Class 2: SPE rule notation review

To do for next time
- Read K&K excerpts of ch.3 (pp. 45-62) and ch. 9 (pp. 331-339), and turn in study questions (see my web page) on Thursday—after this week, just one set of readings per week.
- First assignment (Malagasy) is due Friday, Oct. 1.

1. An example: SPE's main stress rule (p. 240)

$$\text{V} \rightarrow [1\ \text{stress}] \,/\, [\ X\_\_\text{C}_0 \left(\begin{bmatrix} -\text{tense} \\ \gamma\text{stress} \\ \text{V} \end{bmatrix} \text{C}_0^1 (\begin{bmatrix} \alpha\text{voc} \\ \alpha\text{cons} \\ -\text{ant} \end{bmatrix}) \right)$$

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

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.

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*

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

3. A rarity: 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 applying a length-based evaluation metric), 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. 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 abbreviates the features needed to pick it out of the language's phoneme inventory: Ø → i / C __ C#

6. Redundancy

The claimed principle that shorter rules are preferred by learners over longer rules means that unnecessary features should be eliminated from *A* and *B*.

- What is wrong with each of the following rules?

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

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

7. Right side of the slash (context)

X and *Y* are strings made up of

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

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

([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.

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.

Category boundaries (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]$_{\text{N}}$.

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 $1 \le i \le n$.

8. + is special

If + is included in *X* and *Y*, then it is required: V → Ø / __+VC does not apply to *iba<u>uk</u>*, because +V does not match any substring of *ibau*.

But extra +s in the form are always OK: V → Ø / __VC does apply to *iba+<u>uns</u>*, because "__VC" matches any of { __VC, __+VC, __V+C, __+V+C}.

- Which version of the rule is matching here?

does not work this way; it works like any other feature matrix.

9. 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 __)?

Expansion conventions

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.

10. 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]

11. Parentheses

Used to indicate optionality.

For example, the rule schema V → Ø / __(V)C# is expanded into these two rules (in that order—but we'll come back to that another day):

V → Ø / __VC#
V → Ø / __C#

- Do you ever need parentheses in a feature matrix?

The rules that a schema expands into are *disjunctively ordered*. That means, informally, that 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 schema to a single word.

- 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 parts 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 another day too.)

12. Braces, a.k.a. curly brackets

Used to indicate multiple possibilities

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

$\text{i} \rightarrow \text{Ø} \ / \ \_\_\text{V}$
$\text{o} \rightarrow \text{Ø} \ / \ \_\_\text{V}$

- Can you imagine a way to translate parentheses into braces?
  Try it with $\text{V} \rightarrow \text{Ø} \ / \ \_\_\text{V(C)\#}$

13. Super- and subscripts

e.g., C_n means "*n* or more Cs" (most common is C_0)
V^m means "up to *m* Vs"
C_n^m means "anywhere from *n* to *m* Cs"

$\text{C} \rightarrow \text{Ø} / \_\_\text{C}_0\#$ = …
$\text{C} \rightarrow \text{Ø} / \_\_\text{CCCC\#}$
$\text{C} \rightarrow \text{Ø} / \_\_\text{CCC\#}$
$\text{C} \rightarrow \text{Ø} / \_\_\text{CC\#}$
$\text{C} \rightarrow \text{Ø} / \_\_\text{C\#}$
$\text{C} \rightarrow \text{Ø} / \_\_\#$

The tricky thing about this is that we apply the *longest* rule whose structural description matches.

- How would the schema above apply to /tabskt/?

14. Parentheses with star (But see discussion in upcoming Anderson reading)

(…)* means that the material in parentheses can occur zero or more times.

V → [+stress] / #C(VCVC)*__ expands to

V → [+stress] / #C__
V → [+stress] / #CVCVC__
V → [+stress] / # CVCVCVCVC __
etc.

With ()*, disjunctive ordering does *not* apply. Every version of the rule that can apply does apply—simultaneously.

- How would the stress rule above apply to /badupidome/?
- How would C → Ø/ __C*# apply to /tabskt/?

15. Angled brackets

Like parentheses, but when the optional information is in more than one place. A schema with angle brackets expands into two rules: the rule with the information in the angle brackets and the rule without that information.

