Class 20: Retrospective and prospective course wrap-up

To do
- Manam due this Friday (Dec.6)
- Work on presentation (Dec. 10) and paper (due Dec. 13)

Overview: Some summarizing, some stock-taking, some prospect, a little synthesis.
"☼" means you're likely to learn more about the topic in 201A.

1. Learnability

- Review of the Chomskyan basics:
  - an observationally adequate grammar labels the utterances that a typical learner would encounter as grammatical (perhaps trivially, e.g. by listing them)
  - a descriptively adequate grammar captures the psychologically real generalizations—this could be operationalized as 'treats novel utterances the same way real speakers do'
  - the real prize, an explanatorily adequate theory, will, given typical learning data, return an descriptively adequate grammar

- Achieving an explanatorily adequate theory is going to have to involve ☼**learning algorithms**.
  - Interestingly, there was never a good learning algorithm that could induce an ordered list of rules from surface forms, or even from underlying-surface pairs.
  - By contrast, there's a big literature on learning algorithms in OT.

- In OT, *under the assumption of a finite, universal constraint set...*
  - ...and *given input-output pairs, it's easy*: see Tesar & Smolensky 2000, Riggle 2004
  - ...and given inputs and just the audible portion of the outputs (e.g., no foot boundaries): it's harder. See Tesar 2000, Jarosz to appear.
  - ...and given just outputs (with or without their inaudible parts): it's a lot harder. See Tesar et al. 2003, Jarosz 2006.
    - A fair amount of phonotactic learning can be accomplished, which could later be used to learn alternations, though that step remains largely unimplemented (see Hayes 2004).

- There are also learning algorithms for ☼**variable/probabilistic constraint rankings**:
  - Gradual Learning Algorithm: Boersma 1998, Boersma & Hayes 2001, Magri 2012
  - Maximum Entropy OT: Goldwater & Johnson 2003

    You can try out these two (plus a couple of non-variable algorithms) by downloading software from Bruce Hayes's webpage.
  - Harmonic Grammar: Pater, Potts, & Bhatt 2007, Boersma & Pater 2008, Pater 2009

    Try it out, using the same format for input files as in Bruce's software, at web.linguist.umass.edu/~halp/ (Potts, Becker, Bhatt & Pater)

- What if the constraint set isn't universal, and constraints have to be constructed by the learner?
  - This is still fairly uncharted territory—see Heinz 2007, Hayes & Wilson 2006.

1.1 When multiple grammars are consistent with data, which one does learner select?

- This is the evaluation-metric problem that we've seen since the beginning of the course—solving it is part of developing an explanatorily adequate theory.

- The **subset problem**—say you are exposed to the following language:

tagu	'goat'	tagune	'goats'	taguba	'my goat'
ale	'mango'	alene	'mangos'	aleba	'my mango'
siri	'corkscrew'	sirine	'corkscrews'	siriba	'my corkscrew'

  - In a rule framework, what grammar would you learn?

  - How do you think you would then react to the word *sirab*? Is this predicted by the grammar?

  - Same question for OT—what ranking would you learn for the constraints NOCODA, MAX-C, and DEP-V? What does this ranking predict for *sirab*?

- Some learning algorithms have addressed this question of how a learner knows that something they've never seen is forbidden, in the absence of helpful alternations (Prince & Tesar 2004, Hayes 2004).
  - The idea is, force markedness constraints to be ranked as high as is consistent with data.

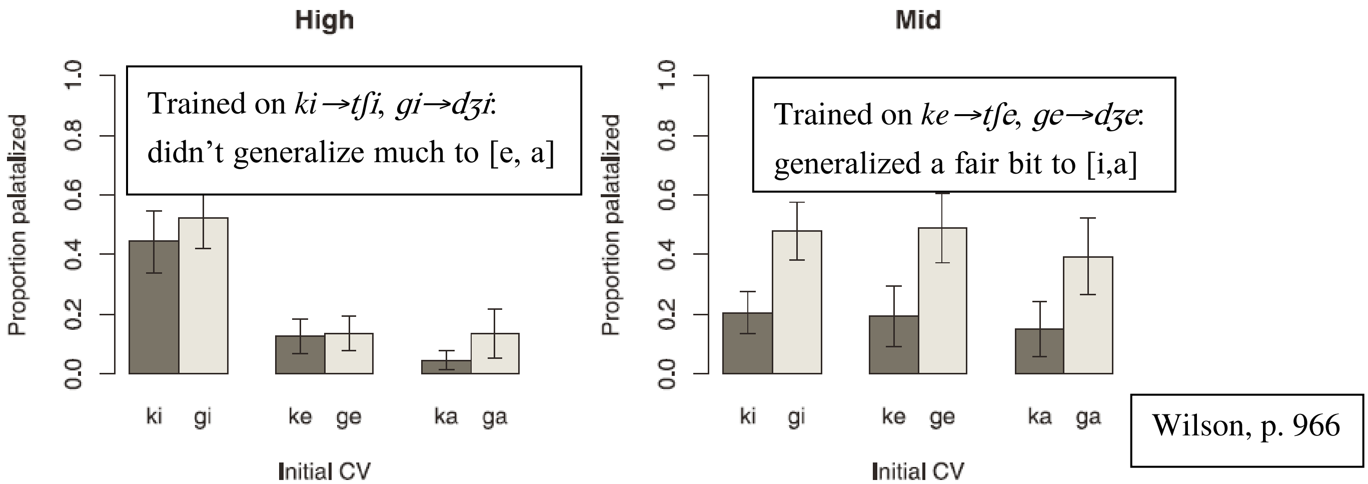
1.2 Ranking bias *within* markedness or faithfulness constraints?

