Talk at National Chiao Tung University
May 11, 2012

Bruce Hayes
UCLA

How predictable is Italian word stress?

1. Project on which this talk is based

- This is a research collaboration in progress with **Erica Cei** of UCLA.
- For more on the project, visit our lab notebook: http://italianstressstudy.blogspot.com

BACKGROUND

2. The debate over arbitrary rules/constraints in phonology

Classical Optimality Theory (Prince and Smolensky 1993; McCarthy 2002)	*Approaches that permit arbitrary rules or constraints*
Only a **single universal inventory of constraints** is available in phonology; only the constraint ranking is language-specific.	Language learners can **induce** (possibly) arbitrary generalizations from the data they encounter.

3. Assessing the universal-constraint research strategy

- Plus:
  - It's a tight, rigorous theory whose predictions are clear (result of all possible rankings should match observed languages)
  - Explaining language learning is relatively easy (simply learn the ranking; Tesar and Smolensky 2001).
- Minus:
  - What if its basic premise is simply not true? We should check.

4. How might we check for the existence of arbitrary constraints?

- Step I: intensive scrutiny of the data pattern of languages.
- Step II: experimentation: are the patterns embodied in arbitrary rules and constraints actually internalized by native speakers?

5. The current state of research in this area

- Quite a bit of evidence has accumulated that **arbitrary phonological generalizations can be learned by humans**: see the literature-review section of Hayes and White (in press).[1]

[1] Downloadable: /www.linguistics.ucla.edu/people/hayes/Papers/HayesWhitePhonologicalNaturalnessAndPhonotacticLearning.pdf

- A strong **holdout position** is held by Becker and colleagues (Becker et al. 2011a, 2011b), whose experimental work asserts the sufficiency of the universal constraint set.
- **Bias theories** (Wilson 2006 and later work; see Hayes and White (in press) for review) are a sort of compromise:
  - Language learners more easily recognize patterns based on universal constraints, but have some purely-inductive capacity as well.

6. Phonology without universal constraints can still be theoretical phonology

- We can explore:
  - Learning algorithms that can locate arbitrary constraints
  - The role of phonetic naturalness in guiding phonological learning

7. Research plan for this project

- Carry out "Step I" (in (4) above) for Italian stress = conduct a thorough study, permitting ourselves to posit arbitrary constraints whenever we like.
- Initial goal is **accurate prediction** of stress, by whatever means necessary.
- Later on we hope to test whether people learn the relevant generalizations.

8. We will have to examine exception-ful patterns

- There are **hardly any exceptionless phonological regularities** concerning Italian stress.
- Therefore, we are willing to seek regularities that involve exceptions.
- Hence the work will be **probabilistic**, trying to achieve a quantitative match to the data.

9. Plan for this talk

- **Data corpus and research methods**
- **Five types of constraints** that can help predict Italian stress.
- **Assess some grammars** and the effect of the five constraint types.

ITALIAN DATA CORPUS

10. Corpus work

- A digital corpus permits us to check many proposed constraints quickly against the full vocabulary.
- We use a small search utility program I wrote, available at
  - www.linguistics.ucla.edu/people/hayes/EnglishPhonologySearch/

11. Goal in constructing the corpus

- Approximate the phonological lexicon of a possible Italian experimental subject (= young adult internet user)
- Method: assess lexical knowledge of one reference speaker— coauthor Erica Cei, a native speaker.

- Such a "living lexicon"[2] should work better than using a dictionary
  - Dictionaries contain many obsolete words.

12. Basis of our corpus

- The most frequent words (> 5 occurrences) 73,573 words used in Italian television subtitles, uploaded to Wiktionary.org by Matthias Buchmeier.
- The complete text of the novel *La coscienza di Zeno* by Italo Svevo
- We removed duplicates, then filtered out loanwords, obsolete words, abbreviations, compounds, words with pronominal clitics.
- Total words in corpus: 33,316

13. Producing the pronunciation entries

- These reflect the pronunciations of Erica Cei, checked with other speakers.
- These pronunciations appear to be rather close to standard Italian.

