On the role of phonotactics in phonology

Kyle Gorman
Graduate Center, City University of New York*

1 Introduction

Phonotactic theory is concerned with generalizations about permissible sound sequences in languages. There are descriptive limitations on possible sound sequences in nearly every language. But what—if any—role do these constraints play in the phonological grammar of that language?

1.1 Empirical scope

- To develop this theory one must first decide what facts the theory should account for.
- Russian for example has a process of anticipatory voice assimilation.[1]

(1) Russian voice assimilation (Halle 1959):

a.	['ʒedʒbɨ]	'were one to burn'	['ʒetʃlʲi]	'should one burn?'
b.	['mogbɨ]	'were (he) getting wet'	['moklʲi]	'was (he) getting wet?'

- Voice assimilation alternations in Russian are an **entrenched** grammatical fact.
- Voice assimilation is *surface-true* in the sense that all Russian obstruent clusters agree in voice; i.e., there are no hetero-voiced obstruent clusters.

(2) $*\begin{bmatrix} +\textsc{Cons} \\ -\textsc{Son} \\ \alpha\textsc{Voi} \end{bmatrix} \begin{bmatrix} +\textsc{Cons} \\ -\textsc{Son} \\ -\alpha\textsc{Voi} \end{bmatrix}$

- What is the grammatical status of (2)?
- I submit the null hypothesis is that it has no grammatical status at all.

*This talk contains materials earlier presented in Gorman 2013, 2014a,b, 2020. Thanks to audiences at CLS 47, CLS 48, and SYNC 2021. Recent conversations with Karthik Durvasula, Jeff Heinz, Jimin Kahng, and Charles Reiss have helped to refine my thinking on these matters.

[1]For simplicity I put aside the complex behavior of Russian [v].

> Even if we, as linguists, find some generalizations in our description of the lexicon, there is no reason to posit these generalizations as part of the speaker's knowledge of their language, since they are computationally inert and thus irrelevant to the input-output mappings that the grammar is responsible for. (Hale and Reiss 2008:17f.).

(3) H_0: Phonotactic generalizations have no grammatical status.

- Alternatively, one might suppose that the presence of surface-true (1) "derives" or "enforces" the phonotactic generalization (2).

(4) H_1: Let A, B, C, D be possibly-null string sets. Then, a phonotactic constraint $*CAD$ is active in G just in the case that G contains a surface-true rule $A \rightarrow B \,/\, C \_\_ D$.

- Gorman (2013) gives a lengthy defense of H_1.
- Charles Reiss (p.c.) proposes the following thought experiment.

> Imagine that you were to ask a naïve non-linguist monolingual English speaker to discern whether a short snippet of spoken language was either, say, Māori or Czech. Let us suppose that this speaker does not know a single word of either language. Would this speaker do better than chance at this task?

- This would instead suggest that speakers' phonotactic knowledge can be acquired indirectly and effortlessly, without acquiring rules underlying H_1.
- Similarly, Oh et al. (2020) find that monolingual English-speaking New Zealanders are able to discriminate between "possible" and "impossible" Māori nonce words.
- Halle (1962) and Chomsky and Halle (1965) claim that speakers easily distinguish between well-formed and ill-formed nonce words (see Appendix A).
- For instance, neither *blick* [blɪk] nor *bnick* [bnɪk] is a word of English, yet speakers recognize that only the former word is a "possible" word of English.[2]
- Chomsky & Halle (ibid.) suggest that the rules underlying this knowledge are *morpheme structure constraints* (MSCs)—specifically what Stanley (1967) calls *sequence structure constraints*—supplying redundant, language-specific information to maximally underspecified lexical entries.

(5) An English sequence structure constraint, after Chomsky & Halle:

$$\left[-\textsc{Cont}\right] \rightarrow \left[+\textsc{Liq}\right] \,/\, \# \left[-\textsc{Cont}\right] \_\_$$

[2] Chomsky and Halle note that this cannot be a language-general preference. For example, [bn] onsets are unobjectionable in Moroccan Arabic (e.g., *bniqa* 'closet').

