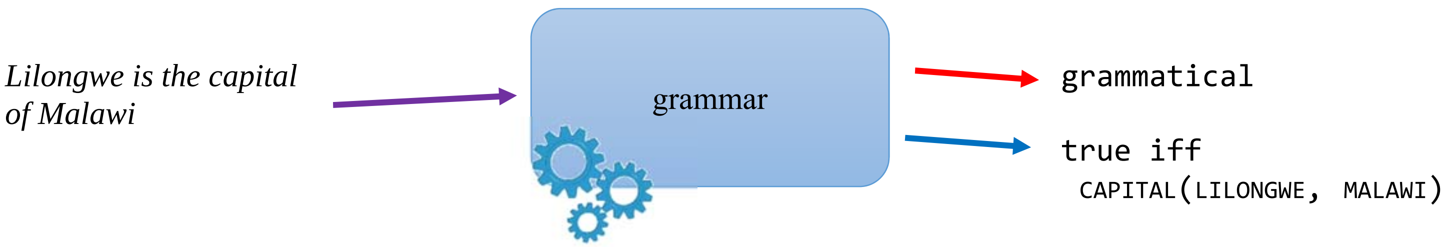
Class 2 (Week 1): SPE review, maybe expansion conventions

0. Business items

- How is Perusall going?
- Anything else?

1. Review of last time

- A **descriptively adequate** grammar not only assigns correct grammaticality judgments and truth conditions, but does so in a way that captures the significant, psychologically real generalizations.



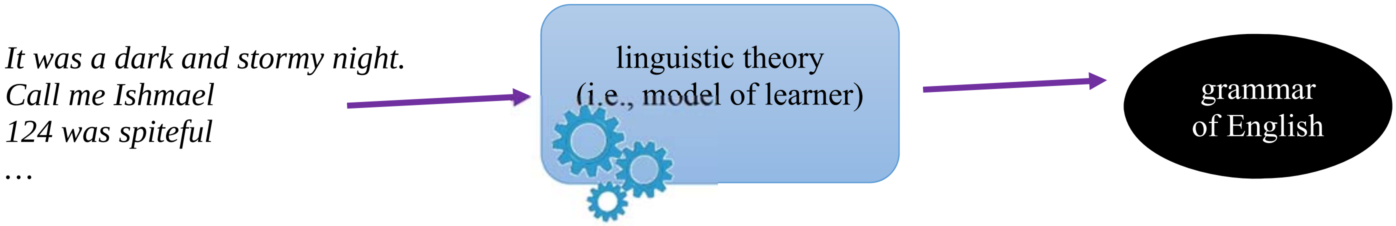
2. Why is it hard to develop a <u>descriptively</u> adequate grammar in phonology?

- Well, first—what does it even mean for a machine like the above to "capture a generalization"?
  - If we want to check whether it "captures" the English generalization that [-ɨz] is the plural after sibilants…ideas? Let me wait for at least 3 hands up.

- If a speaker already knows a word, feeding it into the machine is uninformative!
  - Known words don't tell us anything about what generalizations the speaker has learned—they may have simply memorized those words.)
- Constructing novel phonological situations to put speakers in is a challenge.
  - Contrast this with syntax, where it's easier to construct sentences that—presumably—the speaker has not encountered before.
- We often can't be sure that these novel situations really test what we want them to test.
- In 200A, we'll mostly ignore this problem and proceed as though generalizations that we notice in the data are real to speakers.
  - In 201A there will probably be a more emphasis on methods for determining which generalizations are real.

3. Why is it hard to develop an <u>explanatorily</u> adequate theory?

- Recall: an **explanatorily adequate linguistic theory** is a function that, given a realistic set of utterances, produces a descriptively adequate grammar..