BASIC FACTS ABOUT ITALIAN WORD STRESS

14. Stress is phonemic (unpredictable)

- Minimal pairs:
  - [ˈankora] 'anchor' [anˈkora] 'still'
  - [ˈpagano] 'they pay' [paˈgano] 'pagan'
  - [ruˈbino] 'ruby' [ˈrubino] 'that they steal'
- These are not common, suggesting there is some room for prediction.

15. Three-syllable window

- Stress must fall on one of the three last syllables of the word ("three-syllable window").[3]
- So words can have
  - Antepenultimate stress: [ˈankora]
  - Penultimate stress: [anˈkora]
  - Final stress: [penseˈra] 'will think'

16. Stress on final syllables

- This is uncommon.
- Of the ca. 1500 cases of final stress in our corpus, the great majority are accountable by just a few suffixes:

[2] This term comes from Inkelas's Turkish Electronic Living Lexicon (http://linguistics.berkeley.edu/TELL/).

[3] A small, systematic exception is found in 3 plur. verbs like [ˈindikano] 'they indicate'; these inherit the stress of the corresponding 3 sg. Preantepenultimate stress also arises in encliticized forms, which we ignore here.

[-ˈo] ‘3rd sg. preterite, 1 sg. future’
[-ˈa] ‘3rd sg. future’
[-ˈi] ‘1 sg. preterite’
[-iˈta] ‘-ity’
[-ˈu] ‘-itude’ (4 examples)

17. Simplifying the problem slightly

- Let us only predict *penultimate vs. antepenultimate* stress, in words of at least three syllables.
- This modest simplification helps greatly with the computations.
- Basic numbers: in words of at least three syllables, there are
  - 18,386 words with penultimate stress (81.9%)
  - 4,057 words with antepenultimate stress (18.1%)

GRAMMAR FRAMEWORK AND ILLUSTRATION

18. Maxent grammar

- Very similar to Optimality Theory (= OT, Prince and Smolensky 1993)
  - The grammar is made of a **set of constraints**
  - A **batch of candidates** is created by a candidate-creating component (in OT, called “GEN”).
  - We find the **violations** for each constraint for each candidate and constraint and use them to pick a winner.
- Unlike OT, maxent uses constraint **weights** (real numbers), not ranking.
- Unlike OT, maxent assigns not one single winner but a **probability** to every candidate.
  - This works well for a gradient problem like Italian stress.
- There are **other frameworks** that do the same thing:
  - Stochastic OT; Boersma 1997, Boersma and Hayes 2001; Noisy Harmonic Grammar, Boersma and Pater 2008
  - Only maxent is equipped with a **provably converging algorithm** for finding the weights, so we use it here.

19. Our GEN function is trivial

- It inputs the segmental form of a word of at least three syllables, and produces two candidates:
  - one with penultimate stress
  - one with antepenultimate stress

20. Assigning probabilities to candidates

- For each word we want to calculate its predicted probability for penultimate vs. antepenultimate stress.
- This is done using the computation illustrated in (22)-(26) below.

21. Where do constraint weights come from?

- They are **calculated** as part of doing the analysis — software finds a best fit to the data.
- For the software, see footnote.[4]

22. A very small sample grammar — three constraints

a. STRESS PENULT	Assess a violation to candidates with antepenultimate stress.
b. STRESS ANTEPENULT IN VpV#	Assess a violation if there is penultimate stress in a word ending VpV.
c. STRESS PENULT IF PENULTIMATE VOWEL IS [a]	Assess a violation to candidates with antepenultimate stress when the penultimate vowel is low [a].

