Embedding Grammar in a Quantitative Framework: Case Studies from Phonology and Metrics

a minicourse by

Bruce Hayes
Department of Linguistics
UCLA
bhayes@humnet.ucla.edu

Class 1: Four Kinds of Gradience in Language

1. Welcome

- This is an informal course, without requirements.
  - There are many readings, and the more you read the more you will learn…
  - I've designated quick daily reading for busy people.
- Readings are all downloadable:
  - Web site: http://www.linguistics.ucla.edu/people/hayes/brownminicourse
  - Also: backup handouts, links to relevant software.
- You can talk to me! I will take a break in each class, and can also make appointments to discuss course material and research you are doing.
  - My email: bhayes@humnet.ucla.edu.

2. Prospectus

- As linguistics moves beyond its theoretical infancy, it is becoming more important/fruitul to employ theories that include a quantitative component. Why?
  - To describe data more accurately and thoroughly.
  - To discover new things about languages—it turns out that there are many things the native speaker knows that we were previously unaware of, but discovered through the use of gradient theories (we'll see some here).
- Such theories need not be radically different from the kind that linguists have previously been used to dealing with—the older approaches are fruitfully "embedded" in the quantitative system.

3. Who might find this useful

- Linguistic analysts/theorists: tools that can help you.
- Computationalists: problems that if you can solve them you will make a tremendous contribution.

4. Minisyllabus (also on web site)

For full references, visit the web site.

Day 1. Four Areas of Gradience in Grammar
and the models we can use to analyze them

Papers I will draw upon: Labov (1969), Zuraw (2000), Boersma/Hayes (2001), Goldwater/Johnson (2003), Pater (forthcoming)
Suggested short readings: Goldwater/Johnson

Day 2. The Law of Frequency Matching
exemplification (Hungarian vowel harmony, Turkish and Dutch final devoicing)
deviations
modeling
Papers I will draw upon: Ernestus/Baayen (2003), Becker et al. (forthcoming), Hayes/Londe (2006), Hayes et al. (submitted)
Suggested short reading: Hayes/Londe (2006)

Day 3 Phonotactics
English and other languages
UG and the question of accidentally-true generalizations
Papers I will draw upon: Hayes/Wilson, Coetzee/Pater

Suggested short reading: Hayes/Wilson through section 5.

Day 4 General and Specific Constraints and how they Interact
English past tenses — Minimal Generalization
Challenging an OT assumption: is conflict the only basis for ranking/weighting the constraints?
Papers I will draw upon: Albright and Hayes (2003)
Suggested short reading: Albright and Hayes, skimming the analogical-model section

Day 5 Meter
sung and chanted verse
Shakespeare and Milton
Gerard Manley Hopkins
Papers I will draw upon: Hayes/Kaun (1996), Hayes (in press), Kiparsky (1975, 1977, 1989), Hayes (2008)

Suggested short reading: Hayes (in press)

5. Today

- Four kinds of gradience in phonology
- Start in on possible tools for analyzing them

6. Four kinds of gradience

By “gradience”, I mean “phenomena where an accurate description would require numbers”

- **Gradience in production**: people say the same thing in different ways, following a quantitatively-describable pattern.
- **Gradience in intuition**: people assign varying degrees of well-formedness to words, statistically following various phonological factors.
- **Gradience in lexical behavior**: the words of the lexicon show different phonological behaviors, statistically following various phonological factors.
- **Gradience in learning**: the data available to the child at any given time support differing hypotheses to varying degrees

GRADIENCE IN PRODUCTION

7. Some simple examples

- You can say *envelope* as ['?nv?lo?p] or ['? nv?lo?p].
- More interesting: systematically patterned: /æ/ as either [Uə] or [æ] before /n/ or /m/: *man*, *Dan*, *plan*, *Sam*, *spam*, etc., etc.
- This is ubiquitous—try doing consultant work and watching for it; or querying consultants with questions like “is this acceptable”?

8. Sociolinguistics

- A classic study:
  - ➢ William Labov studies hundreds of New Yorkers in the 1960’s: how often do they omit syllable final /r/?

