Class 5: Rule+constraint theories; taking stock

To do
- Forget about Sommerstein reading on syllabus—get started on Prince & Smolensky excerpt instead (SQs due Thursday)
- Metaphony assignment is due Tuesday

Overview: First we'll try to make the framework for rule/constraint interaction more explicit. Then, we'll turn to some big-picture issues (reviewing levels of adequacy) and consider some of the questions you've been asked to ponder in your study questions.

1. Implementing triggering: Sommerstein's (1974)[1] proposal (underlining is mine)

"A P-rule R is positively motivated with respect to a phonotactic constraint C just in case the input to R contains a matrix or matrices violating C AND the set of violations of C found in the output of R is null or is a proper subset of the set of such violations in the input to R." (p. 74)

"A rule, or subcase of a conspiracy, positively motivated by phonotactic constraint C does not apply unless its application will remove or alleviate a violation or violations of C." (p. 75)

Later modified: "a rule applies if its application will remove or alleviate a violation of AT LEAST ONE of its motivating constraints" (p. 87)

2. Latin example (Sommerstein p. 87)

genitive sg.	*nominative sg.*	*UR*
lakt-is	lak	/lakt/
kord-is	kor	/kord/

$$\begin{bmatrix}+\text{continuant}\\ \langle -\text{voice}\rangle\end{bmatrix} \rightarrow \emptyset \,/\, \begin{bmatrix}+\text{consonantal}\\ \langle -\text{sonorant}\\ -\text{continuant}\rangle\end{bmatrix} \_\_\ \#$$

positively motivated by two constraints that are **surface-true** in the language:

"no final voiced in cluster" $\quad * [+\text{consonantal}] \begin{bmatrix}+\text{consonantal}\\ +\text{voice}\end{bmatrix} \#$

"final obst. restrictions" $\quad \text{if} \underset{1}{\begin{bmatrix}-\text{sonorant}\\ \langle -\text{continuant}\rangle\end{bmatrix}} \underset{2}{[-\text{sonorant}]} \ \#\ \text{then 2 is} \begin{bmatrix}+\text{coronal}\\ \langle +\text{continuant}\rangle\end{bmatrix}$

paraphrase: if you have a word-final cluster of two obstruents, the second must be coronal, and if the first is a stop the second must be a fricative (/st ps ks/ are allowed).

- With those constraints, how much can we simplify the deletion rule?

[1] Sommerstein, Alan H. (1974). On Phonotactically Motivated Rules. *Journal of Linguistics* 10: 71-94.

A derivation might look like this:

	/lakt/	/kord/	/re:ks/
violates no final voiced in cluster*?*	no	yes	no
violates final obstruent cluster restrictions*?*	yes	no	no
if so, tentatively apply deletion			NA
is the violation alleviated/eliminated?			NA
if so, accept the change (else don't)			NA

we'll have to fill in this part according to how we formulate the rule.

3. Multiple available repairs

Imagine a hypothetical language, "Matin", that is just like Latin except that it has this rule too:

[] → [–voice]

- How does our derivation change? Do we need to add more information to the grammar?

Imagine a hypothetical language, "Natin", that is just like Latin except that it has this rule too:

[] → [+continuant]

- How does our derivation change? Do we need to add more information to the grammar?

4. Partial violation, violation alleviation

Under Sommerstein's conception, a constraint doesn't have to be surface-true to be part of the grammar.

> The DEGREE OF VIOLATION $V_{M,C}$ to which a matrix M violates a phonotactic constraint C is equal to the cost of the minimal structural change necessary to turn M into a matrix satisfying C.
> The application to a matrix M of operation A ALLEVIATES a violation in M of phonotactic constraint C just in case the output M´ of such application is such that $0 < V_{M',C} < V_{M,C}$ (p. 76)

- Can you invent a case where a violation could be alleviated without being eliminated? (It's OK if it's silly—it's hard to think of plausible cases, and Sommerstein himself introduces this idea just to keep the possibility open, not because he has any data that require it.)

5. Implementing blocking: taking inspiration from Sommerstein...

A P-rule R is negatively motivated with respect to a phonotactic constraint C just in case the tentative output of R contains a matrix or matrices violating C AND the set of violations of C found in the input to R is null or is a proper subset of the set of such violations in the tentative output of R.

A rule that is negatively motivated by phonotactic constraint C does not apply if its application will create or worsen a violation or violations of C.

