1. **Assignments**

- Modern Hebrew homework is due in class Thursday Feb. 2.
- New reading for Tues. 1/22:
  - Do a half-page summary
  - Due Tues. Jan 31.

2. **Near-neutralization in adult phonology**

- **Near-neutralization** is by now a widely-studied topic in adult phonology:
  - “[A]” derived from /B/ is not the same as [A] derived without change, or derived from /C/.
- Some familiar processes that are likely non-neutralizing:
  - Final Devoicing in German\(^1\)
  - 3rd Tone Sandhi in Mandarin\(^2\)
  - North American English Tapping\(^3\)
  - The empirical literature can be accessed with a Google Scholar search on “incomplete neutralization phonology”.
  - A nice theoretical approach, with MaxEnt phonetics, is in the work of Aaron Braver at Texas Tech. https://www.aaronbraver.com/
- Methodologically: it pays to measure, if you think you have a phonological neutralization.
- It’s not clear that a near-neutralizing process in an opacity configuration is actually opaque.

3. **Near-neutralization in children**

- It is probably ubiquitous — diary studies mostly can’t detect it.

---


• Smith, using just his ear, sometimes was able to spot it.

4. A classic experimental study of near-neutralization in children: Macken and Barton on VOT in children


• Several kids played with a bunch of stop-initial toys in a recording booth,\(^4\) in various sessions, as they got older.
• General age range was 1;5 to 2;4.
• Researchers measured Voice Onset Time for all the word-initial stops.

5. Results

• Early on: vegetative values, including reflection of “more voicing in fronter places,” which has an articulatory explanation (Keating and Westbury, *J. Linguistics* 1986).
• Gradually: the clouds of data for the categories voiced/voiceless part, leaving an ever more perceptible distinction.
• During the middle stages: difference is statistically significant, but *transcribers can’t hear it*.

6. Example: Little Tessa gradually gets it right

```
+200
+160
+120
+80
+40
0
-40
-80
```

7. This happens, though less often, for adults

• Here is a possible case.
• The Null Parse candidate means, “shut up, say nothing”. No constraints other than *NULL PARSE* are violated.

\(^4\)”If your family has a Piglet cuddly, please bring it.”
Linguistics 201A  
Class 6, 1/26/2023: Learnability; Token Variation  
p. 3

<table>
<thead>
<tr>
<th>/happy + ly/</th>
<th>*[ı] + LY</th>
<th>*VCₐCₓV̆</th>
<th>*NULL PARSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ˈ[ʰæpəli]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Null Parse</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[ˈhæpəli]</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

vs.

<table>
<thead>
<tr>
<th>/silly + ly/</th>
<th>*[ı] + LY</th>
<th>*VCₐCₓV̆</th>
<th>*NULL PARSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ˈ[sɪləli]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[ˈsɪləli]</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Exceptions to *VCₐCₓV: canonization, classicist, diocesan, probable, indescribable; all are in derived forms. I don’t know how to allow these but forbid *sillily.

8. A reference for NULLPARSE

- There are other papers that take this on but my favorite is:

9. Avoidance is probably more common for little kids

- To prove it you have to show that the kid knows a lot of words with the avoided sound or sequence; this has been done.

10. Jacob Hankamer’s velar stops (Menn reading p. 18)

    #k → #k
    #g → don’t try to say these words
    k# → k#
    g# → k (or null)

- Faithfulness is characteristically very high in initial position (often though to be for psycholinguistic reasons; see Noah Elkins UCLA M.A. or Becker et al. 2021⁵)
- Avoidance might be a way to avoid both markedness and faithfulness violations.

---

LEARNING THE PARENTAL SYSTEM

11. (Some of) the data we are responsible for

- Infancy: head-turn preference and similar results
- Early childhood: Passive data still relevant, but we can also wait till Junior speaks clearly enough to inform us about her intuitions:
  ➢ Wug-test, as in Berko (1958)
  ➢ Just plain elicitation can be very informative! See Do (2018) for some nice methodology.\(^6\)
  ➢ I’m not sure we canblick-test, but perhaps …

12. Research methods

- Write grammars that make sense of the infants and little kids’ intuitions and behavior.
- Collaborate with computationalists in trying to model the learning process itself.
  ➢ This is going on here at UCLA; e.g. Hunter with Perkins, Sundara with Hayes and Ph.D. Breiss\(^7\)

---


\(^7\) The latter case shows that the computationalists can be consumers, rather than inventors.
• It seems, learnability, especially early learnability, has been a really fruitful area for computational participation in linguistic research.
  ➢ Why so? The early stuff is amenable to simply tracking distributions, something computationalists are good at.