C → Ø / V<C>__<C>V (silly example) expands to

C → Ø / VC__CV
C → Ø / V__V

- Expand the following schema and apply it to *putod*, *luged*, and *fesil*.

$$\begin{bmatrix} +\text{syll} \\ <+\text{back}> \end{bmatrix} \rightarrow [-\text{hi}] \,/\, \_\_\ C < \begin{bmatrix} +\text{syll} \\ +\text{back} \\ -\text{hi} \end{bmatrix} C >\#$$

You can also subscript angle brackets to show which ones go together:

C → Ø / V<$_1$C>$_1$__<$_2$s>$_2$<$_1$C>$_1$V<$_2$h>$_2$# (even sillier rule) expands to

C → Ø / VC__sCVh#
C → Ø / V__sVh#
C → Ø / VC__CV#
C → Ø / V__V#

16. Applying the SPE stress rule

- S means 'stem', P means 'prefix'
- [0stress] = stressless; [1stress] = primary-stressed; [2stress] = secondary-stressed; etc.
- Vowels in English *beat, bait, boat, boot* are [+tense]. Those in *bit*, *bet*, *lot*, *foot* are [–tense].

tense		**lax**	
spelling	**typical IPA**	**spelling**	**IPA**
beat	i	*bit*	ɪ
boot	u	*put*	ʊ
bait	e/ej/ei/eɪ	*bet*	ɛ
boat	o/ow/ou/oʊ		
bought	ɔ	*bat*	æ
		but	ʌ

- Divide into groups; I'll give each group one sub-part of the schema to apply to all of these words. Then, we'll see whose sub-part should take precedence for each word.

capital letter=tense vowel

1. [V imagine]V	2. [N consensus]N	3. [A repugn+ant]A
4. [V mAIntAIn]V	5. [N larynx]N	6. [A fratern+al]A
7. [V collapse]V	8. [N effigy]N	9. [A vertebr+al]A
10. [N javelin]N	11. [A person+al]A	12. [V inter=sect]V
13. [N arOma]N	14. [A desIr+ous]A	15. [V per=mit]V

- After Main Stress Rule, need to apply the Alternating Stress Rule:

$$V \rightarrow [1\ \text{stress}] / \_\_\ C_0\ (=)\ C_0VC_0\ [1\ \text{stress}]\ C_0\ ]_{NAV}$$

...with the convention: when a V is assigned [1stress] by rule, all other nonzero stresses are increased by 1 (p. 64)

17. [V im=plic+Ate]V	18. [N hurricAne]N
19. [V exorcIse]V	20. [N chickadEE]N

- Further complication when a word has multiple pairs of []: "**cyclicity**"

Take the innermost bracketed string, apply the rules to it, and erase its brackets. Repeat till all brackets gone.

- Let's try it for these items

21. [N [V per=mit]V]N *cf.* Kermit—*do they rhyme perfectly for you?*
22. [N mono [S graph]S]N
23. [N mono [N #genes+is#]N]N (watch out for the extra #s)
24. [N [V soliloquIz#]V ing]N
25. [N [P parel [Sle]S lo]P [S gram]S]N[1]

[1] When you plug in **[N parallélo grám]N**, the rule will re-primary stress the *é*, which demotes the stress of *grám* (=> **parallélogràm**). This is unlike what usually happens when a segment is already [αF] and a rule tries to make it [αF] (i.e., nothing changes). Another reason why stress doesn't look much like a feature...

17. Transformational rules

Useful for metathesis, coalescence…anything where more than one segment is affected at once.

In SPE, these were given in two parts:

Structural description: $\begin{bmatrix}+\text{syll}\\+\text{low}\end{bmatrix}_1 , \begin{bmatrix}+\text{syll}\\+\text{hi}\\\alpha\text{round}\end{bmatrix}_2$

Structural change: $1\ 2 \rightarrow \begin{bmatrix}1\\-\text{lo}\\+\text{long}\\\alpha\text{round}\\\alpha\text{back}\end{bmatrix}, \begin{bmatrix}2\\\emptyset\end{bmatrix}$

- What does this rule do?

- It may seem arbitrary to say that 1 changes and 2 deletes rather than the reverse. Try writing the rule the other way. Is one more concise?

We'll use a simplified notation instead that collapses the structural description and structural change:

$$\underset{1}{\begin{bmatrix}+\text{syll}\\+\text{low}\end{bmatrix}} \underset{2}{\begin{bmatrix}+\text{syll}\\+\text{hi}\\\alpha\text{round}\end{bmatrix}} \rightarrow \begin{bmatrix}1\\-\text{lo}\\+\text{long}\\\alpha\text{round}\\\alpha\text{back}\end{bmatrix}$$

- What's wrong with just saying the following?

$$\begin{bmatrix}+\text{syll}\\+\text{low}\end{bmatrix} \begin{bmatrix}+\text{syll}\\+\text{hi}\\\alpha\text{round}\end{bmatrix} \rightarrow \begin{bmatrix}-\text{lo}\\+\text{long}\\\alpha\text{round}\\\alpha\text{back}\end{bmatrix}$$

- Formulate a transformational rule for the metathesis seen in the warm-up problem from last week.

- In the directions for the warm-up problem, I said that you could use a rule of the form AB → BA, but that this was not quite adequate. Rewrite your transformational rule from above as a AB → BA rule. What's wrong with it?