- But only a few years later, Halle rejects the assumption of maximally underspecified lexical entries; he attributes it to an overenthusiastic misapplication of information theory.

> In the 1950's I spent considerable time and energy on attempts to apply concepts of information theory to phonology. In retrospect, these efforts appear to me to have come to naught. (Halle 1975:532)

- Without MSCs, it is difficult to derive the distinction between *blick* and *bnick* under H_1, though one might appeal to the processes of syllabification, etc.
- Yet an even-strong alternative is possible.

(6) H_2: Let g be a phonotactic constraint. Then g is active in phonological grammar G if it is a statistically robust generalization (in a sense to be defined) over lexemes in G.

- Or, as Brown puts it:

> ...the patterns outlined above are statistically significant. Given this, it stands that these sound patterns should be explained by some linguistic mechanism. (Brown 2010:48)

- As Gorman (2013:39f.) notes, studies which assume H_2 have become something of a cottage industry in phonology (e.g., Alderete and Bradshaw 2013, Anttila 2008, Berkley 1994a,b, 2000, Brown 2010, Buckley 1997, Coetzee 2008, Coetzee and Pater 2008, Colavin et al. 2010, Colavin 2013, Davis 1989, Elmedlaoui 1995, Frisch et al. 2004, Graff and Jaeger 2009, Grimes 2010, Hayes and Wilson 2008, Kawahara et al. 2006, Kinney 2005, MacEachern 1999, Martin 2007, 2011, McCarthy 1988, Mester 1988, Pozdniakov and Segerer 2007, Padgett 1991, 1992, Pierrehumbert 1993, 1994, Yip 1989...you get the idea).

Outline

In this talk I will attempt to problematize received wisdom on

- the role of phonotactic generalizations in phonemic analysis and
- the role of phonotactic generalizations in language change

as a way of illustrating the different predictions of H_0, H_1, and H_2.

2 Phonotactics in phonemics

Phonologists often turn to phonotactic generalizations to motivate phonemic analyses. Without such phonotactic motivations, many phonemic puzzles will have to remain unsolved.

2.1 The Latin labiovelars

- In Classical Latin, *qu* is pronounced [k^{w}].[3][4]
- Latin also has [g^{w}], but its spelling *gu* is also used for [gu].

  (7) *an**gu**is* [aŋ.**g**^{w}is] 'snake', *exi**gu**us* [ek.si.**gu**.us] 'strict'

- Curiously, *gu* is read as a labiovelar if and only if it is preceded by *n*, in that position allophonically [ŋ]. It is not clear if this is
  - a phonemic generalization,

    (8) g → g^{w} / [+Nas] __
  - a phonotactic generalization,

    (9) *[−Nas]g^{w}
  - an orthographic principle, or
  - merely an accident of history.
- Focusing on *qu* for the moment, two possible analyses suggest themselves.[5]

  (10) Unisegmental: /k^{w}/

  (11) Bisegmental: /kw/
- Much ink (e.g., Allen 1978:16–20, Cser 2013, Devine and Stephens 1977, Gouskova and Stanton 2021, Sturtevant 1939, Touratier 2005, Watbled 2005, Zirin 1970:29–40) has been spilled to decide between (10–11).
- Cser (2020:§2.2.2) provides a recent summary of the arguments, and concludes that "the question cannot be decided definitively".
- Let us now consider how many of Cser's observations are phonotactic generalizations.[6]

  1. **Frequency:** according to Devine and Stephens (1977), henceforth D&S, the lexical frequency of *qu* greatly exceeds that of *c* [k] and [w]; this is judged evidence for (10).

[3] Cser (2020:23) refers to *testimonia* suggesting the "[w] element in ⟨qu⟩ was less consonant-like than other [w]s", which also suggests labialization is properly treated as a secondary articulation. Since nothing here depends on this phonetic detail I stipulate to Cser's proposal.

