

## Class 1: Introduction and course overview

### To do for next time

- Read K&K chs 1 & 2 and turn in study questions on Tuesday—if you are experienced in phonology you can skim (but still answer the questions)
- Check out course web page, especially **feature** links
- Do warm-up problem (ungraded) if you didn't do it already and leave it in my mailbox

### 1. What's phonology?

*Definition I:* Phonology = study of the sound patterns of human language.

For any given language, this includes answering the questions (Goldsmith 1995<sup>1</sup>)...

- What are the legal words of the language?
  - phone inventory (set of legal sounds)
  - phonotactics (set of legal sound sequences)
- What alternations occur?
- Which phonetic differences are contrastive? (we will not look at this one very much)

*Definition II:* Phonology = study of humans' knowledge of linguistic sound patterns

### 2. What is our job as phonologists?

Again, there are different answers...

- To discover the answers to the questions in Definition I above (i.e., to describe phonologies).
- To develop a theory of “what tools we need in order to provide adequate descriptions of individual languages”<sup>2</sup> (i.e., to develop a descriptive theory).
- To explain why phonologies are the way they are by constructing...
  - a theory of what people's knowledge of linguistic sound patterns is and how they learn, store, and use that knowledge
  - plus a theory of how linguistic sound patterns change over time

This will be our focus

### 3. How do we know if our explanation is a good one?

- 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: the learner should choose the grammar with the fewest symbols (and we hope for no ties).
- If that's right, and if we've got the notation right too, then you can tell which grammar, if any, is correct for a given language, out of some set of candidate grammars.
- More plausibly, we want to find independent evidence as to which grammar is right, and work backwards from there to how/why the learner chose that one—this is a lot harder!

<sup>1</sup> Goldsmith, John (1995). Phonological theory. In John Goldsmith (ed.) *Handbook of Phonological Theory*. Cambridge, MA & Oxford: Blackwell. Pp. 1-23.

<sup>2</sup> Dryer, Matthew S. (2003). Descriptive theories, explanatory theories, and basic linguistic theory. Ms., U.of Buffalo.

4. Example: pre-Spanish-contact Tagalog

p	t	k	ʔ	i	u
b	d	g		a	
	s		h		
m	n	ŋ			
	l				
w	j (glide)				

[e], [o] occur as allophones of /i/, /u/

[r] occurs as intervocalic allophone of /d/ (does not occur word-initially)

<i>bare stem</i>		<i>paN+stem</i>	
poʔók	‘district’	pampoʔók	‘local’
túbig	‘water’	pantúbig	‘used in making broth’
katawá	‘body’	paŋkatawán	‘physical’
n			
ʔabála	‘interruption’	paŋʔabála	‘distraction’
baʔák	‘split in tow’	pambaʔák	‘tool for splitting’
dagán	‘weight placed on’	pandagán	‘weight placed on top’
gíliŋ	‘grinding’	paŋgíliŋ	‘grindstone’
sigáʔ	‘blaze’	pansigáʔ	‘funeral pyre’
húli	‘catch’	paŋhúli	‘trap’
madláʔ	‘people’	paŋmadláʔ	‘public’
	(no examples for n or ŋ)		
litó	‘confused’	panlitó	‘red herring’
wakás	‘conclusion’	paŋwakás	‘final’
jári	‘power’	makapaŋjári	‘to dominate’

- Formulate a (set of) rule(s) to account for the different shapes of the prefix. Let’s see which grammar is best according to the evaluation metric of brevity.

Convention for handouts in this course: open bullets mean that this is a question for you to answer or discuss.

- What does our rule predict for loans that begin with /r/? What about possible longer rules—what do they predict?

It turns out that there is variation:<sup>3</sup>  
 rá:dyo ‘radio’ paŋrá:djo ~ panrá:djo ‘for the radio’

- Your thoughts on what we might conclude from this?

**5. Excursus: does the learner really have/need an evaluation metric?**

The idea of an evaluation metric suggests that the learner constructs multiple grammars and then chooses one.

Another way of looking at it: the learner follows an algorithm that assembles the grammar step by step, never considering alternatives.

- Possibility #1: We can state an evaluation metric such that the grammar arrived at by the algorithm always scores at least as well on the metric as any alternatives compatible with the data.
- Possibility #2: We are unable to state such an evaluation metric (except the rather uninteresting one that simply runs the algorithm and then assigns a good score to the resulting grammar and a bad score to any other grammar).

These aren’t the only possibilities. Suppose the learner constructs a preliminary grammar, then considers making a limited change to it; accepts the change under certain circumstances; then considers making another change to the result, etc. (i.e., “hill-climbing”)

- What status might an evaluation metric have under this scheme?

