GENERATIVE PHONOTACTICS

Kyle Gorman

A DISSERTATION

in

Linguistics

Presented to the Faculties of the University of Pennsylvania in Partial

Fulfillment of the Requirements for the Degree of Doctor of Philosophy

2013

Supervisor of Dissertation

____

Charles Yang

Associate Professor, Linguistics &

Computer and Information Science

Graduate Group Chairperson

____

Robin L. Clark

Associate Professor, Linguistics

Dissertation Committee

Stephen R. Anderson

Dorothy R. Diebold Professor,

Linguistics, Psychology, & Cognitive

Science, Yale University

Mark Yoffe Liberman

Trustee Professor of Phonetics,

Linguistics & Computer and

Information Science

Eugene Buckley

Associate Professor, Linguistics

Rolf Noyer

Associate Professor, Linguistics

To my parents, Mark and Bev, who never once said **bnick* or **forwent*

Acknowledgements

While many of my fellow students at Penn have had a hand in this document, none so much as my cohort: Toni Cook and Laurel MaKenzie. Without Joe Fruehwald, Marielle Lerner, Constantine Lignos, Brittany McLaughlin, and Hilary Prichard, it wouldn't have been nearly as much fun. Carly Fisher and Katherine McEldoon have provided me with unflagging support from afar. John Murdock and Hilary Prichard graciously proofread the entire document numerous times.

The faculty, staff, and students of the Institute For Research in Cognitive Science provided me with not merely an office, but something close to a home. Thanks to my officemates, Jessica Marcus, Chandan Narayan, John Trueswell, and especially Laurel Sweeney.

The ideas I first encountered as an undergraduate at the University of Illinois under the tutelage of Jennifer Cole and Richard Sproat continue to guide my research. I am grateful to the members of the committee, Gene Buckley, Mark Liberman, and Rolf Noyer, for creating a nurturing environment where quantitative and theoretical approaches can meet. Insofar as I have acknowledged and understood the arguments of my predecessors, it is thanks to Steve Anderson, who graciously agreed to serve on my committee and showed me the wonders of New Haven. I do not expect to ever again have such a fearless advocate in the incomporable Charles Yang; the scientist he is is the one I strive to be. While they were not involved in the preparation of this dissertation, it is hard to imagine that it would exist without Lila Gleitman, Morris Halle, and Bill Labov.

ABSTRACT

GENERATIVE PHONOTACTICS

Kyle Gorman

Charles Yang

Associate Professor, Linguistics and Computer and Information Science

This dissertation outlines a program for the theory of phonotactics—the theory of speakers' knowledge of possible and impossible (or likely and unlikely) words—and argues that the alternative view of phonotactics as stochastic, and of phonotactic learning as probabilistic inference, is not capable of accounting for the facts of this domain. Chapter 1 outlines the proposal, precursors, and predictions.

Chapter 2 considers evidence from wordlikeness rating tasks. It is argued that intermediate well-formedness ratings are obtained whether or not the categories in question are graded. A primitive categorical model of wordlikeness using prosodic representations is outlined, and shown to predict English speakers' wordlikeness judgements as accurately as state-of-the-art gradient wellformedness models. Once categorical effects are controlled for, these gradient models are largely uncorrelated with wellformedness.

Chapter 3 considers the relationship between lexical generalizations, phonological alternations, and speakers' nonce word judgements with a focus on Turkish vowel patterns. It is shown that even exception-filled phonological generalizations have a robust effect on wellformedness judgements, but that statistically reliable phonotactic generalizations may go unlearned when they are not derived from phonological alternations.

Chapter 4 investigates the role of phonological alternations in constraining lexical entries, focusing specifically on medial consonant clusters in English. Static phonotactic constraints previously proposed to describe gaps in the inventory of medial clusters are shown to be statistically unsound, whereas phonological alternations impose robust

restrictions on the cluster inventory. The remaining gaps in the cluster inventory are attributed to the sparse nature of the lexicon, not static phonotactic restrictions.

Chapter 5 summarizes the findings, considers their relation to order of acquisition, and proposes directions for future research.

Contents

List of Tables

List of Figures

Chapter 1

A program for phonotactic theory

This dissertation has two distinct but complementary aims. The first is to outline the empirical scope of *phonotactic theory*, the theory of speakers' knowledge of possible and impossible sounds, sound sequences, and words. The second is to suggest that the core core facts of this domain are compatible with a traditional view of phonotactic knowledge as independent of the lexicon, categorical, and closely related to phonological processes, but inconsistent with an increasingly popular view of phonotactics and phonotactic learning as a type of probabilistic inference over the lexicon, and therefore gradient and independent of phonological processes.[1]

Despite a recent surge of interest in phonotactic theory, the empirical scope of the theory remains poorly defined. The first task in developing a theory of phonotactic knowledge, then, is to outline the types of facts that the theory should account for.

[1]Throughout, the term *lexicon* is used in a specific sense of the set underlying representations in some language; this is not meant to imply a position on the possibility that larger, composite linguistic representations are also stored in lexical memory. For a review of recent experimental evidence on this question, see Lignos and Gorman in press.

1.1 The empirical scope of phonotactic theory

Phonotactic knowledge is evidenced in quite different ways than phonological knowledge. In Russian, for instance, a process of anticipatory assimilation ensures that derived clusters of obstruents agree in voice.[2]

(1) Russian voice assimilation alternations (adapted from Halle 1959:22f.):

a.	[ˈʒedʒbɨ]	'were one to burn'	[ˈʒetʃlʲi]	'should one burn?'
b.	[ˈmogbɨ]	'were (he) getting wet'	[ˈmoklʲi]	'was (he) getting wet?'

There is only one concise explanation for the [dʒ]-[tʃ] and [k]-[g] alternations, namely that Russian speakers have internalized a process of voice assimilation. It is also the case that voicing in underlying representations in Russian is "nondistinctive in all but the last member of a cluster of obstruents" (Anderson 1974:283). It is possible to deny, however, that this restriction on underlying hetero-voiced obstruent clusters (and many other constraints on underlying representations that have been proposed) is part of the grammatical knowledge of Russian speakers, as some linguists have done.

> MSCs are merely artifacts of the grammar, and thus play no part in the phonological component of a language. (Clayton 1976:302)

> 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.)

Hale and Reiss go on to label constraints on underlying representations as "extensional",

[2] For sake of discussion, the complex behavior of [v] has been ignored here.

and thus irrelevant to generative grammar, in contrast to "intensional" statements like phonological rules. Regarding the facts about Russian obstruents discussed so far, it is clear that the alternation facts are in some sense more privileged. Were it the case that Russian speakers had not internalized a process of obstruent voicing assimilation, the only alternative explanation for the forms in (1) is a massive system of phonologically conditioned suppletive allomorphy. In contrast, were there no constraint on underlying obstruent voicing, however, it is not immediately obvious that anything would be different. Presumably, this is what Hale and Reiss mean when they refer to constraints on underlying representations as "computationally inert". There are also many cases where restrictions on underlying representations are specific to underlying representations and do not apply to surface representations. For instance, hiatus is exceptionally rare in native Turkish roots; of the handful of examples in the Turkish Electronic Living Lexicon (Inkelas et al. 2000), many appear to be compounds (e.g., *ısıalan* 'endothermic', cf. *ısı* 'heat', *alan* 'taker'). Yet numerous phonological processes give rise to hiatus in derived environments (e.g., Kabak 2007). Silverman (2000) observes that operations like reduplication and truncation also tend to introduce violations of restrictions that hold of underlying representations. However, there are many facts which strongly suggest that speakers internalize phonotactic constraints.

1.1.1 Wordlikeness judgements

As Halle (1962) and Chomsky and Halle (1965) note, speakers can distinguish between a well-formed and an ill-formed word, neither of which is an actual word. Neither [blɪk] *blick*[3] nor [bnɪk] *bnick* is a word of English, yet speakers immediately report that only the

[3] In fact, philospher R.M. Hare (1955) uses the term *blik* to describe *a priori*, unfalsifiable frames of reference through which experiences are filtered; for instance, materialism can be thought of as the *blik* interpreting all experience in terms of the properties of matter. Since this term has some currency in the philosophy of science and religion, this is likely a "real word" rather than a lexical gap for a significant English-speaking minority. Thanks to Steve Anderson for brining this to my attention.

former is “possible”, an *accidental* gap in the lexicon, whereas the latter is judged to be impossible (i.e., structurally excluded). There can be no question that it is part of speakers’ knowledge. Elicited in a controlled fashion, these *wordlikeness judgements* are perhaps the most important (and least controversial) source of phonotactic evidence, and they play a major role in Chapters 2–3 of this dissertation.

1.1.2 Word production and recognition

Speakers have difficulty producing (e.g., Davidson 2006a, 2010, Rose and King 2007, Vitevitch et al. 1997) and perceiving (Dupoux et al. 1999, Kabak and Idsardi 2007, Massaro and Cohen 1983) certain types of nonce words judged to be phonotactically illicit. For instance, Gallagher (in press) finds that speakers of Quechua have difficulty repeating nonce words with multiple ejectives (e.g., [k’ap’i]), which do not occur in the language; in other words, multiple ejective sequences are not merely absent, but also difficult for these speakers to perceive or produce.

In English, sequences of adjacent obstruents which do not also agree in voice (e.g., *a*[b.s]*inth*) are quite rare within a word, and therefore a hetero-voiced obstruent cluster is a clue to the presence of a word boundary in running speech. Infants (e.g., Mattys and Jusczyk 2001) and adults (McQueen 1998, Norris et al. 1997) are thought to use this heuristic for word recognition in experimental settings.[4]

Berent et al. (2001) and Coetzee (2008) claim that non-word recognition latencies in lexical decision tasks reflect speakers’ phonotactic knowledge, the hypothesis being that a phonotactically illicit nonce word will be rejected more quickly than a well-formed nonce word. However, phonotactic constraints are often confounded with independent predic-

[4] It is important to distinguish the computations involved in using this word segmentation heuristic from those implicated by other phonotactic behaviors. Whatever the locus of [bnɪk]’s ill-formedness, for example, no segmentation into multiple words or morphs renders it a well-formed sequence. See §3.1.1 for further discussion.

tors of lexical decision latencies. For instance, Coetzee (2008) finds that English speakers recognize [sp…p] and [sk…k] nonce monosyllables faster than [st…t] nonce monosyllables in an auditory lexical decision task. Coetzee attributes this to ad hoc phonotactic constraints against the former sequences, but another explanation is available. Even at an early stage of recognition, [stVt] is distinguished from [spVp, skVk] by its higher *cohort density*: there are far more English words starting with initial [st] than with [sp] or [sk]. High cohort density is known to inhibit auditory processing of non-words (e.g., Marslen-Wilson and Welsh 1978) and this alone could account for the processing difference.