- Wilson 2006, drawing on Guion 1996: Cross-linguistically, velar palatalization (k→tʃ, g→dʒ) before one front vowel implies palatalization before a higher front vowel—that is, we see languages *ki, ke* and *tʃi, ke* and *tʃi, tʃe* but not *ki, tʃe*.

  - If we simply have these three constraints, what's the predicted typology: *ki, *ke, IDENT(place) (I'm leaving out *ka to keep things simple)

    - One approach is to build more structure into the constraint inventory: *k[+hi], *k[–lo], IDENT(place).
  - What typology do we get now?

- Another approach, for which see Wilson (who has experimental evidence for it):
  - In a ranking system where each constraint is associated with a weight (this is different from Classic OT's strict ranking), the learning problem involves discovering the weights.
  - We can start with each weight at zero—that is, all constraints are without effect—and promote them in response to the data.
  - Each constraint *i* is also associated with a value σ_i that determines how willing the constraint is to change its weight. (Wilson derives these from Guion's confusion rates.)
  - If we give *ke a smaller σ than *ki, then we require more evidence in order to promote *ke than *ki.
  - So it's <u>possible</u> to learn the typologically anomalous *ki, tʃe* language, but it's a lot easier (requires less evidence) to learn the other possibilities.
  - See White under review for approach where constraints have same σ, but different default weights.



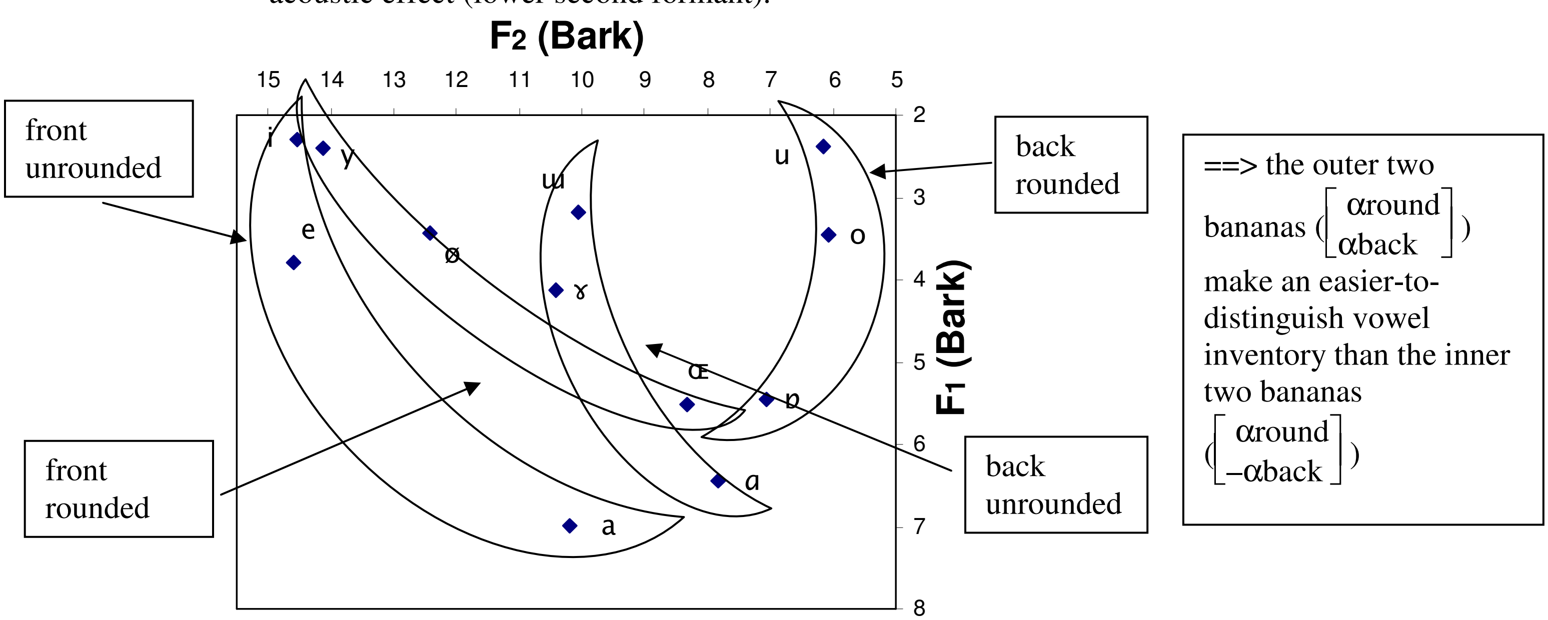
1.3 Constraint learning

- What about constraints themselves?
  - If the learner has to construct constraints, are all possibilities equally good?
  - There might be a criterion of formal simplicity, but, as with rules, that's probably not enough.

