
Week 4 – Optimality Theory — Richness of the Base and Phonotactics

March 10 and 12, 2008

1 Phonotactics

1.1 What is Phonotactics?

- (1) Phonotactics are the principles/generalizations that allow us to determine phonological legality:
 - a. Chomsky and Halle (1965): [brik] is possible and exists [blɪk] is possible and doesn't exist *[bnɪk] is impossible
 - b. More precisely, phonotactics consists of the set of legal surface forms in a language, regardless of whether they actually are dictionary entries in the language (i.e. are in the OED).
- (2) Phonotactic well-formedness is gradient (?[pɔ̃ɪk]), but we will idealize for the moment to a grammatical/ungrammatical distinction.
- (3) Phonotactics is learned in the absence of negative evidence: from what we do hear, we figure out what we never expect to hear.
- (4) OT: What is legal results from how high Faithfulness is ranked
 - a. The higher the Faithfulness constraints, the more things emerge from the grammar.

1.2 Case I: Phonemicness

- (5) As an exercise, assume the following constraints:

IDENT([VOICE])	(Faithfulness)
*[-SONORANT, +VOICE]	(Markedness; cf. Hawaiian)

- ★ Consider both rankings and state the distribution of [p] and [b].

★ What would happen if a Hawaiian speaker foolishly internalized the UR /ba/.

- (6) Lexicon Optimization (hypothesis of Prince and Smolensky 1993): All else being equal, learned morphemes are represented in the lexicon in whatever way minimizes their constraint violations (necessarily, of Faithfulness).

1.3 Case II: Contextual Markedness

(7) Assume:

*[-SON, -VOICE] / V _____V (cf. Korean, Yidj, various Australian)
 *[-SONORANT, +VOICE]
 IDENT([VOICE] IN OBSTRUENTS)

★ Consider all six rankings and determine the pronunciability/phontactic legality under each of [ba], [pa], [apa], [aba].

1.4 Case III: Contextual Faithfulness

(8) Assume:

*[-SONORANT, +VOICE]
 IDENT([VOICE] IN OBSTRUENTS)
 IDENT([VOICE] IN OBSTRUENTS / _____[+SONORANT])

★ Consider all six rankings and determine the pronunciability under each of [ba], [pa], [ap], and [ab]. What would traditional linguistics say about the “phonemicness” of voicing under each grammar?

1.5 Earlier approaches to Phonotactics

- (9) Chomsky and Halle (1968), Sound Pattern of English:
- a. Place constraints on possible underlying representations.
 - b. Morphology concatenates these
 - c. Rules apply to the result
 - d. Hence, surface phonotactics are theorems.

- (10) At Least Some Phonotactics are Post-UR
- a. Shibatani (1973) Japanese coda conditions
 - b. Chung, Sandra. 1983. Transderivational relationships in Chamorro phonology. *Language* 9:35-66.

- (11) Chamorro vowels:

i	u
e	o
æ	ɑ

- (12) In the initial syllable, after one of a particular set of front-voweled prefixes

'gumaʔ	'house'	i 'gimaʔ	'the house'
'haga	'daughter'	i 'haga	'the daughter'
'soŋsuŋ	'village'	i 'seŋsuŋ	'the village'

- ★ What would be a traditional rule-based analysis of this pattern? Assume that [u] is [+round, +labial, -front, +back] and that [i] is [-round, -labial, +front, -back].

- ★ What happens if you fail to change [round]? [labial]?

- (13) In general:
- a. rules often create not exactly what you would get by changing the crucial feature, but rather the closest legal phoneme.
 - b. Experiments in rule-writing indicate that, if you try to be rigorous, you often need a great number of features on the right side of the arrow.
 - c. Such cases are multipliable—this is one aspect of the conspiracy problem.
- (14) More precisely: a statement of the underlying Chamorro phoneme inventory (true of Chamorro morphemes) is recapitulated in the statement of Fronting.

1.6 OT rationalizes such cases by providing a complete grammar

- (15) The absence of *[y] from Chamorro is something that an adequate phonology should account for in any event.