23. Constraint weights — obtained by calculation in R over whole corpus

STRESS PENULT	1.20
STRESS ANTEPENULT IN VpV#	2.91
STRESS PENULT IF PENULTIMATE VOWEL IS [a]	1.16

[4] For multicandidate problems we use the Maxent Grammar Tool (http://www.linguistics.ucla.edu/people/hayes/MaxentGrammarTool/). For two-candidate problems we use the bayesglm() function in R (www.**r**-project.org/). For why logistic regression can be substituted for full-scale maxent in such cases, see www.linguistics.ucla.edu/people/hayes/ItalianStress/Text/AboutLogisticRegression.pdf.

24. Total up the penalties for constraint violations

- Four representative forms; these happen to have antepenultimate stress.

		STRESS PENULT	ANTEP. VpV#	PENULT [a]	**Total**
Input	*Candidates*	**1.20**	**2.91**	**1.16**	
Penelope	peneˈlope[5]		*		**2.91**
‘Penelope’	☞ peˈnɛlope	*			**1.20**
canapa	caˈnapa		*		**2.91**
‘hemp’	☞ ˈcanapa	*		*	**2.35**
algebra	alˈgebra				**0**
‘algebra’	☞ ˈalgebra	*			**1.20**
ippopɔtamo	ippopoˈtamo				**0**
‘hippopotamus’	☞ ippoˈpɔtamo	*		*	**2.35**

25. Take *e* to the negative of the result

		STRESS PENULT	ANTEP. VpV#	PENULT [a]	Total	***e* to neg of previous column**
		1.20	2.91	1.16		
Penelope	peneˈlope		*		2.91	**0.055**
	☞ peˈnɛlope	*			1.20	**0.302**
canapa	caˈnapa		*		2.91	**0.055**
	☞ ˈcanapa	*		*	2.35	**0.095**
algebra	alˈgebra				0	**1**
	☞ ˈalgebra	*			1.20	**0.302**
ippopotamo	ippopoˈtamo				0	**1**
	☞ ippoˈpɔtamo	*		*	2.35	**0.095**

[5] We would obtain [e], not [ɛ], in this form because of Vowel Reduction; see (35) below. Same for [o] and [ɔ] in *ippopotamo*.

26. Compute the share of the total for each candidate — this is predicted probability (of penultimate stress)

		STRESS PENULT	ANTEP. VpV#	PENULT [a]	Total	*e* to neg of previous column	**Probability of penultimate stress**
		1.20	2.91	1.16			
Penelope	pene'lope		*		2.91	0.055	**.153**
	☞ pe'nɛlope	*			1.20	0.302	(=.055/(.055+.302))
canapa	ca'napa				2.91	0.055	**.365**
	☞ 'canapa	*	*	*	2.35	0.095	(=.055/(.055+.095))
algebra	'algebra				0	1	**.768**
	☞ 'algebra	*			1.20	0.302	(=1/(1+.302))
ippopɔtamo	ippopo'tamo				0	1	**.913**
	☞ ippo'pɔtamo	*		*	2.35	0.095	(=1/(1+.095))

27. There are words in Italian that have exactly the same constraint violations yet have penultimate stress

		STRESS PENULT	ANTEP. VpV#	PENULT [a]
Eurɔpa	☞ Eu'rɔpa		*	
'Europe'	'Europa	*		
daccapo	☞ dac'capo		*	
'from the beginning'	'daccapo	*		*
amore	☞ a'more			
'love'	'amore	*		
toccata	☞ toc'cata			
'touched'	'toccata	*		*

28. Goal of the grammar

- The stress patterns of individual words are almost certainly memorized (Zuraw 2001).
- What we seek is the native speaker's sense of *what is likely in general in Italian word stress*—which is why we need probabilities.

