Linguistics 251 Spring 2013
Variation in Phonology Hayes/Zuraw

Class 8, 4/25/13: Substantive Bias

1. Assignments etc.

- MaxEnt exercise with priors; on website

2. Where are we in the course?

- The question is whether language learning is subject to **biases**.
- Kie, Tuesday: **conservativity bias**; don't commit to strong hypotheses without strong data.
  - basis of the Martinian "leakage" effect
  - … of the Ryanian "project from minimal data" effect
- Today, **substantive bias** — content of phonological representations, related to mainstream phonological theory, especially OT.

3. Readings: if Moreton and Pater are right, we are about to waste two hours!

- I.e. they suggest that all substance-based effects in phonology have a diachronic origin.
- Flesh this out a bit:
  - Prevalence of "phonetic precursors" (Moreton/Pater)
  - Propensity for "innocent misapprehension", leading the precursors over the line into phonology (Ohala various, Blevins 2004)
  - I'm also taken with the idea that spontaneous phonology of little kids is natural; and that (through the influence of DBC[1] children, seen elsewhere in language change), things like θ → f, t̪ percolate to adult phonology (Hayes 2004)[2]
- BUT I think Moreton/Pater have not drawn an important distinction; viz.:

4. Substantive bias can be taxonomized following the natural architecture of OT

- **Faithfulness**: biases discourage the learning of phonological alternation, particularly phonetically salient alternation.
- **Markedness**: biases discourage the learning of patterns that create marked surface forms.
  - Research so far suggests, to me, that Faithfulness-based naturalness patterns are extensive; Markedness-based ones dubious and hard to find.
- Let's try Faithfulness first.

---

[1] = dumb but charismatic. Here, articulatorily inept but charismatic.
[2] (2004) "Phonological acquisition in Optimality Theory: the early stages," in René Kager, Joe Pater, and Wim Zonneveld, eds., *Fixing Priorities: Constraints in Phonological Acquisition.* Cambridge: Cambridge University Press, pp. 158-203.

SUBSTANTIVE BIAS I: AVOID PHONETICALLY-SALIENT ALTERNATION

5. The best (to me) substantive bias

- A preference for similarity among the allomorphs of a single morpheme: minimize alternation.
- There is some evidence (e.g. Zuraw (2007)[3]; Kie will mention others) that this preference is based on raw phonetic similarity; not abstract feature-based similarity.
- If this is so, then the Faithfulness we should consider is Output-Output Faithfulness, not Input-Output Faithfulness (i.e. to an abstract underlying representation).

6. Theoretical apparatus

- Steriade's P-map theory
  - Steriade, Donca. 2001. Directional asymmetries in place assimilation: A perceptual account. In Hume, Elizabeth, and Keith Johnson (eds.) 2001. *The role of speech perception in phonology*. San Diego, CA: Academic Press.
  - Steriade, Donca. 2008. The phonology of perceptibility effects: The P-map and its consequences for constraint organization. In *The nature of the word: Studies in honor of Paul Kiparsky*, ed. by Kristin Hanson and Sharon Inkelas, 151–80. Cambridge, MA: MIT Press
- Zuraw's P-map-based *MAP constraint family
  - Zuraw, Kie. 2007. The role of phonetic knowledge in phonological patterning: Corpus and survey evidence from Tagalog. *Language* 83: 277-316.

7. *MAP constraints (Zuraw 2007)

- ***MAP(*x*, *y*)** assesses a violation to a candidate if a segment belonging to natural class *x* in the input is mapped to a corresponding segment in natural class *y* in the output.
- This kind of constraint is **more flexible** than classical OT correspondence, since there is no requirement that *x* and *y* differ by just one feature.

8. Default rankings of *MAP

- Appealing to phonetic substance, *MAP constraints are assigned a default ranking as follows:
  - *MAP constraints banning changes that cover a **larger perceptual distance** are assigned a default ranking higher than constraints banning smaller changes.
- This ranking preference is taken to be a **learning bias** in UG.
  - Given sufficient evidence in the ambient language, it is possible for learners to subvert the default rankings, but this is harder.

---

[3] Zuraw, Kie (2007). The role of phonetic knowledge in phonological patterning: Corpus and survey evidence from Tagalog. Language 83. Pp. 277-316.

9. Rankings in *MAP are based on the P-map

- Zuraw draws on earlier work by Steriade (2001, 2008) on the **P-map**.
  - = a compilation of the language learner's phonetic experience concerning the perceptual distance between forms.
  - From a Steriade ms. on line; font size = perceptual distance.

Obstruent voicing	V_V	C_V	V_R	V_]	V_T	C_T
p/ b	**p/b**	p/b	p/b	p/b	p/b	p/b
t/ d	t/d	t/d	t/d	t/d	t/d	t/d
k/ g	k/g	k/g	k/g	k/g	k/g	k/g
s/z	s/z	s/z	s/z	s/z	s/z	s/z

- It is assumed that language learners consult the P-map in assessing preferred rankings for *MAP constraints.

