Longitudinal Phonetic Variation in a Closed System

Max Bane\textsuperscript{a}
(bane@uchicago.edu)
In collaboration with: Morgan Sonderegger,\textsuperscript{b}
Peter Graff.\textsuperscript{c}

\textsuperscript{a}Linguistics, U of C \hspace{1cm} \textsuperscript{b}Computer Science, U of C \hspace{1cm} \textsuperscript{c}Linguistics, MIT

November 20, 2009
CAS Workshop on Language, Cognition, and Computation
Variation is pervasive in language.

- Nondeterministic realization of linguistic form.

At many levels:

- Syntactic: *give it to her* ~ *give her it*.
- Morpho-lexical: *pérfume* ~ *perfúme*, *eating* ~ *eatin’*.
- Phonological: *button* [ˈpʌŋt] ~ [ˈpʌtʃt].
- Fine-grained phonetic: voice onset time, pitch usage, speech rate, formant location and trajectories, etc.
Variation is pervasive in language.

- Nondeterministic realization of linguistic form.

- At many levels:
  - Syntactic: *give it to her* ∼ *give her it*.
  - Morpho-lexical: *pérfume* ∼ *perfúme*, *eating* ∼ *eatin’*.
  - Phonological: *button* [ʰpʰʌɳ] ∼ [ɾpʰʌɳ].
  - Fine-grained phonetic: voice onset time, pitch usage, speech rate, formant location and trajectories, etc.

- Often structured non-randomly, with grammatical, functional, social correlates.

- Between and within individuals.
The dynamics of variation

- **Stable variation**
  - Variation may persist over the long term (Brunberg 2002).
  - E.g., hundreds of years of variation between -*ing* ~ -*in’*.

- **Unstable variation**
  - One of the variants can “win out” over the others.
  - Structure of variation can shift over time.
The dynamics of variation

- Stable variation
  - Variation may persist over the long term (Brunberg 2002).
  - E.g., hundreds of years of variation between *-ing* vs *-in’*.

- Unstable variation
  - One of the variants can “win out” over the others.
  - Structure of variation can shift over time.

- How do these (non)dynamics arise?
- What courses can they follow?
Artificial dynamics: manipulating variation in the lab

- **General paradigm:**
  - Participants speak; measure some phonetic variables.
  - Participants listen to speech in which those variables have been manipulated.
  - Participants speak; measure same variables.
  - Assess whether participants’ variables have changed after exposure, compared to control.

- **Phonetic imitation** (also convergence, accommodation)
Artificial dynamics: manipulating variation in the lab

- Goldinger 1998, Shockley et al 2004, “shadowing” studies:
  - Participants produce words more similar to what they’ve recently heard, according to gestalt AXB ratings.
- Nielsen 2007, 2008:
  - English speakers unconsciously imitate manipulated voice onset times, but not if doing so would endanger a contrast.
  - Japanese speakers imitate exaggerated high-vowel devoicing/reduction.
  - Imitation is generalized to novel environments.
Natural dynamics: short term convergence in conversation

- Pardo 2006:
  - Participants cooperating in a shared task converge in their pronunciations of task-salient words, according to gestalt AXB ratings.

- Conversational convergence of many variables (e.g., Giles et al 1987):
  - Mean vocal intensity.
  - $F_0$ and pitch range, pitch contours.
  - Speaking rate, rate and duration of pauses.
  - Other non-acoustic variables (lexical usage, etc.)
Questions

“These results suggest that... phonetic convergence... can form the basis for phenomena such as accent change and dialect formation.” (Pardo 2006)

Is there a link between laboratory studies of imitation and community-level linguistic change?

How long lasting, persistent are convergence effects?