29. The overall lexicon of Italian is well matched by the assigned probabilities

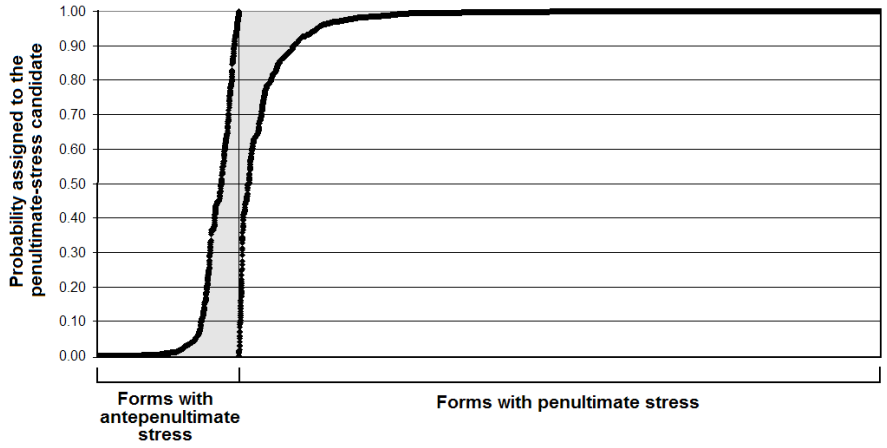
	Total words of each type	*Observed ratio*	*Probability assigned by grammar to penultimate candidate*
Words like *Euˈropa* Words like *Peˈnɛlope*	5 26	0.161	0.153
Words like *dacˈcapo* Words like *ˈcanapa*	1 2	.333	0.365
Words like *aˈmore* Words like *ˈalgebra*	11047 3334	0.768	0.768
Words like *tocˈcata* Words like *ippoˈpɔtamo*	7334 697	0.913	0.913

30. A way to evaluate such grammars

- If the grammar guesses penult with 0.8 probability and the stress is penultimate, we will count this as **0.8 correct, 0.2 error**.
- If the grammar guesses 0.8 penult and the stress is antepenultimate, we will count this as **0.2 correct, 0.8 error**.
- We total the correct values and divide by total forms, obtaining percent correct.
- The above grammar (three constraints) gets **71.5%** correct.
- Our best-performing grammar (97 constraints; see next item) gets **94.7%** correct

31. Distribution of guesses in a large grammar (97 constraints)

- We sort the data first by antepenultimate > penultimate, then by ascending prediction of the grammar for penultimate stress.
- Gray area is error, as defined above; white area is 94.7% of total.



WHERE CAN WE GET GOOD CONSTRAINTS?

32. Five sources

I. Classical generative phonology
II. Phonetically natural constraints
III. Telescoped constraints
IV. Lexical stratum constraints } excellent source: historical linguistics
V. Morphology

I. CLASSICAL GENERATIVE PHONOLOGY

33. Two good sources for classical generative analysis of Italian stress

- den Os and Kager (1986) / D'Imperio and Rosenthall (1999)
  - These provide analyses that achieve the three-syllable window.
  - They express the generalizations in a more sophisticated form than we will use here.

34. Heavy penults

- HEAVY PENULT: stress must be penultimate when the penult is a heavy syllable.
- Database has **6455** words with heavy penult and penultimate stress
  - E.g. *lambrusco* [lam.ˈbrus.ko] 'grape variety', *andante* [an.ˈdan.te] 'walking', *speranza* [spe.ˈran.tsa]
- Database has **2** words with heavy penult and penultimate stress
  - They are: *mandorla* [ˈman.dor.la] 'almond', *polizza* [ˈpo.lit.tsa] 'little note'

35. Vowel reduction and "backwards" constraints

- Italian vowel inventory:

i u
e o
ɛ ɔ
a

- In Italian, /ɛ/, /ɔ/ → [e, o] when stressless (exceptionless)

[ˈl**ɛ**tto]	'bed'	[l**e**tˈtino]	'bed-diminutive'
[ˈk**ɔ**ro]	'choir'	[k**o**ˈrale]	'choral'

- We **deduce** that if a syllable has /ɛ/ or /ɔ/ it is stressed.
- This is a "backwards" constraint but necessary to our particular project.
- So, two exceptionless constraints:

- IF PENULTIMATE LAX MID VOWEL, THEN PENULTIMATE STRESS.
- IF ANTEPENULTIMATE LAX MID VOWEL, THEN ANTEPENULTIMATE STRESS.