11. Backdrop to *MAP-cum-P-map theory

- Essential prediction: alternation prefers to be phonetically minimal; proposed long ago (Vennemann 1972, Kiparsky 1978/1982:65)

12. There's a phonological literature backing up a bias for phonetically-minimal alternation

- Kiparsky (1982)
  - a 15-year research program on "analogical" change
  - languages can even create new phonemes when this reduces the distance of phonological alternation
- Language acquisition
  - I'm not sure how common this is, but here is one relevant case:

    Children are able to innovate sequences that are illegal in the target language, in the interest of maintaining output to-output correspondence. This was observed by Kazazis (1969) in the speech of Marina, a four-year-old learning Modern Greek. Marina innovated the sequence *[xe] (velar consonant before front vowel), which is illegal in the target language. She did this in the course of regularizing the verbal paradigm: thus [ˈexete] 'you-pl. have' (adult [ˈeçete]), on the model of [ˈexo] 'I have'.[4]

- Fleischhacker (2001), Shademan (2002) [both on our website]
  - epenthesis favors location that changes the stem least saliently
- Zuraw (2007, cited above)
  - infixation favors location that changes the stem least saliently
- Lofstedt (2010, UCLA diss., our website)
  - allomorphic (irregular) phonological variation in Swedish; leveling happens where
- Ancient history
  - the widely-accepted-but-vague "phonetic similarity" requirement in phonemicization ([h]-[ŋ], [p^h]-[t]/[t^h]-[k]/[k^h]-[p], Hayes 2009 textbook)

13. A famous paper on bias

- Wilson, Colin. 2006. Learning phonology with substantive bias: an experimental and computational investigation of velar palatalization. *Cognitive Science* 30:945–982
- This was
  - a pioneering effort to study naturalness though artificial-language learning (palatalization of velars preferred before high vowels)
  - an early use of artificial-language experiments to explore learning bias
  - the first formal implementation of Steriade's concept of the P-map
  - an influential polemic for bias-based UG
  - the source of an effective new tool for modeling biased acquisition: maxent grammars with Gaussian priors
- We're not assigning it because we want to keep to our focus on variation; but it's worth reading for phonologists.
- For critiques of whether Wilson's conclusions are valid, see the readings.

14. Jamie White's dissertation work

- on the learning of saltatory alternations uses the P-map
- This pursues the line Wilson has taken but I think the results are more convincing.
- I will let Jamie tell you about it in his colloquium (Fri. May 3 11-1 Haines A26).
- I will do a couple simpler papers and use Wilsonian/Whitian analysis with them.

SKORUPPA ET AL. ON ALTERNATION-DISTANCE

[4] Quoting my own summary of Kazazis from Hayes (1999), "Phonological Acquisition in OT: The early stages." Kazazis ref: Kazazis, Kostas (1969). Possible evidence for (near-)underlying forms in the speech of a child.CLS 5: 382-388.

15. Reference

- Katrin Skoruppa, Anna Lambrechts and Sharon Peperkamp (2011) The Role of Phonetic Distance in the Acquisition of Phonological Alternations. In: S. Lima, K. Mullin & B. Smith (eds.), *Proceedings of the 39th Annual Meeting of the North East Linguistic Society*. Volume 2. Amherst, MA: GLSA, 717-729. 2011.

16. A pure experiment

- They mean to test alternation distance, and alternation distance alone.
- So, they make up some really crazy alternations that could have no authentic markedness basis in what we think we know about phonology.
- Controlling for distance using features (where features are simply the very traditional categories of place, manner, and voicing)
  - 1 feature, 2 features, 3 features.

Language	*Alternating phrases* *pair 1*	*pair 2*
S1	ʁe **p**amu – nø **t**amu	ʁe **z**afam – nø **ʒ**afam
S2	ʁe **ʃ**amu – nø **s**amu	ʁe **d**afam – nø **b**afam
M1	ʁe **p**amu – nø **s**amu	ʁe **d**afam – nø **ʒ**afam
M2	ʁe **ʃ**amu – nø **t**amu	ʁe **z**afam – nø **b**afam
L1	ʁe **p**amu – nø **z**amu	ʁe **t**afam – nø **ʒ**afam
L2	ʁe **ʃ**amu – nø **d**amu	ʁe **s**afam – nø **b**afam

17. The cases examined

- 1 feature distance ("small"): p → t ʃ → s z → ʒ d → b
- 2 feature distance ("medium"): p → s d → ʒ z → b ʃ → t
- 3 feature distance ("large"): p → z t → ʒ s → b ʃ → d

18. Context of the change

- Base form occurs after [ʁe].
- Changed form occurs after [nø]
- Unlikely that rounding on the vowel could induce any of the changes above (except, conceivably, z,s,d → b

