Introduction

- Phonetic variation extensive, yet usually does not lead to change. How do V & C coexist?

Actuation problem (restated):

1. Why does language change occur?
2. Why does it arise from variation?
3. What determines whether variation is stable (→ change) or not?

Two approaches to language change:

- 1. Diachronic datasets: Historical linguists (Wang), sociolinguists (Labov)
- 2. Math. models of linguistic populations: Dynamical systems for language [4, 3]

Here use appropriate models (2) to understand empirically observed trends (1).

Claim: Bifurcations in linguistic systems are possible explanation for the actuation problem: Long-term stability and sudden change coexist when dyn. sys. models contain bifurcations, correspond to learners with “ambiguity”.

Background: Dynamical systems

- Describe system state $\alpha$ at $t + 1$ as function of state at $t$.
- Fixed points: $\alpha_{t+1} = \alpha_t$, can be stable or unstable (when system perturbed).
- Ex: Pendulum has 1 of each.

Stress in English N/V pairs

- English 2-syllable noun/verb pairs have variable stress:

<table>
<thead>
<tr>
<th>N</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1, 1)</td>
<td>$\sigma \sigma$</td>
</tr>
<tr>
<td>(2, 2)</td>
<td>$\sigma \sigma$</td>
</tr>
<tr>
<td>(1, 2)</td>
<td>$\sigma \sigma$</td>
</tr>
<tr>
<td>(2, 1)</td>
<td>$\sigma \sigma$</td>
</tr>
</tbody>
</table>

- Ongoing variation: perfume, research, ally...

Variation within individuals: National Public Radio (US):

<table>
<thead>
<tr>
<th>Word</th>
<th>Research Only</th>
<th>Only Var</th>
<th>Var Spans</th>
</tr>
</thead>
<tbody>
<tr>
<td>perfume (N)</td>
<td>0.22</td>
<td>0.44</td>
<td>0.33</td>
</tr>
<tr>
<td>address (N)</td>
<td>0.4</td>
<td>0.4</td>
<td>0.2</td>
</tr>
</tbody>
</table>

- Bifurcations: stability of f.p.(s) changes suddenly as system parameter changes continuously.
- E.g. phase transitions in physics: as temp passes 100°C, stable state of water liquid → gas.

Data collection

- (1, 1), (1, 2), (2, 2) are stable states.
- In random subset (List 2), most do not change over time.
- But some do (Sherman [5]): ~149 pairs (List 1) changed since 1600, pron. data from dictionaries → lexical diffusion to (1,2)?

 Pronunciation trajectories

- Plot moving average of N, V prons of each pair, 50 year window.
- Complete changes observed:
  - (1, 1) → (1, 2) (A)
  - (1, 2) → (1, 1) (B)
  - (2, 1) → (2, 2) (C)
  - (2, 2) → (2, 1) (D)

- Short-term variation common (c), long-term variation rare (p).
- (2, 1) never occurs.
- Multidirectional diffusion
- What causes sudden loss of stability?

What is driving change?

Mishearing [2]:
- (English) N occur more often than V in trochaic-biasing ($\sigma \sigma$) contexts.

- Biases perception: $N \rightarrow \sigma \sigma$, $V \rightarrow \sigma \sigma$

- Similar trajectories for words sharing a prefix, e.g. con-
- Effect stronger for larger classes.
- Low-frequency words in a class change first, not true cross-class.

Model notation

- Dynamical systems models, assume:
  - Infinite population
  - Discretized generations: gen $t + 1$ learns from gen $t$.
  - For each N/V pair, each speaker keeps:

  \[
  \alpha_t, \beta_t \in [0, 1] = \text{prob of producing the 2 form.}
  \]
  \[
  \text{Let } \alpha_t = \text{probability N example at } t \text{ produced as } 2, \beta \text{ same for V.}
  \]
  \[
  \text{Mishearing probs } a_{i1}, a_{i2}, b_{i1}, b_{i2}:
  \]
  \[
  a_i = P(N \text{ heard as } 1 | 2 \text{ intended})
  \]

Model 1: No ambiguity

- Batch learner: at $t$ hears $N_t$ noun examples, $V_t$ verb examples.
- Hears $K_t$ N examples as 2, $K_v$ V examples as 2, then sets

  \[
  \alpha_t = \frac{K_v}{K_v + K_t}, \beta_t = \frac{K_t}{K_v + K_t}
  \]
  \[
  \text{Then take expectations: } \alpha_{t+1} = E(\alpha_t), \beta_{t+1} = E(\beta_t)
  \]
  \[
  \text{Find fixed points: } (\alpha_{t+1}, \beta_{t+1}) = (\alpha_t, \beta_t)
  \]

- Unique stable N,V freqs, depend on $a_i/b_i$ ratios. Long-term stability, but doesn’t explain sudden change.

Model 2: With ambiguity

- Try another error type: no mishearing, but each example can be heard as 1, 2 or ambiguous $\Rightarrow$ discarded.
- Let $r_i = \text{prob heard as ambiguous} | j \text{ intended}$ ($i = 1, 2$).
- Learner hears $K_t$, $K_v$, $N_t - K_t - K_v$ as 1, 2, ambiguity, sets

  \[
  \alpha_t = \frac{K_v}{K_v + K_t}
  \]
  \[
  \text{if } K_v + K_t > 0
  \]
  \[
  \text{(and 1/2 if } K_t + K_v = 0)
  \]
  \[
  \text{For large } N, \text{ find}
  \]

  \[
  \alpha_{t+1} = \alpha(1 - r_2) \frac{r_1}{(1 - r_1) + \alpha(r_1 - r_2)}
  \]

- Get fixed points

  \[
  x_1 = 1 \text{ stable for } r_1 > r_2 \quad x_2 = 0 \text{ stable for } r_1 < r_2
  \]

- Bifurcation at $r_1 = r_2$, sudden change explained as loss of stability of a f.p.

Variations

- Frequency effects: Make N finite, get bifurcation-like behavior, low-frequency words change first.
- Prefix effects: Couple $\alpha, \beta$ variables for words with same prefix.
- S-shaped curves: Mixture of ambiguity and mishearing, $R(\%)$ errors mishearing.
- Determines how bifurcation-like curve is.

Conclusions

- Observe sudden change in N/V stress between multiple states following long-term stability.
- Model as bifurcations in nonlinear dynamics of linguistic populations.
- Models have bifurcations $\iff$ have ambiguity.
- Bifurcations possible explanation for actuation.

References

[1] Abel Boyte. The royal dictionary: in two parts... The royal dictionary: in two parts...