36. Performance of the classical generative model

- A grammar with the three constraints of this section, plus STRESS PENULT (22b) gets 79.9% right.

II. PHONETICALLY NATURAL CONSTRAINTS

37. Phonetically natural constraints: Ryan on weight

- Ryan, Kevin (2011) *Gradient weight in phonology*, UCLA diss.
  - studied quantitative meter in various languages; corpus study and wug test in English
  - focus was on areas where the effect of weight is not absolute, but statistical
- Outcome:
  - Where the effect of weight is absolute, weight criteria are **sharp** (e.g. *all* CV:, CVC count as heavy)
  - Where the effect of weight is statistical, weight criteria get **blurry**, and invoke a blend of phonetically natural criteria, for example:

Vowel length	longer is heavier
Vowel height	lower is heavier
Coda consonant count	more is heavier
Onset consonant count	more is heavier

38. Where Italian fits in

- It has one razor-sharp criterion (heavy = CVC), used in the Heavy penult constraint (34).
- Elsewhere (when penult is light), gradience sets in, following Ryan's principle.

39. Some Ryan-inspired constraints that work well for Italian

Constraint	*Penult/Antep in corpus*	*Weight*[6]
STRESS PENULT WHEN ITS VOWEL IS LOW.	7335/699	3.16
STRESS PENULT WHEN IT HAS A TRIPLE ONSET (CCC)	18/0	0.92
STRESS PENULT WHEN IT HAS (AT LEAST) A DOUBLE ONSET (CC)	2596/178	0.84

III. TELESCOPED CONSTRAINTS

[6] Hereafter, all weights reported are for Grammar II in (52) below.

40. Latin stress and telescoping

- Latin, the ancestor of Italian, stressed the penult when the penult was heavy, else the antepenult.
- This pattern was altered — "warped" —by later sound changes.
- When we look at the Italian pattern, we see Latin "through a telescope"; hence the constraints we write are "telescoped constraints".
    - For telescoping see: Kenstowicz and Kisseberth (1977, 64-65), Iverson and Salmons (2005), Hayes and White (in press)

41. Example of telescoping in Italian stress: Latin penultimate [u]

- Historical vowel reduction: [u] in stressless penults evolved into [o].
- So, any surviving penultimate [u] are likely to be stressed!
- From database: 866 penult, 88 antepenult.
- **Telescoped constraint**: STRESS PENULTIMATE [u]

42. Historical handbooks are useful for finding telescoped constraints

- Procedure:
    - Read Grandgent's (1927) handbook about Italian historical phonology.
    - Try to deduce which sound changes would give rise to telescoped constraints.

43. Example of telescoping in Italian stress: Latin penultimate [u]

- Historical vowel reduction: [u] in stressless penults evolved into [o].
- So, any surviving penultimate [u] are likely to be stressed.
- From database: 866 penult, 88 antepenult.
- **Telescoped constraint**: STRESS PENULT IN PRESENCE OF PENULT [u]

44. Some additional telescoped constraints we found

Constraint	*Penult/Antep in corpus*	*Weight*
STRESS ANTEPENULT IN WORDS ENDING IN [olV][7]	5/298	5.73
STRESS ANTEPENULT IN WORDS ENDING IN [erV][8]	62/496	4.28
STRESS PENULT WHEN THE PENULT HAS A RISING DIPHTHONG[9]	1887/56	2.21
STRESS PENULT WHEN THE FINAL HAS A RISING DIPHTHONG[10]	164/0	2.19

[7] Because unstressed penultimate vowels frequently reduced to [o] before [l]; Grandgent 61.
[8] Because unstressed penultimate vowels frequently reduced to [e] before [r]; Grandgent p. 61.
[9] Because CV'CiVCV went to CV'CjVCV, with accent shift, after gliding; Grandgent p. 12.
[10] Because CV'CVCiV went to CV'CVCjV; Grandgent p. 12.