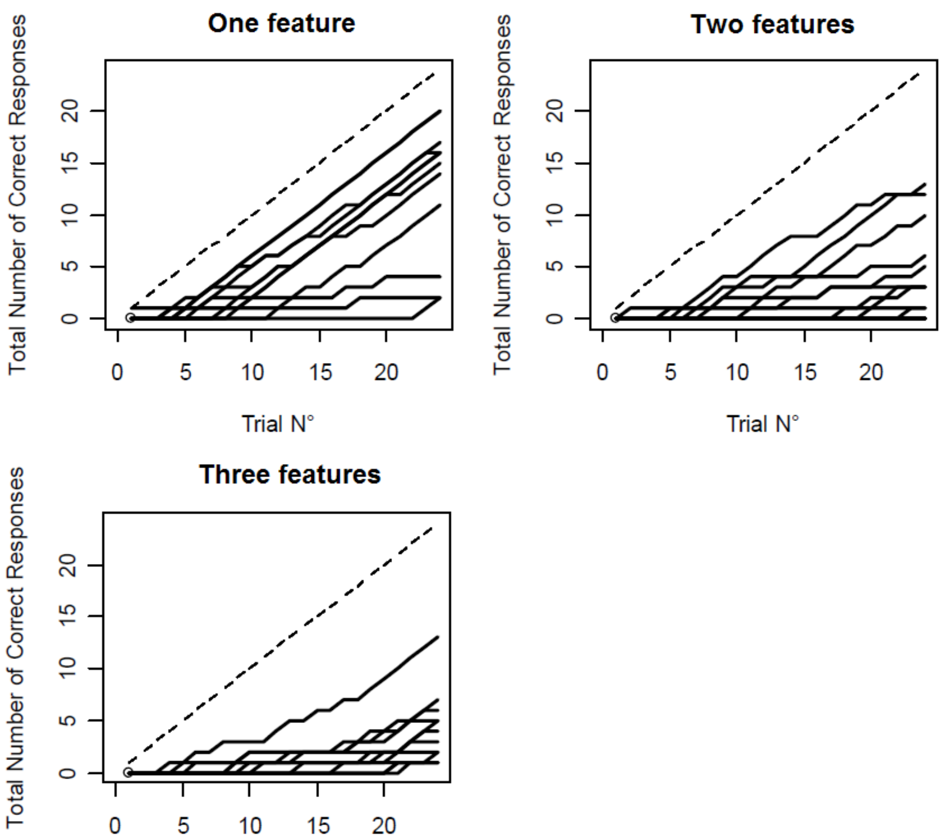
19. Elicitation method

- Quoting:

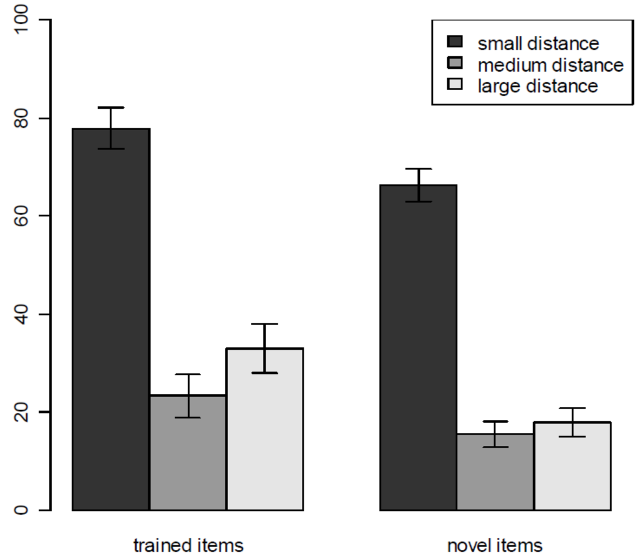
- "Step 1: After a blank screen of one second, the word *Moi*: ('I:') appeared on the screen. One second later, the first phrase of a pair was played (e.g. [ʁe pibut]).
- Step 2: One second later, the word *Vous*: ('You:') appeared on a new line below the former, and the participant could give a response. She had to press a button when finished.
- Step 3: The word *Correct*: ('Correct:') appeared on a new line, and one second later the second phrase of the pair was played (e.g. [nø tibut]).
- Step 4: The experimenter coded via the mouse whether the participant's response was correct. If so, she received positive feedback visually and moved on to the next trial; if not, she was told to try again and the trial was repeated."

- Then more training, without feedback.
- Then testing, on both trained and novel items.

20. Progress during training (total correct responses)



21. Final results (testing phase)



- Significance testing goes as you might imagine.
- Odd that the 2 vs. 3 difference showed up during training but not testing…

22. Their conclusion

- "We have focused on a heretofore unexplored aspect of phonetic naturalness, that is, phonetic distance. The results of our experiment show that alternations that involve phonetically minimal distant sounds are learned faster and better than other alternations."

23. Following up (illustration of Wilsonian/Whitian methods)

- Skoruppa et al. ponder whether a phonetic similarity metric might help.
- Discussion below borrows a scheme from White (in progress).

  Error rates in a perception-in-noise experiment →
  μ values for *MAP() constraints →
  train a maxent model as if it were an experimental subject →
  Model predictions →
  Compare with experimental data

- This is not really going to get us any closer to predicting the observed facts, but I hope it will be useful as a pedagogical example.