1.1.3 Loanword adaptation

Loanword adaptation may provide further evidence for the grammatical relevance of phonotactic knowledge (e.g., Fischer-Jørgensen 1952). In Desano (Kaye 1971, 1974), for instance, all underlying representations (URs) are either totally oral (e.g., [jaha] 'to hear') or totally nasal (e.g., [ñãhã] 'to enter'), and loanwords are made to conform to this generalization: Portuguese *martelo* 'hammer' is adapted as [barateru] and Spanish *naranja* 'orange' as [nãnãnã].[5] It is not difficult to imagine that some component of the synchronic grammar is responsible for the fact that the restriction over native vocabulary is extended to loanwords. Furthermore, many Desano affixes have have distinct oral and nasal allomorphs. For instance, a noun classifier suffix for round objects takes the form [ru] after oral roots (e.g., [goru] 'ball') and [nũ] after nasal roots (e.g., [sẽnãnũ] 'pineapple'). As Kaye (1971:38) notes, this distribution implies the existence of a process of nasal harmony, and it is not implausible that the same process is responsible for the abovementioned adaptations.

However, it is not possible to equate all patterns of adaptation with phonological alternations targeting native vocabulary. Peperkamp (2005) highlights several cases where

[5] Kaye's transcriptions have been converted to IPA notation.

native alternations are distinct from loanword adaptations. For instance, in Korean, [s] does not surface in codas. As Kenstowicz and Sohn (2001) report, native /s/ is realized as [t] in codas (e.g., [nat]-[nasɨl] 'sickle NOM.-ACC.'). In loanwords, however, final [s] becomes an onset by epenthesis (e.g., *boss* > [posɨ]). Evidence of this sort has lead many (e.g., Dupoux et al. 1999, Peperkamp and Dupoux 2003, Peperkamp 2005) to suggest that loanword adaptations are the result of speech perception, not phonological computations.

There are other cases in which loanword adaptations are not easily identified with any synchronic process. Consider the case of loanwords which begin in onset clusters and which are borrowed into languages which do not permit complex onsets. In the Wikchamni dialect of Yokuts, for instance, Spanish loanwords with complex onsets are adapted either via anaptyxis or deletion of the first consonant.

(2) Wikchamni Yokuts adaptations (Gamble 1989):

a.	*cruz*	>	k^{h}uluʃ	'cross'
	frijoles	>	pilho:liʃ	'beans'
b.	*plato*	>	la:to	'plate'
	clavo	>	la:wu	'nail'

However, there is no reason to regard these adaptations as a product of the synchronic phonology: no Yokuts root begins with a consonant, and there is no way to derive an initial CC cluster. Furthermore, as Gamble (1989) notes, there is no reason to believe that the adaptations actually occurred within Yokuts at all, since there was at best quite limited contact between Spanish and Yokuts speakers; the adaptations may have occurred to conform to the phonotactic restrictions of another language, also in contact with Yokuts.

Finally, there are many cases where putative phonotactic restrictions are not extended to loanwords (e.g., Clements and Sezer 1982, Davidson and Noyer 1997:75, Fries and Pike

1949, Holden 1976, Itô and Mester 1995a,b, Shibatani 1973:95, Ussishkin and Wedel 2003, Vogt 1954; additional examples can be found throughout this dissertation). For instance, native words in San Mateo Huave all end in a consonant, but Davidson and Noyer (1997) note that final unstressed syllables in Spanish loanwords are never repaired by epenthesis (e.g., *verde* 'green' > [beɾde], *[beɾdej]). Given the considerable disagreement about the nature of loanword adaptation at the present juncture, it may be premature to regard this inertness as strong evidence against the constraints in question, though it may be ultimately be a useful diagnostic.

1.1.4 Alternate phonologies

Language games or speech disguises may also provide evidence for phonotactic knowledge (e.g., Vaux 2011), assuming that these "alternative phonologies" are not qualitatively different from naturally-occurring language processes (e.g., Bagemihl 1995:697). An example of language game evidence bearing on phonotactic questions can be found in §3.1.1, where a language game is used to argue that root-internal vowel sequences in Turkish are subject to vowel harmony.

1.1.5 Lexical statistics

Finally, phonotactic gaps or tendencies in the lexicon are often taken as evidence for phonotactic knowledge, under the hypothesis that grammatical constraints are the cause of these lexical generalizations. Chapters 3–4 consider in detail the evidentiary status of these lexical tendencies.

1.2 The grammatical basis of phonotactic knowledge

With the primary evidence for phonotactic theory now established, it is possible to consider the grammatical architecture that underlies this knowledge.

1.2.1 The insufficiency of morpheme structure constraints

Early generative phonologists posited that phonotactic ill-formedness derives solely from *morpheme structure constraints*, restrictions on underlying representations (Chomsky and Halle 1965, 1968, Halle 1959, 1962). Stanley (1967) distinguishes between two types. *Segment structure constraints* impose restrictions on the underlying segment inventory. For example, voicing is non-contrastive for /ts, tʃ, x/ in Russian (Halle 1959:22): [dz, dʒ, ɣ] appear in surface, but not underlying, representations. This dissertation will have little to say about the nature of segment structure constraints. Of more interest here are *sequence structure constraints*, restrictions on underlying phoneme sequences. An example is given below.

(3) An English MSC (adapted from Chomsky and Halle 1965:100):

$$\begin{bmatrix}+\textsc{Cons}\end{bmatrix} \longrightarrow \begin{bmatrix}+\textsc{Liquid}\end{bmatrix} \;/\; \# \begin{bmatrix}-\textsc{Cont}\end{bmatrix} \;\text{—}$$

This sequence structure constraint specifies that a consonant immediately after a word-initial stop is a liquid; this would preclude underlying */bnɪk/, for example (though it would erroneously rule out attested *[sn], *[sm] onsets).

However, Shibatani (1973) argues that not all wordlikeness contrasts can be expressed as constraints on URs. In German, for instance, obstruent voicing is contrastive, but neutralizes finally: e.g., [gʀaːt]-[gʀaːtə] 'ridge(s)' vs. [gʀaːt]-[gʀaːdə] 'degree(s)'. By hypothesis, the latter root ends in /d/, so the constraint against final voiced obstruents is specific to surface representations. Shibatani claims, however, that German speakers judge voiced obstruent-final nonce words to be ill-formed.[6]

1.2.2 The duplication problem

Whereas Shibatani argues that morpheme structure constraints are insufficient to account for speakers' phonotactic knowledge, other authors observe the tendency for structural descriptions of phonological processes to reappear among the morpheme structure constraints on the same language (e.g., Chomsky and Halle 1968:382, Hale 1965:297, Kisseberth 1970, Postal 1968:212f., Stanley 1967:401). Russian obstruent voice assimilation, discussed above, provides an example of this type: there are no hetero-voiced obstruent clusters in either underlying or surface representations. These two facts are tantalizingly similar, but are treated as separate if a distinction between morpheme structure constraints and phonological processes is drawn. This is a special case of what Kisseberth (1970) calls *conspiracies* and what Kenstowicz and Kisseberth (1977:136) term the *duplication problem*. Dell (1973:205f.) and Stampe (1973:28f.) argue that the distinction between constraints on URs and alternations is artificial, and that these different levels of description are related by a principle now known as *Stampean occultation* (Prince and Smolensky 1993:54). In a language like Russian, in which surface obstruent clusters exceptionlessly agree in voicing, there is simply no reason for the language acquisition device to posit underlying hetero-voiced obstruent clusters: obstruent voice assimilation "occults" underlying */kb/, for instance. Were such an underlying form posited, it would surface as [gb] in all contexts.[7] In Chapter 4, it is argued that constraints described in terms of non-contrastive

[6]Furthermore, voicing of final obstruents is usually lost in German loanword adaptation: e.g., English *hot dog* becomes [hɑt dɔk] (Ussishkin and Wedel 2003:506).

[7]*Lexicon Optimization* (Prince and Smolensky 1993:209) implements a form of Stampean occultation notable in that it projects all non-alternating surface segments directly into URs. For instance, in English, Lexicon Optimization demands underlying /ŋ/ in words like *bank*, though [ŋ] could be otherwise be analyzed as an allophone of /n/ before velar consonants (e.g., Borowsky 1986:65f., Chomsky and Halle 1968:85, Halle and Mohanan 1985:62), eliminating /ŋ/ from the phoneme inventory. However, this is not core to Dell and

prosodic structures can also be derived by Stampean occultation. In an architecture like Lexical Phonology, it is even possible to apply a process to individual underlying representations (i.e., at the "lexical level") to enforce constraints which are not surface-true. Consequently, it is impossible to construct an argument which would distinguish morpheme structure constraints from other types of phonological computations.

This is in ways similar to Halle's famous argument (1959) against the morphophonemic/phonemic distinction.[8] The principle of biuniqueness in vogue at that time separates neutralizing (morphophonemic) and non-neutralizing (phonemic) processes. In Russian, obstruents participate in voice assimilation whether this neutralizes a phonemic distinction (1a) or not (1b): recall that there is no underlying /dʒ/ in Russian. Under biuniqueness, there must be separate, non-overlapping variants of this process, one applying in neutralizing contexts and another in non-neutralizing contexts. From this, Halle argues that biuniqueness (and the distinction between the morphophonemic and phonemic levels that follows from it) entails "a significant increase in the complexity of the representation…an unwarranted complication which has no place in a scientific description of language" (24). While Anderson (1985:110f., 2000) correctly observes that it is in principle possible to sketch an analysis of Russian obstruent voicing which preserves biuniqueness without the morphophonemic/phonemic duplication Halle objects to (see also Kiparsky 1985), this requires further contested assumptions—contrastive underspecification (against which, see Steriade 1995) and a Duke-of-York derivation. This does not seem inconsistent with

Stampe's insight about the relationship between surface and underlying sequence structure restrictions. For instance, the hypothetical {bæŋk} posited by Lexicon Optimization could be revised to /bænk/, and take a *free ride* (in the sense of Zwicky 1970) on the process of nasal place assimilation found elsewhere in English (see §4.2.2). Indeed, this seems desirable, since Lexicon Optimization forces a duplication between underlying and surface constraints. For instance, [ŋ] does not appear word-initially and English speakers have considerable difficulty producing it in this position (Rusaw and Cole 2009). The allophonic analysis of [ŋ] predicts this fact, since there is no way to derive the [ŋ] allophone in onset position. Future work will consider the how purely allophonic relationships might be acquired.

[8] This is not to imply that Halle was the first to make this argument, or to otherwise disregard this distinction: similar ideas can be found in work by Bloch (1941), Bloomfield (1926:160, 1962:5f.), Chao (1934), Hamp (1953:244f.), and Sapir (1930:47f.), among others (see Anderson 1985 *passim*).

Halle's claim that biuniqueness imposes unnecessary additional complexities: under Anderson's analysis the complexity is not two variants of a single process, but rather a dependency on contentious theoretical assumptions.