IV. CONSTRAINTS BASED ON LEXICAL STRATA

45. Lexical strata

- Many languages have a "learned" vocabulary stratum, composed of loan words from a prestigious source language.
- Such strata often have **special phonology**.
  - See e.g. Chomsky and Halle (1968); Ito and Mester (1994)
- Source languages for such strata:
  - for English: Latin (and to some extent Greek)
  - For Japanese and Korean: Chinese
  - For Persian: Arabic
  - For Italian: Latin (and to some extent Greek)

46. The resistance of Latinate words to sound change

- The Latin > Italian sound changes have many exceptions in **learned words** — where historically, people felt they should using prestigious Latin-like pronunciations.
- Example: Latin obstruent + [l] usually became obstruent + [j], but not in learned words (Grandgent p. 85):
  - Regular words: L. [klaːmaːre] > It. [kjaˈmare] 'call', L. [florem] > It. [fjore] 'flower'
  - Learned words: L., [platonikus] > It. [platɔniko] 'Platonic'

47. Antepenultimate stress "feels" learned/Latinate in Italian

- Grandgent, p. 20:
  - "all proparoxytones [= words with antepenultimate stress] … have a certain rhetorical ring, suggesting Latin"
- (Why so? Perhaps because many rules of **syncope** were blocked in learned, Latin-based words
  - Loss in normal words: ˈ*computus* > ˈ*conto* 'count'
  - Retention in learned words: ˈ*debitus* > ˈ*debito* 'debt'
  - So, antepenultimate stress survived better in learned words.)

48. The scenario

- Certain phonological configurations — those that represent learned-word-overrides of sound change — become **cues for learned status** (Moreton and Amano 1999)
- Learned status becomes a **cue for antepenultimate stress**.
- Ideally, a good model would have a separate parameter, "degree of Learnedness."
- We will just fold the learnedness constraints directly into the model.

49. Some learnedness-based constraints that work well for predicting stress

Constraint	*Penult/Antep in corpus*	*Weight*
STRESS ANTEPENULT IN WORDS ENDING IN [VpV] = (22b) [11]	5/26	3.48
STRESS PENULT WHEN ANTEPENULT HAS UPPER-MID VOWEL [12]	7926/420	2.09
STRESS PENULT WHEN THE WORD CONTAINS OBSTRUENT + [l] [13]	330/109	0.71
STRESS ANTEPENULT IN WORDS OF AT LEAST 5 SYLLABLES [14]	917/2966	0.50

V. SUFFIX-SPECIFIC CONSTRAINTS

50. Many suffixes have particular stress properties

-aˈtore	'agentive'	always penultimate stress
-ˈissimo	'superlative'	always antepenultimate stress
-ido	'-id'	always presuffixal (hence antepenultimate) stress

- Of course, including constraints individual suffixes in the model greatly improves performance; example constraint:
  - STRESS ANTEPENULT IN WORDS ENDING IN *-ido*
- Is this "predicting" stress?
  - No, because the constraints express lexical properties of morphemes.
  - Yes, because each suffix represents a whole class of forms that behave in parallel.

LINING UP THE MODELS AND TESTING THEM FOR SIGNIFICANCE

51. Overview

- Review: we've looked at five kinds of constraints:
  - Classical generative phonology
  - Phonetically natural constraints
  - Telescoped constraints
  - Lexical stratum constraints
  - Suffix-specific constraints
- The best model appears to be the one with all five.
- Everything matters: take away any of the five and the model performs worse
  - All model comparisons are significant (likelihood ratio test).