[4] Latin transcriptions are given alphabetically except when greater detail is required. Characters *i* and *u* indicate both the high vowels [i, u] and glides [j, w], respectively, and macrons are used to indicate vowel length.

[5] A third possibility would be to analyze *qu* as bisegmental /ku/, but this seems unworkable for various reasons.

[6] I have taken the liberty of lightly renumbering and relabeling Cser's observations.

2. **Phonetic issues:** according to *testimonia* from the ancients, the "[w] element in ⟨qu⟩ was less consonant-like than other [w]s" (p. 23); this is judged irrelevant.
3. **Geminates:** a unisegmental *qu* would be the only stop which does not occur as a geminate; this is judged evidence for (11).

   (12) *[kk^{w}], *[k^{w}k^{w}]

4. **Positional restrictions I:** few logically-possible obstruent-glide sequences occur; this is judged evidence for (10).
5. **Positional restrictions II:** labiovelars do not occur word-finally; this is judged inconclusive.
6. **Verb root restrictions:** verb roots ending in three-consonant sequences are unattested except for sonorant-labiovelar sequences; this is judged evidence for (10).

   (13) to[rk^{w}]ere 'to turn', ti[ŋg^{w}]ere 'to dip'

7. **Voicing contrasts:** voicing contrasts in nasal-consonant-clusters are unattested except in nasal-labiovelar sequences; this is judged evidence for (10).

   (14) li[ŋk^{w}]am 'I will/would leave' (1sg. fut./subj. act.), li[ŋg^{w}]am 'tongue' (acc.sg.)

8. **Alternations:**
   - *qu* alternates with *cū* [kuː] in the perfect participle (ppl.) of two verbs.
   - This pattern superficially resembles that of stems with root-final [w] alternating with [uː] in the ppl., which would seemingly suggest (11).
   - However, in two other verbs, *qu* alternates with plain *c* [k].[7]

   (15) Perfective alternations:

a.	lo**qu**ī	'to speak'	lo**cū**tus	'spoken'
	se**qu**ī	'to follow'	se**cū**tus	'followed'
b.	uo**lu**ere	'to roll'	uo**lū**tus	'rolled'
	ar**gu**ere	'to show'	ar**gū**tus	'shown'
c.	co**qu**ere	'to cook'	co**c**tus	'cooked'
	relin**qu**ere	'to leave'	reli**c**tus	'left'

   This is thus judged inconclusive.

9. ***ad*-assimilation:**
   - The verbal prefix *ad-* variably assimilates to the place and manner of a following stem-initial consonant.
   - Assimilation is rarely found with *qu*-initial stems.

   (16) a**dqu**irere ≫ a**cqu**irere 'to acquire'

[7] Note also the absence of the Indo-European "nasal insert" in the ppl. of *relinquere*, which may suggest suppletion.

- The same is true of [kC]-initial stems.

Thus this is judged evidence for (11).

10. **Diachronic considerations:** *qu* descends from the Proto-Indo-European *k^{w}, which is likely unisegmental; this is judged irrelevant.
11. **Poetic license:**
    - In *diaeresis*, a [w] is read as [u] to make the meter in Latin poetry; according to Cser, diaeresis does not target the offglide of [k^{w}].
    - But *pace* Cser, it does target the offglide of [g^{w}]; *relanguit* 's/he became faint' twice scans as quadrisyllabic [re.laŋ.gu.it] instead of trisyllabic [re.laŋ.g^{w}it].

    (17) *gu*-diaeresis (D&S:32):

    cum bene pertaesum est animōque **relanguit** ardor (Ov., *Am.* 2.27)
    [kum.be.ne|per.taj|suː.sta.ni|moː.k^{w}e.re|laŋ.**gu**.i|tar.dor]

    impositō frātrī moribunda **relanguit** ōrē (Ov., *Met.* 6.291)
    [im.po.si|toː.fraː|triː.mo.ri|bun.da.re|laŋ.**gu**.i|toː.reː]

    Thus this is judged inconclusive.
12. **The question of *gu*:** labiovelar *gu* is restricted to after *n*; this is judged inconclusive.
13. **The question of [sw]:**
    - Word-initial [sw] is found in a few words.