13. There are lots of tasks to be carried out by the child

• Presentational scheme here:
  ➢ learning task
  ➢ representative modeling work (just a sample; with a tendency to UCLA-bias)

• Figure out the phoneme inventory
  ➢ Kristine Yu dissertation

• Perhaps also the allophonic distribution.

• Perhaps, figure out a language-specific feature system

• Divide utterances into words and form a "proto-lexicon" (words that may not have definitions)

• Armed with a set of words, and perhaps meaning, find what affixes are present

• Figure out the phonotactics (of words and other domains)

• Establish the UR’s of morphemes (more than one, if allomorphs)
  ➢ For a partial bibliography, see Yang Wang and Bruce Hayes (in progress, borrowable) “Learning underlying representations: An approach guided by the Kenstowicz-Kisseberth UR Hierarchy”

• Discover the constraints (Markedness, Faithfulness) that govern phonotactics and alternations (unless they are innate).
  ➢ see Yang and Hayes paper for bibliography

• Establish how these constraints are ranked or weighted.
  ➢ Bruce Tesar and Paul Smolensky (2000 Learnability in Optimality Theory) — classical OT
  ➢ Paul Boersma 1998 dissertation, published — Stochastic OT

• Often, the latter tasks will require further hidden structure: syllabification, feet, phrasing

14. The tension of stepwise vs. all-at-once: all at once works better

• Linguists love components and like to work one component at a time
• Goldwater (2018): all at once is in principle, better: don’t ignore data
• The normal term for this is joint learning
• A well-known paper is
  ➢ Jointly learn the vocabulary and the phonemes

15. Another example of the virtue of joint learning: phonotactics and word learning

• Lots of results from Anne Cutler and other psycholinguists show word segmentation is aided by phonotactics.
  ➢ E.g. English iambic words (balloon, believe) are rare.
  ➢ Children tend to split them up in segmenting: the gui | tar is.
  ➢ Not so for other languages.
• Finnish kids can use “vowel harmony breaks”; Suomi, McQueen, & Cutler, 1997
• So, it pays to simultaneously detect words in utterances, and detect phonotactics of words.

16. The trouble with all-at-once

• Modern learning systems typically define a search space and specify a criterion for best member of search space — this guides learning in a systematic way.
• Joint learning multiplies out the candidate hypotheses for each component, eventually making the search space too big to handle.

VARIATION IN PHONOLOGY

17. Classifying phonological variation

• Type variation: different lexical items differ according to how they undergo particular phonological or processes. E.g. serene [səˈriːn] undergoes Trisyllabic Shortening in [səˈreɪnəti], obese does not ([ouˈbɪs], obesity [ouˈbɪsəti])
• Token variation: the very same morpheme can be pronounced differently on different occasions by the very same speaker—often in response to speaking styles (talk [tʊk, tɑk], Labov readings)

---

8 Talk given at the inaugural meeting of the Society for Computation in Linguistics.
11 For Americans this is [æt, ɛt, i; ou] → [i, æ, e, u] when a non-final atonic syllable follows. Many alternations triggered by suffixes, e.g. -ical, -ity.
18. Who are the experts here?

- Sociolinguists
- Some references:
  - *Sociolinguistics*, by Peter Trudgill, a short text
  - *Sociolinguistic Patterns* (source of your reading) and *Language in the Inner City*, early classics by William Labov. I find these very useful for getting oriented; though everything they say was subj to later amendment.
  - Labov’s magnum opus, *Principles of Linguistic Change*, in four volumes
- Sociolinguistics has evolved since the readings paper
  - Massive infusion of statistical modeling
  - Broader range of societal interests and focus on personal identity

19. Readings

- Covers some techniques used to gather forms in free variation
  - Note the extreme contradiction between the scientific goal of controlled study and the goal of accessing the vernacular forms
- Claims considerable systematicity for the patterns thus obtained.