1.2.3 Static constraints

Stampean occultation provides a mechanism for many sequence structure constraints constraints on underlying representations to be expressed as phonological processes. However, not all restrictions on underlying representation have an obvious reflex in the system of alternations. Of the residue that remains once these *derived constraints* are identified, many can be attributed to the language-specific prosodic inventory. For instance, a language which does not permit onset clusters may do so without there being evidence for processes that simplify initial consonant clusters; Yokuts, discussed above, is a clear example. There is a sense in which prosodic parsing operations like syllabification or footing can be thought of as phonological computations, and the boundary between prosodic restrictions and alternations is not always clear. For instance, Latin does not permit a geminate consonant after a long vowel; since geminates always span syllable boundaries, this appears at first glance to be a restriction on the Latin syllable template. However, Latin seems to enforce this restriction via a process of degemination (Gorman in press); assibilation of *t, d* in perfect passive participles produces a geminate *ss* (e.g., *fossus* 'dug', cf. *fodere* 'to dig') except after diphthongs and long monophthongs, where a singleton *s* appears (e.g., *lūsus* 'played', cf. *lūdere* 'to play'). This restriction is simultaneously a component of the prosodic inventory and of the alternation system.

A question which is central to phonotactic theory is whether there are additional types of sequence structure constraints beyond those which are derived from alternations or restrictions on the prosodic inventory. Recent answers to the affirmative have generally been presented with proposals for how such constraints are learned.

One possibility will be carefully considered in this dissertation. According to this view, much of phonotactic learning is accomplished by statistical inference performed over the lexicon, and phonotactic knowledge is the sum of these lexical statistical patterns. Two predictions immediately follow from this position. First, since statistical generalizations may be more or less true, phonotactic knowledge may be gradient, reflecting the strength of the many competing patterns involved. Secondly, lexical statistical patterns need not be reflected in the system of alternations or in the prosodic system at all: they may be "static". These predictions will be taken up in some detail in Chapters 2 and 3, respectively.

Another possibility is posed by traditional thinking in Optimality Theory. According to a standard proposal, at the "initial state" all markedness constraints are ranked above all faithfulness constraints (e.g., Smolensky 1996). If learning proceeds via constraint demotion, markedness constraints which are not implicated by alternations will remain undominated (e.g., Coetzee 2008). It is difficult to make concrete predictions from this proposal in the absence of a complete proposal for the contents of the universal constraint set Con, but it has the potential to blur the distinction between accidental and structural phonotactic gaps, a distinction which is at the heart of phonotactic theory and is the subject of Chapter 4. If the constraint set is in fact universal, any exceptionless gap which corresponds to a markedness constraint *in any langauge* will be accorded the status of an inviolable phonotactic restriction. If Con is sufficiently rich to incorporate constraints like the *[spVp] proposed by Coetzee (2008), it seems quite likely that it will also contain markedness constraints rulling out English syllable contact clusters which are regarded as accidental gaps in Chapter 4. For instance, a constraint *[s.w] seems plausible as a subcomponent of the so-called *syllable contact law* (e.g., Gouskova 2004, Murray and Vennemann 1983) which disfavors syllable contact clusters with increasing sonority. This *[s.w] constraint is without exception in English, yet nonce words like *teeswa* [tes.wa] seem quite unobjectionable to native speakers. Numerous other examples of this type could be adduced, and pose a serious problem for the markedness-over-faithfulness model of phonotactic learning.

The merits of these two models are the subject of the rest of this dissertation. Whatever these merits, there is some value also in the null hypothesis, that there are no static phonotactic constraints at all, if it can be maintained. Specifically, the principle of *no static phonotactics* has interesting ramifications for evaluating certain competing phonological analyses. Consider Sanskrit aspiration alternations such as *bodhati-bhotsyati* 'he wakes-he will wake'. According to one analysis, which has precedents as far back as the grammar of Pāṇini, the root /budh/ undergoes a process shifting aspiration leftward in certain contexts (e.g., Borowsky and Mester 1983, Hoenigswald 1965, Kaye and Lowenstamm 1985, Sag 1974, 1976, Schindler 1976, Stemberger 1980, Whitney 1889:§141f.). An alternative analysis posits an underlying /bhudh/ and a process of aspirate dissimilation, a synchronic analogue of Grassman's Law (e.g., Anderson 1970, Hoard 1975, Kiparsky 1965:§3.2, Phelps and Brame 1973, Phelps 1975, Zwicky 1965:109f.). Under the latter analysis, multiple surface aspirates (e.g., hypothetical **bhodhati*) are phonotactically marked; the former account makes no such prediction. The principle of no static phonotactic constraints, if it can be maintained, could in theory adjudicate between these competing analyses: if the postulated constraint on multiple surface aspirates is active in wordlikeness judgements or loanword adaptation, for instance, this rules out the former account.

1.3 Outline of the dissertation

The remainder of the dissertation consists of three case studies which provide support for the novel and contentious predictions of the minimal phonotactic model sketched above.

Chapter 2 considers evidence from wordlikeness rating tasks. It is argued that intermediate well-formedness ratings are obtained whether or not the categories in question are graded. A primitive categorical model of wordlikeness using prosodic representations is outlined, and shown to predict English speakers' wordlikeness judgements as accurately as state-of-the-art gradient wellformedness models. Once categorical effects are controlled for, gradient models are largely uncorrelated with well-formedness ratings.

Chapter 3 considers the relationship between lexical generalizations, phonological alternations, and speakers' nonce word judgements with a focus on Turkish vowel patterns. It is shown that even exception-filled phonological generalizations have a robust effect on wellformedness judgements, but that statistically reliable phonotactic generalizations may go unlearned when they are not derived from phonological alternations.

Chapter 4 investigates the role of phonological alternations in constraining lexical entries, focusing specifically on medial consonant clusters in English. Static phonotactic constraints previously proposed to describe gaps in the inventory of medial clusters are shown to be statistically unsound, whereas phonological alternations impose robust restrictions on the cluster inventory. The remaining gaps in the cluster inventory are attributed to the sparse nature of the lexicon, not static phonotactic restrictions.

Chapter 5 summarizes the findings, considers their relation to order of acquisition, and proposes directions for future research.

Chapter 2

Categorical and gradient aspects of wordlikeness

Much of the recent work in phonotactic theory has attempted to incorporate the intuition that phonotactic wellformedness is not an "all-or-nothing" matter. Rather, it is alleged, wellformedness judgements have more fidelity than is implied by a simple contrast between "possible" and "impossible", and therefore must be measured and studied at a greater degree of granularity.

This is hardly a novel claim, though it has taken on greater import with the emergence of computational models of wordlikeness. Early generative discussions of wordlikeness (e.g., Chomsky and Halle 1965, Halle 1962) are best remembered for the famous examples [blɪk] and [bnɪk], the former representing a "possible word" of English and the latter representing an "impossible word". A naïve account of this contrast would be to derive it from the assumption that segments must be parsed into syllables or subject to further phonological repair (e.g., Hooper 1973:10f., Kahn 1976:57f., Itô 1989, Wolf and McCarthy 2009:19f.). Unlike some languages (e.g., Morroccan Arabic: *bniqa* 'closet'), English does not permit stop-nasal onsets like [bn], so the latter nonce word cannot surface as such. In other words, [bnɪk] is an impossible surface representation in English. However, in *The Sound Pattern of English* (henceforth, *SPE*), Chomsky and Halle (1968) introduce a third nonce word, [bznk], constructed so as to be even less English-like than [bnɪk].[1] This leads Chomsky and Halle to conclude that wordlikeness intuitions are gradient.

> Hence, a real solution to the problem of "admissibility" will not simply define a tripartite categorization of occurring, accidental gap, and inadmissible, but will define the 'degree of admissibility' of each potential lexical matrix in such a way as to...make numerous other distinctions of this sort (*SPE*:416–417)

This brings the theory of wordlikeness in line with the view of syntactic grammaticality presented by foundational documents like *The Logical Structure of Linguistic Theory* (Chomsky 1955) and *Aspects of the Theory of Syntax* (Chomsky 1965), and reflexes can be found in later work, such as the proposals of Chomsky (1986) and Huang (1982); see Schütze 1996:43f. for a critique.

Chomsky and Halle's claim about the gradient nature of phonotactic wellformedness does not seem to have had much of an impact on practices of the time—as can be seen from discussion in the previous chapter, contemporary critiques focused on other elements of the *SPE* phonotactic theory—but reflexes can once again be found in later work: for instance, Borowsky (1989), Clements and Keyser (1983:50f.), and Myers (1987) all assume a contrast between "peripheral" and absolutely ungrammatical sound sequences.

Recent discussions of gradient grammaticality in wordlikeness attempt to present experimental support for Chomsky and Halle's intuitions:

> When native speakers are asked to judge made-up (nonce) words, their intuitions are rarely all-or-nothing. In the usual case, novel items fall along a

[1] That wordlikeness judgements depend on language-specific knowledge is apparent given that [bznk] is not impossible in all languages: Imdlawn Tashlhiyt Berber permits whole words consisting of a stop-fricative-nasal-stop sequence (e.g., [tzmt] 'it is stifling'; Dell and Elmedlaoui 1985:112). This is clear evidence for the assumption that wordlikeness depends on language-specific knowledge.

> gradient cline of acceptability. (Albright 2009:9)

> In the particular domain of phonotactics gradient intuitions are pervasive: they have been found in every experiment that allowed participants to rate forms on a scale. (Hayes and Wilson 2008:382)

> A defect of current grammatical accounts of phonotactics is that they render simple up-or-down decisions concerning well-formedness and cannot account for gradient judgements. But when judgements are elicited in a controlled fashion from speakers, they always emerge as gradient, including all intermediate values. (Shademan 2006:371)

If the presence of intermediate values in wordlikeness tasks is evidence for the gradient nature of phonotactic wellformedness, then it follows that wordlikeness intuitions should be measured and modeled with a high degree of granularity. For instance, this would be strong evidence against the naïve account of the [blɪk]-[bnɪk] contrast just alluded to, since it cannot easily be extended to account for "numerous other distinctions". However, this chapter argues that there are both theoretical and empirical reasons to doubt the implicit hypothesis linking scalar wordlikeness ratings and gradient wellformedness. First, intermediate ratings are characteristic of all gradient rating tasks, and therefore are irrelevant to the question of whether the internal phonotactic system is categorical or gradient. Secondly, simple baselines better account for gradient well-formedness judgements than current computational models of phonotactic knowledge, suggesting that the gradience observed in these tasks do not derive from known grammatical mechanisms.

2.1 Aspects of the theory of gradient grammaticality

The aforementioned discussions of gradient aspects of wordlikeness judgements takes for granted that intermediate ratings are the product of an internal system of gradient grammaticality. This view is itself an instance of what is known as *common-sense* or *naïve realism*; in the cognitive sciences, this often takes the form of the assumption that experimental data can be taken at face value, without mediation from other sources of information. However, there are several arguments for *a priori* skepticism about the (a) linguistic abilities required for reporting gradient grammaticality judgements, (b) intermediate acceptability ratings as evidence for gradient grammaticality, and (c) the total lack of previous attempts to consider categorical models of wellformedness.