    (18) [sw]āuis 'sweet', [sw]āuium 'kiss'

    - If the labiovelars are unisegmental, [sw] would then be the only instance of a obstruent-[w] onset cluster, unless it too was unisegmental /s^{w}/.

    This is judged evidence for (11).

- Of the thirteen generalizations, nine are essentially phonotactic in nature; only one of the remaining four refers to actual alternations, and it is inconclusive.

2.2 The English velar nasal

- There are no English words with word-initial [ŋ].

(19) *#[ŋ]

- In virtually all nasal-obstruent syllable contact clusters, the nasal is homorganic.

(20) pi[m.p]le, sta[n.z]a, mo[ŋ.k]ey

- Exceptions do occur but are very rare (Gorman 2013:76, Pierrehumbert 1994:175).

  (21) pli[m.s]oll, scri[m.ʃ]aw

- For adjectives ending in *ng*, [ŋ] alternates with [ŋg] in the comparative *-er* or suplerative *-est* (Halle and Mohanan 1985:62f.).

  (22) *ng*-adjectives:

lo[ŋ]	lo[ŋ.g]er	lo[ŋ.g]est
stro[ŋ]	stro[ŋ.g]er	stro[ŋ.g]est
you[ŋ]	you[ŋ.g]er	you[ŋ.g]est

- [ŋ] occurs word-finally and after certain suffixes such as the progressive *-ing* and the agentive *-er* (SPE:85).

  (23) si[ŋ], si[ŋ]ing, si[ŋ]er

- To handle (22–23), let us first posit a Level 2 (Siegel 1974) rule of stem-final *g*-deletion.

  (24) *g*-deletion: $g \rightarrow \emptyset \; / \; \left[+\textsc{Nas}\right] \; \_\_ \;]_2$

  This rule is seemingly necessary under any reasonable phonemicization.

- Suppose that [ŋ] is (merely) an allophone of /n/ after the velar stops /k, g/.

  (25) Nasal place assimilation: $n \rightarrow ŋ \; / \; \_\_ \; \left[+\textsc{Dor}\right]$

  Thus word-final [ŋ] is /ng/, just as etymology and orthography suggest (Sapir 1925:49).

- However, there are a few words in which word-internal [ŋ] is not followed by a velar obstruent (I&S:3).

  (26) dinghy [dɪŋ.i], hangar [hæŋ.ɚ]

- Why do so many linguists—Sapir, Chomsky, Halle, and Mohanan, to name a few—prefer the allophonic analysis, given that it incur the cost of the exceptions in (26)?

- There are no word-initial nasal-stop clusters in English.

  (27) $^*\#\left[+\textsc{Nas}\right] \; \left[-\textsc{Cont}\right]$

- Since English velars are all stops, the conjunction of (25) and (27) makes it impossible to realize word-initial [ŋ] and thus would derive the apparent structural gap (19) as per H_1.

- Is (19) a true structural gap? Iverson and Salmons (2005), henceforth I&S, claim that its structural status can be inferred from naïve English speakers' difficulty pronouncing it.
- But can we truly expect to find that everything which is hard for a speaker of a grammar *G* to pronounce is given a similar structural explanation by the alternations in *G*?
- And what is the ontology of the gross phonotactic constraint (27) in the first place?

3 Phonotactics across time

If phonotactic knowledge is autonomous, one of the things we might expect it to do is inhibit historical change or shape loanword adaptation. However, even robust phonotactic generalizations may fail to assert themselves in the process of historical change.

3.1 *V̄ʃ# in English

- Modern English /ʃ/ is largely a reflex of Old English (OE) /sk/.

  (28) sk > ʃ

- Remaining *sk*-words are mostly borrowings from Dutch (e.g., *skipper*) or Norse (e.g., *sky*).
- OE long vowels—the ancestors of the Modern English tense vowels—are not found before coda clusters.