**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 more concise descriptions of attested/common phenomena than of unattested/uncommon phenomena.

		obstruent- or nasal-initial	liquid-initial	vowel-initial	glide-initial
		/garson/ ‘boy’	/livr/ ‘book’	/enfant/ ‘child’	/wazo/ ‘bird’
obstruent- or nasal-final	/pətit/ ‘small’	pəti_ garsō	pəti_ livr	pətit āfā	pətit wazo
liquid-final	/ʃɛr/ ‘dear’	ʃɛr garsō	ʃɛr livr	ʃɛr āfā	ʃɛr wazo
vowel-final	/lə/ ‘the’	lə garsō	lə livr	l_ āfā	l_ wazo
glide-final	/parej/ ‘similar’, /vjej/ ‘old’	parej garsō	parej livr	vjej āfā	vjej wazo

<sup>3</sup> Actually, there is also some variation between assimilated and non-assimilated for the consonants shown above as assimilating (seems to reflect degree of lexicalization of the word), but the unassimilated variant is much more frequent for [r] than for those consonants.

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

	<i>vocalic</i>	<i>consonantal</i>
obstruents	–	+
nasals	–	+
liquids	+	+
glides	–	–
vowels	+	–

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

- Revise the rules, using Greek-letter variables (Is there anything funny about this use of the variables?)

- Do Greek-letter variables allow us to compress these two rules?

$\left[ \begin{array}{l} +\text{voc} \\ -\text{back} \end{array} \right] \rightarrow \emptyset / \_ \# [-\text{cons}]$  “nonback vowels and liquids delete before vowels and glides”

$\left[ \begin{array}{l} -\text{high} \\ +\text{cons} \end{array} \right] \rightarrow \emptyset / \_ \# [+ \text{nasal}]$  “nonhigh consonants and glides delete before nasals”

- If French-type situations are common (more common than, say, the rule pair above), is it a good idea (according to SPE’s logic) to allow the same Greek-letter variable to apply to different features within a rule?

## 7. Some theoretical ruminations

The reasoning above relies on some assumptions about linguistic typology that are not always made explicit:

- Assume that the reason a rule is cross-linguistically common is that it's favored by learners—i.e., learners tend to mislearn, in the direction of a more-favored grammar.
- Assume that learners favor short rules (or replace “short” by whatever descriptor best captures the evaluation metric).
- 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 a good one if it allows common rules to become shorter than uncommon rules.

This means that we're not really using “short” (or “simple”) in any abstract, domain-independent sense. Rather, we're tailoring the notation to make the rules that we think learners favor appear short.

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? (in which case we can say that 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. What to expect in this course

Especially if you have a fair amount of background already, the first couple of classes (and readings, from the textbook) should feel like review. And most of the topics on the syllabus should be familiar to you.

But, we will be addressing these topics at an advanced level, relying the research literature rather than textbook material where possible, and doing problem sets that are less sanitized than in an undergraduate course (though still somewhat sanitized).

**9. Exercise, in the unlikely event that we have extra time: Kasem Class C noun plurals**

Based on discussion in SPE Ch. 8, data originally from Callow,<sup>4</sup> plus additional discussion in Halle 1978.<sup>5</sup>

	<i>singular</i>	<i>plural</i>	
<i>a</i>	bakada	bakadi	'boy'
<i>b</i>	sada	sadi	'grass mat'
<i>c</i>	mimina	mimini	'thin'
<i>d</i>	fala	fali	'white man'
<i>e</i>	tula	tuli	'granary'
<i>f</i>	kukuda	kukudi	'dog'
<i>g</i>	fana	fani	'knife'
<i>h</i>	čana	čani	'moon'
<i>i</i>	bakala	bakali	'shoulder'
<i>j</i>	kambia	kambi	'cooking pot'
<i>k</i>	pia	pi	'yam'
<i>l</i>	buga	bwi	'river'
<i>m</i>	diga	di	'room'
<i>n</i>	malaa	male	'chameleon'
<i>o</i>	kabaa	kabe	'slave'
<i>p</i>	zizaa	zize	'grass roof'
<i>q</i>	laŋa	le	'song'
<i>r</i>	naga	ne	'leg'
<i>s</i>	pia	pe	'sheep' (cf. 'yam')
<i>t</i>	babia	babe	'brave'
<i>u</i>	nanjua	nanjwe	'fly'
<i>v</i>	yua	ywe	'hair'
<i>w</i>	koga	kwe	'back'
<i>x</i>	čoŋa	čwe	'path'

**Next time:** Notation review and extrinsic rule ordering

<sup>4</sup> Callow, J.C. (1965). Kasem nominals: a study in analyses. *Journal of West African Languages* 2, 29-36.

<sup>5</sup> Halle, Morris (1978). Further thoughts on Kasem nominals. *Linguistic Analysis* 4, 167-185.