    Compare $*\begin{bmatrix}\alpha\text{round}\\ -\alpha\text{back}\end{bmatrix}$ to $*\begin{bmatrix}\alpha\text{round}\\ -\alpha\text{voice}\end{bmatrix}$ –equally simple, but not equally attested
  - Same issue arises with rules: why [αround]→[αback] but not [αround]→[αvoice]?
- Along with constraint-learning itself, this is an open problem.

1.4 ☼The role of phonetics

- Well-known phonetic explanation for above round/back affinity:
  - lip rounding/protrusion and tongue backing, although articulatorily independent, share an acoustic effect (lower second formant).[1]



[1] Thanks to David Deterding's Excel template (http://videoweb.nie.edu.sg/phonetic/vowels/measurements.html)

- Obviously phonetics explains a lot of observed phonology. But...
  - Does the explanatory mechanism lie in learner preferences (Hayes & Steriade 2004, Kawahara 2007) or in pathways of language change (Blevins 2003)?
  - Do grammars make literal reference to phonetic motivation ("don't have a contour tone if the vowel is shorter than 150 msec")
    - or do phonetic motivations get phonologized ("don't have a contour tone except in diphthongs and final syllables"), and if so how?
    - See Hayes 1999 for this question in general; Zhang 2007 for contour tones in particular.

2. Processes and constraints—some typological possibilities

a. languages (and phenomena within a language) are similar in the structures they avoid (constraints), but not in the changes they apply (processes): e.g., *NC̥, diverse repairs
b. similarity in processes but not in constraints? maybe—how many different "problems" is, say, C-deletion a "solution" to?
c. similarity in both: *VOICEDOBSTRUENT#, devoicing only
d. similarity in neither: ?? I guess very idiosyncratic phenomena like Palauan *s→k / __l*

○ What do you think about SPE's and OT's predictions here?

3. Process interaction: extrinsic ordering?

Feeding in Kalinga

	/sin+pajaw/	*o]$_\sigma$	MAX-V	AGREEPLACE	IDENT(place)
a	sin.pa.jaw			*!	
☞ *b*	sim.pa.jaw				*
	/d-in-opa/	*o]$_\sigma$	MAX-V	AGREEPLACE	IDENT(place)
c	di.no.pá	*!			
d	din.pá		*	*!	
☞ *e*	dim.pá		*		*

- We can't get both (b) and (d) [counterfeeding] to win, at least not with these constraints

Bleeding in English:

	/kæt+z/	OBSTRUENTSAGREEVOICE	IDENT(voice)
a	kætz	*!	
☞ *b*	kæts		*

	/bɹæntʃ+z/	OBSTRUENTSAGREEVOICE	*[+strid][+strid]	IDENT(voice)	DEP-V
c	bɹæntʃz	*!	*		
d	bɹæntʃs		*!	*	
e	bɹæntʃɨs			*!	*
☞ *f*	bɹæntʃɨz				*

- The counterbleeding candidate (e) can't win—with these constraints, it's harmonically bounded.

- Opacity is hard for standard OT to deal with, as we've seen! See McCarthy 2007b for a book-length discussion.
- You will probably see some ☼**proposals** in 201 for how to fix this (not all of these proposals were developed with opacity in mind):
  - containment (Goldrick & Smolensky 1999)
  - sympathy (McCarthy 2003)
  - candidate chains (McCarthy 2007b)
  - output-output correspondence (Crosswhite 1998; Benua 1997; Steriade 2000; Burzio 1998; Kenstowicz 1995 and others)
  - targeted constraints (Wilson 2001)
  - local constraint conjunction (Smolensky 1997, Lubowicz 2005, Kirchner 1996)
  - Stratal OT (Kiparsky 2000)
  - distantial faithfulness (Kirchner 1996)
  - *MAP constraints (Zuraw 2007)
  - comparative markedness (McCarthy 2002)
  - harmonic serialism (McCarthy 2000, McCarthy 2010)
- Most don't capture all types of opacity, and whether all claimed types of opacity are learnable is debated in, e.g., Sanders 2002.

4. Process application

4.1 Self-feeding and self-bleeding

- Recall Takelma from Anderson 1974:
  - [a] becomes [i] if followed by [i]: /alxīxamis/ → [alxīx**i**mis] 'one who sees us'
  - and any preceding [a]s follow suit: /ikūmanananink^h/ → [ikūm**i**n**i**n**i**nink^h] 'he will fix it for him' (unless a voiceless C intervenes)
    - This is expected in OT, where self-counterfeeding would be unexpected: Kaplan 2008.