24. A confusion matrix

- From Wang and Bilger (*JASA* 1973)

Stimulus	Response															
	p	t	k	b	d	g	f	θ	s	ʃ	v	ð	z	ʒ	tʃ	dʒ
p	763	23	80	9	3	10	14	6	6	4	4	2	3	5	1	4
t	25	795	31	1	9	6	4	4	12	3	0	2	2	2	22	9
k	66	36	753	2	7	18	7	5	4	10	3	1	3	5	7	6
b	55	12	28	608	19	87	46	9	9	2	43	6	3	3	2	4
d	9	16	7	10	763	13	8	13	12	2	15	33	14	9	3	9
g	16	10	44	22	7	762	13	5	3	4	20	4	3	4	2	12
f	56	0	11	22	9	13	712	61	10	3	25	8	0	1	0	2
θ	11	14	9	9	42	10	73	656	28	7	23	36	13	1	1	3
s	7	15	12	1	3	4	15	23	828	14	1	1	11	0	1	0
ʃ	9	17	10	7	9	6	6	1	18	788	4	2	4	18	19	18
v	5	5	8	38	22	19	48	7	11	4	637	104	12	7	1	5
ð	8	7	5	38	77	4	10	37	6	7	136	556	34	8	1	2
z	4	8	15	4	12	10	8	22	42	8	54	35	687	17	4	3
ʒ	9	3	9	1	13	7	4	7	15	30	24	13	22	752	4	20
tʃ	10	17	14	5	2	7	2	1	5	14	2	1	0	5	727	121
dʒ	16	20	31	2	25	17	4	1	11	10	3	2	8	7	86	689

- = how often was the row header heard as the column header in noise

25. White's method for obtaining μ's from confusion matrices

- Set up a maxent grammar that predicts the number of confusions using *MAP constraints.
  - ➢ here, "Do not hear [x] as [y]"
  - ➢ The whole question of maxent as a framework for perception grammars deserves attention (cf. Boersma 1998, using stochastic OT)[5]
- Underlying premise: the probability that a speaker would accept an allomorphy pair with (e.g.) [p] ~ [b] is similar to the probability that a listener would confuse the two in noise.
  - ➢ I.e. the μs for $*\text{MAP}_{\text{PerceptionGrammar}}$ will be similar to the μs for $*\text{MAP}_{\text{OO}}$
- … a big if, but far more principled than just inventing μ's yourself.

26. My input file for establishing the μ's

- Values from Wang and Bilger

Input	*Candidate*	*Observed Freq.*
d	b	10
	d	763
	ʒ	9
p	p	763
	s	6
	t	23
	z	3
s	b	1
	s	828
ʃ	d	9

---

[5] See his published dissertation (1998), *Functional phonology: Formalizing the interactions between articulatory and perceptual drives.*

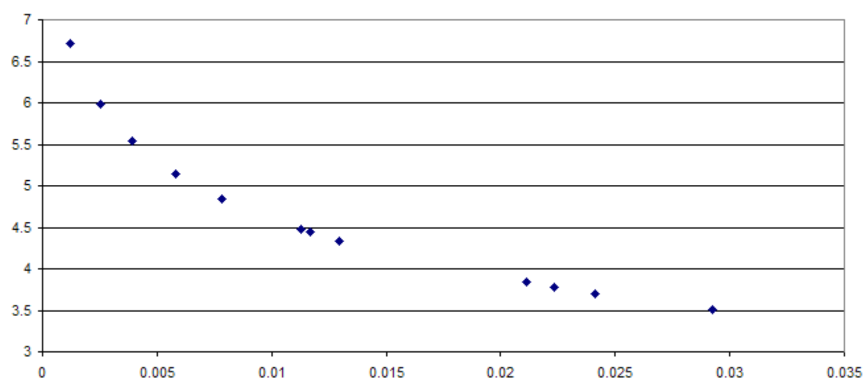
	ʃ	788
	s	18
	t	17
t	t	795
	ʒ	2
z	b	4
	z	687
	ʒ	17

27. The *Map constraints and the weights obtained

Constraint	Weight	Features different	Errors in Wang and Bilger
*Map(p, t)	3.5	1	23
*Map(z, Z)	3.7	1	17
*Map(S, s)	3.8	1	18
*Map(d, b)	4.3	1	10
*Map(S, t)	3.8	2	17
*Map(d, Z)	4.4	2	9
*Map(p, s)	4.8	2	6
*Map(z, b)	5.1	2	4
*Map(S, d)	4.5	3	9
*Map(p, z)	5.5	3	3
*Map(t, Z)	6.0	3	2
*Map(s, b)	6.7	3	1