How do they interact with, or depend on, the social context?
An Orwellian experiment

- An ideal, long-term convergence study, expensive and possibly unethical:
  - Force subjects to live together for months with no outside contact: linguistically and socially closed system.
  - 24-hour audio/video surveillance:
    - Continuous longitudinal phonetic data.
    - Continuous social interaction data.
  - Occasionally perturb the sociolinguistic system, adding or removing subjects, to observe effects.
Leave it to capitalism: a natural Orwellian experiment

- **Big Brother**: reality-TV program
  - Originally created by Endomol, Dutch entertainment company, for Dutch audiences.
  - Later imported to UK, US, where it’s seen the most success.
Big Brother: premise

- Sixteen contestants (“housemates”) live together in a house for 3 months (92 days).
- No outside contact: closed system
- Cameras and microphones in every room: continuous surveillance.
  - During season, 24-hour streams are available on a dedicated television channel (in UK), and online.
  - Additionally, an hour-long episode is produced and aired each day.
Big Brother: premise

- Housemates must complete various cooperative/competitive tasks each week.
- Each week, a housemate is voted off of the show (periodic perturbations).
  - Housemates vote to nominate fellow housemates as eligible for eviction.
  - Public vote by viewers selects one of the nominees to evict.
- Final week: five housemates remain; viewers select a winner among them by call-in vote.
  - Winner receives a cash prize (£100,000 in UK).
Our data come from season 9 of the UK version, aired June 5th to September 5th, 2008.

Two classes of data:
- Phonetic measurements taken at regular intervals.
- Longitudinal measures of degree of social interaction between pairs of housemates.
Phonetic data

- Ideally, would take measurements continuously from 24 hour live feed.
  - Impractically large corpus, and currently unavailable.
- Instead: take measurements from “diary room” clips.
  - Each episode, HMs may enter the diary room, where they talk to “Big Brother” alone.
  - Relatively controlled environment acoustically. Little noise, one speaker, consistent.
  - Speech is still spontaneous.
  - Marginally conversational—responding to queries from Big Brother.
- Example clip.
For pilot study: concentrate on a subset of the housemates.
- Rachel (eventual winner), Michael, Rex, Lisa.
- Each is present for the whole season.

For each of these four HMs, 12–18 diary room clips collected, approximately evenly distributed throughout season.

<table>
<thead>
<tr>
<th>HM</th>
<th>Clips</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rachel</td>
<td>18</td>
</tr>
<tr>
<td>Michael</td>
<td>14</td>
</tr>
<tr>
<td>Rex</td>
<td>12</td>
</tr>
<tr>
<td>Lisa</td>
<td>12</td>
</tr>
</tbody>
</table>
Phonetic data

- Phonetic variables measured in the diary room clips:
  - Means and standard deviations of pitch within “spurts” of connected speech.
    - Semi-automatic spurt detection.
  - Speaking rate (syllables per second) within a spurt.
    - Automatic syllable detection (de Jong and Wempe 2009).
  - Voice onset time (VOT) of voiceless, word-initial plosives (p, t, k).
Social interaction data

- Again, would ideally be based on 24-hour surveillance video.
  - Laborious, how to code?
- The Sun tabloid:
  - During season 9, staff writers continuously monitored the surveillance feeds and blogged about what was happening, minute by minute.
Again, would ideally be based on 24-hour surveillance video.

Laborious, how to code?

The Sun tabloid:

During season 9, staff writers continuously monitored the surveillance feeds and blogged about what was happening, minute by minute.

Example:

Day 24

23.41: Kat is talking to Lisa, Mario and Rachel in the garden. HMs are comforting her. Lisa tells her people are jealous of her. She says: "People are jealous because ..."

23.49: Jen is talking to Luke and Dale in the luxury bedroom. Jen tells Dale: "You’ve got a right face on ya, you have."
“Heaven and Hell” split

- On day 37, housemates split into two groups:
  - Heaven
    - Luxury seats, bedrooms, bathrooms, swimming pool access, unrestricted smoking rights.
  - Hell
    - Crummy seats, bedrooms, bathrooms, cold outdoor showers, no swimming pool, restricted smoking rights, sole access to kitchen (must cook for Heaven).