  (29) *V̄CC#

- The sound change (28) and the phonotactic generalization (29) would seem to conspire against tense vowels followed by word-final /ʃ/.

  (30) *V̄ʃ#

- Indeed, exceptions to (30) are quite rare in Modern English (I&S) and this generalization is statistically robust (Gorman 2014b:85).
- Some exceptions to (30) do exist, but according to I&S, they tend to
  - be markedly foreign, (e.g., *cartouche*),
  - be proper names (e.g., *LaRouche*), or
  - convey an "affective, onomatopoeic quality" (e.g., *sheesh*, *woosh*).
- Yet I&S claim (30) has slowly but systematically eroded since c. 1100 CE.
- But **why** have generations of English speakers ignored (30)?
- And for that matter, **why** does this constraint have no measurable impact on English speakers' wordlikeness judgments (Hayes and White 2013)?

3.2 *VsV in Latin

- Intervocalic *s* merged with *r* in Old Latin.[8]

  (31) s > r / V __ V

- This change was "Neogrammarian" in the sense of Labov (1981), and temporarily eliminated all traces of intervocalic *s* in late Old Latin.

  (32) *VsV

- This change also introduced substantial inflectional allomorphy across the lexicon, for instance in 3rd declension nouns.

  (33) Paradigm of *honōs-honōris* 'honor':

	sg.	pl.
nom.	honō**s**	honō**r**ēs
gen.	honō**r**is	honō**r**um
dat.	honō**r**ī	honō**r**ibus
acc.	honō**r**em	honō**r**ēs
abl.	honō**r**e	honō**r**ibus
voc.	honō**s**	honō**r**ēs

- Thus most analysts (e.g., Albright 2005, Foley 1965, Gruber 2006, Heslin 1987, Kenstowicz 1996, Watkins 1970) posit a synchronic analogue of (31) for early Classical Latin.
- Assuming *r* is predictably [+Voi], this rule can be stated as follows.

  (34) Rhotacism: $[+\textsc{Cor}] \rightarrow [-\textsc{Strid}] \;/\; [+\textsc{Voc}] \_\_ [+\textsc{Voc}]$

- Does (34) "enforce" the phonotactic generalization (32), as H_1 predicts?
- In the 1st c. BCE, degemination of *ss* after diphthongs and long monophthongs reintroduced intervocalic *s* (e.g., *caussa* > *causa* 'cause').

  (35) ss > s / $\mu\mu$ __

- This subsequent change led in part to the morphologization (e.g., Roberts 2012) or restructuring (e.g., Gorman 2014a, 2020) of (34).
- But **why** didn't (32) prevent actuation of (35) in the first place?

[8] The *terminus post quem*—the latest possible date—for the actuation is the 4th c. BCE; see Gorman 2020.

- Loanwords are often made to conform to phonotactic generalizations of the recipient language. For instance, in Desano a process of nasal harmony (Kaye 1971) ensures that every word—whether native and foreign—is either totally oral or totally nasal.

  (36) Desano loanwords (Kaye 1974):

  a. [barateru] 'hammer' (< Port. *martelo*)
  b. [nãnãnã] 'orange' (< Span. *naranja*)

- At roughly the same time time as the actuation of (35), Classical Latin borrowed a number of Greek words with intervocalic *s*.[9]

  (37) ambrosia 'food of the gods', *asōtus 'libertine' (acc.sg. asōtum), basis 'pedestal', basilica 'public hall', casia 'cinnamon' (cf. cassia), cerasus 'cherry', gausapa 'woolen cloth', lasanum 'cooking utensil', nausea 'id.', pausa 'pause', philosophus 'philosopher', poēsis 'poetry', sarīsa 'lance', seselis 'seseli'

- One can also find intervocalic *s* in Germanic and Celtic loanwords (Gorman 2014a:282).

- But **why** then didn't (32) impose itself on (37)?

