When does variation lead to change? A dynamical systems account of an English stress shift
Partha Niyogi∗, Morgan Sonderegger†
Department of Computer Science, University of Chicago
LabPhon 11
Victoria University of Wellington
July 1, 2008

1 Introduction

• Variation between A and B can disappear (A/B → A), persist (A/B → A/B), or lead to change (A/B → B).
• Phonetic variation extensive, yet usually does not lead to change. How do V & C coexist?
• The actuation problem (restated):
  1. Why does language change occur?
  2. Why does it arise from variation?
  3. What determines whether variation is stable (leads to change) or not?
• Explore here through case study of an English stress shift, combine two approaches to language change:
  1. Diachronic datasets: Historical linguists (Wang), sociolinguists (Labov), focus on full description of changes, how they progress.
  2. Mathematical models of linguistic populations: Dynamical systems for language [8, 7], focus on how/whether change can occur in a population under varying assumptions.
• Here use appropriate models (2) to understand empirically observed trends (1).
• Claim: Bifurcations in linguistic systems are a possible explanation for the actuation problem: Long-term stability and sudden change coexist when dynamical systems models contain bifurcations, correspond to learners with “ambiguity.”

2 Preliminaries
Dynamical Systems

• (Discrete version) Describe system state α at t + 1 as function of state at t:
  \[ \alpha_{t+1} = f(\alpha_t) \]
  This describes any iterated map; the dynamical systems viewpoint is to look for steady states, or fixed points, where \( f(\alpha_t) = \alpha_t \), e.g. \( \alpha_{t+1} = \alpha_t \).
  Fixed points can be stable or unstable under perturbations.
  Example: A pendulum has two fixed points, one stable (pendulum arm down) and one unstable (pendulum arm up). A perturbation here means touching the pendulum arm.
  Pendulum fixed point stabilities determined by the direction of gravity: system parameter.
  A bifurcation occurs when the stability of some fixed points changes suddenly as a system parameter changes continuously.
  Example: Phase transitions in physics: as temperature passes 100° C, the gaseous state of water becomes stable and the liquid state becomes unstable.

Stress in English N/V pairs

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

∗niyogi@cs.uchicago.edu
†morgan@cs.uchicago.edu
<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>V</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(1, 1)</td>
<td>δσ</td>
<td>δσ</td>
<td>(elbow, fracture)</td>
</tr>
<tr>
<td>(1, 2)</td>
<td>δσ</td>
<td>σδ</td>
<td>(consort, protest)</td>
</tr>
<tr>
<td>(2, 2)</td>
<td>σδ</td>
<td>σδ</td>
<td>(police, review)</td>
</tr>
</tbody>
</table>

- Variation between varieties, and ongoing within: perfume, research, ally...\(^1\)
- Variation occurs within individuals: Data from speakers in news stories, interviews on National Public Radio (US).\(^2\)

<table>
<thead>
<tr>
<th>Word</th>
<th>1 only</th>
<th>2 only</th>
<th>Var</th>
<th>Spkrs</th>
</tr>
</thead>
<tbody>
<tr>
<td>research (N)</td>
<td>0.53</td>
<td>0.12</td>
<td>0.35</td>
<td>17</td>
</tr>
<tr>
<td>perfume (N)</td>
<td>0.22</td>
<td>0.44</td>
<td>0.33</td>
<td>9</td>
</tr>
<tr>
<td>address (N)</td>
<td>0.4</td>
<td>0.4</td>
<td>0.2</td>
<td>5</td>
</tr>
</tbody>
</table>

The first row means that out of 17 speakers, 53% used only δσ for “research” (n.), 12% only σδ, and 35% both.

- What are the diachronic dynamics of N/V pairs?

3 Data

1. (1,1), (1,2), (2,2) are stable states
   - In random subset (List 2), most do not change over time.
   - But some do: Sherman [9] cites 149 pairs (List 1) changed since 1600, based on pronunciation data from dictionaries. What does change look like diachronically?

2. Dataset:
   - 149 words (List 1), 76 dictionaries, 1550–2007 (most 1700 on), collected by Sherman (1550-1800), us (1800-2007).
   - British and American, only British used here.

3. Pronunciation trajectories:
   - Plot moving average of N, V pronunciations of each pair, 50 year window.\(^3\)
   - Complete changes observed (between “endpoints”):
     (a) (1, 1) → (1, 2)
     (b) (1, 2) → (1, 1)
     (c) (2, 2) → (1, 2)
     (d) (1, 2) → (2, 2)
   - Short-term variation common (e), long-term variation rare (f).
   - Examples:

\[^1\]From American English-speaking students.
\[^2\]Only speakers with ≥ 5 tokens were included.
\[^3\]A point was recorded for \(w\) at \(t\) if ≥ 2 dictionaries in \((t - 25, t + 25)\) listed pronunciation data for \(w\).
• $(2, 1)$ never occurs.
• Multidirectional diffusion: $(1, 1) \leftrightarrow (1, 2) \leftrightarrow (2, 2)$
• What causes sudden loss of stability?