2.1.1 A model of gradient intuitions

Current research into gradient wellformedness is concerned with specifying a function from sound sequences to scalar judgements, and thus describes the wellformedness system at a level of some abstraction, corresponding roughly to what Marr (1982) calls the "computational" level of description. This is only one part of any model of gradient grammaticality, however; further assumptions are needed to articulate the internal representations and algorithms by which this function computes.

First, consider the architecture which is implied by any system of gradient grammaticality. It is essential that a system of gradient grammaticality have access to a relatively faithful representation of stimuli in a wellformedness task, and therefore it must be able to parse an enormous range of linguistic structures, including many which are not generated by the grammar itself; independent perception and production grammars may be necessary. A scalar value, representing wellformedness, must then be assigned to this parse. To report wellformedness, the speaker must transform this scalar value in accordance with the numerical scale chosen by the experimenter.

Each step of this procedure merits scrutiny, however. First, speakers have difficulty perceiving (Berent et al. 2007, Brown and Hildum 1956, Dupoux et al. 1999, Kabak and Idsardi 2007) and producing (Davidson 2005, 2006a,b, 2010, Gallagher in press, Rose and

King 2007, Vitevitch and Luce 1998, 2005) phonotactically illicit non-words, suggesting that speakers' ability to faithfully parse illicit representations is at best quite limited. Secondly, the computation of a scalar value serves no further purpose than to provide for gradient well-formedness judgements, so an objection might be made here on evolutionary grounds. It is quite mysterious why the human language endowment includes constraints on pronominal binding, for instance, but no one can doubt that these constraints are implicated in everyday language use; far more bizarre is the suggestion that the linguistic endowment includes mechanisms only implicated in certain experimental tasks. Next, speakers must be able to consciously access and report the magnitude of this value (it must be *cognitively penetrable* in the sense of Pylyshyn 1984), an ability which is limited in many other domains. Finally, there is some evidence that speakers do not (or perhaps, cannot) respect the numerical scales chosen by experimenters (Sprouse 2011).

It is informative to compare this baroque model to the architecture implied by a binary well-formedness judgement. When presented with a linguistic item in a judgement task, the grammar attempts to assign a parse. Speakers then access whether or not parsing was successful. There are reasons to think that parsing of ungrammatical structures does in fact result in a "crash": whereas syntactic priming increases the acceptability of grammatical structures (Luka and Barsalou 2005), ungrammatical structures show no priming effects (Sprouse 2007). As priming of linguistic structures is thought to implicate shared representations in memory, this suggests different memory mechanisms for grammatical and ungrammatical linguistic objects. Finally, the fact that requests for repetition and clarification are ubiquitous in spontaneous speech illustrates further that speakers are frequently aware when parsing has failed. Consequently, a great deal of evidence is needed to reject this simple model in favor of the gradient grammaticality architecture.

2.1.2 What some linguistic intuitions might not be

As first noted by Chomsky and Miller (1963), speakers experience difficulty processing sentences with multiple center embeddings. Gibson and Thomas (1999) perform a controlled experiment which reveals that speakers rate sentences like (4a), which is well-formed, less grammatical than (4b), despite the fact that a moment of reflection reveals that the latter sentence is nonsensical.

(4) A well-formedness illusion:

a. The patient who the nurse who the clinic had hired admitted met Jack.

b. *The patient who the nurse who the clinic had hired met Jack.

It is informative to consider that this well-known result has had no effect on the theory of syntactic representations, only on the theory of linguistic memory; it is recognized as the product of cognitive restrictions found in non-linguistic domains, as a *task effect*. This contrasts with the argument made by Hayes (2000), that gradient wordlikeness judgements demand a considerable and unmotivated revision to the grammatical architecture, as is discussed below.

The results of controlled experiments are often biased by subtle details that seem orthogonal to the task: for instance, certain types of duration judgements are systematically biased by consumption of caffeine (Gruber and Block 2005). It should come as no surprise, then, that a highly salient aspect of a judgement task, the scale used for responses, also influences the results obtained. Armstrong, Gleitman, and Gleitman (1983) argue that rating tasks using many-valued scales may induce intermediate ratings as a task effect.

Armstrong et al. (1983) are concerned with experimental evidence for the nature of cognitive concepts. While they do not attempt to dispute that certain concepts (e.g., *fruit*) have a family-resemblance structure (e.g., Rosch 1975), they assert that it is apparent that

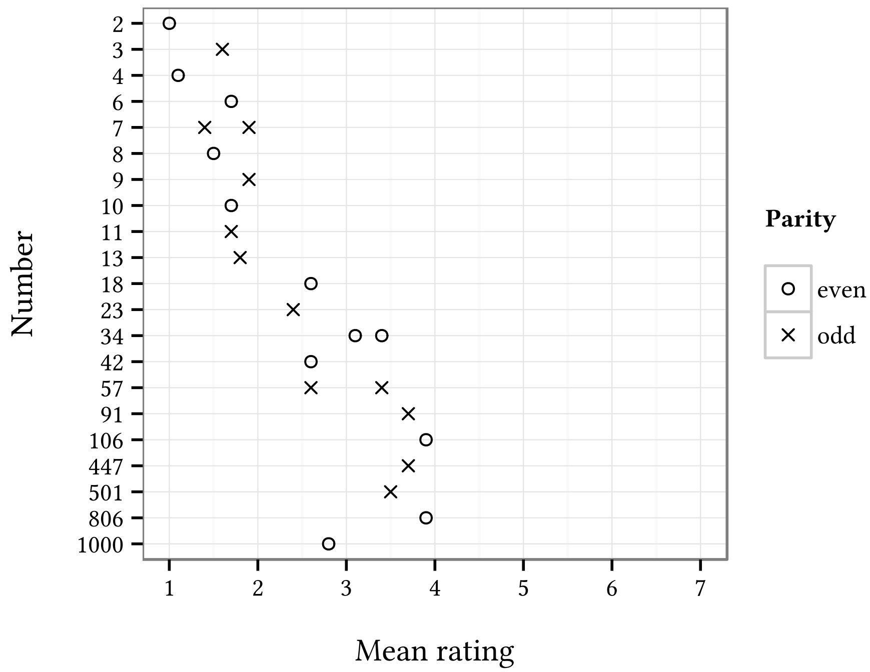


Figure 2.1: When asked to rate how representative even and odd numbers were of “even” and “odd”, respectively, subjects use intermediate ratings (from Armstrong et al. 1983; “1” indicates “most representative”)

other concepts are “definitional” (i.e., all-or-nothing), a notion which they illustrate with *odd number*.

> No integer seems to sit on the fence, undecided as to whether it is quite even, or perhaps a bit odd. No odd number seems odder than any other odd number. (Armstrong et al. 1983:274)

However, when subjects are asked to rate, using a 7-point Likert scale, how representative individual odd counting numbers are of the concept *odd number*, they freely use intermediate ratings; the ratings they obtain with instances of *odd number* and *even number* are shown in Figure 2.1.

This suggests that the gradience observed is primarily an artifact of the task itself. Schütze suggests that the nature of this effect might be understood as the result of speakers' attempts to reconcile bizarre experimental tasks with their knowledge.

> Putting it another way, when asked for gradient responses, participants will find some way to oblige the experimenter; if doing so is incompatible with the experimenter's actual question, they apparently infer that she must have really intended to ask something slightly different. (Schütze 2011:24)

As Armstrong et al. observe, these results show that the scalar judgement tasks provide no evidence as to whether the category being rated is categorical or gradient.

> ...we hold that *fruit* and *odd number* have different structures, and yet we obtain the same experimental outcome for both. But if the same result is achieved regardless of the concept structure, then the experimental design is not pertinent to the determination of concept structure. (Armstrong et al. 1983:284–5)

It might be said that these results reveal something about the representation of odd numbers. Armstrong et al. anticipate this objection.

> Some have responded to these findings very consistently, by asserting that the experimental findings are to be interpreted as before: that, psychologically speaking, odd numbers as well as birds and vegetables are graded concepts... We reject this conclusion just because we could not explain how a person could compute with integers who believed that 7 was odder than 23. We assert confidently that the facts about subjects being able to compute and about their being able to give the definition of odd number, etc., are the more important, highly entrenched, facts we want to preserve and explain... (Armstrong et al. 1983:284)

No scientist has risen to the challenge of constructing a theory that might account for the fact that 447 is rated more odd than 3, and as Armstrong et al. suggest, it is unclear whether such a theory could preserve more central facts about oddness. Here it is possible to draw an analogy to phonotactic theory. According to current orthodoxy, the wellformedness of a sequence is closely related to its type frequency (i.e., frequency in the lexicon). Is it then the case that [bl] is a significantly "better" onset than [kl], simply because the former is approximately twice as frequent, and if so, is it possible to also ensure that these sequences are treated the same with respect to syllabification?

2.1.3 Evidencing gradience

Hayes (2000) argues that it is "uninsightful" to attribute gradience to task effects, insofar as these effects implicate grammatical representations.

> ...patterns of gradient well-formedness often seem to be driven by the very same principles that govern absolute well-formedness... I conclude that the proposed attribution of gradient well-formedness judgments to performance mechanisms would be uninsightful. Whatever "performance" mechanisms we adopted would look startlingly like the grammatical mechanisms that account for non-gradient judgments. (Hayes 2000:99)

The logic of this implication is indisputable. However, there is little empirical support for the claims that absolute and gradient well-formedness derive from similar principles; indeed, there have been no prior attempts to evaluate categorical and gradient models of wordlikeness on an equal footing. In light of the complexities of gradient models, such an evaluation requires strong quantitative evidence for the superiority of gradient grammatical models. The evaluation below represents a first attempt to fill this gap.

It is not that categorical models have been ignored by the literature on wordlikeness modeling, but rather that they have not been compared. Frisch et al. (2000) and Vitevitch et al. (1997) find that speakers' wordlikeness ratings of multisyllabic words are correlated with a probabilistic measure of the well-formedness of the constituent syllables. Unfortunately, no attempt is made to control for the well-formedness of syllable contact clusters in these words: some of the stimuli have medial consonant clusters containing both voiced and voiceless obstruents (e.g., [gaɪbsaɪk]), something which is exceptionally rare in English simplex words (see §4.2.2). Similarly, Hayes and Wilson (2008), who compare their gradient model of wordlikeness against a set of English phonotactic constraints proposed by Clements and Keyser (1983), first transform these constraints, several of which are without exception, into probabilities. While this is consistent with their claim, that "the ability to model gradient intuitions [is] an important criterion for evaluating phonotactic models" (Hayes and Wilson 2008:382), little insight can be gained by annotating an exceptionless rule "$p = 1.0$". Hayes and Wilson's principle precludes any attempt to test the hypotheses that underlies it, and therefore must be rejected.

2.2 Evaluation

As Newmeyer (2007) writes "the idea that categoricity is not represented in the data itself is a truism. Whether distinctions of grammaticality (as opposed to acceptability) are binary is a difficult question." (398) The mere presence of gradience in judgements cannot falsify the claim of gradient grammaticality; another method is needed to evaluate this claim. As a first step towards a falsifiable theory of gradient wordlikeness, the remainder of this chapter considers intermediate ratings in gradient wordlikeness tasks are reliably predicted by the computational models that have been proposed. If a model is incapable of accounting for intermediate ratings, there are two possibilities: either the model itself is improperly specified and therefore at fault, or the inputs and outputs of the model are unrelated to the actual causes of the intermediate ratings, *contra* Hayes (2000).