20. The crucial variable phonological phenomena of New York City English

**R Dropping**

“(r-0)” *beard* is [biəd], r-less;

“(r-1)” *beard* is [bi̯ad], r-ful

- Notes:
  - Labov thinks /r/ is **underlyingly present**, learned from nondeleted tokens in the ambient language.
  - i.e. language learners can distinguish “always [sə]” from “sometimes [st], sometimes [sə]” (saw/sore)
  - This would not not be true, I suspect, of people like Standard British speakers, who hear no variation in these forms during childhood.

**/æ/ Diphthongization**

“eh-1” *bad* [biad]

“eh-4” *bad* [bæd]

**/a/ Diphthongization**

“oh-1” *coffee* [ˈkɒfɪ]

“oh-5” *coffee* [ˈkɔfi]
/θ/ and /ð/ Hardening

“th-3”  \( \text{thin} [\theta\text{n}] \)  “dh-3”  \( \text{this} [\theta\text{t}\text{s}] \)
“th-2”  \( [\theta\theta\text{m}] \)  “dh-2”  \( [\theta\delta\text{s}] \)
“th-1”  \( [\theta\text{m}] \)  “dh-1”  \( [\delta\text{s}] \)

Claim: these often vary in lockstep by speaking style. It’s like there is a knob, controlling all of the grammar at once.

21. The informal meter of style

A  free conversation with interviewer  
B  interview  
C  read paragraph  
D  read words from a list

22. Phonological free variation in the speech of Miriam

- Miriam is 35 years old, graduated Hunter College and St. John’s law school, works as lawyer.

23. Variation in the speech of Doris

- Doris is 39, homemaker, African-American.
- She doesn’t have perfect lockstep
• Labov thinks that for Doris, and others, r-dropping is more sensitive to style than other processes.

24. Free variation in society is structured as well
• Fig. 4.2 from William Labov (1972) Sociolinguistic Patterns

![Fig. 4.2 from William Labov (1972) Sociolinguistic Patterns](image)

**Fig. 4.2.** Class stratification of a linguistic variable in process of change. (r) in guard, ear, bare, beard, board, etc. SEC (Socio-economic class) scale: 0-1, lower class; 2-4, working class; 5-6, 7-8, lower middle class; 9, upper middle class A, casual speech; B, careful speech; C, reading style; D, word lists; D', minimal pairs.

• from 81 native speakers of New York City English
• Vertical axis: what percentage of underlying /ɪ/ are retained in the output?
• An independent investigation sorted the speakers into their social classes.
• The “leaping up” of the lower-middle-class speakers in the formal styles is found in other studies, and is claimed to reflect a social insecurity independently diagnosed by other tests.  

25. More on knobs

• It is sensible, at least metaphorically, to see the linguistic system as a whole is attached to “knobs” — external factors that cause its output to vary.
• Here are some knobs:
  ➢ style (casual-formal, per Labov reading)
  ➢ speaking rate
  ➢ frequency of words being said

26. The claim of true randomness

• Claim: even when we include the knobs, speakers are essentially stochastic devices — we can predict the distribution of outputs in a large sample, but not the outcome in each speaking occasion.
• Here, we will start with just the stochastic-device idea, and later discuss the rather small amount of theoretical work that has been done with knobs.

27. A case of speakers acting as stochastic devices

• Hayes and Londe (Phonology 2009)
• Hungarian stems whose last vowels are [+back], then [eː] go both ways with harmony.
  ➢ Mostly, stem-by-stem
  ➢ A few individual stems are vacillators.

• We wug-tested [haːdɛl] and [koleːn]
  ➢ Options for dative: [haːdɛl-ːnɔk, haːdɛl-ːnek], [koleːn-ːnɔk, koleːn-ːnek]
• “In a series of chi-square tests, we found that consultants who gave [haːdɛl-ːnɔk] were no more likely to give [koleːnnɔk] than consultants who gave [haːdɛl-ːnek]. We obtained similar results for all other pairs where enough data were available for testing.”

28. An informal observation

• Speakers who make such a random decision often feel that their choice was correct.
• [Try this yourself on past tense of spling]
• Psychologists might call this “self-priming”.

---
12 For example: series of questions: “how do you say this word? … how should this word be said?”, total cases of difference.