The application to a matrix M of operation A worsens a violation in M of phonotactic constraint C just in case the output M´ of such application is such that $V_{M',C} > V_{M,C}$

6. What a derivation might look like

syncope rule $V \rightarrow C\_\_C$

cluster constraint $* \begin{Bmatrix} \# \\ C \end{Bmatrix} C \begin{Bmatrix} \# \\ C \end{Bmatrix}$

	/abito/	/ildoku/	/uda/
tentatively apply syncope	(abto)	(ildku)	NA
does this create/worsen violation of cluster constraint?	no	yes	NA
if not, accept the change (otherwise reject)	abto	ildoku	NA
	[abto]	[ildoku]	[uda]

7. Blocking vs. triggering: Myers 1991's [2] persistent rules

Zulu: prenasalized affricates, but no prenasalized fricatives.
We might propose a constraint:[3]

$$* \begin{bmatrix} +\text{continuant} \\ +\text{nasal} \end{bmatrix}$$

Here is a prefix that creates prenasalized consonants (p. 329):

singular	*plural*	
uː-bambo	izi^{m}bambo	'rib'
uː-p^{h}ap^{h}e	izi-^{m}pap^{h}e	'feather'
ama-t^{h}at^{h}u	ezi-^{n}tat^{h}u	'three'
uː-k^{h}uni	izi-^{ŋ}kuni	'firewood'

- Assume the underlying form of the prefix is /izin/. Formulate a prenasalization rule.

[2] Myers, Scott (1991). Persistent rules. *Linguistic Inquiry* 22: 315-344.
[3] Myers actually uses autosegmental representations.

Here's what happens when the prefix attaches to a fricative-initial stem:

singular	*plural*	
eli-ʃa	e-ⁿt͡ʃa	'new'
uː-fudu	izi-ᵐp͡fudu	'tortoise'
uː-sizi	izi-ⁿt͡sizi	'sorrow'
uː-zwa	izi-ⁿd͡zwa	'abyss'
uː-zime	izi-ⁿd͡zime	'walking staff'
uː-ɮubu	izi-ⁿd͡ɮubu	'groundnut'
uː-ʃikisi	izi-ⁿt͡ʃikisi	'quarrelsome person'

- What would happen if prenasalization were subject to blocking by the constraint above?

Myers proposes instead a "persistent rule"—it tries to apply at every point in the derivation, so that any time its structural description is created, it immediately gets changed.

$$\begin{bmatrix} \text{+nasal} \\ \text{+continuant} \end{bmatrix} \rightarrow \begin{bmatrix} \text{+delayed release} \\ \text{–continuant} \end{bmatrix}$$ i.e., nasal fricative → affricate

- Let's spell out what the derivation would look like.

- Can we recast this as a simpler rule that is triggered by the constraint?

Reflecting on big-picture issues

How do humans learn, store, and use linguistic sound patterns? Chomsky lays out a useful framework for investigating this question for language in general (see Chomsky 1965[4] pp. 25-27—but what it is below is an amalgam of various works, slightly simplified and colored by my own views).

[4] Chomsky, Noam (1965). *Aspects of the theory of syntax*. Cambridge: MIT Press.

8. Preliminaries

Let a **grammar** consist of (at least)[5]

- a function that labels any utterance as **grammatical** or **ungrammatical**. We can call such labelings **grammaticality judgments.**
- a function that assigns truth conditions to any utterance

The grammar might be implemented as a lexicon and a list of rules, or a set of constraints, or something else.

Let a **linguistic theory** be a function that, given a (finite) set of utterances (the **learning data**), produces a grammar.[6]

These functions should be accompanied by algorithms for calculating them.

Let's use a concrete example, English noun plurals.

cat	kʰæt	kʰæt**s**
sack	sæk	sæk**s**
dog	dɑg	dɑg**z**
grub	gɹʌb	gɹʌb**z**
dish	dɪʃ	dɪʃ**əz**
fudge	fʌd͡ʒ	fʌd͡ʒ**əz**
pea	pʰi	pʰi**z**
cow	kʰaʊ	kʰaʊ**z**
man	mæn	m**ɛ**n
foot	fʊt	f**i**t
wife	waɪf	waɪ**v z**
whiff	wɪf	wɪf**s**
...		

9. Observational adequacy

A grammar that accepts all the forms that a typical speaker would have been exposed to and assigns the right truth conditions to them (we don't care what the grammar says about other forms), is an **observationally adequate grammar**.

Note that there are infinitely many observationally adequate grammars for any (finite) set of learning data (why?).

---

[5] We probably want the grammar to do much more. It could, given an utterance, return a gradient "goodness score" rather than a simple binary judgment. Given one utterance and some instruction, it could return some other utterance (e.g., *cat* + PLURAL = *cats*). And of course there's a lot more to meaning than truth conditions. (Chomsky also requires a grammar to assign a structural description to an utterance, but I wonder if this is begging the question: the structural description can be used to explain more-observable properties of a sentence like its truth-conditions, but we don't know *a priori* that it's necessary.)