- Average weights for the three categories: 3.8, 4.6, 5.7

28. Weight of *Map is a smooth (but nonlinear) function of error percentage



29. Next step: a grammar intended to predict the experimental results

- Constraints:
    - ➢ the set of *MAP constraints, as above

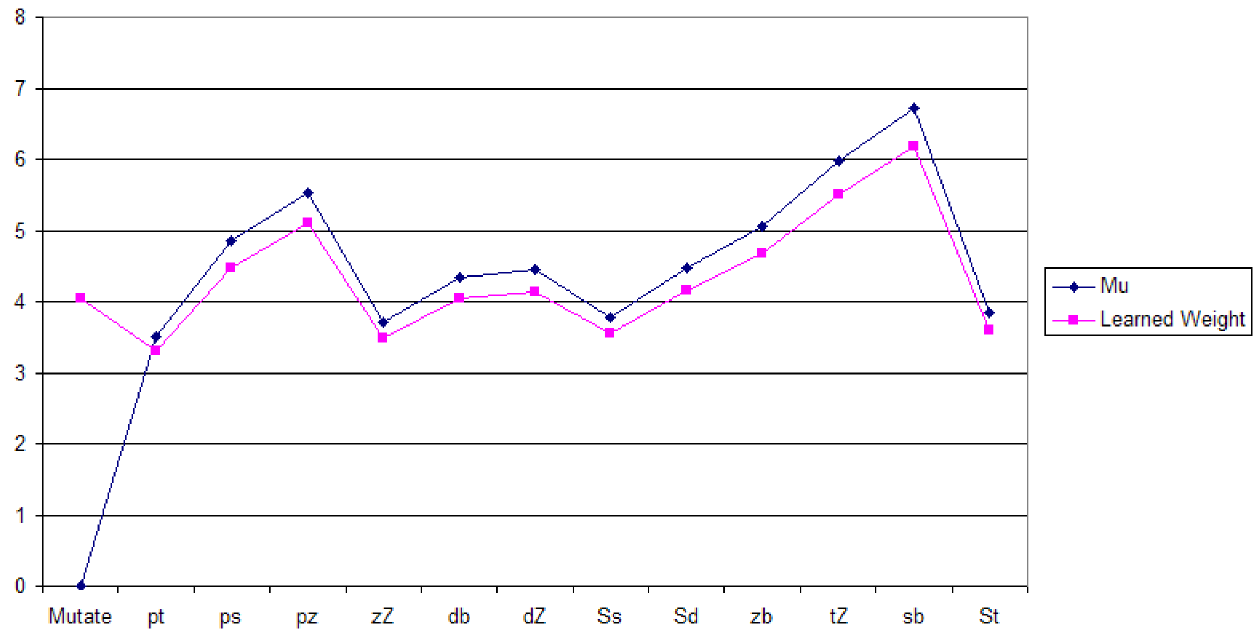
- ➢ a bogus, cover constraint MUTATE, which militates in favor of change — there is no legitimate markedness constraint that would prefer the changes in the data, so let's use a stupid one instead

Alternation			MUTATE	pt	ps	pz	zʒ	db	dʒ	ʃs	ʃd	zb	tʒ	sb	ʒt
d ~ b	d	0	1												
	b	24						1							
d ~ ʒ	d	0	1												
	ʒ	24							1						
p ~ s	p	0	1												
	s	24			1										
p ~ t	p	0	1												
	t	24		1											
p ~ z	p	0	1												
	z	24				1									
s ~ b	s	0	1												
	b	24												1	
ʃ ~ d	ʃ	0	1												
	d	24									1				
ʃ ~ s	ʃ	0	1												
	s	24								1					
ʃ ~ t	ʃ	0	1												
	t	24													1
t ~ ʒ	t	0	1												
	ʒ	24											1		
z ~ b	z	0	1												
	b	24										1			
z ~ Z	z	0	1												
	ʒ	24					1								

- This grammar, learned from the data, assigns a large weight to MUTATE and zero to all others, achieving a perfect match.

30. A biased-learning simulation

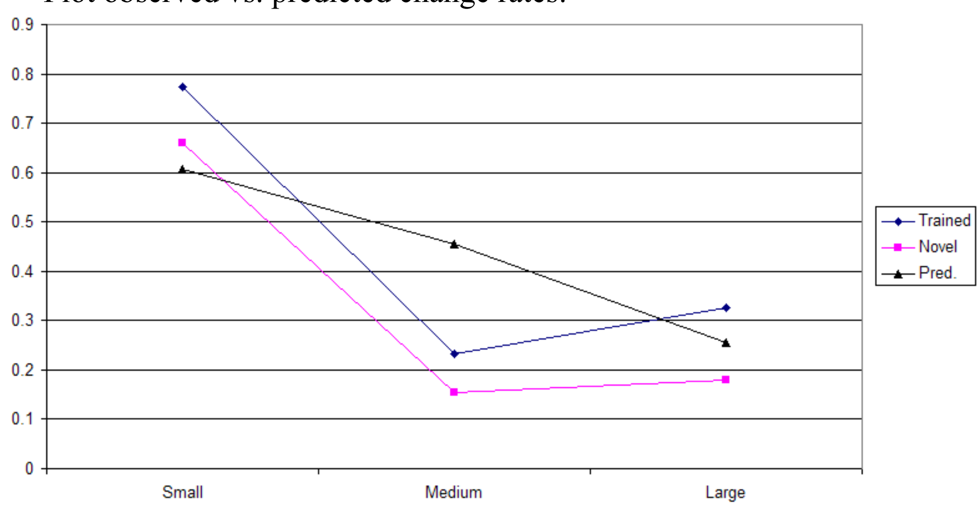
- The μ's for the *MAP constraints were set as above
- μ for Mutate was set at zero; which seems appropriate for an unmotivated constraint.
- σ for all was set at 0.05, fitting the data roughly by eye.
- The weights at the end of learning were a compromise between μ and what the learning data were demanding.