It is plausible that speakers might differentiate, in a regular fashion, between different types of "impossible" words, and a gradient model should reliably predict the distinctions that speakers make. There are also claims that speakers distinguish between different types of "possible" words, so that, for instance, [stɪn] *stin* is rated more English-like than [blɪn] *blin* (e.g., Albright 2009), because the former onset is more frequent in the English lexicon. Even if wordlikeness judgements can be effectively modeled with a gross contrast between possible and impossible words, a gradient model might show a correlation with the residual ratings. All of these possibilities are considered below.

2.2.1 Materials

This evaluation uses a large sample of three previously published studies on English wordlikeness comprising 125 subjects and 187 items. Two criteria were used to select these three studies. First, the stimuli must be presented aurally so as to eliminate any possibility of orthographic effects (e.g., Berent et al. 2001, Berent 2008). Secondly, the data must be sufficiently "phonotactically diverse": that is, it must include both items like *blick* and *bnick*. This excludes studies like that of Bailey and Hahn (2001), in which few if any items contain gross phonotactic violations of the type represented by *bnick*. In the absence of phonotactic violences, little variance in wordlikeness ratings can be attributed to phonotactic wellformedness, making it difficult to determine the coverage of gradient wellformedness models. The data used here is summarized in Table 2.1.

Albright 2007

Albright (2007) administers a wordlikeness task in which 68 adult speakers rate 40 monosyllabic nonce words, presented aurally, on a 7-point Likert scale with endpoints labeled

	subjects	items	trials
Albright	68	40	2,720
Albright and Hayes	24	86	2,064
Scholes	33	63	2,178
TOTAL	125	187	6,962

Table 2.1: Subject and item counts for the wordlikeness study

"completely impossible as an English word" and "would make a fine English word". Albright's study is primarily concerned with the effects of different onset types (e.g., well-formed /bl/, marginal /bw/, unattested /bn, bd, bz/), and there is less diversity among the choice of rimes, none of which are obviously ill-formed.

Albright and Hayes 2003 (norming experiment)

Albright and Hayes (2003) have 24 adult speakers rate 87 aurally presented monosyllabic nonce words on a 7-point Likert scale with endpoints labeled "completely bizarre, impossible as an English word" and "completely normal, would make a fine English word". This task was administered to establish phonotactic norms for a later nonce word inflection task. Their item [raɪf] is excluded in this study, since this is an actual word of English, *rife*. Albright (2009) uses this data to compare computational models of wordlikeness.

Scholes 1966 (experiment 5)

Scholes (1966) conducts several wordlikeness tasks with students in 7th grade (approximately 12–13 years of age). The data used here is his experiment 5, in which 33 speakers provide a "yes" or "no" as to whether each of the 63 items, presented aurally, are "likely to be usable as a word of English". Like the study by Albright (2007), the focus is on onset well-formedness and there is minimal diversity in rime type. Two items, [klʌŋ] *clung* and [bɹʌŋ] *brung* (a dialectal past participle of *bring*), are excluded here as actual words of English. Albright (2009) and Hayes and Wilson (2008) also use this data for the purposes of model evaluation; following Frisch et al. (2000), they use the proportion of "yes" responses for each item so as to derive a continuous measure of well-formedness.

2.2.2 Method

Models are evaluated by comparing their scores to the average rating of each word using four correlation statistics. Each of these range between $[-1, 1]$, where 1 indicates a perfect positive correlation and -1 denotes a perfect negative correlation. Hayes and Wilson (2008) evaluate their model using the Pearson ("product-moment") r, a parametric correlation measure. It has long been argued (e.g., Stevens 1946) that parametric statistics are inappropriate for analysis of Likert scale data, like those used by Albright (2007) and Albright and Hayes (2003), since the Pearson r makes a *linearity assumption*. That is, it assumes that nonce words rated "1" and "3", for instance, are just as different as are those rated "4" and "6". A weaker assumption, more appropriate for Likert scale data, is the *monotonicity assumption*: that "1" is less English-like than "3", which is less English-like than "4", and so on. However, it also has been claimed that r is particularly robust to violations of the linearity assumption (e.g., Havlicek and Peterson 1976). Pearson r is reported here, but this should not be taken to imply an endorsement of its use for Likert scale data.

Hayes and Wilson also report Spearman ρ; this statistic requires only the weaker assumption of monotonicity, but it is difficult to give a simple interpretation to the coefficient. Two other non-parametric statistics, the Goodman-Kruskal γ and the Kendall τ_b are much easier to interpret, as follows (Noether 1981). These statistics are computed by comparing every model score/wordlikeness rating pair to every other: a comparison is counted as *concordant* if the greater of the two model scores is the one associated with the greater of the two wordlikeness ratings (that is, the model ranks these two nonce words in accordance with speakers' ratings), and as *discordant* otherwise. These two statistics differ only in the treatment of “ties”, pairs where either the model score or wordlikeness rating are identical. For γ, ties are ignored, and the coefficient is

$$\gamma = \frac{c-d}{c+d}$$

where c and d represent the number of concordant and discordant pairs, respectively. The τ_b statistic uses a similar formula, but also incorporates a penalty for ties in model score which are not also paired with ties in wordlikeness ratings, or vice versa. Albright (2009) uses a variant of this statistic to evaluate wordlikeness models.

2.2.3 Models

The nonce word stimuli from these three studies are scored automatically using four computational models. The first two models represent baselines for comparison to the latter two state-of-the-art gradient models. The scores are reproduced in Appendix A.

Gross phonotactic violation

A simple baseline is constructed by separating nonce words into those which contain a phonotactic violation and those which do not. As all nonce words here are monosyllabic, this task can be localized to two subcomponents of the syllable, the onset and the rime. This is not to imply that these are the only domains over which phonotactic violations might be stated, but there are prior claims that onset and rime are particularly important domains for stating phonotactic constraints (e.g., Fudge 1969, Kessler and Treiman 1997, Treiman et al. 2000). Speakers are adept at separating syllables into these units (Treiman 1983, 1986, Treiman et al. 1995), and they are implicated by patterns of speech errors (Fowler 1987, Fowler et al. 1993).

Operationalizing “phonotactic violation” is somewhat more difficult. The simplest

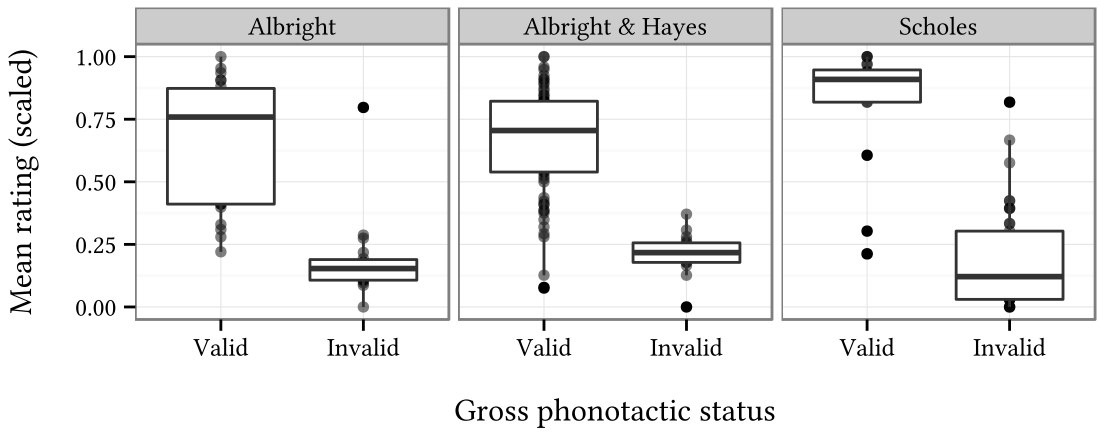


Figure 2.2: Gross phonotactic status and item-averaged wordlikeness ratings

possible mechanism is chosen here: an onset or rime is identified as well-formed if it occurs with non-zero frequency in a representative sample, and is identified as ill-formed otherwise. This is not to imply that all unattested onsets or rimes should be regarded as ill-formed, or that all onsets or rimes with non-zero frequency in this data are well-formed. For instance, Albright (2009) judges [dɹɛsp] *dresp* to be phonotactically well-formed, despite the total lack of [ɛsp] rimes in English; similar observations have been made concerning English onsets (e.g., Cairns 1972, Moreton 2002).

The representative sample used to define the phonotactic baseline is derived from those entries of the CMU pronunciation dictionary which occur at least once per million words in the SUBTLEX-US frequency norms; these norms are known to be particularly strongly correlated with other behavioral measures (Brysbaert and New 2009). These pronunciations are then syllabified, and individual syllables parsed into onset and rime, according to a process described in detail in Appendix B.

Wordlikeness ratings from the three studies are plotted against this gross contrast in Figure 2.2. While there are several outliers, there can be little doubt that gross phonotactic status accounts for a considerable amount of variance in wordlikeness judgements.

Lexical neighborhood density

A second baseline is provided by measures of similarity to existing English words, which has long been applied to model wordlikeness judgements (e.g., Bailey and Hahn 2001, Greenberg and Jenkins 1964, Kirby and Yu 2007, Ohala and Ohala 1986, Shademan 2006, 2007, Vitevitch and Luce 1998, 1999). Chomsky (1955: 151, fn. 27) suggests that grammaticality judgements in general might be influenced by similarity to existing grammatical structures, and Chomsky and Halle (1968:417f.) outline a similarity-based wordlikeness model. More recently, it has been observed (e.g., Coleman and Pierrehumbert 1997:51, Hay et al. 2004) that nonce words which flagrantly violate English sonority restrictions but which bear common affixes (e.g., **mrupation*) are rated highly English-like.

A wide variety of lexical similarity measures were considered, including a variant of the Generalized Neighborhood model (Bailey and Hahn 2001), PLD20 (Suárez et al. 2011), and a set of measures provided by the Irvine Phonotactic Calculator (Vaden et al. 2009). The measure best correlated with wellformedness judgements is also the most venerable measure of lexical similarity: Coltheart's *N* (Coltheart et al. 1977), which is defined as the number of words in some representative sample which can be changed into a target nonce word by a single insertion, deletion, or substitution of a phoneme. Greenberg and Jenkins (1964) find a correlation between wordlikeness ratings and a variant of this measure which only counts words differing by a single substitution. This is plotted against ratings from the three studies in Figure 2.3 with a superimposed local regression (LOESS; Cleveland and Devlin 1988) curve; neighborhood density accounts for much of the variance in ratings.