Abbreviations used

D&S Devine and Stephens 1977

I&S Iverson and Salmons 2005

MSC Morpheme structure constraint

OE Old English

Ov. P. Ouidius Naso: *Am.* (*Amores*), *Met.* (*Metamorphoses*)

Port. Portuguese

Span. Spanish

SPE Chomsky and Halle 1968

[9] Many Greek borrowings into Latin exhibit Greek-like inflectional endings, but with the possible exception of the very early borrowing *tūs-tūris* 'incense' (Thiselton-Dyer 1911), even Greek borrowings that adapt Latin inflectional affixes preserve intervocalic *s*.

References

Albright, Adam. 2005. The morphological basis of paradigm leveling. In *Paradigms in Phonological Theory*, ed. Laura J. Downing, Tracy A. Hall, and Renate Raffelsiefen, 17–43. Oxford University Press.

Alderete, John, and Mark Bradshaw. 2013. Samoan root phonotactics: Digging deeper into the data. *Linguistic Discovery* 11:1–21.

Allen, W. Sidney. 1978. *Vox Latina: A Guide to the Pronunciation of Classical Latin.* Cambridge University Press, 2nd edition.

Anttila, Arto. 2008. Gradient phonotactics and the complexity hypothesis. *Natural Language and Linguistic Theory* 26:695–729.

Berkley, Deborah M. 1994a. The OCP and gradient data. *Studies in the Linguistic Sciences* 24:59–72.

Berkley, Deborah M. 1994b. Variability in Obligatory Contour Principle effects. In *Papers from the 30th meeting of the Chicago Linguistic Society*, 1–12.

Berkley, Deborah M. 2000. Gradient Obligatory Contour Principle effects. Doctoral dissertation, Northwestern University.

Brown, Jason. 2010. *Gitksan Phonotactics.* LINCOM.

Buckley, Eugene. 1997. Tigrinya root consonants and the OCP. *Penn Working Papers in Linguistics* 4.3:19–51.

Chomsky, Noam, and Morris Halle. 1965. Some controversial questions in phonological theory. *Journal of Linguistics* 1:97–138.

Chomsky, Noam, and Morris Halle. 1968. *Sound Pattern of English.* Harper & Row.

Coetzee, Andries W. 2008. Grammaticality and ungrammaticality in phonology. *Language* 84:218–257.

Coetzee, Andries W., and Joe Pater. 2008. Weighted constraints and gradient restrictions on place co-occurrence in Muna and Arabic. *Natural Language and Linguistic Theory* 26:289–337.

Colavin, Rebecca S. 2013. Phonotactic probability in Amharic: a psycholinguistic and computational investigation. Doctoral dissertation, University of California, San Diego.

Colavin, Rebecca S., Roger Levy, and Sharon Rose. 2010. Modeling OCP-Place in Amharic with the Maximum Entropy Learner. In *Proceedings of the Forty-sixth Annual Meeting of the Chicago Linguistic Society: The Parasessions*, 27–42.

Cser, András. 2013. Segmental identity and the issue of complex segments. *Acta Linguistica Hungarica* 60:247–264.

Cser, András. 2020. *The Phonology of Classical Latin.* Wiley-Blackwell.

Davis, Stuart. 1989. Cross-vowel phonotactic constraints. *Computational Linguistics* 15:109–111.

Devine, Andrew M., and Laurence D. Stephens. 1977. *Two Studies in Latin Phonology.* Anma Libri.

Elmedlaoui, Mohamed. 1995. Géométrie des restrictions de cooccurrence de traits en sémitique et de berbère: synchronie et diachronie. *Revue canadienne de linguistique* 40:39–79.

Foley, James. 1965. Prothesis in the Latin verb *sum*. *Language* 41:59–64.

Frisch, Stefan A., Janet Pierrehumbert, and Michael Broe. 2004. Similarity avoidance and the OCP. *Natural Language and Linguistic Theory* 22:179–228.