- Recall French (optional) schwa deletion from Anderson, following Dell 1973:

  ə → Ø / VC__C(r)V
  - /ty#dəvəne/ → [ty#dəvəne] or [ty#d_vəne] or [ty#dəv_ne]
  - but not *[ty#d_v_ne] 'you were becoming'
    - Again, expected in OT, where self-counterbleeding (Kikuyu??) would be unexpected.

4.2 Directional application

- If there is such a thing as directional rule application...
  - in the sense that the left/rightmost eligible site has priority for undergoing the rule, regardless of whether it's stressed/unstressed, word-initial/word-final...
  - then standard OT doesn't have much to say about it (see Hyman & VanBik 2004)

- **Hypothetical** case (pseudo-French):
  - *only one target*: /dəv**ə**ne/ → [dəvne]
  - *multiple targets:* /ty#d**ə**vəne/ → [ty#d_vəne], *[ty#dəv_ne]
  - /...vudre#k**ə**#sə#k**ə**#lə#pɔlisje…/→[...vudre#k_#sə#k_#lə#pɔlisje],*[...vudre#kə#s_#kə#l_#pɔlisje]

- Eisner's (2002) directional constraint evaluation (proposed for computational reasons, not because of data like this):
  - Index a copy of *SCHWA to each position (counting by segments, though other constraints might count differently) in the output string.
  - Left-to-right version:

	/ty#dəvəne/	*CCC	*ə-1	*ə-2	*ə-3	*ə-4	*ə-5	*ə-6	*ə-7	*ə-8
☞ *a*	[ty#d_vəne]						*			
b	[ty#dəv_ne]					*!				
c	[ty#dəvəne]					*!		*		
d	[ty#d_v_ne]	*!								

4.3 Modes of variation claimed to exist (see details and references in Class 7/8 handout)

- **Global**: in Warao, a word has either all [p]s or all [b]s—no mixing
- **Local**: Vaux's *[maɹkətʰəbɪlətʰi]* ~ *[maɹkərəbɪləri]* ~ *[maɹkətʰəbɪləri]* ~ *[maɹkərəbɪlətʰi]*
- **Iterational**: Vata */ɔ̀ ká zā pī/* → *ɔ̀ ká zā pī* ~ *ɔ̀ ká zʌ̄ pī* ~ *ɔ̀ kʌ́ zʌ̄ pī* ~ ***ɔ̀*** *kʌ́ zʌ̄ pī*
- **At-most-one-target**: Dominican Spanish *hablar fisno* style *a**s**.bo.ga.do* ~ *a.bo**s**.ga.do* ~ *a.bo.ga**s**do* ~ *a.bo.ga.do**s***, but **a**s**.bo.ga**s**.do*, (*a.bo**s**.ga.do**s***), etc.
- **At-least-one-target**: Munro & Riggle 2004
  - In Pima [Uto-Aztecan, Arizona; *Ethnologue* groups it with Tohono O'odham, for 9,600 speakers], reduplication marks plurality, but in compounds plurality is expressed by reduplicating <u>any non-empty subset of the conjuncts</u>:

(5)

gloss and etymology	singular	plural forms
'bridge' (tree-road)	*'ùs-vóog*	*'ù'us-vópog, 'ù'us-vóog, 'ùs-vópog*
'church' (mass-house)	*mìish-kíi*	*mìmsh-kíik, mìmsh-kíi, mìish-kíik*
'dish' (baskety.thing-jar)	*hòas-há'a*	*hòahas-háha'a, hòahas-há'a, hòas-háha'a*
'onion soup' (onion-soup)	*sìvol-sóoba*	*sìsvol-sósba, sìsvol-sóoba, sìvol-sósba*
'peso' (Mexican-dollar)	*Jùukam-píish*	*Jùujkam-píipsh, Jùujkam-píish, Jùukam-píipsh*
'peyote' (coyote-plant.type)	*bàn-nód:adag*	*bàban-nónd:adag, bàban-nód:adag, bàn-nónd:adag*
'saltbush' (salt-grass)	*'ònk-váshai*	*'ò'onk-vápshai, 'ò'onk-váshai, 'ònk-vápshai*
'tamarack' (salt-tree)	*'ònk-'ús*	*'ò'onk-'ú'us, 'ò'onk-'ús, 'ònk-'ú'us*
'uvula' (throat-bell)	*bà'itk-kámpañ*	*bàba'itk-kákampañ, bàba'itk-kámpañ, bà'itk-kákampañ*
'wagon' (tree-car)	*'ùs-kálit*	*'ù'us-káklit, 'ù'us-kálit, 'ùs-káklit*

(3rd page of manuscript version)

5. Derivational look-ahead

- Crowhurst & Michael 2005, Nanti [Arawakan, Peru, 480 speakers]:
  - an iterative rule shifting stress within a foot can be triggered by a violation of *CLASH:
    (o.kò)(ri.k**ʃì**)(**tá.**ka) → (ò.ko)(**rì**.kʃi)(**tá**.ka) 'she wore a nose-disk'
  - but stress can't shift to a less-prominent (e.g., higher) vowel:
    (i.kà)(tsi.**tò**)(**ká**.kse) 'he held (it) in his talons'
  - o What do you think of this form? How could it be analyzed with rules? OT?
    (no.tà)(me.**sè**)(**tá**.kse) 'I scraped (it)'

- OT may go too far with its look-ahead ability (see Kaplan 2011 for discussion)...
  - The problematic predictions usually seem to involve two different phenomena (instead of a single phenomenon, stress, as in Nanti)
    - e.g., does any language add or subtract syllables in order to get stress onto a more-prominent vowel??
  - The problem here may be not look-ahead, but which processes can solve which kinds of problems.
  - See Blumenfeld 2006 for examples and a theory.