Forces driving change
1. Mishearing [4, 5, 2]:
   • Known that English N occur more often than V in trochaic-biasing contexts, same for V in iambic-basing contexts.
   • Biases perception such that iambic N misheard as trochaic, trochaic V misheard as iambic.
2. Relative N/V frequency (not on poster):
   • Let $r$ be $Vfreq/(Nfreq+Vfreq)$, $y$ distance along $(1, 1) \leftrightarrow (1, 2) \leftrightarrow (2, 2)$ axis.
   • At fixed points in time, $y$ positively correlated with $r$, e.g. higher Nfreq relative to Vfreq shifts pronunciation of N and V towards the default N pattern (trochaic).
3. Prefix similarity:
   • 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.

4 Models
Dynamical systems models, assume:
1. Infinite population
2. Discretized generations: generation $t + 1$ learns from generation $t$.
3. For each N/V pair, each speaker keeps $\tilde{\alpha}, \tilde{\beta} \in [0, 1]$ = probabilities of producing the N, V forms (resp.) as 2.4

\[^{4}\text{We assume learners store probabilities } \alpha \in [0, 1] \text{ because of evidence from the NPR data that variation} \]
4. Let $\alpha_t = \text{probability a N example at } t \text{ produced as 2, } \beta_t \text{ same for V. These probabilities are across the whole population, and model the input for a learner at } t$.

5. To model misperception effects described above, assume \textit{mishearing probabilities}:

$$a_1 = P(N \text{ heard as 1 } | \text{2 intended}), \quad b_1 = P(N \text{ heard as 2 } | \text{1 intended})$$

$$a_2 = P(V \text{ heard as 1 } | \text{2 intended}), \quad b_2 = P(V \text{ heard as 2 } | \text{1 intended})$$

6. Then the probabilities a N/V example at $t$ is heard as 2 are:

$$N: \quad P_1(t) = \alpha_t (1 - a_1) + (1 - \alpha_t)b_1$$

$$V: \quad P_2(t) = \beta_t (1 - a_2) + (1 - \beta_t)b_2$$

Model 1: no ambiguity

- Each learner hears $N_1$ noun examples, $N_2$ verb examples, of which $K_1, K_2$ have final stress.
- $K_1, K_2$ are random variables, each learner one sample.$^5$
- Batch learner: After hearing all examples, each learner sets

$$\tilde{\alpha} = \frac{K_1}{N_1}, \quad \tilde{\beta} = \frac{K_2}{N_2}$$

- Expectation of learners’ values gives $\alpha$ and $\beta$ for the next generation:

$$\alpha_{t+1} = E\left(\frac{K_1}{N_1}\right), \quad \beta_{t+1} = E\left(\frac{K_2}{N_2}\right)$$

- To take these expectations, note $K_1$ and $K_2$ binomial random variables (parameters $P_1, P_2$), get iterated maps:

$$\alpha_{t+1} = f_1(\alpha_t) := \alpha_t (1 - a_1) + (1 - \alpha_t)b_1$$

$$\beta_{t+1} = f_2(\beta_t) := \beta_t (1 - a_2) + (1 - \beta_t)b_2$$

- Want fixed points: $f_1(\alpha^*) = \alpha^*, f_2(\beta^*) = \beta^*$:

$$\alpha^* = \frac{b_1}{a_1 + b_1}, \quad \beta^* = \frac{b_2}{a_2 + b_2}$$

- Unique, stable fixed points, depend on ratios of mishearing probabilities... but this \textit{doesn’t explain sudden change.}

Model 2

- Try another type of error: no mishearing, but an example can be heard as 1, 2, or \textit{ambiguous}, in which case discarded.
- Consider just one form, same population assumptions as in Model 1, $\alpha_t$ probability a random example produced as 2 at $t$.
- Ambiguity:

$$r_i = P(\text{heard as ambiguous } | \text{i intended}) \quad (i = 1, 2)$$

occurs \textit{within} individuals. If learners instead always used one form or the other, we would take $\alpha \in \{0, 1\}$, and if variation only seemed to occur above some threshold percentage $r$, we would take $\alpha \in (r, 1 - r)$. These choices lead to very different model dynamics.$^7$.

$^5 K_1$ and $K_2$ are random variables for learners at $t$, and would be new random variables for learners at $t + 1$. To make this explicit we could write $K_{1,t}$ and $K_{2,t}$. 