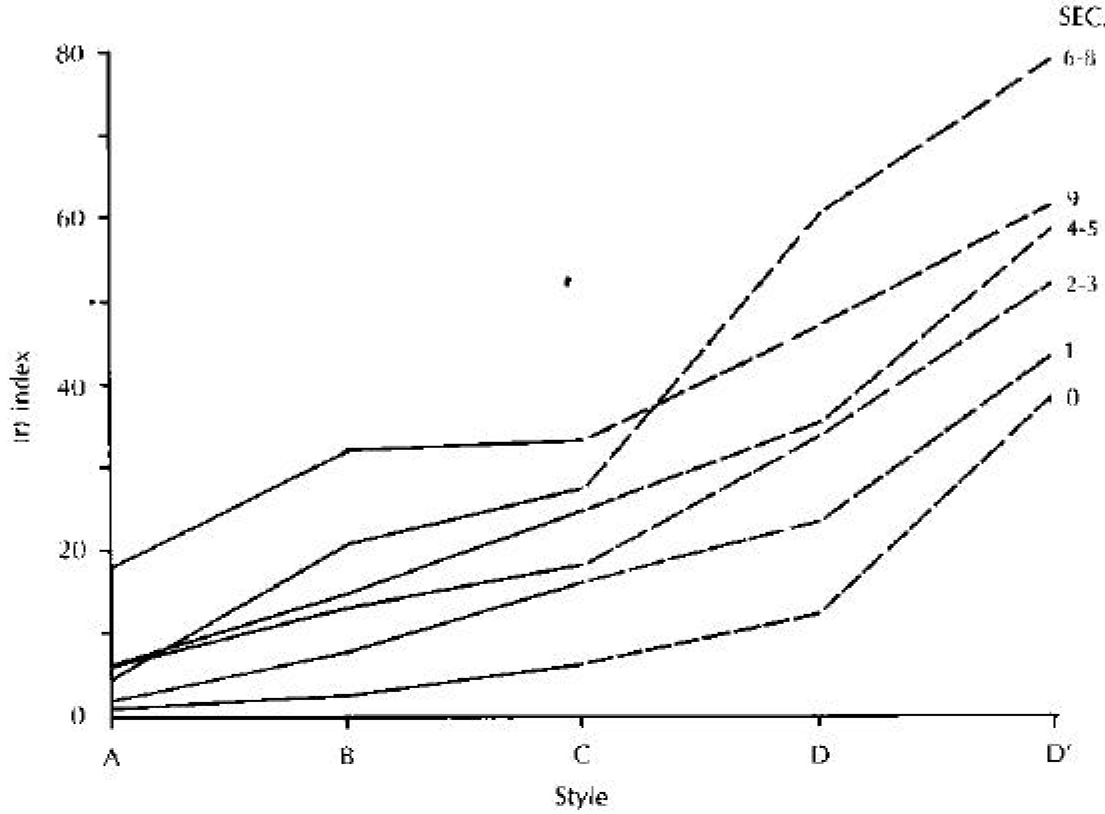


Fig. 4.2. Class stratification of a linguistic variable in process of change: (r) in *guard, car, beer, beard, board,* etc. SEC (Socio-economic class) scale: 0–1, lower class; 2–4, working class; 5–6, 7–8, lower middle class; 9, upper middle class. A, casual speech; B, careful speech; C, reading style; D, word lists; D', minimal pairs.

- Answer: it's totally systematic
  - The more attention is paid to speech, the more often /r/ is dropped.
  - The lower the social class you belong to, the more likely you are to drop /r/, unless you are lower-middle class and attention is being paid to your speech.
  - Later work developed very systematic accounts of other optional rules that are sensitive to segmental contexts, such as the dropping of /t/ in e.g. *past time.*
  - The numerically-annotated rules were an important formalization of gradience in outputs.

GRADIENCE IN INTUITION

9. Past tense formation

- Albright and Hayes (2003, *Cognition*) asked college students to rate imaginary English past tenses on a 1-7 scale, 1 = "completely bizarre, impossible", 7 = "completely normal, would make a fine past tense"
- Here is a range of representative irregular forms:

fleep	flept	6.09	compare *sleep*
nold	nold	6.05	compare *tell*
gleed	gled	6	compare *bleed*
scoil	scoilt	5.15	compare dialectal *spoil*
bize	boze	4.57	compare *rise*
shurn	shurnt	4.22	compare dialectal *burn*
ry'nt	rount	3.71	compare *find*
shy'nt	shount	3.39	compare *find*
stin	stan	2.74	compare *swim*
pum	pame	2.71	compare *come*
gezz	gozz	2.52	compare *get*

We were intrigued to find that a similar range is found for the regular past tenses:

gezz	gezzed	6.61
murn	murned	6.57
shurn	shurned	6.57
zay	zayed	6.39
tesh	teshed	6.22
stire	stired	6
chake	chaked	5.74
glit	glitted	5
chind	chinded	3.89

- We think much of this variation is explainable with a suitable grammar—Class 4.