- In short:
  - MUTATE rises in response to observed alternations.
  - *MAP weights fall in response to observed alternation.
  - The learned *MAP() weights turn out to be a linear function of their μ's; $y = 0.8885x + 0.1859$; $R^2 = 0.9998$

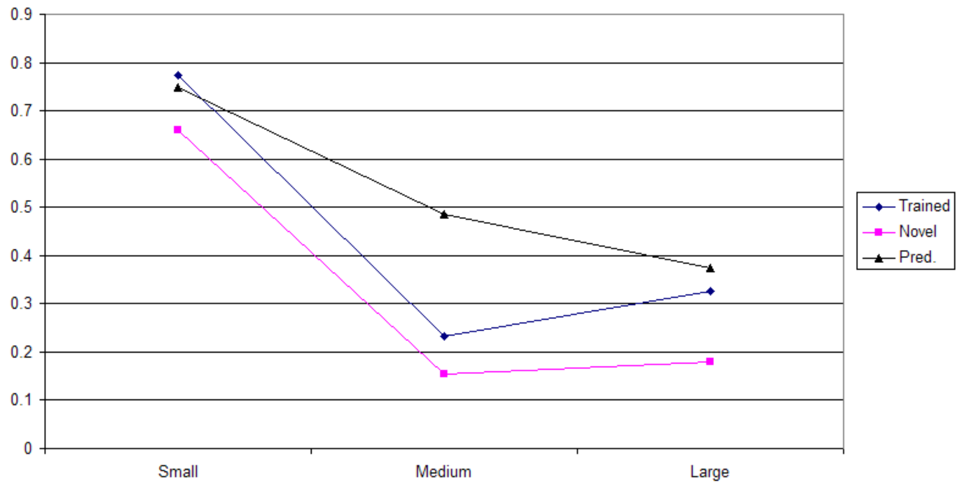
31. End result: no miracles

- Plot observed vs. predicted change rates:



- The error rates in Wang and Bilger are such that the model is predicting a difference between Medium and Large; just like a simple feature-counting model.

32. Using feature differences as the prior; σ = .03



- Presumably this is a slightly better fit; but both models face a mystery in explaining the lask of the expected difference between Medium and Large.

SUBSTANTIVE BIAS II: AVOID MARKED FORMS

33. A commonly addressed typological asymmetry

- Phonologists expect assimilation, which is ubiquitous, to be normal.
  - Many, many languages have vowel harmony
  - Local assimilation of consonants in place, voicing, nasality
  - Total assimilation of vowels when adjacent or separated by a glottal.
- Dissimilation is rare enough to be interesting whenever we encounter it.
  - Common among liquids (Latin/English -al showing up as -ar in words like *tubular, modular, lunar*)
  - A nice paper by Junko Ito from 1984 discusses some cases of vowel dissimilation

34. Ainu

- Source: Junko Ito (1984) Melodic dissimilation in Ainu. *Linguistic Inquiry* 15: 505-513.
- Ainu transitivizing suffix -V can be any of the five vowels in Ainu.
- Often, it is just a copy of the stem vowel:

CaC+a	mak-a	'to open'	tas-a	'to cross'
CeC+e	ker-e	'to touch'	per-e	'to tear'
CiC+i	pis-i	'to ask'	nik-i	'to fold'
CoC+o	pop-o	'to boil'	tom-o	'to concentrate'
CuC+u	tus-u	'to shake'	yup-u	'to tighten'

- On other occasions (not predictable), it is a high vowel that dissimilates to the stem vowel in backness:

CuC+i	hum-i	'to chop up'	mus-i	'to choke'
CoC+i	pok-i	'to lower'	hop-i	'to leave behind'
CiC+u	pir-u	'to wipe'	kir-u	'to alter'
CeC+u	ket-u	'to rub'	rek-u	'to ring'

- And if the stem vowel is /a/ (i.e. central, neither front nor back), a non-identical suffix vowel must be high but can be either [i] or [u]:

CaC+i	kar-i	'to rotate'	sar-i	'to look back'
CaC+u	ram-u	'to think'	rap-u	'to flutter'

- Ito has a clever 1980's-style analysis with floating autosegments.
- This might be fun to try to do in OT…

	i	u	e	o	a
i	✓	✓	*	*	*
u	✓	✓	*	*	*
e	*	✓	✓	*	*
o	✓	*	*	✓	*
a	✓	✓	*	*	✓

- Odd bit: dissimilation amidst total assimilation! Ito cites two further examples from Tzeltal and Ngbaka.

35. Farther afield

- I haven't searched much.
- Oceanic languages dissimilate aCa to VCa, where V is variously ə, e, i, o.[6]

36. The consensus result on experiments that test learnability of assimilation and dissimilation

- No difference; see Moreton and Pater's survey in the readings.

37. A specific case: Skoruppa and Peperkamp

- Katrin Skoruppa and Sharon Peperkamp. 2011. Adaptation to Novel Accents: Feature-Based Learning of Context-Sensitive Phonological Regularities. Cognitive Science 35: 348–366

[6] Low Vowel Dissimilation in Vanuatu Languages. Lynch, John, Oceanic Linguistics, Volume 42, Number 2, December 2003, pp. 359-406

38. An interesting kind of artificial-language experiment

- An artificial **dialect** (i.e. of French)
- Dialect recognition is presumably facilitated by the development of perceptual rule systems.