While there is nothing inherently "phonotactic" about Coltheart's *N*, it indirectly incorporates much of the information present in the gross phonotactic baseline. Consider [blɪk]: since there is nothing marked about any part of this nonce word, a "neighbor" might be found by modifying any phone: e.g., *click*, *brick*, *bloke*, *bliss*. However, since [bn] onsets are unattested in English, a neighbor of [bnɪk] must somehow modify this

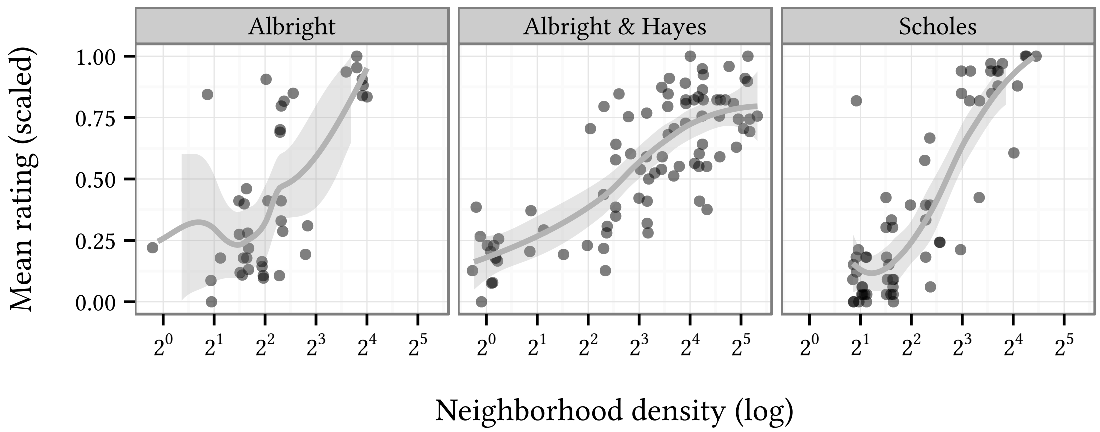


Figure 2.3: Correlation between Coltheart's *N* and item-averaged wordlikeness ratings, with LOESS curve

cluster: this leaves only *brick* and *nick*. Bailey and Hahn (2001) and Frauenfelder et al. (1993) note that neighborhood density is also strongly correlated with measures like bigram probability, but it has been argued elsewhere that phonotactic measures and neighborhood density have distinct effects (e.g., Berent and Shimron 2003, Pitt and McQueen 1998, Vitevitch and Luce 1998, 1999).

Segmental bigram probability

Faciliatory effects of bigram probabilities (i.e., shorter latencies) are reported for other nonce word tasks conducted with adults, including single-word shadowing (Vitevitch et al. 1997, Vitevitch and Luce 1998), same/different judgements (Lipinski and Gupta 2005, Luce and Large 2001, Vitevitch and Luce 1999, 2005), and lexical decision (Pylkkänen et al. 2002). Albright (2009) applies bigram probability as a model of wordlikeness judgements. The bigram probability of a sequence *ijk*, for instance, is defined as

$$\hat{p}(ijk) = p(i|\textsc{start}) \cdot p(j|i) \cdot p(k|j) \cdot p(\textsc{stop}|k)$$

	Pearson r	Spearman ρ	G-K γ	Kendall τ_b
featural bigrams	.71	.64	.45	.45
segmental bigrams	.74	.67	.48	.47
segmental bigrams with smoothing	<u>.75</u>	<u>.70</u>	<u>.50</u>	<u>.50</u>

Table 2.2: Correlation between item-averaged wordlikeness ratings for the Albright and Hayes (2003) norming study and three variants of bigram probability

That is, it is the product of sequence-initial *i*, the probability of *j* following *i*, the probability of *k* following *j*, and the probability of the sequence ending after *k*.

Albright (2009) compares two variants of this model, the first operating over segments, the second over sets of features. Unfortunately, the latter model is not described in sufficient detail to allow it to be implemented directly, and there is no publicly available implementation. However, Albright's evaluation, which includes the Scholes (1966) and Albright and Hayes (2003) data, finds an advantage for the segment-based model. In implementing this model, it was found that a slight improvement could be made by preventing any phoneme-to-phoneme transition from having zero probability. This is accomplished by adding 1 to the count of every transition, a technique used in natural language processing and known as Laplace, or "add one", smoothing. As can be seen in Table 2.2 this results in a slight increase in the correlation between the scores from this model and well-formedness ratings This smoothed segmental bigram score is adopted below. In Figure 2.4, it is plotted against wordlikeness ratings from three studies.

Maximum entropy phonotactics

Hayes and Wilson (2008) present a model which uses the principle of maximum entropy to weigh a large number of competing phonotactic constraints (e.g., Goldwater and Johnson 2003, Jäger 2007). Hayes and Wilson use a complex method to evaluate their model. First, they extract onset sequences from the CMU pronunciation dictionary, and use these to train the model. The model is then used to score the onsets of the Scholes (1966) nonce

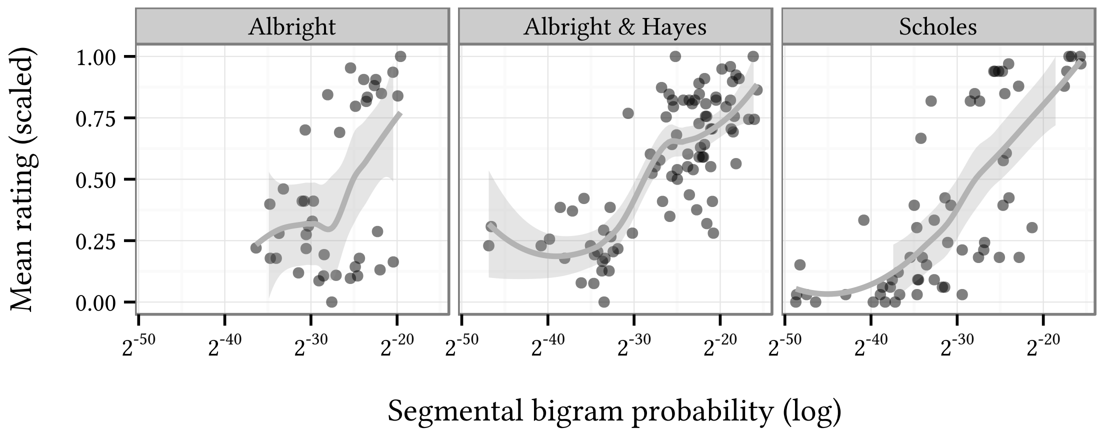


Figure 2.4: Correlation between smoothed segmental bigram score and item-averaged wordlikeness ratings, with LOESS curve

words. Then they compute a parameter for transforming their model scores so as to maximize the correlation between these transformed scores and wordlikeness ratings, then report the resulting correlation.[2] Albright (2009) reports that the maximum entropy model, training and testing only on onsets, performs well on the Scholes (1966) data, but does not generalize well to the Albright and Hayes (2003) sample. Consequently, the model was trained to score whole words, not just onsets, using the subset of the CMU dictionary described above.

Since this model has numerous experimenter-defined parameters, a close replication of Hayes and Wilson's original paper is attempted: both their implementation and phonological feature specifications are used here. Following Hayes and White (in press), dictionary entries are syllabified using the procedure described in Appendix B, and a novel feature [±Coda] is added to allow the model to distinguish coda and onset consonants. Also, following Hayes and Wilson, constraints are limited to those spanning as many as

[2]This is contrary to standard practices in natural language processing, in that the data used for evaluation is also used to fit the model (namely, the transformation's parameter); when this is the case, there is reason to suspect the parameter values will not generalize to new data. No transformation is used here; this only has an effect on the Pearson *r* coefficient, since they use a transformation that preserves monotonicity.

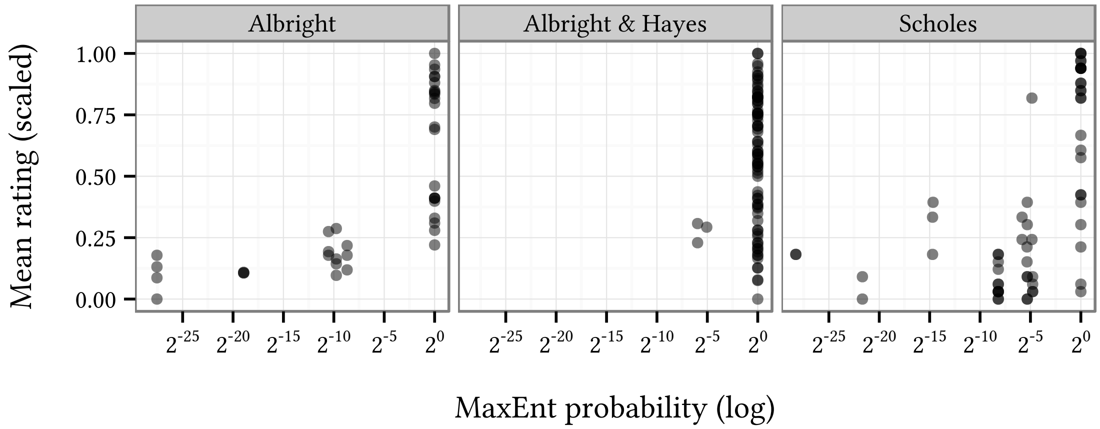


Figure 2.5: Correlation between MaxEnt score and item-averaged wordlikeness ratings

three segments and an "accuracy schedule" of [.001, .01, .1, .2, .3] is used. Since the maximum entropy model produces slightly different scores on each run, the worst-performing of 10 runs is reported here, following Hayes and Wilson. The resulting scores are plotted against wordlikeness ratings in Figure 2.5; it can be seen that the model assigns the highest possible score to a large variety of nonce words, though many words with a low rating receive the highest MaxEnt probability score. It appears that this model is still not robust enough to reliably extracting phonotactic generalizations from monosyllabic words.

2.2.4 Results

Table 2.3 displays the full set of correlation coefficients, for each of the three data sets, and for each of the four models. The first observation is that in general, there is a positive correlation between model score and ratings in each pair. The two baselines, gross phonotactic status and neighborhood density, are by far the strongest models across statistics and studies, with gross phonotactic status performing the strongest under the Goodman-Kruskal γ and on the Albright (2007) data, and neighborhood density performing strongly under nearly all other statistics and data sets.