10. Phonotactic well-formedness

- A control experiment for the Albright/Hayes study simply asked the subjects whether the word sounded like it could be an English word. We, like previous experimenters, got a continuous range of values in the responses.

bzarshk, 1.5; pwuds, 1.74; sprarf, 2.05; smeenth, 2.06; plonth, 2.26; thaped, 2.26; dwoge, 2.29; ploamph, 2.42; smeelth, 2.47; smairf, 2.47; thweeks, 2.53; smairg, 2.58; krilg, 2.58; trilb, 2.63; throiks, 2.68; frilg, 2.68; shwouge, 2.68; smeerg, 2.79; skluned, 2.83; pwip, 2.89; raint, 2.89; sfuned, 2.94; snoiks, 3; smum, 3.05; twu, 3.17; squalk, 3.26; nung, 3.28; gwenge, 3.32; shrooks, 3.32; chool, 3.42; shaint, 3.42; zapes, 3.47; blig, 3.53; flidge, 3.79; drice, 3.84; scoil, 3.89; bredge, 3.95; queed, 3.95; skick, 4; flet, 4; nold, 4; kive, 4.05; skride, 4.11; zay, 4.16; drit, 4.16; fleep, 4.16; blafe, 4.21; ghez, 4.21; gude, 4.32; plim, 4.37; chind, 4.37; glip, 4.53; bize, 4.58; grell, 4.63; tesh, 4.63; teep, 4.63; spling, 4.72; lum, 4.79; pum, 4.79; squill, 4.83; dize, 4.84; tunk, 4.84; nace, 4.84; shilk, 4.89; preak, 5; graint, 5; chake, 5.05; gleed, 5.05; gare, 5.11; shurn, 5.11; tark, 5.11; dape, 5.11; skell, 5.11; glit, 5.11; spack, 5.16; trisk, 5.21; rask, 5.21; stin, 5.28; snell, 5.32; plake , 5.39; murn, 5.42; mip, 5.47; stire, 5.47; rife, 5.53; stip, 5.53; pank, 5.63; pint, 5.67; kip, 5.84; whiss, 5.84

- We will discuss a model that aspires to predict these values in Class 3.

GRADIENCE IN LEXICAL BEHAVIOR

11. Gradience in lexical behavior

- Much phonology is lexically idiosyncratic.
- I.e. it involves a general pattern seen across the language, but for which you have to memorize the items to which the pattern is applicable.

12. An example of gradience in lexical behavior

- Zuraw, Kie (2000) *Patterned exceptions in phonology*, UCLA dissertation.

13. Zuraw: lexical study of percent application of Nasal Substitution in Tagalog

- Basics: take a prefix that ends with [ŋ]
- Attach to a stem that begins with a stop or fricative: [ptkbdgs]
- The stop or fricative disappears, leaving a trace: the [ŋ] takes on its place of articulation.

hukˈboʔ	'army'
paŋ-hukˈboʔ	'military'
kamˈkam	'usurpation'
pa<u>ŋ+k</u>amˈkam → pa**ŋ**amˈkam	'rapacious'
daˈla:ŋin	'prayer
ʔi-pa<u>ŋ+d</u>aˈla:ŋ-in → ʔi-pa**n**aˈla:ŋ-in	'to pray'

- This is taught as if it were exceptionless in beginning phonology classes.

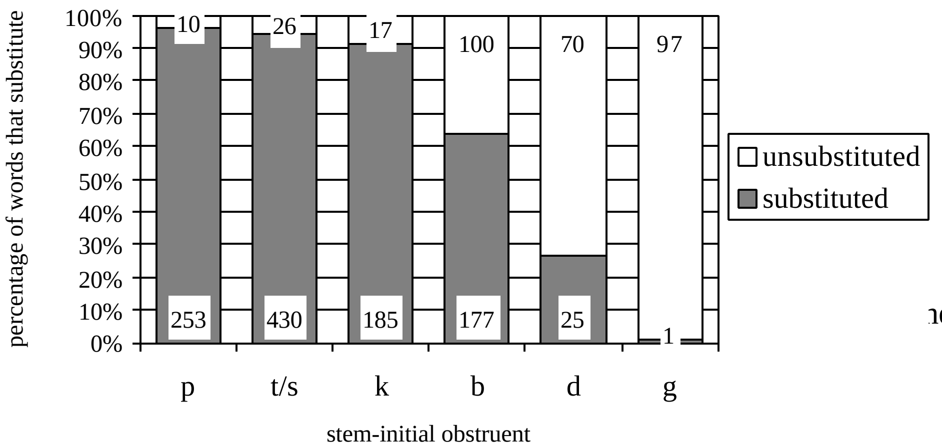
14. The pattern is riddled with exceptions