39. Dialect recognition in the forward direction

- r-ful American-English-speaking kids learning to recognize r-dropping speech would do well to internalize a form of "R Dropping in other people"
  - "r → ∅ in syllable coda when speaking as a Brit" (etc.)
  - Put this in the "forward model" for Bayesian recognition of British English.
  - Result: you have a second candidate (*sore*) for when you hear [sɔ] from an Brit.

40. Dialect recognition in the backward direction: using lexical incidence statistics

- The vast majority of British [ɑː] / ___ CV corresponds to American [ɑr].
  - harbor, party, ardor, harbor, artist, starboard, Harvard, Harvey, etc.
  - father, Mata Hari, Mali, tomato
- So [pɑːli] sounds to Americans like *parley*, not *Pali*; at least to me.

41. Less anecdotally: Sumner and Samuel

- "A priming study by Sumner and Samuel (2009)[7] suggests that Americans who have never lived in New York find it difficult to process English r-final words (e.g., *baker*) pronounced in the New York dialect, in which word-final [r] is deleted. By contrast, listeners who were raised in New York had less difficulty understanding r-less variants, even if they did not speak the New York dialect themselves and thus did not produce r-less words."

42. Dialect recognition in artificial-language learning

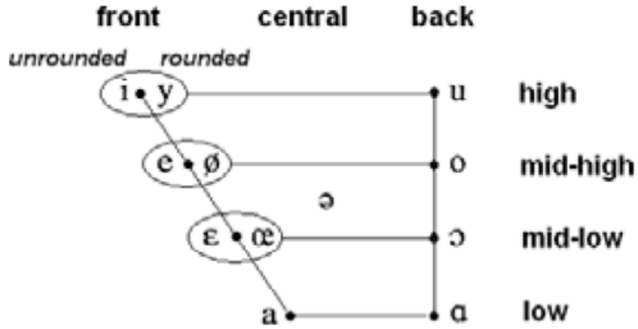
- S+P cite various experiments where a phoneme is mutilated in the training set, changing perceptions in the testing set.
- "Maye et al. (2008)[8] exposed American participants to a 20-min-story in which all front vowels were systematically lowered. For example, the verb *live* was pronounced as [lev].[9] In a subsequent lexical decision task, participants accepted new items with altered front vowels (e.g., [ked]—altered from *kid*) more readily as English words than during a pretest, suggesting that they adapted their phonetic categories to the new accent."

---

[7] Sumner, M., & Samuel, A. G. (2009). The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language*, 60, 487–501.
[8] Maye, J., Aslin, R., & Tanenhaus, M. (2008). The weckud wetch of the wast: Lexical adaptation to a novel accent. *Cognitive Science*, 32, 543–562.
[9] Do they mean [ɛ]? cf. article title.

43. Inventing pseudo-rounding harmony/disharmony

- French has no rounding harmony but its vowel system is nicely set up for it:



- P+S are frustratingly vague about their rules (why not use undergraduate notation?) but I think they are something like:

Rounding Harmony

$$\begin{bmatrix} +\text{syllabic} \\ -\text{back} \\ -\text{low} \end{bmatrix} \rightarrow [\alpha\text{round}] \;/\; \begin{bmatrix} +\text{syllabic} \\ -\text{back} \\ -\text{low} \\ \alpha\text{round} \end{bmatrix} C_0 \,\_\_\_$$

"A nonlow front vowel takes on the rounding of a preceding non-low front vowel."

Rounding Disharmony

$$\begin{bmatrix} +\text{syllabic} \\ -\text{back} \\ -\text{low} \end{bmatrix} \rightarrow [-\alpha\text{round}] \;/\; \begin{bmatrix} +\text{syllabic} \\ -\text{back} \\ -\text{low} \\ \alpha\text{round} \end{bmatrix} C_0 \,\_\_\_$$

"A nonlow front vowel takes on the opposite rounding from a preceding non-low front vowel."

44. A note on Rounding Disharmony

- In classical phonology, this is a **self-bleeding rule** and therefore yields different outcomes by directionality when you consider inputs like /y y y/.
- Perhaps the authors avoided words of this form …

45. Teaching the languages

- Balanced set of 152 words that can undergo Rounding Harmony (e.g. *liqueur* [likœʁ][10])
- Balanced set of 152 words that can undergo Rounding Disharmony (e.g. *pudeur* 'modesty' [pydœʁ][11], becoming [pydɛʁ])

---

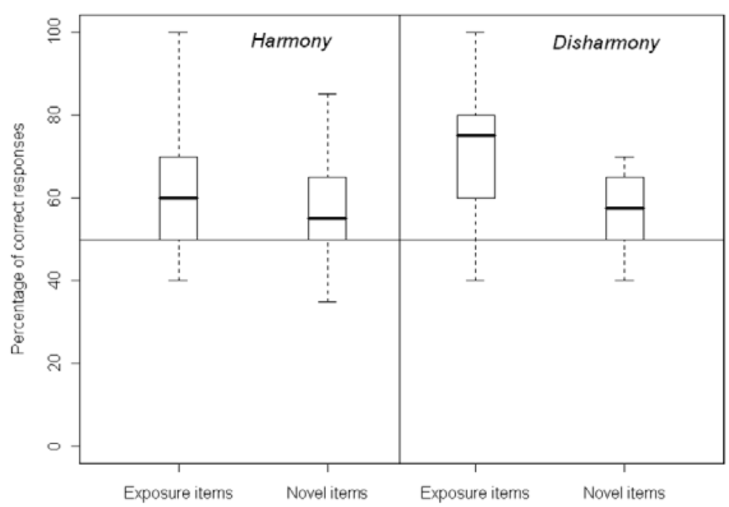
[10] All IPA transcriptions are my guess.

