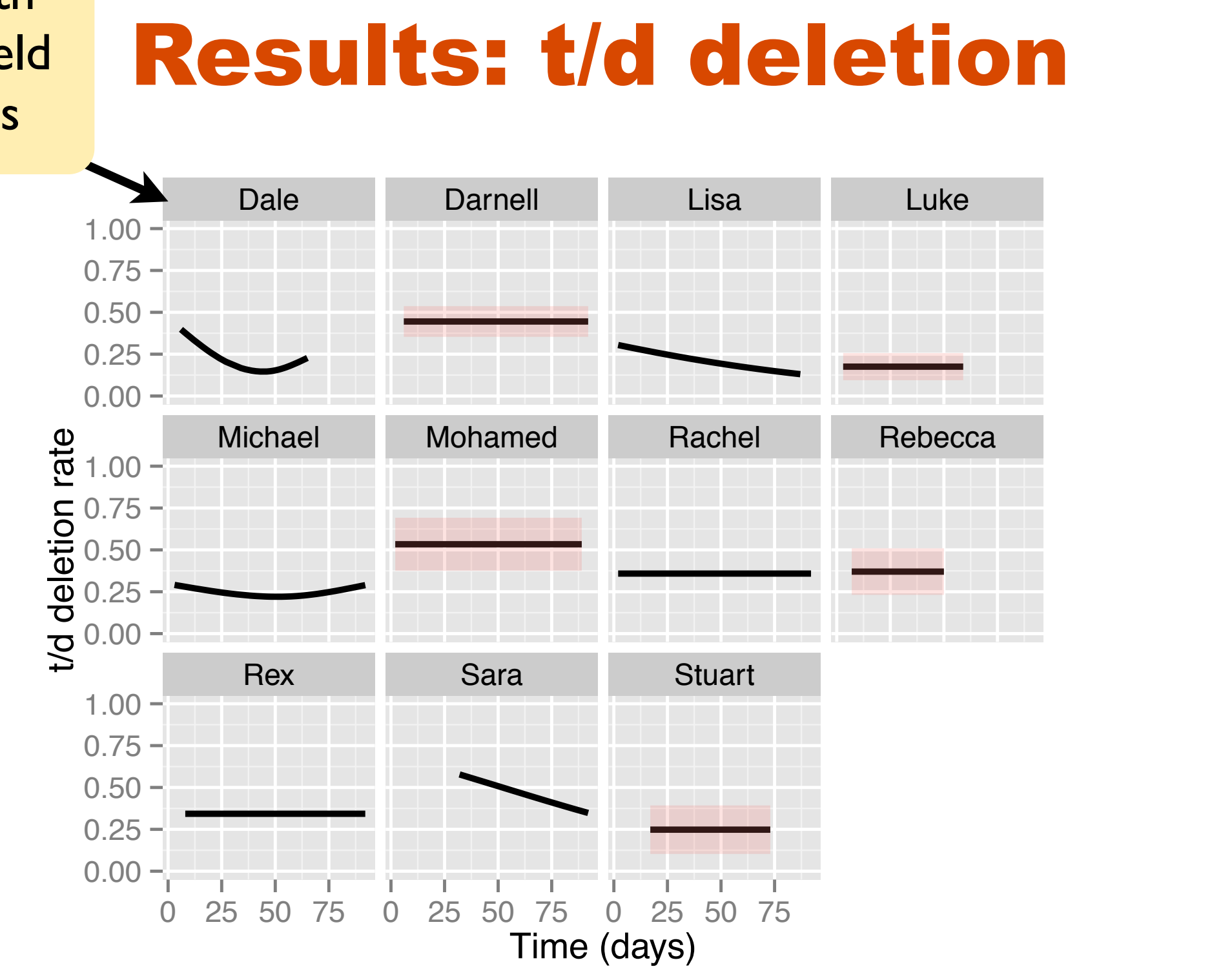
Best model's predictions with static factors held at mean values

- Similar time trajectories for voiced & voiceless VOT (contrast maintenance?)
- Variability over time: 79% of cases**
  - By-day variability: 63%
  - Time trends: 50%
- Magnitude of BDV (2σ): 14-37% voiceless, 17-41% voiced**
  - comparable to other sources of variability:
    - Nielsen (2011) VOT imitation: median 10%
    - static factor effect sizes (with  $p < 0.05$ ): voiceless: 5-20%, voiced: 3-64%

shading = by-day variability ( $\pm 1 \cdot \sigma$ ). (not errorbars)

i.e. for most variable speaker, VOT on a +1σ day is 41% more than on a -1σ day

## Results: t/d deletion



- Variability over time: 82% of cases**
  - By-day variability: 45%
  - Time trends: 36%
- Magnitude of BDV (2σ):** odds of t/d deletion increase by 1.18-1.37
  - small compared to other sources of variability:
    - Static factor effect sizes: 1.12-8.0
    - no imitation/accommodation studies (?)
- Differences from VOT, vowel formants
  - Continuous vs. categorical?

## Discussion

- Variability over time of phonetic variables in individuals is the norm: reject null hypothesis
- By-day variability is common
  - Relatively large magnitude: relevance for panel studies, forensic phonetics
  - Similar magnitude to imitation experiments
  - Consistent with hypothesis that imitation effects persist on a timescale of hours to days => could accumulate
- Time trends are (less) common, but not systematic: no clear convergence
  - Consistent with hypothesis that by-day fluctuations often don't accumulate => relative rarity of long-term change in individuals (Sankoff, 2005 et seq)

## Acknowledgements

- Annotation:** Liam Bassford, Thea Knowles, Maggie Labelle, Marion Coumel, Maria Nelson, Natalie Rothfels
- Programming:** Thea Knowles, Misha Schwartz
- Funding:** FRQSC #183356, SSHRC #430-2014-00018

## References (selected)

Babel, M. 2011. Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics* 39, 437-456. / Delvaux, V. & Soquet, A. The influence of ambient speech on adult speech productions through unintentional imitation. *Phonetica* 64, 145-173. Nielsen, K. 2011. Specificity and abstractness of VOT imitation. *Journal of Phonetics* 39, 132-142. / Rosenfelder, I., Frushwald, J., Evanimi, K., & Yuan, J. (2011). FAVE Program Suite [Computer program]. / Sankoff, G. & Blondeau, H. 2007. Language change across the lifespan: /r/ in Montreal French. *Language* 83, 560-588. / Sonderegger 2012. Phonetic and phonological dynamics on reality television. Ph.D. Thesis, University of Chicago.