p	**p**ighatîʔ	'grief'	pa-**mi**-**m**ighatîʔ	'being in grief'
	poʔók	'district'	pa**m**-**p**oʔók	'local'
t	paq-**t**ú:loj	'staying as guest'	ka:-pa-**n**ulú:j-an	'fellow lodger'
	tabój	'driving forward'	pa**n**-**t**abój	'to goad'
k	**k**amkám	'usurpation'	ma-pa-**ŋ**amkám	'rapacious'
	kaliskís	'scales'	pa**ŋ**-**k**aliskís	'tool for removing scales'
b	mag-**b**igáj	'to give'	ma-**m**igáj	'to distribute'
	bigkás	'pronouncing'	ma**m**-**b**i-**b**igkás	'reciter'
d	**d**alá:ŋin	'prayer'	ʔi-pa-**n**aláŋ-in	'to pray'
	diníg	'audible'	pa**n**-**d**iníg	'sense of hearing'
g	**g**indáj[15]	'unsteadiness on feet'	pa-**ŋ**i-**ŋ**indáj	'unsteadiness on feet'
	gá:waj	'witchcraft'	ma**ŋ**-**g**a-**g**áwaj	'witch'
s	**s**ú:lat	'writing'	ma:-**n**u-**n**ulát	'writer'
			pa**n**-**s**ú:lat	'writing instrument'

15. Zuraw's further study

- Used a lexical+internet-gathered corpora
- Check for generalizations about *how often* the process applied.
- It's systematic.

16. Frequency of Nasal Substitution varies in the lexicon according to the stem-initial consonant

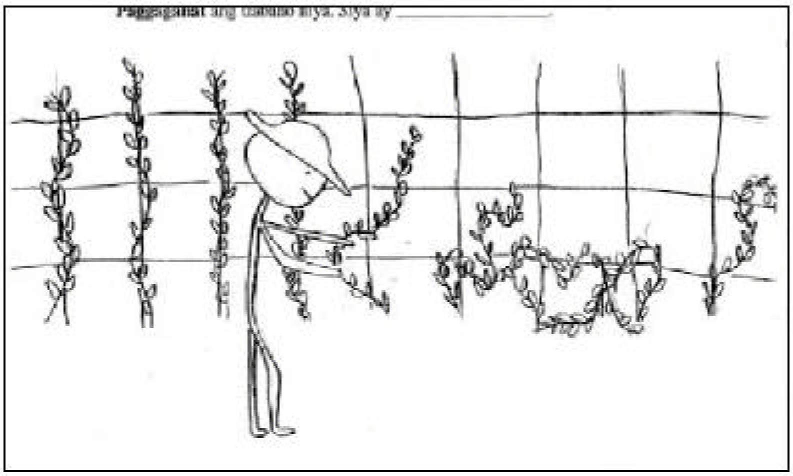


ıe

17. Is this result meaningful?

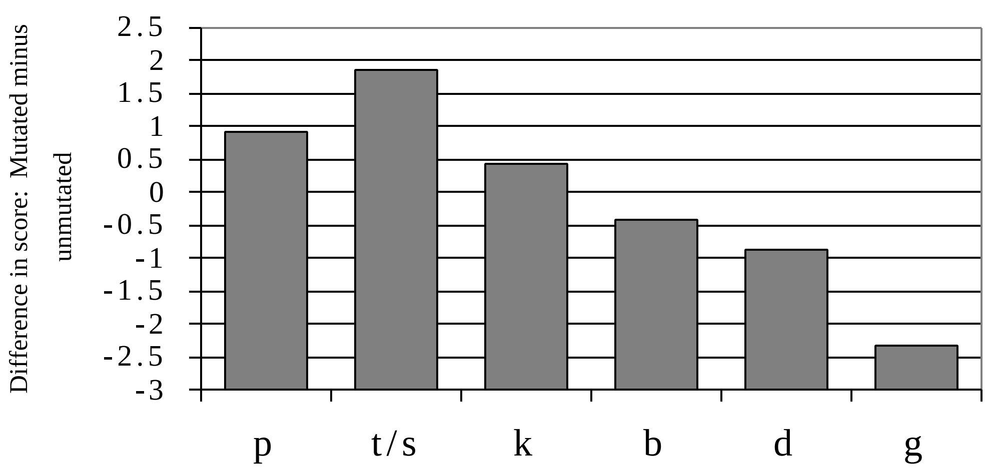
- Since speakers must memorize everything on a case-by-case basis, couldn't you just stop there? Or do they process the lexicon and detect the pattern?