- Example sentences from the stories they wrote, given in orthography.
  - a. Standard French: *Sans pudeur, il se versa un verre de liqueur.*
    'He shamelessly poured himself a glass of liquor.'
  - b. Harmonic French: *Sans pudeur, il se versa un verre de **liquère**.*
  - c. Disharmonic French: *Sans **pudère**, il se versa un verre de liqueur.*
- Presumably, there were also cases where the Standard French word had *unrounded + rounded*.
- This method is nice because the listeners were asked to memorize the stories, not particular words; increasing naturalness of the context.

46. Testing procedure

- Play [likɛʁ] and [pydɛʁ].
- "Which one was pronounced with the French accent you just heard?"
  - [Socrates: correct answers?]

47. Results



- Statistics: linear mixed-effects models, now widely used in phonological studies…
- They are doing better than chance.
- They are not doing any better on harmony than disharmony
  - Indeed, disharmony looks better, though not significantly (p = .12).

48. Bottom line

- "These findings suggest, then, that there are no abstract linguistic biases favoring harmony over disharmony in perceptual phonological learning. Provided young children show the same behavior, the preponderance of vowel harmony in the world's languages

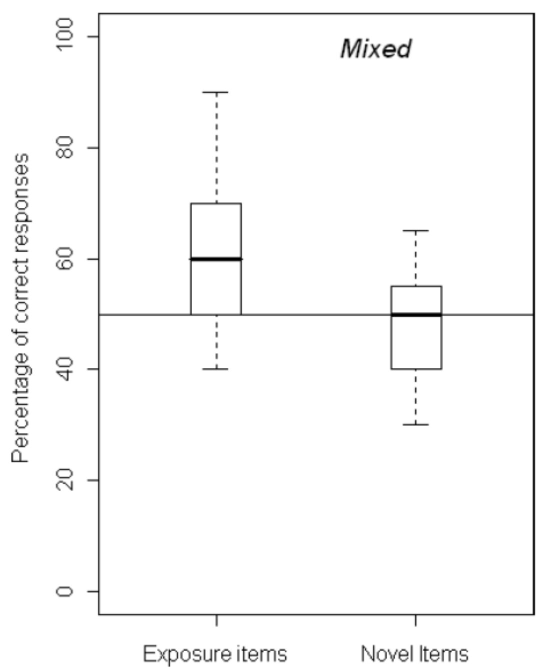
[11] All IPA transcriptions are my guess; psychology journals don't seem to care about this aspect of replicability.

must stem from a different source, most likely from phenomena of diachronic sound change or from an ease-of-articulation constraint favoring harmony that affects speech production only.”

49. Experiment 2: a simplicity-bias study

- Instead of Rounding Harmony and Disharmony, a “crazy” rules that is sometimes harmonic, sometimes disharmonic, depending on vowel height.
- “Mixed French, in which the high vowels [i] and [y] disharmonize to preceding front vowels, whereas the mid vowels [e], [ø], [ɛ], and [œ] harmonize to it.”
- *If* you believe in features as a basis for phonological learning, this is a very complex rule (two rules, really).
- Same method

50. Results: no effect



- The only effect seen is attributable to **memorization**.
  - ➢ Cf. the idea of a **accent-perception lexicon**; I know that British English has *father* [ˈfɑðə] and *bother* [ˈbɔðə], *master* [ˈmɑːstə] and *Newcastle* [ˈnjukæsəl] but I cannot predict this by rule (neutralization of proto /ɑː/, /ɔ/ in history of American English; irregular /æ/ fronting in British before /s/). This is memorization.

51. What does it all mean?

- As with other experiments cited by Moreton/Pater, no assimilation-dissimilation difference is found, contradicting typology.
- So nothing is gained by assuming a phonetic naturalness bias in learning.
- N.B. this bias is one of **Markedness** constraints, not Faithfulness (cf. bifurcation given above).
  - What constraints? The usual ones here would be AGREE(round), OCP(round).
- Other theoretical winners, from 2nd experiment
  - validity of features in guiding learning
  - simplicity bias (Moreton/Pater)

52. Are we done with Markedness bias?

- Perhaps not; I'd like to see experiments exploring *highly* marked configurations.
- E.g. nasal assimilation vs. dissimilation, played to listeners whose language has no consonant clusters.

53. Implications of the absence of a Markedness bias

- Classical OT is in big trouble! It embodies a tacit huge Markedness bias (non-gradient), embodied in the universal constraint set.
- If this is correct, phonological theory will have to redirect itself to algorithms for generalization-discovery, presumably with simplicity bias (Moreton/Pater).
  - I.e. the framework itself may be fine, but the constraints that are weighted/ranked will have to be discovered by the learner[12]

[12] My own struggles, with able collaborators: Albright/Hayes (*Cognition* 2001), Hayes/Wilson (2008, *LI*).