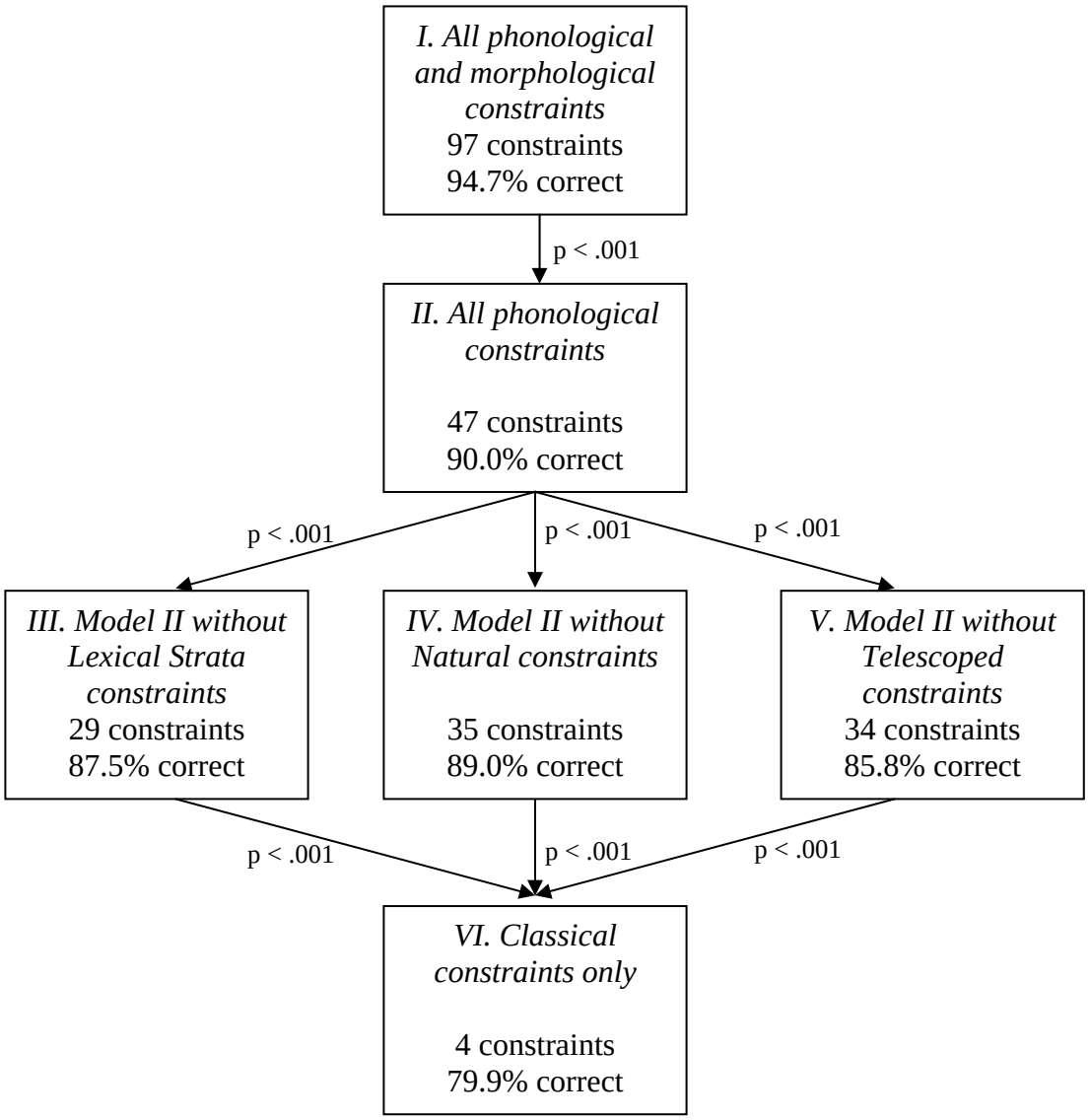
[11] For why see (47); [p] became [v] intervocalically in non-learned words.

[12] Because of the conventions for reading Latin. Orthographic *e* and *o* were customarily rendered as [ɛ] and [ɔ] in words — especially antepenultimately stressed words — felt to be Latin; Grandgent p. 20.

[13] Because [l] became [j] after obstruents in non-learned words; Grandgent p. 73.

[14] Because learned words tend to be long, perhaps because they resist otherwise general syncope processes; Grandgent pp. 53-59.