18. Zuraw's Tagalog wug test



- Subjects given: Pag-ga-ganat aŋ trabaho niya. Siya ay [maŋganat/maŋanat]
- "His job is to *ganat*. He is a *ganat*er."
- Native speakers of Tagalog rate both mutated and unmutated forms on a 1-10 scale.

19. Results



20. The Law of Frequency Matching

- Speakers of languages with variable lexical patterns respond stochastically when tested on such patterns. Their responses aggregately match the lexical frequencies.
- Some other phonological experiments whose results support this law are reported in Eddington (1996, 1998, 2004), Berkley (2000), Coleman and Pierrehumbert (1997), Zuraw (2000), Bailey and Hahn (2001), Frisch and Zawaydeh (2001), Albright (2002), Albright and Hayes (2003), Pierrehumbert (2006), and Jun and Lee (2007).

- Sociolinguistic study demonstrates frequency matching by children during real-life phonological acquisition (Labov 1994, Ch. 20).

- **The Law in (much) broader perspective**

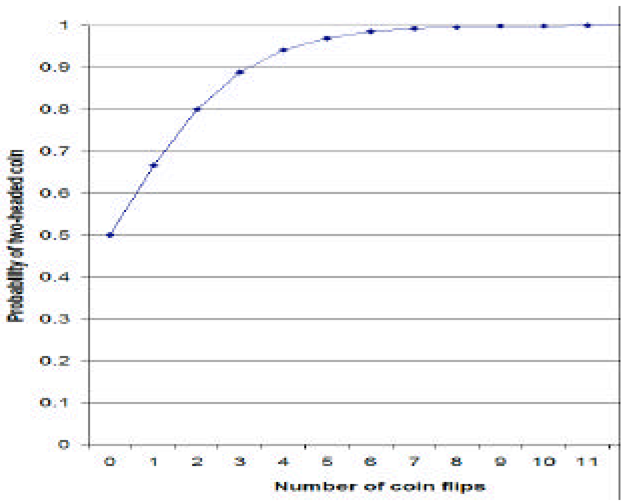
  - Frequency-matching is known to be a common ability in animals (Gallistel 1990, ch. 11); and in humans for nonlinguistic tasks (Hasher and Zacks 1984).

GRADIENCE IN LEARNING

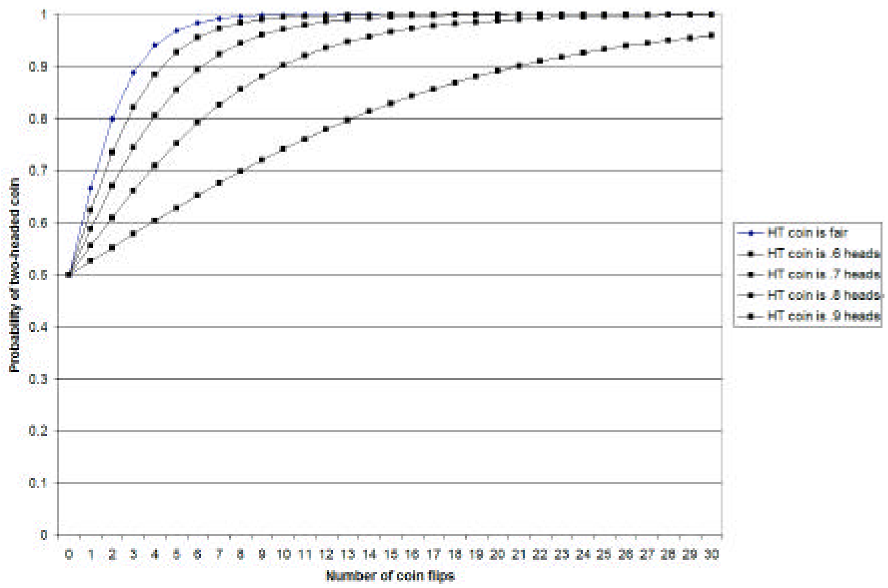
- **The Parable of the Possibly-Bad Coin**

  - I have in my pocket a normal coin and a two headed coin.
  - Without looking, I pull a coin at random from my pocket and flip it *n* times.
  - You must infer which one it is.