	Pearson r			Spearman ρ			G-K γ			Kendall τ_b		
	A	AH	S	A	AH	S	A	AH	S	A	AH	S
Gross status	.73	.60	.80	.82	.66	.80	.87	.93	.91	.67	.47	.62
Density (N)	.67	.79	.86	.61	.74	.82	.49	.57	.74	.45	.56	.67
Bigram p	.46	.75	.74	.34	.70	.79	.25	.50	.63	.25	.50	.61
MaxEnt p	.70	.21	.53	.66	.39	.58	.85	.61	.56	.68	.16	.48

Table 2.3: Correlation between item-averaged wordlikeness ratings and model scores

	"Valid" items			"Invalid" items		
	A	AH	S	A	AH	S
Bigram score	.65	.34	.60	.03	$-.17$	.47
MaxEnt score	.00	$-.15$	$-.32$	$-.42$	.29	$-.16$

Table 2.4: Kendall τ_b correlation between model scores and item-averaged wordlikeness ratings, sorted according to gross phonotactic status

It is also possible to consider whether there is any residual correlation between bigram and MaxEnt model scores, and wordlikeness ratings within the "valid" and "invalid" groups defined by the gross phonotactic status measure. Kendall τ_b correlations within these subgroups for each data set are shown in Table 2.4. The only reliable positive correlation is present among the "valid" items as rated by the smoothed segmental bigram model. This model is somewhat capable of accounting for contrasts between different "possible" nonce words: for instance, it favors [plin] *plean* over [brɛlθ] *brelth* just as subjects in the Albright (2007) study do; this can also be seen in the top three panels of Figure 2.6. Within the set of "invalid" items, however, neither grammatical model reliably distinguishes among items; both models, for instance, rate [ptʌs] *ptus* more well-formed than [bnʌs] *bnus*, but speakers have the opposite preference.

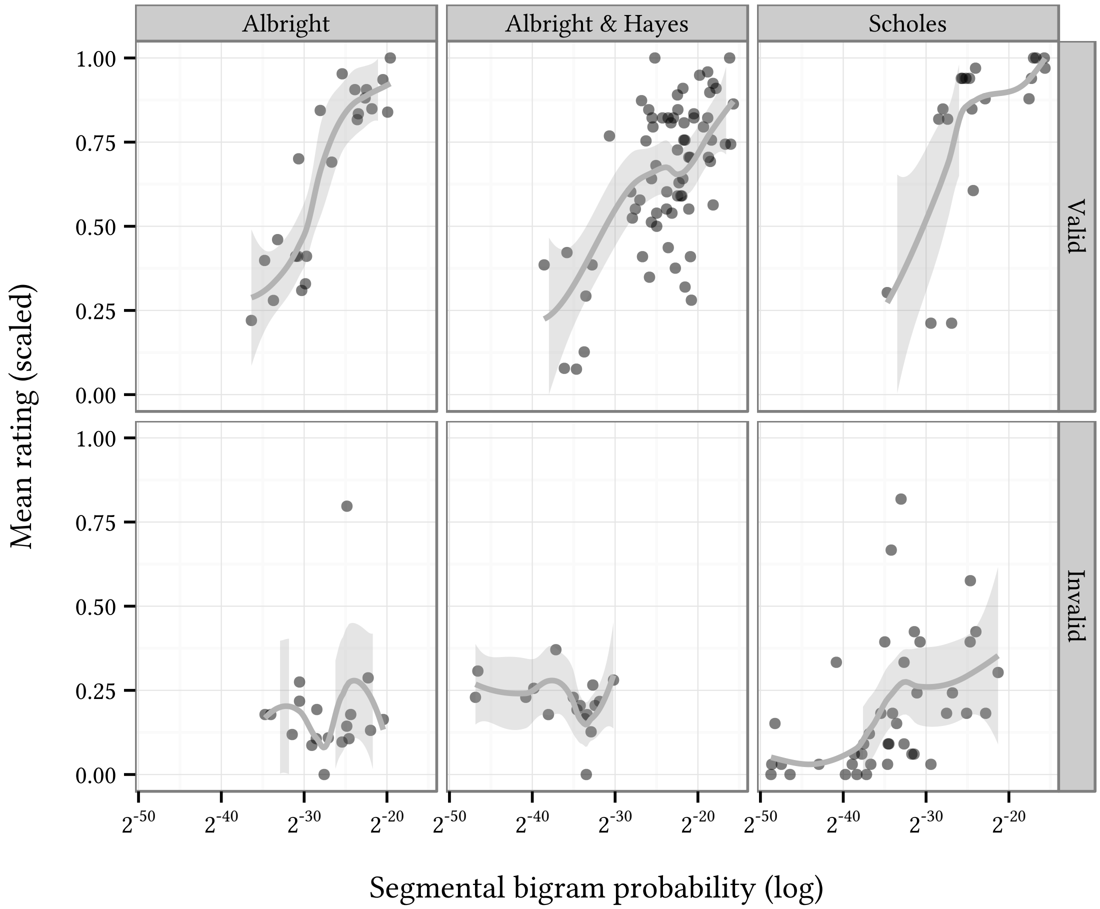


Figure 2.6: Correlation between smoothed segmental bigram score and item-averaged wordlikeness ratings, sorted according to the gross phonotactic status, with LOESS curve

2.2.5 Discussion

The bigram and MaxEnt models do not reliably outperform simple baselines. From this it can be inferred that the gradient models do not reliably predict intermediate ratings. Nor do these models reliably distinguish within classes of "valid" and "invalid" words in a way that conforms with wordlikeness ratings.

A serious limitation of this evaluation is the primitive nature of the gross phonotactic status baseline. It does not allow for any way to state constraints on onset-nucleus sequences, which have been proposed for some languages (e.g., Kirby and Yu 2007 on Cantonese), or constraints spanning whole syllables (e.g., Berkley 1994a,b, Clements and Keyser 1983, Coetzee 2008, Fudge 1969).[3] Furthermore, the gross phonotactic baseline does not have any mechanism for generalizing the wellformedness of [ɛsp] rimes from *clasp*, *lisp*, and other rimes consisting of a lax vowel followed by [sp] found in English, but Borowsky (1989), for instance, proposes a theory of possible rimes in English which makes the correct prediction regarding [ɛsp]. This is not embedded in an acquisitional model, but many models of syllable type acquisition have been proposed (e.g., Fikkert 1994, Levelt et al. 2000, Pan and Snyder 2003, 2004). As observed by Smith (1973) in a careful study of a single child acquiring English, children's productions are at first highly restricted but progress systematically to stages with fewer and fewer restrictions. Assuming productive competence is an appropriate measure of syllable acqusition, this suggests that syllable types are acquired like many other linguistic phenomena in that the child progresses from subset to superset. The difficulty here is that the typology of syllables must be delineated so that, for instance, the robust presence of [æsp] and [ɪsp] implies acceptance of [ɛsp].

The gross phonotactic baseline could also be extended so as to recognize more than two levels of wellformedness, without introducing the infinite amount of contrast implied by fully gradient models. While the bigram and MaxEnt models do not appear to be able to reliably distinguish intermediate levels of well-formedness, it might be desirable to encode the intuition that, for example, [ʒlɪk] *zhlick*, is more English-like than [bnɪk], though both have unattested onsets (e.g., Clements and Keyser 1983:50f.). It is also possible to imagine that phonotactic violations would have a cumulative effect on well-formedness.

[3] It is disputed whether English in particular exhibits onset-nucleus restrictions. Clements and Keyser (1983) claim that "cooccurrence restrictions holding between the nucleus and preceding elements of the syllable appear to be just as common as cooccurrence restrictions holding between the nucleus and following elements" (20), but admit that at least some of these generalizations may represent accidental gaps. However, Kessler and Treiman (1997), argue there are no clear restrictions on English onset-nuclei pairs.

For instance, a nonce word with an unattested onset and an unattested rime, like [tsɪlm], might be less English-like than either [tsɪl] or [sɪlm], an ability that could easily be extended to the gross phonotactic baseline. Cumulativity effects are predicted by the bigram and MaxEnt models, among others (e.g., Albright et al. 2008, Anttila 1997), but could easily be incorporated into a simple baseline by counting the number of violations. However, as of yet there is no convincing evidence for cumulativity effects in wordlikeness tasks, and the stimuli used here are not suited to test this hypothesis.

2.3 Conclusions

State-of-the-art computational models of wellformedness do not reliably predict intermediate ratings in wordlikeness tasks. To the degree to which the bigram or MaxEnt models are correlated with speakers' judgements, these judgements are more precisely modeled by similarity to existing words, or by a gross contrast between attested and unattested onsets and rimes. While it remains an open question whether future gradient models will account for intermediate judgements, the current evidence suggests that gradient grammaticality is not crucial for modeling gradient wordlikeness judgements. This does not imply that wordlikeness judgements collected using Likert scales or magnitude estimation are tainted; Sprouse and Almeida (submitted) argue that gradient wellformedness measures are better able to detect syntactic violations thought to be categorical than are binary judgements, and it seems likely this result would also hold for wordlikeness tasks. However, intermediate ratings can no longer be taken at face value.

Chapter 3

Static and derived Turkish phonotactics

It has long been speculated that statistical criteria could distinguish between *accidental phonotactic gaps*, those which could arise without any antecedent cause, and those gaps which are *structural* in nature (e.g., Fischer-Jørgensen 1952, Saporta 1955, Saporta and Olson 1958, Vogt 1954).[1] Following seminal work by Mester (1988) and McCarthy (1988), it is now commonly assumed that the phonotactic patterns in lexical entries "directly determine the mental representation of the phonotactic constraints" (Frisch et al. 2004:180) acquired by speakers, and therefore phonotactic constraints, whether categorical or gradient, can be directly inferred from statistical analysis of the lexicon. There is by now an enormous amount of research that adopts this assumption. Afroasiatic languages, especially Arabic (Coetzee and Pater 2008, Frisch et al. 2004, McCarthy 1988, Pierrehumbert 1993) but also Berber (Elmedlaoui 1995) and Tigrinya (Buckley 1997), and Austronesian languages, including Javanese (Graff and Jaeger in press, Mester 1988), Muna (Anttila 2008, Coetzee and Pater 2008), and Samoan (Alderete and Bradshaw in press), have been of particular interest to linguists adopting this hypothesis; Aymara (Graff and Jaeger in press),

[1]Some of this early work is, however, critical of any attempt to establish a firm distinction between structural and accidental gap. For instance, Fischer-Jørgensen (1952:3) writes that "it is theoretically impossible to fix a non-arbitrary borderline between law and accident."

Cantonese (Yip 1989), English (Berkley 1994b,a, 2000, Coetzee 2008, Davis 1989, Martin 2007, 2011), Gitksan (Brown 2010), Hungarian (Grimes 2010), Japanese (Kawahara et al. 2006), Ju|'hoansi (Kinney 2005), Navajo (Martin 2007, 2011), Ofo (MacEachern 1999:38f.), Russian (Padgett 1991, 1992), Shona (Hayes and Wilson 2008), and Wargamay (Hayes and Wilson 2008) have also received a statistical treatment. Pozdniakov and Segerer (2007) analyze lexical tendencies in a diverse sample of 30 languages.

Brown (2010) presents a strong form of this hypothesis, implying that *any* statistically significant pattern in the lexicon is one that is internalized by speakers:

> ... 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)

It would be a result of great interest were it shown that statistical significance is both necessary and sufficient to identify linguistic generalizations which are internalized by speakers, but there is no reason to grant this assumption with respect to phonotactic knowledge. But as Zwicky and Pullum (1987:330) correctly observe, "[n]ot every regularity in the use of language is a matter of grammar". Nothing demands that the antecedent cause of a statistically reliable lexical tendency be grammatical. On the contrary: there are reasons to suspect that many static phonotactic constraints identified in this manner have no synchronic reality.

