Class 1 (Week 1 Tues.): Deep into SPE

To do for next time
- Check out course web page (on my own web page), especially feature links
- Visit CCLE page, where readings can be downloaded
- Enroll! If you need a PTE number from me, give me your student ID number

Overview: I know you're familiar with rule-based phonological theories, but I want us to dig deeply into the inner workings, be a bit more formal, and (on Thursday) think about how to make decisions as to how those innards should work.

1. An example: SPE's (Chomsky & Halle 1968) main stress rule for English (p. 240)

$$V \rightarrow [1\ \text{stress}] \,/\, [\ X\_\_C_0 \left(\begin{bmatrix} -\text{tense} \\ \gamma\text{stress} \\ V \end{bmatrix} C_0^1 (\begin{bmatrix} \alpha\text{voc} \\ \alpha\text{cons} \\ -\text{ant} \end{bmatrix}) \right)$$

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

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 panic—you'll almost never encounter a rule this complicated.

2. A → B / X __ Y

Example: $\begin{bmatrix} +\text{syll} \\ -\text{low} \end{bmatrix} \rightarrow [+\text{high}]$ / __ 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**

3. *Something we'll skip*: 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.

4. Left side of the arrow

A can be a feature matrix or Ø.

- If *A* is a feature matrix, like $\begin{bmatrix}+\text{syll}\\-\text{low}\end{bmatrix}$, 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 $\begin{bmatrix}+\text{syll}\\-\text{low}\end{bmatrix}$? $\begin{bmatrix}+\text{syll}\\-\text{low}\\+\text{round}\\+\text{back}\end{bmatrix}$, $\begin{bmatrix}-\text{low}\\-\text{round}\end{bmatrix}$, $\begin{bmatrix}-\text{syll}\\-\text{low}\\+\text{high}\end{bmatrix}$
  - 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 sound, we use an IPA (or other transcription system) symbol instead:

  u → [–high] / __ (C)#

  - This is a good idea for readability, but in order to determine how long the rule is (for purposes of using a length-based evaluation metric to compare grammars), you'd have to expand the IPA symbol into a feature matrix.

  - 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*?

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

- 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?

5. An unsolved issue: underspecified targets

- Imagine a rule like $\begin{bmatrix}+\text{coronal}\\-\text{voice}\end{bmatrix} \rightarrow Ø$ / __ #
  - And imagine we've decided that sonorants in the language in question are underlyingly underspecified for [voice] (some later rule will fill in their voicing values).
  - 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? Does this seem right?
- There's an inconclusive discussion on pp. 382-389 of SPE about whether we should...
  - change the definition of when a rule is applicable so that nondistinctness isn't enough
  - 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
  - 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.

6. Right side of the arrow

B also can be a feature matrix or Ø

- If *B* is a feature matrix, 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]?

- If *B* is Ø, 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*:

  $\begin{bmatrix}+\text{syll}\\-\text{low}\end{bmatrix} \rightarrow$ 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 IPA symbol for B abbreviates the features needed to pick it out of the language's phoneme inventory: Ø → i / C __ C#

7. Redundancy

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

- What is suboptimal about each of the following rules?

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

$$\begin{bmatrix}+\text{nas}\\ +\text{voice}\end{bmatrix} \rightarrow [+\text{anterior}] \qquad \text{(assume the phoneme inventory of English)}$$

8. Right side of the slash (context)

X and *Y* are strings of *A*-like objects:

- feature matrices
- IPA symbols, which abbreviate feature matrices
- the boundary types # and +, which in SPE also abbreviate feature matrices

plus, at their outside edges, category boundaries

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

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

  $\#$ is $\begin{bmatrix}-\text{seg}\\ -\text{FB}\\ +\text{WB}\end{bmatrix}$ $+$ 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.
  - 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*). You won't see this one much.

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

- You guessed it—you can also write a rule where *A* or *B* is a boundary symbol, though this might be a bad idea...

- Category boundaries (labeled brackets) like $]_{Noun}$ and $_{Verb}[$ 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#$]_N$

  - By convention, this can be abbreviated as / __ VC$]_N$

We extend the definition of nondistinctness from pairs of units to pairs of strings exactly how you'd think:

- *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 $1 \leq i \leq n$.