Gorman, Kyle. 2013. Generative phonotactics. Doctoral dissertation, University of Pennsylvania.

Gorman, Kyle. 2014a. Exceptions to rhotacism. In *Proceedings of the 48th annual meeting of the Chicago Linguistic Society*, 279–293.

Gorman, Kyle. 2014b. A program for phonotactic theory. In *Proceedings of the 47th annual meeting of the Chicago Linguistic Society: the Main Session*, 79–93.

Gorman, Kyle. 2020. Anatomy of an analogy. Talk presented at Stony Brook University.

Gouskova, Maria, and Juliet Stanton. 2021. Learning complex segments. *Language* 97:151–193.

Graff, Peter, and T. Florian Jaeger. 2009. Locality and feature specificity in OCP effects: evidence from Aymara, Dutch, and Javanese. In *Proceedings from the Main Session of the 45th Annual Meeting of the Chicago Linguistic Society*, 127–141.

Grimes, Stephen M. 2010. Quantitative investigations in Hungarian phonotactics and syllable structure. Doctoral dissertation, Indiana University.

Gruber, James F. 2006. The prosodic domain of roots: output oriented faithfulness in Classical Latin. *Georgetown University Working Papers in Theoretical Linguistics* 6:127–165.

Hale, Mark, and Charles Reiss. 2008. *The Phonological Enterprise*. Oxford University Press.

Halle, Morris. 1959. *Sound Pattern of Russian*. Mouton.

Halle, Morris. 1962. Phonology in generative grammar. *Word* 18:54–72.

Halle, Morris. 1975. Confessio grammatici. *Language* 51:525–535.

Halle, Morris, and K. P. Mohanan. 1985. Segmental phonology of Modern English. *Linguistic Inquiry* 16:57–116.

Hayes, Bruce, and James White. 2013. Phonological naturalness and phonotactic learning. *Linguistic Inquiry* 44:45–75.

Hayes, Bruce, and Colin Wilson. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39:379–440.

Heslin, Thomas P., Jr. 1987. Prefixation in Latin. *Lingua* 72:133–154.

Iverson, Gregory K., and Joseph C. Salmons. 2005. Filling the gap: English tense vowel plus final /š/. *Journal of English Linguistics* 33:1–15.

Kawahara, Shigeto, Hajime Ono, and Kiyoshi Sudo. 2006. Consonant co-occurrence restrictions in Yamato Japanese. In *Japanese/Korean Linguistics 14*, ed. Timothy J. Vance and Kimberly Jones, 27–38. CSLI.

Kaye, Jonathan. 1971. Nasal harmony in Desano. *Linguistic Inquiry* 2:37–56.

Kaye, Jonathan. 1974. Morpheme structure constraints live! *Recherces linguistiques à Montréal / Montreal Working Papers in Linguistics* 1:55–62.

Kenstowicz, Michael. 1996. Base identity and uniform exponence: alternatives to cyclicity. In *Current Trends in Phonology: Models and Methods*, ed. Jacques Durand and Bernard Laks, 355–385. European Studies Research Institute.

Kinney, Ashlynn. 2005. Serial organization of speech sounds in creole languages. Doctoral dissertation, University of Texas at Austin.

Labov, William. 1981. Resolving the Neogrammarian controversy. *Language* 57:267–308.

MacEachern, Margaret. 1999. *Laryngeal Co-occurrence Restrictions*. Garland.

Martin, Andrew. 2007. The evolving lexicon. Doctoral dissertation, University of California, Los Angeles.

Martin, Andrew. 2011. Grammars leak: modeling how phonotactic generalizations interact within the grammar. *Language* 87:751–770.

McCarthy, John J. 1988. Feature geometry and dependency: a review. *Phonetica* 45:84–108.

Mester, Armin. 1988. *Studies in Tier Structure*. Garland.

Oh, Y., T. Simon, C. Beckner, J. Hay, J. King, and J. Needle. 2020. Non-Māori-speaking New Zealanders have a Māori proto-lexicon. *Scientific Reports* 10:22318.