$^7$
Figure 1: Curves showing stable fixed point location as a function of relative error (in hearing form 1 vs. form 2) for different values of $R$ (see text).

- For a random example heard at $t$, let $P_i(t) = P($heard as i$)$:
  \[ P_1 = (1 - \alpha)(1 - r_1), \quad P_2 = \alpha(1 - r_2) \]

- Modified batch learner: learner hears $N$ examples, $k_1$ heard as 1, $k_2$ as 2, $N - k_1 - k_2$ ambiguous, estimates $\alpha$ as
  \[ \tilde{\alpha} = \begin{cases} \frac{k_2}{k_1 + k_2} & \text{if } k_1 + k_2 > 0 \\ 0.5 & \text{if } k_1 + k_2 = 0 \end{cases} \]

- For large $N$, find that
  \[ E(\tilde{\alpha}) = \frac{E(k_2)}{E(k_1) + (k_2)} \implies \alpha_{t+1} = \frac{\alpha(1 - r_2)}{(1 - r_1) + \alpha(r_1 - r_2)} \]

- Get two fixed points:
  \[ \alpha^*_+ = 1 \quad \text{stable for } r_1 > r_2 \]
  \[ \alpha^*_- = 0 \quad \text{stable for } r_1 < r_2 \]

- Bifurcation at $r_1 = r_2$, explains sudden change as loss of stability of a fixed point.

Variations

(Not discussed further for lack of space.)

1. **Frequency effects**: Make $N$ finite, get bifurcation-like behavior, low-frequency words change first.
2. **Prefix effects**: Couple $\alpha, \beta$ variables for words with same prefix.
3. **Parametrized S-shaped curves**: Mixture of ambiguity and mishearing, $R =$% errors mishearing (rest are ambiguity), determines how “bifurcation-like” curve is (Fig. 1).

5 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.\(^6\)
- Bifurcations possible explanation for actuation.

\(^6\)If there is any experimental evidence for or against the idea that learners treat some items as ambiguous (e.g. disregard them), please let us know!
### List 1: Sherman’s word list [9]

<table>
<thead>
<tr>
<th>abstract</th>
<th>compact</th>
<th>contract</th>
<th>discharge</th>
<th>impact</th>
<th>invert</th>
<th>permit</th>
<th>rebel</th>
<th>rehash</th>
<th>torment</th>
</tr>
</thead>
<tbody>
<tr>
<td>accent</td>
<td>compound</td>
<td>converse</td>
<td>discord</td>
<td>import</td>
<td>legeate</td>
<td>pervert</td>
<td>rebound</td>
<td>reject</td>
<td>transfer</td>
</tr>
<tr>
<td>addict</td>
<td>compress</td>
<td>convert</td>
<td>discount</td>
<td>impress</td>
<td>misprint</td>
<td>postdate</td>
<td>recall</td>
<td>relapse</td>
<td>transplant</td>
</tr>
<tr>
<td>address</td>
<td>concert</td>
<td>convict</td>
<td>discourse</td>
<td>imprint</td>
<td>object</td>
<td>prefix</td>
<td>recast</td>
<td>repeat</td>
<td>transverse</td>
</tr>
<tr>
<td>affect</td>
<td>concrete</td>
<td>convoy</td>
<td>egress</td>
<td>incense</td>
<td>outcast</td>
<td>prelude</td>
<td>recess</td>
<td>research</td>
<td>undress</td>
</tr>
<tr>
<td>affix</td>
<td>conduct</td>
<td>decay</td>
<td>incline</td>
<td>incline</td>
<td>outcast</td>
<td>premise</td>
<td>recoil</td>
<td>reprint</td>
<td>traverse</td>
</tr>
<tr>
<td>alloy</td>
<td>confect</td>
<td>decrease</td>
<td>escort</td>
<td>increase</td>
<td>outgo</td>
<td>presage</td>
<td>record</td>
<td>research</td>
<td>upcast</td>
</tr>
<tr>
<td>ally</td>
<td>refine</td>
<td>defect</td>
<td>essay</td>
<td>indent</td>
<td>outlaw</td>
<td>present</td>
<td>recount</td>
<td>reset</td>
<td>upcast</td>
</tr>
<tr>
<td>annex</td>
<td>conflict</td>
<td>defile</td>
<td>excerpt</td>
<td>infux</td>
<td>outleap</td>
<td>produce</td>
<td>redraft</td>
<td>sojourn</td>
<td>upgrade</td>
</tr>
<tr>
<td>assay</td>
<td>conscript</td>
<td>descant</td>
<td>excise</td>
<td>inflow</td>
<td>outlook</td>
<td>progress</td>
<td>redress</td>
<td>subject</td>
<td>uplift</td>
</tr>
<tr>
<td>bombard</td>
<td>conserve</td>
<td>desert</td>
<td>exile</td>
<td>inlay</td>
<td>outpour</td>
<td>project</td>
<td>refil</td>
<td>sublease</td>
<td>upright</td>
</tr>
<tr>
<td>cement</td>
<td>consort</td>
<td>detail</td>
<td>exploit</td>
<td>inlet</td>
<td>outdated</td>
<td>protest</td>
<td>refit</td>
<td>sublet</td>
<td>uprise</td>
</tr>
<tr>
<td>collect</td>
<td>content</td>
<td>dictate</td>
<td>export</td>
<td>insert</td>
<td>outstretch</td>
<td>purport</td>
<td>refund</td>
<td>surcharge</td>
<td>uprush</td>
</tr>
<tr>
<td>combat</td>
<td>contest</td>
<td>digest</td>
<td>extract</td>
<td>inset</td>
<td>outwork</td>
<td>rampage</td>
<td>refuse</td>
<td>survey</td>
<td>upset</td>
</tr>
</tbody>
</table>