9. + is special

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

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

- But—this is the special part—extra +s in the form are always OK: V → Ø / __VC does apply to *iba+uns*,
  - because "__VC" matches any of { __VC, __+VC, __V+C, __+V+C}.

- Which version of the rule is matching here?

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

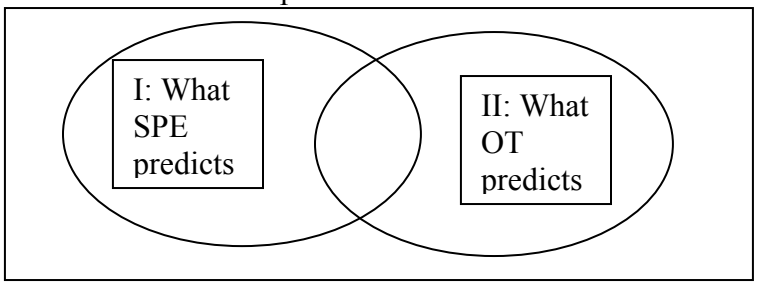
10. Basic rule application

- 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 __)?

11. Interim stock-taking—why are we doing this?

- We've gone into excruciating detail about how a seemingly simple theory works—why?
  - In your introductory courses, you probably were 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 the theory's workings unspecified
  - 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, productive cases that are in I but not II? In II but not I?

Expansion conventions, part I—the most common ones

12. Overview

- You may recall seeing symbols like ( ) { } < > * C_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.

13. Individual rules vs. rule schemata

- Devices like parentheses, curly brackets ("braces"), and angle brackets are used to collapse related rules into a single *rule schema* (whose length is shorter = cost is lower).
- Rather than adjusting the definition of nondistinctness, SPE gives *expansion conventions* to turn those schemata into lists of rules that can then be applied using the simple definition of nondistinctness.

14. Lowercase Greek letters

➢ Variables that stand for +, –, or whatever values the theory says some feature can take (could be 1,2,3 for some features—can you think of any good candidates?).

C → [αvoice] / [αvoice] __ [αvoice] expands into

C → [+voice] / [+voice] __ [+voice]
C → [–voice] / [–voice] __ [–voice]

15. Parentheses

➢ Used to indicate optionality.

- For example, the rule schema V → Ø / __(V)C# is expanded into these two rules:

V → Ø / __VC#
V → Ø / __C#

o Do you ever need parentheses in a <u>feature matrix</u>? Consider both *A/X/Y* and *B*.

- The rules that a parenthesis-schema expands into are ***disjunctively*** *ordered*.
  - Informally: you try to apply the first one
  - if its structural description is met, you apply that first rule and don't try any of the other rules from the same schema
  - if not, move on to the next rule and proceed in the same fashion.
- In other words, you never apply two rules of the same parenthesis schema to a single word.

o How does the rule above apply to /bauk/?

(This is a bit too crude, because it doesn't give the right result for cases where different rules of a schema apply to different <u>parts</u> of a word—in those cases, we want multiple rules of the schema to apply to the same word, just in different places. We'll come back to that in Week 3.)

16. Braces, a.k.a. curly brackets

- Used to indicate multiple possibilities

- For example, the rule schema $\begin{Bmatrix} i \\ o \end{Bmatrix} \rightarrow Ø$ / __V is expanded into these two rules (in this order):

  i → Ø / __V
  o → Ø / __V

- Rules from the same curly-bracket schema apply ***conjunctively*** (apply the first one, then the second, etc.)
  - Thanks to Patrick Jones for de-confusing me on this!
  - SPE gives an example where you do actually need to apply multiple sub-rules (p. 341)—can you devise an input for the rule above where conjunctive and disjunctive order would produce different results?

- Some phonologists think that curly brackets are so powerful that the theory shouldn't allow them—that resorting to them is an admission of failure (either of the analyst or of the theory).

Next time:
- The less-common expansion conventions: super- and subcripts, parentheses with star, angled brackets; plus transformational rules
- According to SPE, how does a learner choose a grammar?
- According to SPE, how do we make decisions about the theory? E.g.,
  - Is it good or bad to allow the same Greek-letter variable to apply to two different features in a rule?
  - Should the expansion conventions for curly brackets produce conjunctive or disjunctive ordering?
- This will lead us to the question of whether constraints are a good idea, and how they could be incorporated into the theory.
- And we'll see some actual data :)

References

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