Padgett, Jaye. 1991. Stricture in feature geometry. Doctoral dissertation, University of Massachusetts, Amherst. Published by CSLI, 1995.

Padgett, Jaye. 1992. OCP subsidiary features. In *Proceedings of the 22rd meeting of the North East Linguistic Society*, 335–346.

Pierrehumbert, Janet. 1993. Dissimilarity in the Arabic verbal roots. In *Proceedings of the 23rd meeting of the North East Linguistic Society*, 367–381.

Pierrehumbert, Janet. 1994. Syllable structure and word structure: a study of triconsonantal clusters in English. In *Phonological Structure and Phonetic Form: Papers in Laboratory Phonology III*, ed. Patricia A. Keating, 168–188. Cambridge University Press.

Pozdniakov, Konstantin, and Guillaume Segerer. 2007. Similar place avoidance: a statistical universal. *Linguistic Typology* 11:307–348.

Roberts, Philip J. 2012. Latin rhotacism: a case study in the life cycle of phonological processes. *Transactions of the Philological Society* 110:80–93.

Sapir, Edward. 1925. Sound patterns in language. *Language* 1:37–51.

Siegel, Dorothy. 1974. Topics in English morphology. Doctoral dissertation, Massachusetts Institute of Technology.

Stanley, Richard. 1967. Redundancy rules in phonology. *Language* 43:393–436.

Sturtevant, Edgar H. 1939. The pronunciation of Latin *qu* and *gu*. *Language* 15:221–223.

Thiselton-Dyer, William T. 1911. Origin of incense. *Nature* 85:507–507.

Touratier, Christian. 2005. Système des consonnes. In *Essais de phonologie latine*, ed. Christian Touratier, 61–134. Publications de l'Université de Provence.

Watbled, Jean-Philippe. 2005. Théories phonologiques et questions de phonologie latine. In *Essais de phonologie latine*, ed. Christian Touratier, 25–57. Publications de l'Université de Provence.

Watkins, Calvert. 1970. A case of non-chronological rule insertion. *Linguistic Inquiry* 1:525–527.

Yip, Moira. 1989. Feature geometry and cooccurrence restriction. *Phonology* 6:349–375.

Zirin, Ronald A. 1970. *The Phonological Basis of Latin Prosody*. Mouton.

A Chomsky & Halle on phonotactic knowledge

Among the redundancies that must be eliminated are those where the appearance of a given feature in a segment is contextually determined. Thus, for instance, /tsaym/, /gnayt/ and /vnig/ are not English words, since English words to [sic] not begin with the sequences /ts/, or /gn/, or /vn/. ...English speakers will regard /vnig/, /tsaym/, and /gnayt/ as not only meaningless, but also totally un-English; impossible by the rules of their language. ...English speakers will accept the equally meaningless /blik/, /θōd/, and /nis/ as possible English words, perhaps as words found in an unabridged dictionary rather than in the vocabulary of the average speaker. (Halle 1962:384f.)

...in English there is a form *brick* (/brik/), but no /blik/ or /bnik/. Nevertheless, a speaker of English knows that /blik is an admissible form in a sense in which /bnik/ is not. This distinction is, furthermore, not a matter of universal phonetics. ... A description of English will achieve the level of observational adequacy, in this case, if it distinguishes /brik/, as an occurring form, from /blik/ and /bnik/, as non-occurring forms. Thus a lexicon—a list of all occurring forms—meets the level of observational adequacy. ...The description will meet the level of descriptive adequacy if it distinguishes /brik/ and /blik/, as admissible forms, from /bnik/, as an inadmissible form. In this case, it will state what the speaker knows...to be true. To meet the level of explanatory adequacy, a linguistic theory must justify the descriptively adequate grammar on internal grounds. That is, it must show on what basis the [language acquisition–KBG] device ...selects a grammar admitting /blik/ and excluding /bnik/. (Chomsky and Halle 1965:101)