- Removed on day 65.
- Relevant to convergence.
Phonetic Data
Collecting the phonetic data

- Each clip transcribed orthographically, annotated for nonspeech and non-target speech.
- Rough spurt boundaries marked.
- Then automatically adjust/subdivide rough boundaries to find good spurts.
  - No more than 0.4 secs of continuous “silence” w/in spurt.
  - Similar to de Jong and Wempe (2009).
Spurts

Yields 1,148 spurts:

<table>
<thead>
<tr>
<th>Housemate</th>
<th>Spurts</th>
</tr>
</thead>
<tbody>
<tr>
<td>HM</td>
<td></td>
</tr>
<tr>
<td>Rachel</td>
<td>314</td>
</tr>
<tr>
<td>Michael</td>
<td>325</td>
</tr>
<tr>
<td>Rex</td>
<td>272</td>
</tr>
<tr>
<td>Lisa</td>
<td>237</td>
</tr>
</tbody>
</table>

![Spurt duration chart]

Seconds

Housemate

Lisa | Michael | Rachel | Rex
Within each spurt:
- Mean pitch
- Standard deviation of pitch
Speech rate

Within a spurt:
- Average syllables per second: number of syllables divided by length of spurt.

Syllable nuclei detected automatically with algorithm by de Jong and Wempe (2009).
- Tends to underestimate number of syllables. May miss unstressed syllables, quiet speech, whisper, etc.
- Okay for comparison as long as consistent.
Weak correlation with mean pitch: Pearson’s $\rho = 0.22$, Spearman’s rank $\rho = 0.24$, Kendall’s rank $\tau = 0.17$. 
Voice onset time (VOT)

- **Duration of aspiration:**
  - Time in milliseconds between onset of high-frequency aperiodicity of stop burst and onset of voicing in following vowel.
- **Manually measured on all voiceless, word-initial stops in a clip.**
  - Coded for stop (p, t, k) and host word.
  - Stop-initial clusters (tr-, pr-, kl-, etc.) included, but coded separately.
  - Also coded for reduced following nucleus (t’her, t’him), and preceding fricative (this token).
Balanced sample of VOTs

<table>
<thead>
<tr>
<th>HM</th>
<th>No. of VOTs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rachel</td>
<td>193</td>
</tr>
<tr>
<td>Michael</td>
<td>223</td>
</tr>
<tr>
<td>Rex</td>
<td>197</td>
</tr>
<tr>
<td>Lisa</td>
<td>207</td>
</tr>
</tbody>
</table>

Number of VOT measurements

- **Lisa**
  - P
  - T
  - K

- **Michael**
  - P
  - T
  - K

- **Rachel**
  - P
  - T
  - K

- **Rex**
  - P
  - T
  - K
Distribution of VOT

<table>
<thead>
<tr>
<th></th>
<th>Lisa</th>
<th>Michael</th>
</tr>
</thead>
<tbody>
<tr>
<td>P 50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T 100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K 150</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P 200</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Rachel</th>
<th>Rex</th>
</tr>
</thead>
<tbody>
<tr>
<td>P 50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T 100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K 150</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P 200</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

VOT (msecs)
Known influences of VOT

- In spontaneous speech, VOT is heavily influenced by (Yao 2009):
  - Place of articulation \((p < t < k)\)
  - Speech rate (faster \(\Rightarrow\) shorter VOT)
  - Word frequency (higher \(\Rightarrow\) shorter VOT)

- Must control for these factors.
- For speech rate, two possibilities:
  - Include as a linear predictor of VOT.
  - Normalize VOTs by speech rate.
Two methods considered:

- **VOTSyls**: Normalize by spurt-level speaking rate:
  
  \[
  \text{VOT (msec)} \times \text{Syllables per second in spurt (syls/msec)}
  \]
  
  - Units are fraction of an average syllable in the spurt.

- **VOTSegs**: Normalize by word-level speaking rate:
  
  \[
  \text{VOT (msec)} \times \text{Segments per second in word (segs/msec)}
  \]
  
  - Units are fraction of an average segment in the word.
  - Based on number of segments in CELEX transcription of host word.
Distribution of normalized VOT

VOTSegs (avg. segments)

1
2
3
P T K
Lisa
P T K
Michael
Rachel
P T K
Rex

VOTSyls (avg. syllables)

0.5
1.0
1.5
P T K
Lisa
P T K
Michael
Rachel
0.5
1.0
1.5
P T K
Rex
Phonetic Trends
General methodology: Change

