Overview: What is stored and what is calculated?

1. Bases of paradigms: do we really need an underlying form?
   - Albright 2002: every paradigm has a base that the other members are derived from
     - N.B. This is different from the “base” in *Richness of the Base* (where it means “input”), or in *base-reduplicant correspondence* (where it means the part of the word that the reduplicant is copied from)
   - First big idea: The base has to be one of the surface forms of the paradigm
     - e.g. Russian noun paradigm: ‘pie’ (from Wiktionary, with phonology added)

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>nominative</td>
<td>pirók</td>
<td>piragí</td>
</tr>
<tr>
<td>genitive</td>
<td>piragá</td>
<td>piragóf</td>
</tr>
<tr>
<td>dative</td>
<td>piragú</td>
<td>piragám</td>
</tr>
<tr>
<td>accusative</td>
<td>pirók</td>
<td>piragí</td>
</tr>
<tr>
<td>instrumental</td>
<td>piragóm</td>
<td>piragámi</td>
</tr>
<tr>
<td>prepositional</td>
<td>piragé</td>
<td>piragáx</td>
</tr>
</tbody>
</table>

   Knowing that Russian has vowel reduction and final devoicing, what would we normally say the underlying form is?
   - In Albright’s model, the learner can’t have a “composite” underlying form, and must settle for one of these surface forms
     - anything not predictable from that surface form must be memorized as exceptional
     - or perhaps covered by a minor rule that applies to a few words

   see Bowers 2015 for arguments in favor of composite underlying forms!

   - Second big idea: Within a language, this base is the same cell of every paradigm
     - e.g., always the genitive singular

   - Third big idea: Learners choose as the base the paradigm member that is most informative
     - implemented as how well a rule system (learned by Alright-Hayes morphological learner, Albright & Hayes 2003) can derive the rest of the paradigm from that cell

To do
- Read *Wagner 2012* for Thursday.
- Have you thought about a project topic?

---

*Ling 219, Phonological Theory III. Fall 2015, Zuraw*
• **Fourth big idea:** We can get evidence about which cell is the base from diachronic change
  - Latin example from Albright 2001
  - Pre-Classical Latin had a rule of approximately $s \rightarrow r / V__V$

<table>
<thead>
<tr>
<th></th>
<th>pre-Classical Latin</th>
</tr>
</thead>
<tbody>
<tr>
<td>nominative</td>
<td>hono:s</td>
</tr>
<tr>
<td>genitive</td>
<td>hono:ris</td>
</tr>
<tr>
<td>dative</td>
<td>hono:ri:</td>
</tr>
<tr>
<td>accusative</td>
<td>hono:rem</td>
</tr>
<tr>
<td>ablative</td>
<td>hono:re</td>
</tr>
</tbody>
</table>

  - What actually happened: *hono:s* changed to *honor* (there was also vowel shortening)
  - Apparently, once learners had to memorize the nominative [s] as a quirk of certain words, they started losing it.

2. **How redundant should an underlying representation be?**

2.1 **A traditional view**

- Chomsky & Halle (1968)’s answer: not redundant at all
  - Strip out anything that could be predicted by the grammar
    - Some tricky ones to ponder in American English: *butter, spot, fear, see*

- Encode exceptional behavior in the underlying representation, where possible. E.g.
  
  $[\text{ɹɑɪt}]$ ‘right’  $[\text{ɹɑɪt}ʃə]$ ‘righteous’
  (exceptional because no trisyllabic shortening)

  - SPE’s solution: /rixt/! Let’s see if we can reconstruct how it would work

- Taking it too far?
  - Coetzee (1999) example: how much do we really need to specify about the first consonant of English *string*?

- This all reflects a view that storage is expensive (and calculation is cheap, I guess)
  - I think cognitive scientists have changed their view on this though
2.2 Richness of the Base (review)—Prince & Smolensky (1993)

- In OT, the grammar is responsible for mapping the set of all possible underlying forms (which is the same for every language) to the set of legal surface forms

\[
\begin{align*}
/k^hiri/ & \rightarrow [k^hiri] \\
/k^hti/ & \rightarrow [k^hti] \\
/k^hidi/ & \rightarrow [k^hidi] \\
/kiri/ & \rightarrow [kiri] \\
/ktti/ & \rightarrow [ktti] \\
/k^hidi/ & \rightarrow [k^hidi]
\end{align*}
\]

- In English, it doesn’t really matter if the UR is /k^hiri/ or /ktti/

2.3 Lexicon optimization—also Prince & Smolensky (1993)

- The idea is to run an output form through the grammar to choose the best input candidate

  o Define “markedness constraint”:

  o Define “faithfulness constraint”:

  o With those definitions in mind, fill in the tableau

<table>
<thead>
<tr>
<th></th>
<th>*#UNASP</th>
<th>*V{t,d}V</th>
<th>IDENT(spread glottis)</th>
<th>IDENT(voice)</th>
<th>IDENT(tap)</th>
</tr>
</thead>
<tbody>
<tr>
<td>/k^hiri/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/k^hti/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/k^hidi/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/kiri/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/ktti/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/k^hidi/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

  o Thoughts on whether this seems like what we want? (Also, how could we know anyway?)

- P&S propose that alongside *STRUC, which we’ve used a couple of times now (“don’t have material in the output”), there is *SPEC (“don’t have material in the input”).
  o What would be the effect of including it in the ranking above? We probably need some less-specified candidates to compare.
2.4 Underspecification

- What if the UR is just missing some feature values?
  - e.g., the first consonant of ‘kitty’ has no value for [spread glottis]
- By the way, in rule-based days, some theories made a distinction between “feature-filling” rules and “feature-changing” rules
- An example where this could be useful: Turkish voicing alternations, Inkelas 1995

(3) a. Alternating root-final plosive:

- kanat ‘wing’
- kanat-lar ‘wing-pl’
- kanat-im ‘wing-1sg.poss’

b. Nonalternating voiceless plosive:

- sanat ‘art’
- sanat-lar ‘art-pl’
- sanat-im ‘art-1sg.poss’

c. Nonalternating voiced plosive:

- etüd ‘etude’
- etüd-lar ‘etude-pl’
- etüd-üm ‘etude-1sg.poss’

- Let’s think how underspecification could help get the three-way distinction

- This is a bit different from underspecification in a output representation, where the idea is that there will be phonetic interpolation. See Steriade (1995) for a survey of underspecification.

2.5 What if we just store surface forms?

- How narrow?
  - How narrow could we get for cat?

- The challenge: what if the representation is so detailed that the details it represents are not reliable ones?
  - Can we come up with some examples for cat?

- This could make it hard to recognize new tokens as instances of that word
- Which leads us to…

3. Making the lexicon do more work: exemplars

- Student presentations of Pierrehumbert 2002
References