Observed data	Probability that the coin is two-headed.
H	2/3
HH	4/5
HHH	8/9
HHHH	16/17
…	
n heads	$2^n/(2^n+1)$



- The "learning curve" gets slower if we suppose that the HT coin is also fraudulent (weighted), so it yields heads more often then tails:



21. Interpreting the parable

- From the viewpoint of the language learner, every single aspect of experience looks gradient.
- The non-gradient aspects of grammar, as learned by the child, become so asymptotically.

22. Is there any way grammar could learned non-gradiently?

- Scenario: you have a default parameter setting (reference), which you deviate from only when you hear data.
- I don't think this will work, because of the problem of speech errors, e.g.
  - "first floor dlorm room", reported in Stemberger (1984, *Language*)

23. Suggestion

- We need to develop predictive and accurate scientific models of all four areas.
- If we find the right basic mathematical approach, it might serve for the analysis of all four.

MODELING I: HARMONIC GRAMMAR

24. A simple empirical example to work with

- What happens when Japanese speakers render foreign words, adapting them to the phonological system of Japanese?
- Source: Shigeto Kawahara (2006) "a faithfulness ranking projected from a perceptibility scale: the case of [+voice] in Japanese," *Language*.
  - I'm inspired in using this example re. gradience by Pater (in press) and earlier work.

25. Facts about Japanese Voiced Obstruents

- /b/, /d/, /dʒ/, /g/, /z/
- They only marginally tolerate appears in double (“geminate”) clusters:
    - ➢ ?[bb], ?[dd], ?[ddʒ], ?[gg], ?[zz] occur only in recently borrowed words.
    - ➢ An earlier stage of the language “fixed” the borrowed words by devoicing to [pp, tt, tt?, kk, ss].
        - In the “core,”, ur-Japanese vocabulary (“Yamato”), you only get one voiced obstruent per word. kata, kada, gata, ?gada
    - ➢ This is called **Lyman’s Law**.
    - ➢ It has noticeable effects elsewhere in the language, not covered here.

26. The striking fact

- Japanese speakers nowadays generally “fix” borrowed words only if they *both have a voiced geminate and violate Lyman’s Law.*
- Devoicing

gepperusu	‘Göbbels’
gutto	‘good’
betto	‘bed’
doretto	‘dreadlocks’
dettobooru	‘dead ball (baseball term)’
batto	‘bad’
deibitto	‘David’
butta	‘Buddha’
dokku	‘dog’
bakku	‘bag’
dorakku	‘drug’
bikku	‘big’

- No devoicing, only one condition met:

webbu	'web'
sunobbu	'snob'
habburu	'Hubble'
kiddo	'kid'
reddo	'red'
heddo	'head'
suraggaa	'slugger'
eggu	'egg'
furaggu	'flag'

bagii	'buggy'	bogii	'bogey'
bobu	'Bob'	bagu	'bug'
dagu	'Doug'	daibu	'dive'
daiyamondo	'diamond'	doguma	'dogma'
giga	'giga (10^9)'	gaburieru	'Gabriel'
gibu	'give'	gaidansu	'guidance'

27. An old-fashioned linguistic rule

[+geminate] → [–voice] / ___ … [+voice]
/ [+voice] … ___

- This is unsatisfactory since it doesn't link up what Japanese speakers already knew (semi-badness of Lyman's Law violation, semi-badness of voiced geminates.)
- But the loan behavior is spontaneous…

28. Constraint-based linguistics

- Posit simple, often directly observable constraints, and work out ways for changes to follow from them (e.g. Optimality Theory (Prince/Smolensky)).
- In OT, you use GEN, a formal component that makes up all possible outputs that could derive from an input.
- So whereas rules must generate, the constraints need only select (which can often be a lot simpler).

29. Constraints needed here

- LYMAN'S LAW: *[+voice] … [+voice]
- NO VOICED GEMINATE: *[–sonorant, +long, –voice]
- FAITHFULNESS: Do not change the voicing of an input form.

30. Constraints are functions

- They take as their argument an input-output pair, and output an integer, the number of violations.

31. The core theoretical question

- How do the constraints interact to choose the correct output?