52. Comparing a hierarchy of models for predicting Italian stress



CONCLUSIONS

53. What we were trying to do

- Establish a high-detail analysis of Italian stress —
- Long term goal is testing it psycholinguistically, and thus bring evidence to bear on the controversy described above in (1)-(5) above.

54. What we learned about discovering effective constraints

- Unsurprisingly, the constraints discovered in classical phonology work very well.
- Phonetically natural constraints (Ryan) help somewhat.

- History helps: the lexicon is sufficiently stable that telescoped generalizations left over from Latin are still helpful in the analysis today.
- Lexical-strata constraints help: antepenultimate stress is learned, and thus correlates with the historically-derived cues to learnedness.

55. In conclusion

- There is much more to say in attempting to predict Italian stress than is given in traditional generativist accounts.

References

Becker, Michael, Lauren Eby Clemens, and Andrew Nevins (2011a) A richer model is not always more accurate: the case of French and Portuguese plurals. On line: www.phonologist.org/projects/FrenchPortuguese/#text

Becker, Michael, Nihan Ketrez, and Andrew Nevins (2011b) The Surfeit of the Stimulus: Analytic biases filter lexical statistics in Turkish laryngeal alternations. *Language* 87: 84–125.

Boersma, Paul. 1998. Functional phonology: Formalizing the interactions between articulatory and perceptual drives. The Hague: Holland Academic Graphics.

Boersma, Paul and Bruce Hayes (2001) Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32: 45-86.

Boersma, Paul and Joe Pater (2008) Convergence properties of a gradual learning algorithm for Harmonic Grammar. Rutgers Optimality Archive 970.

Chomsky, Noam and Morris Halle (1968) *The Sound Pattern of English*. New York: Harper and Row.

den Os, Els and René Kager (1986) Extrametricality and stress in Spanish and Italian. *Lingua* 69:23-48.

D'Imperio, Mariapaola and Sam Rosenthall (1999) Phonetics and phonology of main stress in Italian. *Phonology* 16: 1-28.

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

Grandgent, Charles H. (1927) *From Latin to Italian: An historical outline of the phonology and morphology of the Italian language*. New York: Russell and Russell.

Hayes, Bruce and James White (in press) Phonological naturalness and phonotactic learning. To appear in *Linguistic Inquiry*. [www.linguistics.ucla.edu/people/hayes/Papers/HayesWhitePhonologicalNaturalnessAndPhonotacticLearning.pdf]

Ito, Junko and Armin Mester (1994). Japanese phonology. In Goldsmith, J. (ed.), *The Handbook of Phonological Theory*, Blackwell, 817-838.

Iverson, Gregory K., and Joseph C. Salmons. 2005. Filling the gap. *Journal of English Linguistics* 33:207-221.

Kenstowicz, Michael and Charles Kisseberth (1977) *Topics in phonological Theory*. Academic Press.

McCarthy, John J. (2002) *A Thematic Guide to Optimality Theory*. Cambridge: Cambridge University Press.

Moreton, Elliott, and Shigeaki Amano (1999) Phonotactics in the perception of Japanese vowel length: Evidence for long-distance dependencies. *Proceedings of the 6th European Conference on Speech Communication and Technology*, Budapest.

Moreton, Elliott (2008). Analytic bias and phonological typology. *Phonology* 25:83-127.

Prince, Alan and Paul Smolensky (1993/2004) *Optimality Theory: Constraint Interaction in Generative Grammar*. Oxford: Blackwell.

Ryan, Kevin (2011) *Gradient weight in phonology*, UCLA diss.

Tesar, Bruce and Paul Smolensky (2000) *Learnability in Optimality Theory*. Cambridge, Mass.: MIT Press.

Wilson, Colin. 2006. Learning phonology with substantive bias: an experimental and computational investigation of velar palatalization. *Cognitive Science* 30:945–982

Zuraw, Kie (2000). *Patterned Exceptions in Phonology*. Ph.D. dissertation, UCLA.