6. Constraint violability

- In a rules+constraints analysis of Nanti, for instance, we could have *CLASH
  - it's frequently violated, though, so we have to restrict its power, either by giving it a limited set of rules to trigger, or by stipulating that some other constraint can block its triggered rules.
- In OT, at least the theory makes it clear how this kind of interaction works:

*CLASH >> RHTYPE=IAMB...

okorikʃitaka	NONFINALITY	PROMINENCE INFOOT	*CLASH	RHTYPE= IAMB
a (o.kò)(ri.kʃì)(tá.ka)			*!	*
b (o.kò)(rì.kʃi)(tá.ka)			*!	**
☞ *c* (ò.ko)(rì.kʃi)(tá.ka)				***
d (o.kò)(ri.kʃì)(ta.ká)	*!			

...but PROMINENCEINFOOT >> *CLASH

nosamerejaka	NONFINALITY	PROMINENCE INFOOT	*CLASH	RHTYPE= IAMB
e (nò.sa)(mè.re)(já.ka)		*!		***
f (no.sà)(mè.re)(já.ka)			*	**!
☞ *g* (no.sà)(me.rè)(já.ka)			*	*
h (no.sà)(me.rè)(ja.ká)	*!			

7. Representations

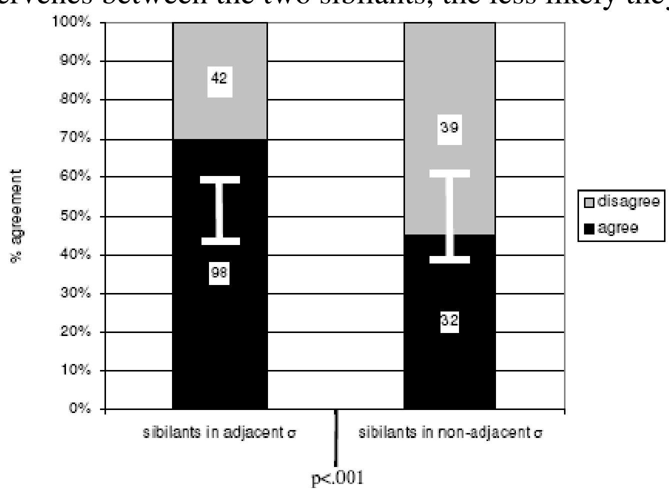
7.1 Autosegmentalism

We saw

- features' independence from segments (especially tone)
- long-distance interactions between certain types of segments (e.g., sibilant harmony)
- group behavior of certain features (e.g., place)

Open questions

- Is locality really all-or-nothing? Recall Martin's Navajo sibilant harmony case:
  - The autosegmental account predicts that it doesn't matter how much material intervenes between the two stridents—they are still adjacent as far as the [anterior] tier is concerned.
  - But Martin found that, in compounds, agreement is *gradient*: the more material intervenes between the two sibilants, the less likely they are to agree:



Martin 2004, p. 23

  - (Additional twist, explored further in Martin 2007: most of the agreement in compounds is already there in the underlying forms!)
- Do we need a geometry to group features, or do we include in the evaluation metric principles that decide which features are favored to be referred to together in a spreading rule or an AGREE constraint?

7.2 Metrical stress theory

We saw...

- that stress is not like "real" features, not even autosegmental ones
- to deal with this, grids+feet

We didn't get to (among other things)

- proposals for additional hierarchical structure in phonological representations: feet grouped into prosodic words, then phonological phrases, then larger intonational phrases... (e.g., Selkirk 1978; Nespor & Vogel 1986; Hayes 1989).

8. The role of morphology

We looked at matters like...

- **Cyclicity**: derived words sometimes retain characteristics of their morphological predecessors
- **Non-derived environment blocking**: some processes apply only when triggered by morphology or (perhaps) other phonology
- **Levels**: within a language, subsets of the phonological processes are associated with subsets of the word-formation rules
- and relatedly, **Lexical vs. post-lexical**: there seem to be two syndromes—productive vs. not as much, gradient vs. categorical, carrying over into L2 vs. not, applying across word boundaries vs. not...