### List 2: Sample of words in use 1700–2007

<table>
<thead>
<tr>
<th>abuse</th>
<th>bottom</th>
<th>contest</th>
<th>envy</th>
<th>harbour</th>
<th>measure</th>
<th>proceed*</th>
<th>repeal</th>
<th>table</th>
<th>whistle</th>
</tr>
</thead>
<tbody>
<tr>
<td>accent</td>
<td>breakfast</td>
<td>contract</td>
<td>exile*</td>
<td>harbour</td>
<td>measure</td>
<td>protest*</td>
<td>repose</td>
<td>rally</td>
<td>witness</td>
</tr>
<tr>
<td>advance</td>
<td>buckle</td>
<td>convict</td>
<td>express</td>
<td>import*</td>
<td>merit</td>
<td>purchase</td>
<td>reserve</td>
<td>thunder</td>
<td></td>
</tr>
<tr>
<td>afront</td>
<td>bundle</td>
<td>cover</td>
<td>favour</td>
<td>increase*</td>
<td>motion</td>
<td>puzzle</td>
<td>review</td>
<td>title</td>
<td></td>
</tr>
<tr>
<td>ally*</td>
<td>butter</td>
<td>decrease*</td>
<td>ferret</td>
<td>interest</td>
<td>murder</td>
<td>quarry</td>
<td>rival</td>
<td>torment</td>
<td></td>
</tr>
<tr>
<td>anchor</td>
<td>cement*</td>
<td>degree</td>
<td>flourish</td>
<td>iron</td>
<td>muster</td>
<td>reason</td>
<td>saddle</td>
<td>travel</td>
<td></td>
</tr>
<tr>
<td>arrest</td>
<td>challenge</td>
<td>diet</td>
<td>forecast*</td>
<td>journey</td>
<td>order</td>
<td>redress</td>
<td>second</td>
<td>treble</td>
<td></td>
</tr>
<tr>
<td>assault</td>
<td>channel</td>
<td>digest*</td>
<td>forward</td>
<td>level</td>
<td>outlaw</td>
<td>reform</td>
<td>shiver</td>
<td>triumph</td>
<td></td>
</tr>
<tr>
<td>assay</td>
<td>command</td>
<td>dispatch</td>
<td>gallop</td>
<td>levy</td>
<td>pepper</td>
<td>regard</td>
<td>shoulder</td>
<td>trouble</td>
<td></td>
</tr>
<tr>
<td>attack</td>
<td>concern</td>
<td>dissent</td>
<td>glory</td>
<td>licence</td>
<td>plaster</td>
<td>relapse*</td>
<td>squabble</td>
<td>value</td>
<td></td>
</tr>
<tr>
<td>bellow</td>
<td>conduct</td>
<td>distress</td>
<td>hammer</td>
<td>license</td>
<td>premise*</td>
<td>relish</td>
<td>stable</td>
<td>visit</td>
<td></td>
</tr>
<tr>
<td>blunder</td>
<td>consort</td>
<td>double</td>
<td>handle</td>
<td>matter</td>
<td>present</td>
<td>remark</td>
<td>stomach</td>
<td>vomit</td>
<td></td>
</tr>
</tbody>
</table>

---

72007 pronunciations from Cambridge Advanced Learner’s Dictionary, OED. Words were randomly chosen from all N/V pairs which (a) have both N and V frequency of at least 1 per million in the BNC [6] (b) have both N and V forms listed in a dictionary from 1700 [1] (c) have both N and V forms listed in a dictionary from 1847 [3].

* indicates the pronunciations listed in the 1700, 1847, and 2007 dictionaries are not identical.