- ★ Fill in the violations for these tableaux:

/i + u/	AGREE(BACK)	*[-BACK, +ROUND]	ID(BACK)	ID(ROUND)
☞ i + i				
*i + y				
*i + u				

/i + o/	AGREE(BACK)	*[-BACK, +ROUND]	ID(BACK)	ID(ROUND)
☞ i + e				
*i + ø				
*i + o				

/i + /	AGREE(BACK)	*[-BACK, +ROUND]	ID(BACK)	ID(ROUND)
☞ i + æ				
*i + ɑ				

★ Hypothetical candidate: would underlying /y/ survive to the surface?

/y/	AGREE(BACK)	*[-BACK, +ROUND]	ID(BACK)	ID(ROUND)
☞ i				
*u				
*y				

- (16) The ranking $*[-BACK, +ROUND] \gg ID(BACK)$ or $ID(ROUND)$ is needed in any event to get the legal inventory of Chamorro; hence comes “for free” in analysis of the morphophonemics.
- (17) A brave leap: If some phonotactics are post-UR, then make them all surface phonotactics; hence the Richness of the Base Doctrine.

1.7 The OT view put another way

- (18) A Faithfulness constraint says “Surface Form, resemble the Underlying Form in this way, irrespective of what Markedness says.”
- The set of legal structures are those for which the Faithfulness constraints are ranked high enough (relative to Markedness) to enforce the structure in question.

- b. The set of phonemic contrasts is defined primarily by the set of legal structures (but see later, on free variation). In general, the higher ranked the Faithfulness constraints are, the more things are possible (and phonemic). *Nothing in this depends on constraining the set of inputs.*

2 The Doctrine of the Richness of the Base

2.1 Basic Idea

- (19) Prince and Smolensky (1993). There are no constraints on underlying forms, so all patterns of well-formedness are due to constraint ranking. The phonology can thus be construed as a filter, inputting the set of logically sayable things, and outputting the set of things sayable in the target language.
- (20) Richness of the base is not rigidly affiliated with OT
 - a. In principle, a Rich Base could have been imposed on a rule-based theory. The legal forms are those which emerge when the full set of logically possible inputs has been submitted to the rules.
 - b. This was never pursued, to my knowledge, though BH has flirted with this idea for undergraduate teaching.

2.2 Richness of the Base in Ilokano

- (21) Given the grammars you developed, /abut/ (and quite a few other, imaginative things) map onto [ʔabut]. But Ilokano children never hear forms like [abut]; i.e. the grammar is a rich-base approach to this phonotactic restriction.

2.3 Richness of the Base and allophones

- (22) Good Old Allophones
 - a. Beginners to this day are taught to collect sounds that are phonetically similar and in complementary distribution (e.g. for /t/: [t^h], [t], [t̚], [t̚̚]).
 - b. Standard OT provides no basis for this actually happening. Why?
 - (i) Old theory: constraints on underlying forms included phoneme inventory. I.e. English Phonology mapped English Phonemes to English Allophones.
Phonology: EnglishPhonemes* → EnglishAllophones*
 - (ii) But if the constraints all apply at the surface there is no such thing, only a phone inventory.
Phonology: IPA* → IPA*
- (23) Should phonemes be unified in this way?
 - a. For experiments suggesting that native speakers sometimes hear allophones as “the same sound”, see

- Jaeger, Jeri J. 1980. Testing the psychological reality of phonemes. *Language and Speech* 23, 233-253
 - Bruce Derwing, Terrance M. Nearey, and Maureen L. Dow. 1986. On the phoneme as the unit of the ‘second articulation’. *Phonology Yearbook* 3:45-70.
- b. Likewise, experiments teaching people how to read: what letter do they pick for an allophone untaught to them? (Gudschinsky et al. 1970) got a untutored speaker to spell [ɪk] as *kk*, [i]e being the contextual allophone of /k/.