9. ☼The role of syntax—which we didn't talk about

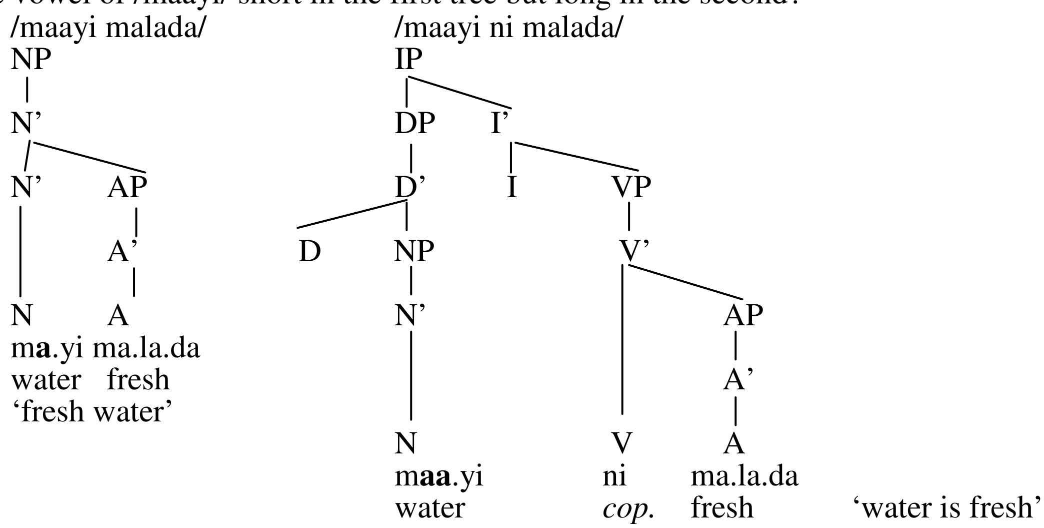
9.1 Syntax influencing phonology

Kisseberth 2000, Chimwiini (dialect of Swahili formerly w/ 40,000 speakers in Somalia; most have emigrated to Kenya)

- Long vowels allowed only in the penult and antepenult of a "phonological phrase".

- Under Kisseberth's analysis, in Chimwiini the end of an XP (DP, NP, AP, VP...) ends a phonological phrase (but the beginning of an XP is irrelevant): ALIGN(XP,R,PPhrase,R)

o Why is the vowel of /maayi/ short in the first tree but long in the second?



- Most approaches to syntax's influence on phonology focus on how syntactic structure defines domains like the phonological phrase, which phonology then refers to.

9.2 Phonology influencing syntax? Or at least word order...

- Embick & Noyer 2001, Latin: the clitic *–que* 'and', attaches after 1st word of 2nd conjunct:

[bonī	puerī]	[bonae–**que**	puellae]		
good	boys	good–and	girls	'good boys and good girls'	(p. 575)

- But when the second conjunct begins with a preposition, its syllable count matters:

circum–**que** ea loca	in rēbus–**que**
around-and those places	in things-and
contrā–**que** lēgem	dē prōvinciā–**que**
against-and law	from province-and (p. 576)

- For more cases, and literature reviews, see Schütze 1994, Shih et al. to appear (among others)

10. Some of my favorite things to think about in phonology, besides the above

- What is stored in the lexicon and what is computed online?
- How detailed is the lexical representation (Bybee 2001; Pierrehumbert 2002; Gahl 2008)? Can it contain redundant information?
- What is the phonology-processing interface like?
  - How does lexical retrieval for production influence pronunciation, e.g. single word vs. morpheme string (Hay 2003, but see Fiorentino 2006) priming and competition from other words (Baese-Berk & Goldrick 2009 and refs. therein, Martin 2007, Smolensky, Goldrick, & Mathis to appear)?
  - How does word recognition influence perception and lexicalization?
- What are the limits of learnability? Within the learnable, are some patterns more learnable than others?
- How can we get good data about competence? Especially, how can we tell what's lexicon and what (if anything) is grammar?

11. Phonological things you can do after this course

- Take Ling 201A (Phonological Theory II) with Bruce Hayes next quarter (required for most of you)
- Check the phonology seminar (261ABC) schedule and feel free to drop in for whatever talks interest you: *www.linguistics.ucla.edu/talksaevents*
- Ling 205, Morphology, is not a yearly event, so take advantage when it comes around
- Same goes for Ling 236, Computational Phonology
- Same goes for Ling 111/211, Intonation, an in-depth look at the highest levels of the prosodic hierarchy—offered this winter
- Look out for phonetics and phonology proseminars (251). These are courses that focus on a special topic