- Detecting significant change:
  - Linear mixed models
  - Phonetic variable as response:
    - MeanPitch, log(MeanPitch), StdDevPitch, SpurtSylsPerSec, VOT, VOTSyls, VOTSegs
  - Fixed effect predictors include Episode and HM plus controls appropriate to response. Categorical variables centered and scaled.
    - Possible nonlinear Episode term (linear tail-restricted cubic spline).
  - Random effect of Word for VOT.
- Significant effect of Episode indicates change over time.
General methodology: Convergence

- Detecting convergence:
  - Contrast coding
  - Fixed effect contrast regressors
    - RA\text{vMI}: Rachel vs Michael
    - RE\text{vLI}: Rex vs Lisa
    - RA\text{vLI}: Rachel vs Lisa
  - Allows us to test for trends in relative differences between pairs of HMs.
  - Must choose three pairs.

- Significant contrast variable means HMs are different from each other.

- Significant interaction with Episode indicates possible convergence/divergence
  - Sign of estimated coefficient gives likely direction
Pitch
Trends in mean pitch: by episode

![Graph showing trends in mean pitch for different characters over episodes.](image-url)
Trends in mean pitch: by fifth
Trends in mean pitch: predictors

- **Fixed effects:**
  - $rcs(Episode, 5)$
    - Nonlinear restricted cubic spline of Episode, with 5 inflection points.
  - HM
  - Split
    - Whether episode is before, during, or after Heaven/Hell split.

- **Interactions:**
  - $rcs(Episode, 5) \times HM$
  - Split $\times$ HM

- **Random effect:**
  - HM
Trends in mean pitch: Best regression

- **Response:**
  - \( \log(\text{MeanPitch}) \)

- **Model:**
  - \( \log(\text{MeanPitch}) \sim rcs(\text{Episode}, 5) \times \text{HM} + \text{Split} \times \text{HM} + (1|\text{HM}) \)

- **Achieves** \( R^2 = 0.497 \).
### ANOVA table

<table>
<thead>
<tr>
<th>Factor</th>
<th>Df</th>
<th>F-value</th>
<th>p (mcmc)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rcs(Episode, 5)</td>
<td>4</td>
<td>8.7227</td>
<td>0.00000</td>
</tr>
<tr>
<td>HM</td>
<td>3</td>
<td>69.3979</td>
<td>0.00000</td>
</tr>
<tr>
<td>Split</td>
<td>2</td>
<td>1.0817</td>
<td>0.33937</td>
</tr>
<tr>
<td>rcs(Episode, 5):HM</td>
<td>12</td>
<td>8.1098</td>
<td>0.00000</td>
</tr>
<tr>
<td>HM:Split</td>
<td>6</td>
<td>10.1189</td>
<td>0.00000</td>
</tr>
</tbody>
</table>
Some selected fixed predictors

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Estimate</th>
<th>t value</th>
<th>p value (MCMC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rcs(Episode, 5)</td>
<td>-0.0001237</td>
<td>-0.06</td>
<td>0.95</td>
</tr>
<tr>
<td>rcs(Episode, 5)'</td>
<td>-0.0259620</td>
<td>-1.13</td>
<td>0.25</td>
</tr>
<tr>
<td>rcs(Episode, 5)''</td>
<td>0.1113481</td>
<td>1.77</td>
<td>0.07</td>
</tr>
<tr>
<td>rcs(Episode, 5)'''</td>
<td>-0.2562371</td>
<td>-2.59</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Splitc1v2</td>
<td>0.2597354</td>
<td>2.75</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Splitc2v3</td>
<td>-0.5672421</td>
<td>-5.35</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>rcs(Episode, 5):HMRAvMI</td>
<td>0.0144272</td>
<td>2.19</td>
<td>0.02</td>
</tr>
<tr>
<td>rcs(Episode, 5)''':HMREvLI</td>
<td>-1.6486158</td>
<td>-3.31</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>rcs(Episode, 5)''':HMRAvLI</td>
<td>1.6144417</td>
<td>3.71</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>HMREvLI:Splitc1v2</td>
<td>1.7850555</td>
<td>5.08</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>HMRAvMI:Splitc2v3</td>
<td>1.6594709</td>
<td>5.77</td>
<td>&lt;0.01</td>
</tr>
</tbody>
</table>
**Smoothed fitted values**