32. What might we have said if the problem were new to us?

- Give every constraint a number, its **weight**, reflecting its strength.
- Every candidate gets a penalty score: every violation of a constraint adds the weight of the constraint to the penalty.
- The winner is the candidate with the lowest score.
- The score (or in some approaches, the logarithm of the score) is called its **harmony**.

33. Intellectual pedigree of Harmony Grammar

- Work of Smolensky and colleagues, starting in the 1980's; now updated collected in two big volumes—harmonic grammar is strikingly implementable with connectionist networks.
- Linear OT of Frank Keller (2000)
- Recent work by Pater and colleagues.

34. References

- Smolensky, Paul, and Geraldine Legendre. 2006. *The Harmonic Mind*. Cambridge: MIT Press. (summarizing work of two decades, including early Harmonic Grammar)
- Pater, Joe; Rajesh Bhatt; and Christopher Potts. 2007. Linguistic optimization. Amherst: University of Massachusetts ms.
- Pater, Joe. 2008. Linguistic Optimization. Amherst: University of Massachusetts ms. (on course web site)

35. Harmonic grammar is a family of theories

… because you can opt to do more with the harmony scores. See below.

36. Harmonic grammar works great for the Japanese example

… because the problem seems rather blatantly to involve addition.

37. Specifically

- … We give FAITHFULNESS a weight that is higher than either LYMAN'S LAW or *VOICEDGEM alone.

- Technical note: [pː] etc. are single segments that are [+long]—just one Faithfulness violation.

input: /bobu/	FAITHFULNESS	LYMAN'S LAW	*VOICED GEM	**Harmony**
weights:	*3*	*2*	*2*	
☞ bobu	0	1	0	**2**
bopu	1	0	0	**3**
pobu	1	0	0	**3**
popu	2	0	0	**6**

input: /eggu/	FAITHFULNESS	LYMAN'S LAW	*VOICED GEM	**Harmony**
weights:	*3*	*2*	*2*	
☞ egːu	0	0	1	**2**
ekːu	1	0	0	**3**

input: /dogːu/	FAITHFULNESS	LYMAN'S LAW	*VOICED GEM	**Harmony**
weights:	*3*	*2*	*2*	
☞ dokːu	1	0	0	**3**
dogːu	0	1	1	**4**
togːu	1	0	0	**3**
tokːu	2	0	0	**6**

38. Back to Theories I: are there other ways to get constraint interaction?

Optimality Theory is utterly different.

- Rank the constraints (intuitively: strictest constraint at the top)
- *Filter* the candidates using the ranking, as follows:
  - ➢ Consider top constraint. Find the candidates that perform best on it (fewest violations). Throw out all others.
  - ➢ Repeat with next-ranked constraint, continuing until you've got just one candidate left.
- Pure theory says, no tradeoffs (unless you "cheat", adding a constraint that says, "don't violate these two constraints"; Smolensky 1995 et seq.)
- So the Japanese case makes a nice poster child for Harmonic Grammar.[1]
- Also, we have a clear basis for evaluating the rival theories: are there "ganging" effects?
  - ➢ See Pater paper in readings for further careful evaluation.

[1] Poster children often have complicated back stories. Kawahara (2006, cited above) jiggers the constraints so that the geminate effect is made part of Faithfulness, and thus succeeds with conventional OT.

ON TO GRADIENCE

39. The Japanese data again

- For conceptual clarity, Kawahara and Pater analyze an idealized version of the Japanese data, stripping away the gradience.
- But suppose we try to get the gradience, too? Let's look at the data more closely.
  - ➢ Rather often, words violating both LYMAN'S LAW and *VOICEDGEM *don't* get devoiced.
  - ➢ Occasionally, a word that merely violates *VOICEDGEM *does* get devoiced.
  - ➢ Words that merely violate LYMAN'S LAW appear to be stable.
- Here are approximate numbers:

doggu-type words:	57.4	"fix" their voicing
eggu-type words:	3.7	"fix" their voicing
bobu-type words:		apparently always keep the input voicing

- What sort of grammar could generate these numbers?

40. Task: amplify Harmonic Grammar so that it generates gradient predictions

- There are at least two ways to do it.
- We'll cover one of them today.

41. Maximum Entropy Grammars ("Maxent")

- Reference: Goldwater, Sharon, and Mark Johnson. 2003. Learning OT constraint rankings using a maximum entropy model. *Proceedings of the Stockholm Workshop on Variation within Optimality Theory*, ed. by Jennifer Spenader; Anders Eriksson, and Osten Dahl, 111–120. Stockholm: Stockholm University Department of Linguistics.

