Class 2 (Week 1 Thurs.): Deep into SPE, part II

To do for next time
- Prince & Smolensky study questions due Tuesday (get reading from CCLE; get study questions from course web page, to be found at my own web page)
- I'll post a homework problem (probably Palauan) tonight. Due Fri., Jan 16.

Overview: The less-common expansion conventions from Chomsky & Halle (1968). Making theoretical choices. The dawn of constraints. (Sorry, still only one chunk of real data.)

1. Super- and subscripts

➢ X_n^m means from *n* to *m* Xs
- C_n : "*n* or more Cs" (most common is $\mathbf{C_0}$)
- V^m : "up to *m* Vs"
- C_n^m : "anywhere from *n* to *m* Cs"

C → Ø/ __C_0# = …
C → Ø/ __CCCC#
C → Ø/ __CCC#
C → Ø/ __CC#
C → Ø/ __C#
C → Ø/ __#

- The tricky thing is that we apply the *longest* rule whose structural description matches (**disjunctive** ordering again).

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

2. Parentheses with star (but see discussion in Week 3 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 ( )*, there is yet a third type of ordering: **simultaneous**—identify targets of any sub-rule and change them simultaneously

o How would the stress rule above apply to /badupidome/?

o How would C → Ø/ __C*# apply to /tabskt/?

3. Angled brackets—we'll just skim this part, because they're rarely used

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

$C \rightarrow \text{Ø} / V\langle C\rangle\_\_\langle C\rangle V$ (silly example) expands to

$C \rightarrow \text{Ø} / VC\_\_CV$ *both*
$C \rightarrow \text{Ø} / V\_\_V$ *neither*

- An exercise for later if you like: expand the following schema and apply it to *putod*, *luged*, and *fesil*.

$$\begin{bmatrix} +\text{syll} \\ \langle +\text{back} \rangle \end{bmatrix} \rightarrow [-\text{hi}] / \_\_\ C \left\langle \begin{bmatrix} +\text{syll} \\ +\text{back} \\ -\text{hi} \end{bmatrix} C \right\rangle \#$$

- You can also <u>subscript</u> angle brackets to show which ones go together:

$C \rightarrow \text{Ø} / V \langle_1 C\rangle_1 \_\_ \langle_2 s\rangle_2 \langle_1 C\rangle_1 V \langle_2 h\rangle_2 \#$ (even sillier rule) expands to

$C \rightarrow \text{Ø} / VC\_\_sCVh\#$ *both the 1s and the 2s*
$C \rightarrow \text{Ø} / V\_\_sVh\#$ *just the 2s*
$C \rightarrow \text{Ø} / VC\_\_CV\#$ *just the 1s*
$C \rightarrow \text{Ø} / V\_\_V\#$ *neither*

4. Not an expansion convention, an extension of the theory's power: 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: $\underset{1}{\begin{bmatrix} +\text{syll} \\ +\text{low} \end{bmatrix}},\ \underset{2}{\begin{bmatrix} +\text{syll} \\ +\text{hi} \\ \alpha\text{round} \end{bmatrix}}$

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

- 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 too.

- You'll often see a simplified notation instead that collapses the structural description and structural change:

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

  - What's wrong with just saying this (same rule without the numerical indices):

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

  - Say you want to write a metathesis rule that changes *s*-stop into stop-*s*. What's wrong with writing $\text{s}\begin{bmatrix}-\text{sonorant}\\-\text{continuant}\end{bmatrix} \rightarrow \begin{bmatrix}-\text{sonorant}\\-\text{continuant}\end{bmatrix}\text{s}$?

  - Re-write the defective rule with transformational notation.

5. How does the learner choose a grammar?

- SPE proposed that if more than one grammar can generate the observed linguistic data, the learner must have some ***evaluation metric*** for choosing one.
- The evaluation metric tentatively proposed in SPE is brevity: learner chooses the grammar with the fewest symbols. (What about ties?)
- If that's right, and if we've got the notation right too, then you can tell which grammar, out of some set of candidate grammars, the learner would choose.
- More plausibly, we want to find independent evidence as to which grammar is right, and then make sure our theory explains how/why the learner chose that one—this is a lot harder!

6. *Example:* French elision/liaison (SPE p. 353 ff.)

- By the logic above, a theoretical innovation is held, in SPE, to be a good one if it allows <u>more-concise descriptions of attested/common phenomena</u> than of unattested/uncommon phenomena.

		obstruent- or nasal-initial	liquid-initial	vowel-initial	glide-initial
		/garson/ 'boy'	/livr/ 'book'	/ɛnfant/ 'child'	/wazo/ 'bird'
obstruent- or nasal-final	/pətit/ 'small'	pəti_ garsõ	pəti_ livr	pətit ãfã	pətit wazo
liquid-final	/ʃɛr/ 'dear'	ʃɛr garsõ	ʃɛr livr	ʃɛr ãfã	ʃɛr wazo
vowel-final	/lə/ 'the'	lə garsõ	lə livr	l_ ãfã	l_ wazo
glide-final	/parej/ 'similar'	parej garsõ	parej livr	parej ãfã	parej wazo

For the sake of reconstructing the argument, use the archaic feature [vocalic] and the still-current feature [consonantal]:

	vocalic	*consonantal*
obstruents and nasals	–	+
liquids	+	+
glides	–	–
vowels	+	–

- Propose rules to account for the C- and V- deletions, without using Greek-letter variables.