References

Anderson, Stephen R. 1974. *The Organization of Phonology*.. New York: Academic Press.
Baese-Berk, Melissa & Matthew Goldrick. 2009. Mechanisms of interaction in speech production. *Language and Cognitive Processes* 24(4). 527–554.
Benua, Laura. 1997. Transderivational Identity: Phonological Relations between Words.. University of Massachusetts, Amherst.
Blevins, Juliette. 2003. *Evolutionary phonology. The emergence of sound patterns*.. Cambridge: Cambridge University Press.
Blumenfeld, Lev. 2006. Constraints on Phonological Interactions.. Stanford ph.d. dissertation.
Boersma, Paul. 1998. *Functional Phonology: Formalizing the Interaction Between Articulatory and Perceptual Drives*.. The Hague: Holland Academic Graphics.
Boersma, Paul & Bruce Hayes. 2001. Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* 32. 45–86.
Boersma, Paul & Joe Pater. 2013. Convergence properties of a Gradual Learning Algorithm for Harmonic Grammar.. Manuscript. University of Amsterdam and University of Massachusetts, Amherst, ms.
Burzio, Luigi. 1998. Multiple correspondence. *Lingua* 104. 79–109.
Bybee, Joan. 2001. *Phonology and language use*.. Cambridge: Cambridge University Press.
Crosswhite, Katherine. 1998. Segmental vs. prosodic correspondence in Chamorro. *Phonology* 15(3). 281–316.
Crowhurst, Megan J & Lev D Michael. to appear. Iterative footing and prominence-driven stress in Nanti. *Student Workshop in Optimality Theory 2002*. Austin, TX: University of Texas at Austin.
Dell, Francois. 1973. *Les regles et les sons*.. Paris: Hermann, Collection Savoir.
Eisner, Jason. 2002. Comprehension and Compilation in Optimality Theory. *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL)*, 56–63. Philadelphia. http://cs.jhu.edu/jason/papers/#acl02-ot.
Embick, David & Rolf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32(4). 555–595.
Fiorentino, Robert. 2006. Lexical structure and the nature of linguistic representations.. University of Maryland, College Park ph.d. dissertation.
Gahl, Susanne. 2008. "Thyme" and "time" are not homophones: word durations in spontaneous speech. *Language* 84(3). 474–496.
Goldrick, Matthew & Paul Smolensky. 1999. Opacity, turbid representations, and output-based explanation.. Edmonton, Alberta.
Goldwater, Sharon & Mark Johnson. 2003. Learning OT Constraint Rankings Using a Maximum Entropy Model.. In Jennifer Spenader, Anders Eriksson, & Östen Dahl (eds.), *Proceedings of the Stockholm Workshop on Variation within Optimality Theory*, 111–120. Stockholm: Stockholm University.
Guion, Susan. 1996. Velar palatalization: coarticulation, perception and sound change.. University of Texas at Austin ph.d. dissertation.
Hay, Jennifer. 2003. *Causes and consequences of word structure*.. Routledge.
Hayes, Bruce. 1989. The prosodic hierarchy in meter.. In Paul Kiparsky & G. Youmans (eds.), *Rhythm and Meter*, 201–260. Orlando: Academic Press.
Hayes, Bruce. 1999. Phonetically driven phonology: the role of Optimality Theory and inductive grounding.. In Michael Darnell, Frederick J Newmeyer, Michael Noonan, Edith Moravcsik, & Kathleen Wheatley (eds.), *Functionalism and Formalism in Linguistics, Volume I: General Papers*, 243–285. Amsterdam: John Benjamins.
Hayes, Bruce. 2004. Phonological acquisition in Optimality Theory: The early stages.. In René Kager, Joe Pater, & Wim Zonneveld (eds.), *Constraints in Phonological Acquisition*. Cambridge: Cambridge University Press.
Hayes, Bruce & Donca Steriade. 2004. Introduction: the phonetic basis of phonological markedness.. In Robert Kirchner, Donca Steriade, & Bruce Hayes (eds.), *Phonetically based phonology*. Cambridge University Press.
Hayes, Bruce & Colin Wilson. 2006. A Maximum Entropy Model of Phonotactics and Phonotactic Learning.
Heinz, Jeffrey. 2007. Inductive learning of phonotactic patterns.. University of California, Los Angeles ph.d. dissertation.
Hyman, Larry M & Kenneth L VanBik. 2004. Directional rule application and output problems in Hakha Lai tone. *Language and Linguistics* 5(4). 821–861.
Jarosz, Gaja. to appear. Naive parameter learning for Optimality Theory: the hidden structure problem. *NELS Proceedings*.
Jarosz, Gaja. 2006. Rich Lexicons and Restrictive Grammars - Maximum Likelihood Learning in Optimality Theory.. Johns Hopkins University ph.d. dissertation. http://roa.rutgers.edu/files/884-1206/884-1206-7-0.PDF.
Kaplan, Aaron F. 2008. Noniterativity is an emergent property of grammar.. University of California, Santa Cruz.

