Class 6 (Week 3, T): Sideways interfaces I, phonology and the lexicon

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

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)

	singular	*plural*
nominative	pirók	piragí
genitive	piragá	piragóf
dative	piragú	piragám
accusative	pirók	piragí
instrumental	piragóm	piragámi
prepositional	piragé	piragáx

  - 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

- 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 → r / V__V

	pre-Classical Latin
nominative	hono:s
genitive	hono:ris
dative	hono:ri:
accusative	hono:rem
ablative	hono:re

By Albright's algorithm, ablative is the best choice for Latin over all
- What could be the diachronic consequence?

  - 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.

  [ɹaɪt] 'right' [ɹaɪtʃəs] 'righteous'
  (exceptional because no trisyllabic shortening)

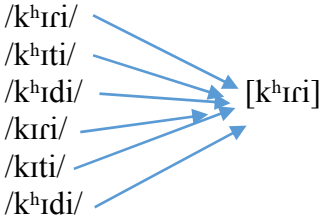
  - SPE's solution: /rixt/! Let's see if we can reconstruct how it would work

- 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



- In English, it doesn't really matter if the UR is /kʰɪɾi/ or /kɪti/

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
    - Define "markedness constraint":
    - Define "faithfulness constraint":
    - With those definitions in mind, fill in the tableau

[kʰɪɾi]	*#UNASP	*V{t,d}V̆	IDENT(spread glottis)	IDENT(voice)	IDENT(tap)
/kʰɪɾi/					
/kʰɪti/					
/kʰɪdi/					
/kɪɾi/					
/kɪti/					
/kʰɪdi/					

    - 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").
    - 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'	kanad-ɨ	'wing-Acc'
kanat-lar	'wing-pl'	kanad-ɨm	'wing-1sg.poss'
b. Nonalternating voiceless plosive:			
sanat	'art'	sanat-ɨ	'art-Acc'
sanat-lar	'art-pl'	sanat-ɨm	'art-1sg.poss'
c. Nonalternating voiced plosive:			
etüd	'etude'	etüd-ü	'etude-Acc'
etüd-ler	'etude-pl'	etüd-üm	'etude-1sg.poss'

(p. 3)

  - 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

Albright, Adam. 2001. Modeling the Latin honor analogy with a computational learner. Washington, DC.

Albright, Adam. 2002. The identification of bases in morphological paradigms. UCLA Ph.D. dissertation.

Albright, Adam & Bruce Hayes. 2003. Rules vs. analogy in English past tenses: a computational/experimental study. *Cognition* 90(2). 119–161.

Bowers, Dustin. 2015. A system for morphophonological learning. UCLA PhD dissertation.

Chomsky, Noam & Morris Halle. 1968. *The Sound Pattern of English*. Harper & Row.

Coetzee, Andries W. 1999. *Tiberian Hebrew Phonology: Focussing on Consonant Clusters*. Assen, The Netherlands: Van Gorcum.

Inkelas, Sharon. 1995. The Consequences of Optimization for Underspecification. *Proceedings of the Twenty-fifth Northeastern Linguistics Society*. GLSA.

Lloret, Maria-Rosa. 2004. The phonological role of paradigms: the case of insular Catalan. *Contemporary approaches to Romance linguistics: selected papers from the 33rd Linguistic Symposium on Romance Languages (LSRL), Bloomington, Indiana, April 2003*, 275–298. John Benjamins Publishing.

Pertsova, Katya. 2004. Distribution of Genitive Plural Allomorphs in the Russian Lexicon and in the Internal Grammar of Native Speakers. UCLA.

Pierrehumbert, Janet. 2002. Word-specific phonetics. *Laboratory Phonology VII*. Berlin: Mouton de Gruyter.

Prince, Alan & Paul Smolensky. 1993. *Optimality Theory*. Blackwell.

Steriade, Donca. 1995. Underspecification and Markedness. *The Handbook of Phonological Theory*, 114–174. Blackwell.

Wagner, Michael. 2012. Locality in phonology and production planning. In A McKillen & J Loughran (eds.), *Proceedings of the Montreal-Ottawa-Toronto (MOT) Phonology Workshop 2011. Phonology in the 21st Century: In honour of Glyne Piggott. McGill Working Papers in Linguistics 22(1).*