2.4 Richness of the Base and paradigms

- (24) Goal: We don’t just want to say what is pronounceable; we need to say what is a possible pattern of alternation.
- a. This is the distinction we maintain between what constitutes a legal SR, and what constitutes a legal SR to UR mapping.r
 - b. What is striking is that to determine the correct SR to UR mapping, we may need to include reference to certain SR to SR mappings (i.e. among certain members of a paradigm)
- (25) Toba Batak and Paradigmatic Constraints (Hayes 1986)
- a. No Batak root ends in a voiced obstruent.

[marisap]	‘smoke’
[majihut]	‘according to’
[pɛddɛk]	‘short’
 - b. Many Batak roots may appear with a suffix, but *in no case does a putative underlying /b,d,g/ ever manifest itself*: [marisap] ~ *[marisaban], only [marisapan]. I.e., Batak is not like German, Dutch, Russian, Catalan ...
- (26) One Answer That Works Here
- a. Most OTists would rule out [marisap] [marisaban] by imposing constraints on paradigmatic alternation: derived forms must be faithful to the voicing of their (surface) morphological bases.

$$\text{IDENT-OUTPUT/OUTPUT(VOICE)} \gg \text{IDENT(VOICE)}$$
- (27) For output-to-output correspondence, see for instance Benua (2000).
- (28) Terminology:
- a. When people say Output, they are referring to the Surface Form.
 - b. When people say Input, they are referring to Underlying Form.
 - c. Hence OO = “output-to-output”
 - d. and IO = “input-to-output” = standard faithfulness
- (29) Toba Batak with OO-Correspondence
- a. Goal is to allow [pa], [ba], [marisap] ~ [marisap-an], but not *[marisap] ~ [marisab-an]

b. Assume:

- *[-SON,+VOICE] ban on voiced obstruents
- ID-IO(VOICE)/____V special IO-Faithfulness constraint protecting prevocalic voicing
- ID-IO(VOICE) general IO-Faithfulness constraint for voicing
- ID-OO(VOICE) forbids disagreement in voicing with base

/pa/	ID-OO(VOICE)	ID-IO(VOICE)/____V	*[-SON,+VOICE]	ID-IO(VOICE)
☞ pa				
*ba				

/ba/	ID-OO(VOICE)	ID-IO(VOICE)/____V	*[-SON,+VOICE]	ID-IO(VOICE)
☞ ba				
*pa				

/marisab/	ID-OO(VOICE)	ID-IO(VOICE)/____V	*[-SON,+VOICE]	ID-IO(VOICE)
☞ marisap				
*marisab				

/marisab+an/	ID-OO(VOICE)	ID-IO(VOICE)/____V	*[-SON,+VOICE]	ID-IO(VOICE)
with base [marisap]				
☞ marisapan				
*marisaban				

(30) Some Papers on Constraining Paradigms

- McCarthy, John. 1998. Morpheme structure constraints and paradigm occultation. To appear in M. Catherine Gruber, Derrick Higgins, Kenneth Olson, and Tamra Wysocki, eds., CLS 32, Vol. II: The Panels. Chicago: Chicago Linguistic Society.
- McCarthy, John. 2003. Richness of the Base and the Determination of Underlying Representations. Rutgers Optimality Archive 616. <http://roa.rutgers.edu>

(31) What is a morphological base of a form?

- a. See Adam Albright’s work, e.g. (Albright 2001)
- b. and Paul Kiparsky’s work on “Stratal OT”; e.g. <http://www.stanford.edu/~kiparsky/Papers/helsingfors.new.pdf>

2.5 Rich base conundrums in Ilokano phonology

- (32) Scenario I
- Suppose we have an underlying form / eaba / ‘sheep’. Given the grammar prepared for the homework, what could be its surface form(s)?
 - If (as Bruce Hayes believes), such cases are illegal in Ilokano, what should be said about the free variation case where / masahe-an / \rightarrow [masahjan , $\text{masah}\text{e}\text{an}$]?
- (33) Scenario II
- Suppose there were an underlying form / babawi? /.
 - Since Ilokano allows no glottal-stop codas, it will surface as [babawi], provided $*\text{?}]_{\sigma}$ outranks $\text{MAX}(?)$.¹
 - But in suffixed forms, [$\text{ba.ba.wi.}\text{?en}$] will defeat [ba.baw.jen]
 - [$\text{ba.ba.wi.}\text{?en}$] is violation-free, while [ba.baw.jen] violates $\text{MAX}(?)$ and $\text{IDENT}(\text{SYLLABIC})$.