- Suppose we could magically achieve descriptive adequacy for all real languages.
  - o That only tells us which generalizations people have extracted for existing sets of data
  - o We don't know what people *would do* if faced with a language with different generalizations

This problem tends to get ignored

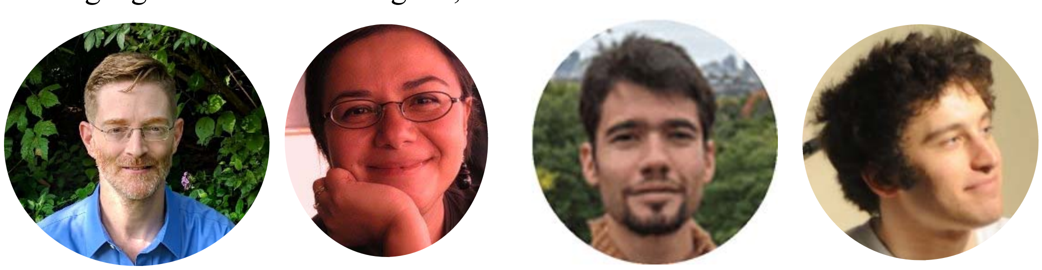
- To build our linguistic theory, we need to know which generalizations people can extract or tend to extract from all kinds of learning data, not just attested learning data.
  - o Are some generalizations preferred to others?
  - o Are there hard limits on learnability?

- In the English example…
  - o Suppose we're convinced by the wug test that English speakers' grammar includes the rule "use the [ɨz] form of the plural after sibilants".

  → Exposed to the English data, learners choose a grammar with that rule

  - o But we still know nothing about the learnability of "use the [ɨz] form of the plural after **non-**sibilants".
    - ▪ If the data had somehow reflected this rule instead, would children be able to learn it?

- Again, this won't be our focus this quarter, but some interesting things you could read:
  - o Becker, Ketrez & Nevins 2011 and Becker, Nevins & Levine 2012 tackle this problem in a very interesting way, by comparing potential generalizations that exist within the same language—Turkish and English, here.



- Bowers 2012 argues that a sudden, one-generation change in Odawa happened because the data changed into something that children couldn't learn.

So much for big picture. Now let's get into the nuts and bolts of what one theory does and predicts.
Material that's new to you: you can focus on the mechanics
Material that's familiar to you: you can focus on the theoretical predictions

SPE Chomsky & Halle 1968 rule notation review
SPE = *The Sound Pattern of English*

4. An example: SPE's main stress rule (p. 240)—let's just admire it for a minute

$$V \rightarrow [1\text{ stress}] \;/\; \Big[\, X \_\_ C_0 \left(\begin{bmatrix} -\text{tense} \\ \gamma\text{stress} \\ V \end{bmatrix} C_0^1 \left(\begin{bmatrix} \alpha\text{voc} \\ \alpha\text{cons} \\ -\text{ant} \end{bmatrix} \right) \right)$$

$$/ \_\_ \left\langle \left(\begin{Bmatrix} (fik)At \\ [+D]C_0 \end{Bmatrix} \right) \begin{Bmatrix} <_1 +C_0 >_1 \begin{bmatrix} -\text{stress} \\ -\text{tense} \\ -\text{cons} \end{bmatrix} [+\text{cons}]_0 \\ <_1 \begin{bmatrix} -\text{seg} \\ <_2 -\text{FB} >_2 \end{bmatrix} >_1 C_0\, [\beta\text{stress}]\, C_0 <_2 V_0 C_0 >_2 \end{Bmatrix} \right\rangle \Big]_{<NSP<_1 VA>_1>}$$

Conditions: $\beta = \begin{Bmatrix} 1 \\ 2 \end{Bmatrix}$

$\gamma \leq 2$ [in another version, says γ is 2 or weaker]

X contains no internal #

(Not much is said in SPE about these "conditions", except that they are truth-functional. It makes a big difference to the theory's computational power what restrictions we place on them.)

- Don't worry—you'll almost never encounter a rule this complicated!!!
- Let's step through the crucial elements of rule notation.

5. Features

- You can think of a feature as a function that takes a phone and emits a value like + or –
- E.g., *voice* assigns + to [b, d, m, o, a] and – to [p, t, s, h]
- Features usually have some phonetic definition, e.g…
    - *voice* (spoken languages) means "vocal folds are vibrating"
    - *spread* (sign languages) means "fingers are abducted, away from middle finger" (Ormel et al. 2017)
- So what does [+voice] means? As we'll see, it depends on where it appears in a rule

6. A → B / X __ Y

Example: $\begin{bmatrix}\text{+syll}\\ \text{–low}\end{bmatrix}$ → [+high] / __ CC#

➢ means "*XAY* is rewritten as *XBY*", or, to put it another way, "*A* is rewritten as *B* when preceded by *X* and followed by *Y*".