![Graph showing fitted log(MeanPitch) for different characters over episodes.](image-url)

- **Episode**
- **Fitted log(MeanPitch)**
  - Rachel
  - Rex
  - Michael
  - Lisa
Trends in mean pitch: conclusions

- Definite nonlinear change over time
- Possible convergence among Rex, Michael, Lisa
  - Rachel goes off on her own.
  - No obvious convergence in beginning, when there are many HMs
  - Apparent convergence begins as number of other HMs decreases
    - Toward the end it’s just five HMs in the house
Trends in pitch variance: by episode
Trends in pitch variance: by fifth of season

![Graph showing trends in pitch variance by fifth of season for different individuals]
Trends in pitch variance: predictors

- **Fixed effects:**
  - `rcs(Episode, 3)`
  - Nonlinear restricted cubic spline of Episode, with 3 inflection points.
  - HM
  - Split

- **Interactions:**
  - `rcs(Episode, 3) × HM × Split`

- **Random effect:**
  - HM
Response:
- StdDevPitch

Model:
- \( \text{StdDevPitch} \sim \text{rcs(Episode, 3)} \times \text{HM} \times \text{Split} + (1|\text{HM}) \)

Achieves \( R^2 = 0.472 \).
Trends in pitch variance: smoothed fitted values

![Graph showing trends in pitch variance for different characters.](image-url)
Trends in pitch variance: conclusions

- Definite individual differences in pitch variance
  - Rex > Rachel > \{Lisa, Michael\}.
- No clear convergence.
- Rex shows some longitudinal dynamics: fall/rise.
Speech rate
## Trends in speech rate: by episode

<table>
<thead>
<tr>
<th>Episode</th>
<th>Syllables per sec in a spurt</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 20 40 60 80</td>
</tr>
<tr>
<td>Lisa</td>
<td>0 2 4 6 8</td>
</tr>
<tr>
<td>Michael</td>
<td>0 2 4 6 8</td>
</tr>
<tr>
<td>Rex</td>
<td>0 2 4 6 8</td>
</tr>
<tr>
<td>Lisa</td>
<td>0 2 4 6 8</td>
</tr>
<tr>
<td>Michael</td>
<td>0 2 4 6 8</td>
</tr>
<tr>
<td>Rex</td>
<td>0 2 4 6 8</td>
</tr>
</tbody>
</table>

The diagram above illustrates the trends in speech rate for each episode, showing the number of syllables per second in a spurt for characters Lisa, Michael, Rex, and Rachel.
Trends in speech rate: by fifth of season

![Chart showing speech rate trends by fifth of the season for different individuals.](chart.png)
Trends in speech rate: best regression

Model:
- SpurtSylsPerSec $\sim$ rcs(Episode, 3) $\times$ HM + (1|HM)
- Achieves $R^2 = 0.135$. 
Trends in speech rate: ANOVA table

<table>
<thead>
<tr>
<th></th>
<th>Df</th>
<th>F-value</th>
<th>p (mcmc)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rcs(Episode, 3)</td>
<td>2</td>
<td>7.2685</td>
<td>0.00073027</td>
</tr>
<tr>
<td>HM</td>
<td>3</td>
<td>10.0405</td>
<td>0.00000157</td>
</tr>
<tr>
<td>rcs(Episode, 3):HM</td>
<td>6</td>
<td>6.3936</td>
<td>0.00000123</td>
</tr>
</tbody>
</table>
## Trends in speech rate: Fixed predictors