- Combine the rules into a schema, using Greek-letter variables

- Greek-letter variables don't allow us to compress these two rules:

$\begin{bmatrix} +\text{voc} \\ +\text{lo} \end{bmatrix} \rightarrow \text{Ø} \;/\; \_\_ \;\#\; [+\text{voc}]$ "low vowels delete before a vowel or glide"

$\begin{bmatrix} +\text{voc} \\ +\text{cons} \end{bmatrix} \rightarrow \text{Ø} \;/\; \_\_ \;\#\; [+\text{cons}]$ "liquids delete before a non-glide consonant"

With that in mind, how should the typology guide us in deciding whether to allow the same Greek-letter variable to apply to different features within a rule?

7. (*skip if no time*) Reasoning above relies on assumptions about linguistic typology:

- Assume a rule is cross-linguistically common only if it's favored by learners—i.e., learners tend to mislearn, in the direction of a more-favored grammar.
- Assume that learners favor short/simple/whatever rules.
- Therefore, rules that are cross-linguistically common should tend to be short.
- Therefore, our theory of rules, which determines what type of notation length is calculated on, should make common rules shorter than uncommon ones.
- Therefore, a theoretical innovation is good if it makes common rules shorter than uncommon ones.

=> We're not really using "short" (or "simple") in any fixed sense. Rather, we're tailoring the notation to make the rules that we think learners favor appear short. [And of course, that first assumption is questionable...]

This leads us into slippery territory in deciding whether shortness is the right criterion:

  - Are learners innately endowed with a certain notation, which they use to calculate grammar length? (i.e., shortness really is the evaluation criterion)
  - Or is it the case that learners employ some other evaluation metric entirely, but we've created a system of notation that makes goodness according to the real evaluation metric translate into shortness in our notation?

Something for you to think about, though no answers will be forthcoming: We've seen how to evaluate a particular description or even a theoretical innovation, given a framework like SPE.

- But how do you evaluate the framework itself—in particular, how can we evaluate a principle such as "if more than one grammar can generate the observed linguistic data, the learner chooses the grammar with the fewest symbols"?

8. Shortening the (previously) unshortenable: constraints

- Kisseberth (1970) introduced the following problem, using Yawelmani Yokuts as a case study.
- These rules can't be collapsed into a schema:

  ∅ → V / C __ CC
  C → ∅ / CC + __

- And yet, they seem to have something in common—can you guess what? It will help to invent a form that each rule can apply to and see what that rule does.

- Cases like this became known as *conspiracies*, and their widespread existence as the *conspiracy problem*.

9. Constraints as rule blockers

- Kisseberth proposes using a constraint to make the rules of Yawelmani Yokuts simpler:

Instead of $V \rightarrow \varnothing$ / V C $\underset{[-\text{long}]}{\_\_\_\_\_}$ C V

use $V \rightarrow \varnothing$ / C $\underset{[-\text{long}]}{\_\_\_\_\_}$ C subject to the constraint *CCC (or *{C,#}C{C,#})

  - The constraint can *block* the rule: the rule applies **unless** the result violates the constraint.

o Let's try to lay out, step by step, what an algorithm would have to do to implement the rule and its blocking constraint.

10. Constraints as rule triggers

- Kisseberth also proposes that constraints can *trigger* rules: a rule applies **only if** it gets rid of a constraint violation.

o What happens if the rule $\varnothing \rightarrow$ i (context-free) applies only when triggered by the constraint *CC? Let's try to break this down into simple steps too.

11. Why is this good?

- In a system without constraints, these two grammars have equal length and should be equally plausible:

Yokuts

C → ∅ / CC + __
∅ → i / C __ CC
V→ ∅ / V C ___ C V
[–long]

imaginary and implausible

C → ∅ / CV + __
∅ → i / V __ CC
V→ ∅ / V C ___ C C
[–long]

- But in Kisseberth's system the Yokuts grammar is shorter than the "implausible" grammar

Yokuts

C → ∅ / + __
∅ → i
V→ ∅ / C ___ C
[–long]
*{C,#}C{C,#}

imaginary and implausible

C → ∅ / CV + __
∅ → i / V __ CC
V→ ∅ / V C ___ C C
[–long]

- If we're right that the language on the right is less plausible than Yokuts, Kisseberth's theory is better because it captures that difference.

12. Rule+constraint theories

- Many more conspiracies were identified, giving rise to more constraints.
- People liked constraints, because they solved the conspiracy problem and also gave theoretical status to the idea of "markedness", which had been floating around.
    - Everyone knew languages don't "like" CCC sequences (they are "marked"), but this was not directly encoded in grammars until constraints like *CCC (or a syllable-based equivalent) came along.
- On the other hand, using constraints introduces some problems into the theory—I invite you to invent cases!
    - What if there's more than one rule (or more than one way of applying a single rule) that could fix a constraint violation?
    - What if one rules makes a constraint violation worse, but feeds another rule that makes it better?
    - What if one constraint wants to trigger a rule, but a different constraint wants to block it?

Next time: Deep into OT and how it deals with the above problems.

References

Chomsky, Noam & Morris Halle. 1968. *The Sound Pattern of English*. Harper & Row.
Kisseberth, Charles. 1970. On the functional unity of phonological rules. *Linguistic Inquiry* 1. 291–306.