- A is the **affected segment, focus**, or **target** of the rule.
- B is the **structural change** that the rule requires
- X__Y is the **context** for the rule
- XAY is the **structural description**

We'll use *A*, *B*, *X*, and *Y* to stand for these positions throughout this handout.

7. *Something we'll skip, but for your reference*: A → B / X __ Y / P__Q

➢ Means "*PXAYQ* is rewritten as *PXBYQ*".

- I.e., A → B / PX __ YQ.
    - Except that ordering for "expansion conventions" (which we haven't discussed yet) is affected—see SPE pp. 72-77.

8. Left side of the arrow, "A"

A can be a feature matrix or Ø.

➢ <u>If *A* is a feature matrix</u>, like $\begin{bmatrix}\text{+syllabic}\\ \text{–low}\end{bmatrix}$, then the rule looks for any segment that is **nondistinct** from that matrix.

- Two feature matrices are **distinct** iff there is some feature F whose value is different in the two matrices.

    ? Which of the following are distinct from $\begin{bmatrix}\text{+syll}\\ \text{–low}\end{bmatrix}$?

    A: $\begin{bmatrix}\text{+syll}\\ \text{–low}\\ \text{+round}\\ \text{+back}\end{bmatrix}$ B: $\begin{bmatrix}\text{–low}\\ \text{–round}\end{bmatrix}$ C: $\begin{bmatrix}\text{–syll}\\ \text{–low}\\ \text{+high}\end{bmatrix}$

    Zoom poll

    - This means that if *A* doesn't mention some feature F, it doesn't care about it—that part of the rule matches segments that are +F, or –F, or even fail to have a value for F.

- Sometimes, if *A* is meant to pick out a single phone, we use a phonetic symbol instead:

  u → [–high] / __ (C)#

  - This is a good idea for readability, but in order to determine how long the rule is (e.g., if you think learners prefer short rules), you'd have to expand the IPA symbol into a feature matrix.

Annotation poll before we answer the next question**:** put a stamp next to whether you want a quick review of the features [high/low/front/back/round]

yes, I want a quick review no, I don't need a review

- What's the smallest feature matrix that "u" could abbreviate if the language's vowel inventory is *i*, *a*, *u*? If it's *i*, *a*, *u*, *o*? If it's *i*, *y*, *a*, *u*, *o*?

  Write down your answer, then I'll ask for **raised hands**

- Sometimes we also use C to abbreviate [–syllabic] or V to abbreviate [+syllabic].
  - Again, this is good for readability.
  - Be careful when *you* read, though, because some authors, following SPE, use C and V to abbreviate {[–vocalic], [+consonantal]} and [+vocalic, –consonantal].

➢ If *A* is Ø, you've got an insertion rule (the idea is that insertion changes "nothing" into something):

  Ø → i / C __ C#

- Why don't we use the empty matrix [ ] instead of Ø? Take 2 minutes to think about this alone, then I'll give you 2 minutes to discuss it with a partner in a break-out room. If you're stuck, try applying the rule [ ] → i / C__C# to the word /potek/ and the word /bamk/.

9. An unsolved issue: underspecified targets (if we have time)

- Imagine a rule like $\begin{bmatrix}\text{+coronal}\\ \text{–voice}\end{bmatrix} \rightarrow Ø\ /\ \_\_\ \#$
  - o And imagine we've decided that sonorants in the language in question are underlyingly underspecified for [voice] (meaning they have no value for this feature—some later rule will fill in their voicing values).
  - o E.g., feature matrix for /n/ doesn't contain any kind of [voice], either [+voice] or [–voice].

? How should the rule apply to /bil/ according to our definitions?
  A: produce [bi]
  B: produce [bil]

Zoom poll

? Does this seem right? Answer with thumbs-up or thumbs-down emoji and I'll call on one or more people to make their case.