[6] Chomsky's definition of a linguistic theory is weaker: it need only define the set of possible grammars, independent of learning data. This allows Chomsky to define the term **descriptively adequate theory**, which is a theory that includes, as possible grammars, a descriptively adequate grammar for every language—but does not necessarily return that grammar given learning data for that language.

Examples of observationally adequate grammars for English noun plurals

I. (just list every word you know)

kʰæt	kʰæt**s**	dɪʃ	dɪʃ**əz**	mæn	m**ɛ**n
sæk	sæk**s**	fʌd͡ʒ	fʌd͡ʒ**əz**	fʊt	f**i**t
dɑg	dɑg**z**	pʰi	pʰi**z**	waɪf	waɪ**vz**
gɹʌb	gɹʌb**z**	kʰaʊ	kʰaʊ**z**	wɪf	wɪf**s** ...

I.e., the grammar's judgment function accepts utterances containing those items in positions where a plural is required (*I like cats*); its truth-condition-assigning function assigns the appropriate truth-conditions to utterances containing the items in the right column (*I like cats* is true iff I like members of the cat group—it has nothing to do with whether I like members of the dog group).

II. Add –s to everything, except for these exceptions:

dɑg	dɑg**z**	fʌd͡ʒ	fʌd͡ʒ**əz**	mæn	m**ɛ**n
gɹʌb	gɹʌb**z**	pʰi	pʰi**z**	fʊt	f**i**t
dɪʃ	dɪʃ**əz**	kʰaʊ	kʰaʊ**z**	waɪf	waɪ**vz** ...

III. Add –z to everything, except for these exceptions:

kʰæt	kʰæt**s**
sæk	sæk**s**
dɪʃ	dɪʃ**əz**
fʌd͡ʒ	fʌd͡ʒ**əz**
mæn	m**ɛ**n
fʊt	f**i**t
waɪf	waɪ**vz**
wɪf	wɪf**s**
...	...

*IV. Add –**ə**z after "sibilant" sounds, –s after non-sibilant [–voice] sounds, and –z otherwise, except for these exceptions:*

mæn	m**ɛ**n
fʊt	f**i**t
waɪf	waɪ**vz**
...	...

*IV. Change final /f/ to [v], and then add –**ə**z after "sibilant" sounds, –s after non-sibilant [–voice] sounds, and –z otherwise, except for these exceptions:*

mæn	m**ɛ**n
fʊt	f**i**t
wɪf	wɪf**s**
...	...

10. Descriptive adequacy

A grammar that not only is observationally adequate, but also gives the same treatment to novel utterances that a real speaker of the target language gives is a **descriptively adequate grammar**.

The idea is that a descriptively adequate grammar captures the generalizations that real learners extract from the learning data—I think it makes the most conceptual sense to operationalize this in terms of novel utterances.

In a famous early study of children, Berko (1958)[7] found that English-speaking adults (all highly educated, in her sample, FWIW) consistently give the following plurals when presented with invented words (pp. 155-158):

wʌg	wʌg**z**	lʌn	lʌn**z**
gʌt͡ʃ	gʌt͡ʃ**əz**	nɪz	nɪz**əz**
kæʒ	kæʒ**əz**	kɹa	kɹa**z**
toɹ	toɹ**z**	tæs	tæs**əz**

- Which of the grammars above could be descriptively adequate, given these data?
- The adults disagreed about this word—what might we conclude?

heaf	hif**s**, hi**vz**

11. Descriptive adequacy is hard!

Achieving descriptive adequacy is often spoken of as though it were easy or could happen through inspection of basic data, but under Chomsky's definition it is actually a huge challenge.

Words that the speaker already knows are uninformative! (They don't tell us anything about what generalizations the speaker has learned—she may have simply memorized that word.)

Constructing novel phonological situations to put speakers in is difficult. Contrast this with syntax, where it's easy to construct sentences that—presumably—the speaker has not encountered before.

- Remember the K&K discussion of Russian devoicing. They point out some observationally adequate accounts that don't include a rule of final devoicing:
  - list two allomorphs for the stems that alternate (/...b/&/...p/ vs. /...p/ for stems that don't alternate)
  - have a devoicing rule but characterize the environment in morphological terms (e.g., "end of nominative singular") instead of /__#
  - have separate rules for b → p, d → t, etc.

  They had some arguments that these accounts are not descriptively adequate. You were asked to consider those arguments' satisfactoriness. So what do you think?

[7] Berko, Jean (1958). The child's learning of English morphology. *Word* 14: 150-177.

12. Explanatory adequacy

A theory that, when given a typical set of learning data, returns a grammar that is descriptively adequate, is an **explanatorily adequate theory**.

Obviously, developing an explanatorily adequate theory is a huge challenge!