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Estimate</th>
<th>t value</th>
<th>p value (MCMC)</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>rcs(Episode, 3)</code></td>
<td>0.0094973</td>
<td>3.46</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td><code>rcs(Episode, 3)'</code></td>
<td>-0.0091205</td>
<td>-2.60</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td><code>HMRAvMI</code></td>
<td>0.5244286</td>
<td>1.75</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td><code>HMREvLI</code></td>
<td>-0.5371566</td>
<td>-1.46</td>
<td>0.14</td>
</tr>
<tr>
<td><code>HMRAvLI</code></td>
<td>1.1074772</td>
<td>2.76</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td><code>rcs(Episode, 3):HMRAvMI</code></td>
<td>0.0001595</td>
<td>0.02</td>
<td>0.98</td>
</tr>
<tr>
<td><code>rcs(Episode, 3)':HMRAvMI</code></td>
<td>-0.0258581</td>
<td>-2.33</td>
<td>0.02</td>
</tr>
<tr>
<td><code>rcs(Episode, 3):HMREvLI</code></td>
<td>0.0029338</td>
<td>0.26</td>
<td>0.79</td>
</tr>
<tr>
<td><code>rcs(Episode, 3)':HMREvLI</code></td>
<td>0.0068192</td>
<td>0.50</td>
<td>0.62</td>
</tr>
<tr>
<td><code>rcs(Episode, 3):HMRAvLI</code></td>
<td>0.0093294</td>
<td>0.79</td>
<td>0.43</td>
</tr>
<tr>
<td><code>rcs(Episode, 3)':HMRAvLI</code></td>
<td>-0.0295398</td>
<td>-1.96</td>
<td>0.05</td>
</tr>
</tbody>
</table>
Trends in speech rate: smoothed fitted values

Episode
Fitted Spurt Syls Per Sec
Rachel
Rex
Michael
Lisa
Trends in speech rate: conclusions

- Major individual differences and nonlinear change over time
  - Michael’s late-season plunge in speech rate.
- Possible convergence between Rex, Michael, Lisa, but inconclusive.
  - Interesting that again Rachel is off on her own.
Voice Onset Time
Trends in VOT

- Reminder: controlling for speech rate.
  - Can include speech rate as a linear predictor for VOT
  - Or can normalize by speech rate
    - VOTSyls: spurt-level speech rate
    - VOTSegs: word-level speech rate

- Here, just look at model of VOTSyls.
Trends in VOT: by episode

![Graph showing VOT trends by episode for different names including Rex, Lisa, Michael, and Rachel. The graph displays VOT values in milliseconds (msecs) against episode numbers, with lines indicating trends for each name.]
Trends in VOTSyls: by episode
Trends in VOTSyls: predictors

- **Fixed effects:**
  - `rcs(Episode, 3)`
  - HM
  - `log(CelexFreq)`
    - Logarithm of word frequency in CELEX (scaled, centered)
  - Phone
    - Which phoneme: p, t, k.
  - CCluster
    - Consonant cluster?
  - NoNuc
    - Reduced following nucleus?

- **Interactions:**
  - `rcs(Episode, 3) × HM`
  - CCluster × Phone

- **Random effect:**
  - Word
Trends in VOTSyls: best regression

- \( \text{VOTSyls} \sim \text{rcs(Episode, 3)} \times \text{HM} + \log(\text{CelexFreq}) + \text{NoNuc} + \text{CCluster} \times \text{Phone} + (1|\text{Word}) \)
- Achieves \( R^2 = 0.328 \).
## Trends in VOTSyls: ANOVA table

<table>
<thead>
<tr>
<th>Effect</th>
<th>Df</th>
<th>F-value</th>
<th>p (mcmc)</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>rcs(Episode, 3)</code></td>
<td>2</td>
<td>2.7028</td>
<td>0.0676</td>
</tr>
<tr>
<td><code>HM</code></td>
<td>3</td>
<td>12.1004</td>
<td>0.0000</td>
</tr>
<tr>
<td><code>scale(log(CelexFreq))</code></td>
<td>1</td>
<td>10.4878</td>
<td>0.0012</td>
</tr>
<tr>
<td><code>NoNuc</code></td>
<td>1</td>
<td>4.7707</td>
<td>0.0292</td>
</tr>
<tr>
<td><code>CCluster</code></td>
<td>1</td>
<td>30.7502</td>
<td>0.0000</td>
</tr>
<tr>
<td><code>Phone</code></td>
<td>2</td>
<td>7.0075</td>
<td>0.0009</td>
</tr>
<tr>
<td><code>rcs(Episode, 3):HM</code></td>
<td>6</td>
<td>4.6855</td>
<td>0.0001</td>
</tr>
<tr>
<td><code>CCluster:Phone</code></td>
<td>2</td>
<td>4.8160</td>
<td>0.0083</td>
</tr>
</tbody>
</table>
### Trends in speech rate: Fixed predictors