- There's an inconclusive discussion on pp. 382-389 of SPE about whether we should...
  - o change the definition of when a rule is applicable so that nondistinctness isn't enough
  - o or impose a condition that segments always have to be specified for all the features that a rule's structural description mentions, by the time the rule applies
  - o or impose conditions on lexical entries that will rule out some of these cases

In practice, this won't come up much. If it does, you'll need to decide how the rule should apply and **be explicit** about your decision.

10. After all that, we're finally ready for… the right side of the arrow, "B"

- *B* also can be a feature matrix or Ø. But, it is totally different from *A*—it does not pick out a set of units!! Instead, it prescribes a set of changes to apply.

➢ <u>If *B* is a feature matrix</u>, then any of the affected segment's features that are mentioned in *B* are changed to the value given in *B*. *All other features are left unchanged*.

  ? What does $\begin{bmatrix}\text{+syll}\\ \text{–low}\end{bmatrix} \rightarrow$ [+high] do to [o]? To [u]?
  - o If I were the boss, there would be a whole different notation, like

$$\begin{bmatrix}\text{+syll}\\ \text{–low}\end{bmatrix} \rightarrow \|\ \text{+high}\ \| \qquad \text{(I use this in my undergraduate courses)}$$

➢ <u>If *B* is Ø</u>, then the segment that *A* matched is deleted.

C → **Ø** / C__# (why not [ ]?)

- Again, we sometimes use an IPA symbol as an abbreviation for all the feature changes necessary to change anything that could match *A* into the desired *B*:

  $\begin{bmatrix}\text{+syll}\\ \text{–low}\end{bmatrix} \rightarrow \text{i / \_\_ \#}$

  ? What does the "i" above abbreviate if the language's vowel inventory is *i*, *a*, *u*? If it's *i*, *a*, *u*, *o*? If it's *i*, *y*, *a*, *u*, *o*?

  - If *A* is Ø, then the phonetic symbol for B abbreviates the features needed to pick it out of the language's phoneme inventory: Ø → i / C __ C#

11. Redundancy

➢ The claimed principle that shorter rules are preferred by learners over longer rules (which we'll get to later) means that unnecessary features should be eliminated from *A* and *B*.

  ? What is suboptimal about each of the following rules? Take a minute to jot something down

  $\begin{bmatrix}\text{+syll}\\ \text{–round}\end{bmatrix} \rightarrow \begin{bmatrix}\text{+syll}\\ \text{–high}\end{bmatrix}$

  $\begin{bmatrix}\text{+syll}\\ \text{–round}\end{bmatrix} \rightarrow [\text{+round}]$

  $\begin{bmatrix}\text{+nas}\\ \text{+voice}\end{bmatrix} \rightarrow [\text{+anterior}]$

(assume the phoneme inventory of English for this last rule)

12. Right side of the slash (context), "X__Y"

- *X* and *Y* are strings made up of
  - o feature matrices
  - o IPA symbols, which abbreviate feature matrices
  - o the boundary types # and +, which in SPE also abbreviate feature matrices
  - o at their outside edges, category boundaries

➢ <u>Feature matrices</u> in *X* and *Y* match segments in the same way that *A* does (i.e., they match a segment if not distinct from it). Phonetic symbols also work the same way

- <u>Boundaries</u>, # (word boundary) and + (morpheme boundary), are treated in SPE as feature matrices that happen to be [–segmental]:

  $\# \text{ is } \begin{bmatrix} -\text{seg} \\ -\text{FB} \\ +\text{WB} \end{bmatrix} \qquad + \text{ is } \begin{bmatrix} -\text{seg} \\ +\text{FB} \\ -\text{WB} \end{bmatrix}$

  [FB] is "formative (roughly, morpheme) boundary" and [WB] is "word boundary"

  - There are some complications about #: in SPE, it's not exactly equivalent to the place where you'd write a space in ordinary writing, i.e. the boundary between syntactic words.
  - SPE also proposes a third boundary type, =, which has the features $\begin{bmatrix} -\text{seg} \\ -\text{FB} \\ -\text{WB} \end{bmatrix}$ and is more or less the boundary between nonproductive or nontransparent affixes and stems (e.g., English *per=mit*). We won't use this one much.