42. Basic goal

- Translate the constraint weights and violations into *predicted probabilities*.
- Here, probabilities of loan-adaptation for a given form.

43. The math is simple

Terminology varies, we're using terminology from Hayes and Wilson (2008)

(1) *Definition: score*

The *score* of a phonological representation x, denoted $\mathrm{h}(x)$, is:

$$\mathrm{h}(x) = \sum_{i=1}^{N} w_i\, C_i(x)$$

where w_i is the weight of the ith constraint,
$C_i(x)$ is the number of times that x violates the ith constraint, and
$\sum_{i=1}^{N}$ denotes summation over all constraints (C_1, C_2, … C_N).[2]

This is the same as Harmony above.

That is: going constraint by constraint, multiply violations by weights, and take the overall sum.

(2) *Definition: Maxent value*

Given a phonological representation x and its score $\mathrm{h}(x)$ under a grammar, the *maxent value* of x, denoted $P^*(x)$, is:

$$P^*(x) = \exp(-\,\mathrm{h}(x))$$

That is, negate the sum you obtained, and take e to this value.

(3) *Probability*

Given a phonological representation x and its maxent value $P^*(x)$, the *probability* of x, denoted $P(x)$, is:

$$P(x) = P^*(x)\ /\ Z \qquad \text{where } Z = \sum_{y \in \Omega} P^*(y)$$

That is, find the share of the candidate among the overall sum of maxent values.

44. Why is the model called "Maximum Entropy"

- Goldwater and Johnson: "Maximum Entropy models are motivated by information theory: they are designed to include as much information as is known from the data while making no additional assumptions (i.e. they are models that have as high an entropy as possible under the constraint that they match the training data)."

---

[2] Our "scores" are closely related to the *harmony* values explored in Smolensky (1986) and subsequent work (Smolensky and Legendre 2006); hence the abbreviation h(x). The term "score" is also used in Prince (2002). The use of scores, but without their theoretical interpretation under maximum entropy as probability, is the basis of "linear optimality theory" (Keller 2000, 2006).

45. Finding the weights

- I'm going to leave the details to next time. For now:
  - The answer is "whatever best matches the data".
  - So, you set the weights by constructing a data corpus, and letting a machine algorithm learn them.

46. Two purposes of machine learning in linguistics

- More ambitious: modeling how human children might learn a language.
- More mundane: creating a grammar that is more accurate than what you could create by hand.

47. A Maxent analysis

simulation file, from which learning took place
for convenience, I assumed 1000 words of each type, reflecting Kawahara's percentages

			FAITHFULNESS	LYMAN'S LAW	*VOICED GEM
dogùu	dogːu	436		1	1
	dokːu	574	1		
	togːu	0	1		1
	tokːu	0	2		
egːu	egːu	963			1
	ekːu	37	1		
bobu	bobu	1000		1	
	bopu	0	1		
	pobu	0	1		
	popu	0	2		

48. Software

- Weighting was done with the "Maxent Grammar Tool", linked from the course website (Colin Wilson/Ben George).

49. Results: weights

FAITHFULNESS	14.55
LYMAN'S LAW	3.53
*VOICED GEM	11.29

50. Results: predictions

Input:	Candidate:	Observed:	Predicted:
dogːu	dogːu	0.431683	0.431716
	dokːu	0.568317	0.568277
	togːu	0	0.000007
	tokːu	0	0
egːu	egːu	0.963	0.962988
	ekːu	0.037	0.037012
bobu	bobu	1	0.999967
	bopu	0	0.000016
	pobu	0	0.000016
	popu	0	0

- You can hand-verify the right column (as I did) by doing the math above (e.g., in Excel).

51. Intuitive queries about maxent

- Does it work?
  - When the data are consistent with some analysis, I find that maxent gets results accurate within several decimal places (see later on for "smoothing", a deliberate decrease in accuracy).
  - When the data are inconsistent, I find that maxent yields a plausible compromise between conflicting cases.
  - Not so for similar algorithms (next time)
- Does it work in principle?
  - Yes, when you define "work" in the right way, there is a mathematical proof that it works.
  - Uniquely so, as far as I know, for this algorithm.
- What this business with taking the exponent of the negative? (see 43b)
  - A constraint is a decrement of output probability. By coverting to the exponential, we convert addition to multiplication—which is how probabilities are combined are in general.

52. Next

- How maxent weights are learned.
- Other gradient models.
- Applying the models to the Law of Frequency Matching: case studies.