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Estimate</th>
<th>$t$ value</th>
<th>$p$ value (MCMC)</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>rcs(Episode, 3)</code></td>
<td>0.0008501</td>
<td>2.538</td>
<td>$&lt;0.01$</td>
</tr>
<tr>
<td><code>rcs(Episode, 3)'</code></td>
<td>-0.0013735</td>
<td>-2.940</td>
<td>0.01</td>
</tr>
<tr>
<td><code>HMRAvMI</code></td>
<td>0.1090062</td>
<td>3.677</td>
<td>$&lt;0.01$</td>
</tr>
<tr>
<td><code>HMRAvLI</code></td>
<td>-0.1121387</td>
<td>-3.125</td>
<td>$&lt;0.01$</td>
</tr>
<tr>
<td><code>scale(log(CelexFreq))</code></td>
<td>-0.0179985</td>
<td>-2.769</td>
<td>$&lt;0.01$</td>
</tr>
<tr>
<td><code>NoNucYvN</code></td>
<td>-0.0396347</td>
<td>-2.183</td>
<td>0.03</td>
</tr>
<tr>
<td><code>CClusterTvF</code></td>
<td>-0.0497443</td>
<td>-3.675</td>
<td>$&lt;0.01$</td>
</tr>
<tr>
<td><code>newPhonePvT</code></td>
<td>0.0398684</td>
<td>2.984</td>
<td>$&lt;0.01$</td>
</tr>
<tr>
<td><code>newPhoneTvK</code></td>
<td>0.0037097</td>
<td>0.318</td>
<td>0.75</td>
</tr>
<tr>
<td><code>rcs(Episode, 3):HMRAvMI</code></td>
<td>-0.0041242</td>
<td>-3.537</td>
<td>$&lt;0.01$</td>
</tr>
<tr>
<td><code>rcs(Episode, 3)':HMRAvMI</code></td>
<td>0.0044163</td>
<td>2.748</td>
<td>$&lt;0.01$</td>
</tr>
<tr>
<td><code>rcs(Episode, 3):HMREvLI</code></td>
<td>0.0017305</td>
<td>1.498</td>
<td>0.13</td>
</tr>
<tr>
<td><code>rcs(Episode, 3)':HMREvLI</code></td>
<td>-0.0033440</td>
<td>-2.125</td>
<td>0.03</td>
</tr>
<tr>
<td><code>rcs(Episode, 3):HMRAvLI</code></td>
<td>0.0010888</td>
<td>0.778</td>
<td>0.44</td>
</tr>
<tr>
<td><code>rcs(Episode, 3)':HMRAvLI</code></td>
<td>-0.0002978</td>
<td>-0.153</td>
<td>0.87</td>
</tr>
</tbody>
</table>
Trends in VOTSyls: smoothed fitted values
Trends in VOTSyls: conclusions

- Definite nonlinear change.
- Probable convergence.
  - Between everybody this time.
  - Again towards the end of the season, as other HMs leave.
Conclusions and Further Work
Conclusions

- We observe phonetic changes over the course of 92 days.
  - Nonlinear change in time.
  - Not across-the-board. Interesting interactions with HM.
- Possible longterm convergence in:
  - Mean pitch in a spurt (not Rachel).
  - Speech rate in a spurt (not Rachel).
  - VOT as fraction of an average syllable in a spurt.
- Conclusions on convergence must be tentative
  - Only looking at four people out of sixteen.
  - What is convergence in this setting?
(Ongoing) Social data
- Based on Sun logs, quantify how likely two people are to be mentioned together relative to independent chance.
- ⇒ metric of degree of social interaction over time.
- Correlates with phonetic trends?

Replicate conversation-level results
- Can we also detect conversation-level convergence?
- How does it relate to the possible longterm convergence?

Other variables, automating data collection.
Thanks to co-authors (Morgan Sonderegger, Peter Graff), and special thanks to Dan Jurafsky for useful advice.

And thank you!


Stable variation and historical linguistics.

Imitation in shadowing words.

Yao, Yao. 2009.
Understanding VOT variation in spontaneous speech.
In *UC Berkeley phonology lab annual report*, 29–43.