- The term 'unit' is used in SPE to refer to all feature matrices, including true segments and boundaries.

- <u>Category boundaries</u> (labeled brackets) like $]_{\text{Noun}}$ and $_{\text{Verb}}[$ can also be used, but **only at the edges** of *X*__*Y* (and if both edges have labeled brackets, the labels have to match):

  / __ VC#$]_{\text{N}}$

  - By convention, this can be abbreviated as / __ VC$]_{\text{N}}$

    - Discuss: What is the theoretical claim that Chomsky & Halle are making by imposing this only-at-the-edges condition?

13. Nondistinctness of strings

Here's how we extend the definition of nondistinctness from pairs of units to pairs of strings:

- *X* (or *Y*) matches (is nondistinct from) some substring *M* of a form iff *X* and *M* have the same number of units *n*, and the i^{th} unit of *X* matches (is not distinct from) the i^{th} unit of *M* for all $1 \leq i \leq n$.

14. + is special

➢ If + is included in *X* and *Y*, then it is required

- V → Ø / __+VC does not apply to *ibauk*, because V+VC does not match any substring of it.

- But—this is the special part—extra +s in the form are always OK: V → Ø / __VC does apply to *iba+uns*,

    - because "VVC" matches any of { VVC, V+VC, VV+C, V+V+C}.
      A *B* *C* *D*

    ? Which version of the rule is matching here?

Zoom poll

doesn't work this special way; it works like any other feature matrix.

15. Basic rule application

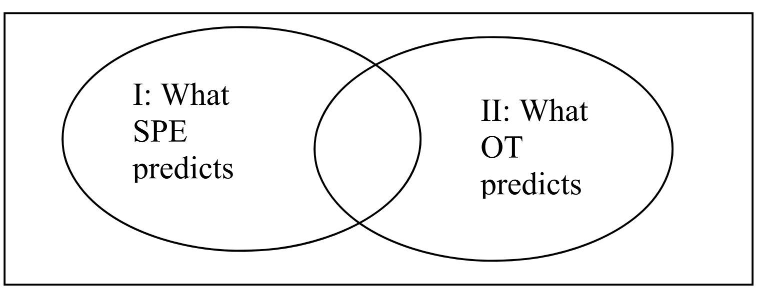
➢ A rule applies to a form if the form contains a string that is nondistinct from *XAY*.

? What if *X* or *Y* doesn't appear in the rule (A → B / __ Y or A → B / X __)?

16. Summary of SPE review

- We've started going into excruciating detail about how a seemingly simple theory works—why?
    - In the past, you've probably been taught a theory of convenience that worked well for the course material.
        - It may have cobbled together elements of various proposals, and left various aspects of its implementation vague.

    - Here we're going to try to be very explicit about what are our 2 base theories and what constitutes a departure from them.



Are there real cases that are
in I but not II?
in II but not I?

Next time

- You may recall seeing symbols like ( ) { } < > * ^{0} and others in rules, and treating them as convenient abbreviatory conventions. We'll review these symbols and see how SPE takes them seriously as theoretical claims.

Students' to do list

- Finish first **reading** (portions of K&K ch. 2, 3, 9), and annotate using instructions on Perusall. Due tomorrow night by 9 PM (Pacific), if possible.

References

Becker, Michael, Nihan Ketrez & Andrew Nevins. 2011. The surfeit of the stimulus: analytic biases filter lexical statistics in Turkish laryngeal alternations. *Language* 87(1). 84–125.

Becker, Michael, Andrew Nevins & Jonathan Levine. 2012. Asymmetries in generalizing alternations to and from initial syllables. *Language* 88(2). 231–268. https://doi.org/10.1353/lan.2012.0049.

Bowers, Dustin. 2012. *Phonological restructuring in Odawa*. UCLA master's thesis.

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

Ormel, Ellen, Onno Crasborn, Gerrit Kootstra & Anne de Meijer. 2017. Coarticulation of Handshape in Sign Language of the Netherlands: A Corpus Study. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 8(1). 10. https://doi.org/10.5334/labphon.45.