★ What OO constraint could keep this from happening?

2.6 Some Quandaries in Richness of the Base

2.6.1 Korean

(34) Korean is like German, only it’s not just voicing that’s neutralized

(35) Korean Neutralizations in Stem-Final Consonants:

Before a vowel-initial suffix	Before pause or oral consonant
{[p], [p ^h], [p’]}	[p]
{[t ^h], [t’], [tʃ], [tʃ ^h], [tʃ’], [s], [s’]}	[t]
{[k], [k ^h], [k’]}	[k]

(36) Here, there are no stems that end in isolation [t] that take [t]² before a suffix. The productive pattern (applied to loans, etc.) is for such stems to take [s].

(37) General point: sometimes morphemes alternate even when they don’t have to.

- Such cases characteristically make it easier for the language learner to predict novel inflected forms.

(38) References on “unnecessary” alternation

- Blevins, Juliette. 1997. Rules in Optimality Theory: two case studies. In I. Roca (ed.), *Derivations and constraints in phonology*, Oxford University Press, Oxford: 227-60.

¹Likewise, $\text{DEP}(V)$ outranks MAX , so no $*[\text{babawi?i}]$.

²Or, more precisely, t’s allophone [d], which is what you get between vowels.

- Hayes, Bruce. 1999. Phonological restructuring in Yidj and its theoretical consequences. In Ben Hermans and Marc van Oostendorp, eds., *The Derivational Residue in Phonological Optimality Theory*, John Benjamins, Amsterdam, pp. 175-205. <http://www.linguistics.ucla.edu/people/hayes/yidjny.htm>
- Ernestus, Miriam and R. Harald Baayen. 2003. Predicting the unpredictable: Interpreting neutralized segments in Dutch. *Language* 79: 5-38.

2.7 Multiple Phonotactics

(39) We recognize Latinate forms even if we haven't studied

★ Exercise: find the non-Latinate material in these words:

prochinity	[prɒ'tsməri]
twellation	[twə'leɪʃən]
anglimnium	[æŋ'limniəm]

Reference: Ito and Mester (1995) and refs therein

- (40) Morphological categories have their own phonotactics; verbs generally can't
- end in schwa (“I’m bletta-ing all of the contents of my attic.”)
 - end in two consecutive stressless syllables.³
- (41) Irregular pasts prefer particular vowels such as [ʌ], explaining innovations like *sneak* ~ *snuck*; also forms from Albright and Hayes’s (2003) wug-test experiment
- shy’nt-shunt, ry’nt-runt, chind-chund, gry’nt-grunt, scride-scrud*
- (42) If speakers can extract the inventory of underlying forms, it is perhaps puzzling that they would not determine their phonotactics.

References

- Albright, Adam. 2001. The identification of bases in morphological paradigms. Ph.D. thesis, University of California, Los Angeles.
- Albright, Adam and Bruce Hayes. 2003. Rules vs. Analogy in English Past Tenses: A Computational/Experimental Study. *Cognition* 90:119–161.
- Benua, Laura. 2000. *Phonological Relations between Words*. New York and London: Garland.

³One clear exception is *jettison*.

- Gudschinsky, Sarah C., Harold Popovic, and Frances B. Popovich. 1970. Native reaction and phonetic similarity in Maxakalí phonology. *Language* 46:77–88.
- Hayes, Bruce. 1986. Assimilation as spreading in Toba Batak. *Linguistic Inquiry* 17:467–499.
- Ito, Junko and Armin Mester. 1995. Japanese Phonology. In *The Handbook of Phonological Theory*, edited by John Goldsmith. Blackwell Publishers, pages 817–838.
- Prince, Alan and Paul Smolensky. 1993. Optimality Theory: Constraint Interaction in Generative Grammar. Technical Report 2, Rutgers University Center for Cognitive Science.