Kaplan, Aaron F. 2011. Harmonic Improvement without Candidate Chains in Chamorro. *Linguistic Inquiry* 42(4). 631–650. doi:10.1162/LING_a_00063.
Kawahara, Shigeto. 2007. The emergence of phonetic naturalness.. University of Massachusetts, Amherst ph.d. dissertation.
Kenstowicz, Michael. 1995. Cyclic vs. non-cyclic constraint evaluation. *Phonology* 12. 397–436.
Kiparsky, Paul. 2000. Opacity and cyclicity. *The Linguistic Review* 17. 351–367.
Kirchner, Robert. 1996. Synchronic chain shifts in Optimality Theory. *Linguistic Inquiry* 27. 341–350.
Kisseberth, Charles. 2000. The phonology-syntax interface: Chimwiini revisited.. Tel Aviv University, ms.
Lubowicz, Ania. 2005. Locality of Conjunction. *Proceedings of WCCFL 24*. Somerville, MA: Cascadilla.
Magri, Giorgio. 2012. Convergence of error-driven ranking algorithms. *Phonology* 29(02). 213–269. doi:10.1017/S0952675712000127.
Martin, Andrew. 2007. The evolving lexicon.. University of California, Los Angeles ph.d. dissertation.
McCarthy, John J. 2000. Harmonic serialism and parallelism.. In Masako Hirotani (ed.), *Proceedings of the North East Linguistics Society 30*, 501–524. Amherst, Mass.: GLSA Publications.
McCarthy, John J. 2002. Comparative Markedness [Long Version].. In Angela C. Carpenter, Andries W Coetzee, & Paul de Lacy (eds.), *Papers in Optimality Theory II (University of Massachusetts Occasional Papers in Linguistics)*, vol. 26, 171–246. Amherst, MA: GLSA.
McCarthy, John J. 2003. Sympathy, cumulativity, and the Duke-of-York gambit.. In Caroline Féry & Ruben van de Vijver (eds.), *The Syllable in Optimality Theory*, 23–76. Cambridge: Cambridge University Press.
McCarthy, John J. 2007a. *Hidden generalizations: phonological opacity in Optimality Theory*.. London: Equinox.
McCarthy, John J. 2007b. *Hidden Generalizations: Phonological Opacity in Optimality Theory*.. London: Equinox.
McCarthy, John J. 2010. An introduction to harmonic serialism.. Mansucript. UMass Amherst, ms.
Munro, Pamela & Jason Riggle. 2004. Productivity and lexicalization in Pima compounds. *BLS*.
Nespor, Marina & Irene Vogel. 1986. *Prosodic Phonology*.. Dordrecht: Foris.
Pater, Joe. 2009. Weighted constraints in generative linguistics. *Cognitive Science* 33. 999–1035.
Pater, Joe, Christopher Potts & Rajesh Bhatt. 2007. Harmonic Grammar with Linear Programming.
Pierrehumbert, Janet. 2002. Word-specific phonetics. *Laboratory Phonology VII*. Berlin: Mouton de Gruyter.
Prince, Alan & Bruce Tesar. 1999. *Learning phonotactic distributions*.. New Brunswick, NJ.
Riggle, Jason. 2004. Generation, recognition, and learning in finite state Optimality Theory.. University of California, Los Angeles ph.d. dissertation.
Sanders, Nathan. 2002. Opacity and Sound Change in the Polish Lexicon.. UC Santa Cruz.
Schutze, Carson. 1994. Serbo-croatian second position clitic placement and the phonology-syntax interface. *MIT Working Papers in Linguistics* 21. 373–473.
Selkirk, Elisabeth. 1978. On prosodic structure and its relation to syntactic structure.. In T. Fretheim (ed.), *Nordic Prosody*, vol. 2, 111–140. Trondheim: TAPIR.
Shih, Stephanie, Jason Grafmiller, Richard Futrell & Joan Bresnan. to appear. Rhythm's role in genitive construction choice in spoken English.. In R Vogel & R van de Vijver (eds.), *Rhythm in phonetics, grammar, and cognition*. Mouton.
Smolensky, Paul. 1997. Constraint interaction in generative grammar II: local conjunction, or random rules in Universal Grammar.. Baltimore, MD.
Smolensky, Paul, Matt Goldrick & Donald Mathis. to appear. Optimization and quantization in gradient symbol systems: A framework for integrating the continuous and the discrete in cognition. *Cognitive Science*.
Steriade, Donca. 2000. Paradigm uniformity and the phonetics-phonology boundary.. In Janet Pierrehumbert & Michael Broe (eds.), *Acquisition and the Lexicon (Papers in Laboratory Phonology 5)*, 313–334. Cambridge: Cambridge University Press.
Tesar, Bruce. 2000. Using inconsistency detection to overcome structural ambiguity in language learning.. Technical Report RuCCS-TR-58. Rutgers Center for Cognitive Science, Rutgers University, ms.
Tesar, Bruce, John Alderete, Graham Horwood, Nazarré Merchant, Koichi Nishitani & Alan Prince. 2003. Surgery in language learning. *Proceedings of WCCFL 22*, 477–490.
Tesar, Bruce & Paul Smolensky. 1998. Learnability in Optimality Theory. *Linguistic Inquiry* 29. 229–268.
White, James. under review. Evidence for a learning bias against saltatory phonological alternations.
Wilson, Colin. 2001. Consonant Cluster Neutralisation and Targeted Constraints. *Phonology* 18(1). 147–197.
Wilson, Colin. 2006. Learning Phonology with Substantive Bias: An Experimental and Computational Study of Velar Palatalization. *Cognitive Science* 30(5). 945–982.
Zhang, Jie. 2007. Contour tone distribution is not an artifact of tonal melody mapping. *Studies in the Linguistic Sciences* 33(1).
Zuraw, Kie. 2007. The role of phonetic knowledge in phonological patterning: Corpus and survey evidence from Tagalog. *Language* 83. 277–316.