By hypothesis, the phonological component may impose constraints on sound segments and sequence on underlying representations. Beyond this, a principled null hypothesis (implied by the principle of No Static Phonotactics) is that there is no synchronic, grammatical explanation for underlying representations a language chooses to instantiate. It is a practical necessity that numerous well-formed underlying representations will be uninstantiated: the lexicon is finite but there are an infinitude of possible URs.

The synchronic grammar cannot reasonably be expected to account for all non-existent underlying representations: for instance, the phonology of English has little to say about absence of */blɪk/ (cf. *flick*, *brick*, *block*, *blink*).

This much seems uncontroversial. However, what has not been appreciated is that even when a phonotactic generalization over underlying representations can be given a phonological characterization, there is a plausible alternative to the assumption that it is part of the synchronic grammar: namely, the generalization may be the result of now-complete diachronic change. Since sound changes often begin as phonological processes, it is no surprise that phonotactic gaps or tendencies can be explained with reference to phonological representations. But this suggests that the structural nature of a gap is not pertinent to determining whether the constraint is synchronically real.

Saussure (1916) presents an example of phonotactic underrepresentation caused by sound change. With only sporadic exceptions, Old Latin intervocalic *s* undergoes a conditioned phonemic merger with *r*. This has two consequences. Second, it introduces many *s*-*r* alternations: e.g., *honōs*-*honōris* 'honor'. The traditional analysis (e.g., Foley 1965:62, Gruber 2006:142, Heslin 1987:134, Kenstowicz 1996:377, Klausenburger 1976:314, Matthews 1972:19, Roberts 2012:88, Watkins 1970:526) treats *r* as the intervocalic allophone of /s/, and derives *honōris* from underlying /hono:s-is/.

(5) Rhotacism:

$$\mathrm{s} \longrightarrow \mathrm{r} \,/\, \begin{bmatrix}+\mathrm{Voc}\end{bmatrix} \;\_\_\; \begin{bmatrix}+\mathrm{Voc}\end{bmatrix}$$

However, subsequent sound changes (e.g., Baldi 1994, Safarewicz 1932), particularly the degemination of Old Latin *ss* after diphthongs and long monophthongs (e.g., *caussa* > *causa* 'cause'), introduce numerous exceptions to Rhotacism. In Classical Latin, intervocalic *s* is found root-internally (*asellus* 'donkey', *casa* 'hut'), in environments derived by inflectional suffixes (*uāsis* 'vase' GEN.SG., *uisēre* 'to view'), prefixation (*dēsecāre* 'to cut off'; cf. *dē* 'from', *secāre* 'to cut'), compounding (*olusātrum* 'parsnip'; cf. *olus* 'vegetable', *ātrum* 'black'), and denominal adjective formation (*uentōsus* 'windy'; cf. *uentus* 'wind'), and is tolerated in nativized loanwords from Celtic (*omāsum* 'tripe'), Germanic (*glaesum* 'amber'), and Greek (*basis* 'pedestal'). These facts lead Saussure to conclude that RHOTACISM is no longer "inhérente à la nature de la langue" (202). While some of the apparent exceptions may be the result of opaque phonological application (Heslin 1987)—an explanation not yet available in Saussure's time—any formulation of rhotacization will admit nearly as many lexical exceptions as there are roots exhibiting *s*-*r* alternations (Gorman in press). Any synchronic account of the underrepresentation of intervocalic *s* must confront the unproductive nature (as indicated by the accumulation of exceptions) of the proximate explanation for this tendency.

In other cases, it is possible to observe once robust phonotactic tendencies induced by sound change become increasingly moribund in real time. Perhaps the most famous example concerns the distribution of short *a* in English. Early Modern English short *a* underwent irregular lengthening and tensing before voiceless fricatives /f, θ, s/ and nasals /m, n/ (Wells 1982:I.203f.), introducing the first traces of the complex pattern now found in many American dialects (as well as in various British dialects, e.g., Jones 1964:74f.). A great deal of variation in Mid-Atlantic speech can be summarized as follows: the low front vowel is lax except before tautosyllabic /f, θ, s, m, n/ (and perhaps other consonants), where the low front vowel is tense (and perhaps raised), with various exceptions in both directions (e.g., Cohen 1970, Ferguson 1975, Labov 1981, Trager 1930, 1934, 1940). Cohen (1970), Labov (1981), and Trager (1940) all argue that this tendency is no longer to be attributed to synchronic allophony, but to a largely-complete lexical/phonemic split. Labov presents three arguments for this analysis. First, there are minimal pairs (tense modal *can* vs. lax noun/verb *can*) and extensive individual variation. Secondly, Payne (1980) finds

	{ɪ, ʊ}__#	{i, u}__#	% long	*p*-value
__ʃ#	46	4	8%	1.67E−06
__s#	40	36	53%	

Table 3.1: Type frequencies of high vowels before word-final [ʃ] and [s]

that children born out-of-state are unable to acquire the short *a* pattern of the Philadelphia metropolitian area in later childhood, though they participate fully in other sound changes specific to the area. Finally, Mid-Atlantic speakers have a greater ability to discriminate tense and lax variants of short *a* than speakers of dialects with allophonic short *a* tensing/laxing. In summary, the historical facts that account for this distribution no longer play a clear role in the synchronic phonology.

Other cases show that many constraints identified by statistical analysis are inert: that is, there is no evidence that they are internalized by speakers. For instance, most instances of Modern English [ʃ] derive from Old English [sk] (e.g., *fisc* 'fish') via sound change. Since Old English does not permit long vowels before complex codas, compared to similar segments like [s], [ʃ] is still rarely preceded by long vowels in word-final syllables (Table 3.1).[2] As can be seen, long vowels are twice as common before [s] as before [ʃ], a significant generalization according to the Fisher exact test. Similarly, Hayes and White (in press) report that this constraint is discovered by the Hayes and Wilson (2008) phonotactic learning model, which uses a related statistical criterion to identify constraints. Despite this, Iverson and Salmons (2005) label the constraint on long vowels before [ʃ] as "phonologically accidental", as a millennium of coinages (e.g., 'affective" *woosh*) and loanwords (e.g., *douche*) disregard this generalization. Hayes and White (in press) find that a variant of this restriction has little or no effect on wordlikeness judgements.

[2]This sample is drawn from the CMU pronunciation dictionary, filtered by excluding words with a token frequency of less than 1 per million words in the SUBTLEX-US frequency norms; similar results can be obtained with less restrictive samples. The four words ending in a long vowel-[ʃ] sequence are *douche*, *leash*, *unleash*, and *woosh*.

This latter example makes it clear that a purely statistical criterion overgenerates in the sense that it predicts phonotactic constraints which speakers do not seem to internalize. To account for the cases like the one above, Hayes and White (in press) propose that speakers are biased in favor of "natural generalizations" in probabilistic phonotactic learning.[3] However, this chapter argues that this overgeneration gives the lie to the broader assumption that phonotactics are extracted from patterns in the lexicon. Rather, the only restrictions which speakers internalize are those which derive from phonological processes in the language. This is independent of "naturalness", since both "natural" and "unnatural" variants of statistically reliable static constraints are shown to be inert.

This chapter focuses on three phonotactic generalizations in Turkish, comparing lexical statistics and the results of a wordlikeness task performed by Zimmer (1969). Both the lexical statistics and Zimmer's experimental results merit reconsideration, because prior discussions do not relate the lexical statistics to experimental data or competing formalizations of the generalizations involved.

3.1 Turkish vowel sequence structure constraints

Lees (1966a,b) proposes three constraints on Turkish vowel sequences; these constraints are the focus of many subsequent studies. In this section, these constraints are formalized, and where possible, related to phonological alternations and to behavioral evidence bearing on speakers' knowlege of the restrictions. The following feature specification for the eight vowels of Turkish is assumed throughout (e.g., Dresher 2009:298).[4]

[3]Hayes and White do not provide an operational definition of "naturalness", so it is difficult to evaluate their specific results despite their relevance to the argument at hand.

[4]Dresher in fact deploys a feature labeled Labial rather than Rnd.

(6) Turkish vowel features:

	[−Back]		[+Back]	
	[−Rnd]	[+Rnd]	[−Rnd]	[+Rnd]
[+High]	i	ü [y]	ı [ɯ]	u
[−High]	e	ö [ø]	a [ɑ]	o

Two less common notations are used in this chapter. First, directional application conditions are assumed, so as to derive the left-to-right, iterative properties of the harmony rules. Since it was first proposed by Johnson (1972), considerable evidence for directional application has been adduced (e.g., Anderson 1974: chap. 9, Gorman in press, Howard 1972:65f., Kavitskaya and Staroverov 2008, Kaye 1982, Kenstowicz and Kisseberth 1977:189f., 1979:319f., Piggott 1975, Sohn 1971; see McCarthy 2003 and Wolf 2011 for recent reviews.) Further, Johnson (1972) and Kaplan and Kay (1994) prove directional application and simultaneous application are equivalent in terms of formal learnability.

Secondly, rather than the use of an unbounded number of Greek-letter variables (α, β, etc.) over feature values $\{+, -\}$, only a single variable, denoted by '=', is used (McCawley 1973). A structural description [=F]...[=F] matches a string S_i...S_j if and only if S_i and S_j are both [+F] or both [−F]. This is more restrictive than Greek-letter notation, in that it prevents the value of one feature being applied to different feature.[5]

3.1.1 Backness harmony

Lees (1966a:35, 1966b:284) models the Turkish vowel harmony system with three rules; the most general of these rules spreads the specification [Back] rightward.

[5] Odden (2012) argues that the two counterexamples against this restriction presented in *SPE* (352-353) are not probative.

(7) Backness Harmony (condition: rightward application):

$$\begin{bmatrix}-\textsc{Cons}\end{bmatrix} \longrightarrow \begin{bmatrix}= \textsc{Back}\end{bmatrix} \;/\; \begin{bmatrix}= \textsc{Back}\end{bmatrix}\ C_0\ \text{—}$$

A vowel becomes [+Back] after a [+Back] vowel, and [−Back] after a [−Back] vowel, ignoring any intervening consonants. The application of this rule proceeds from left to right; no vowel may be skipped.

If permitted to apply in non-derived environments, this rule accounts for the tendency of polysyllabic roots to contain only [+Back] or [−Back] vowels, a tendency which will be quantified below. Backness Harmony also triggers alternations in inflectional suffix vowels. For instance, the nominative plural (nom.pl.) suffix is *-ler* when the final root vowel is [−Back], and *-lar* when it is [+Back].

(8) The Turkish nominative:

	nom.sg.	*nom.pl.*		
a.	ip	ipler	'rope'	(Clements and Sezer 1982:216)
	köy	köyler	'village'	
	yüz	yüzler	'face'	
	kız	kızlar	'girl'	
	pul	pullar	'stamp'	
b.	neden	nedenler	'reason'	(Inkelas et al. 2000)
	kiler	kilerler	'pantry'	
	pelür	pelürler	'onionskin'	
	boğaz	boğazlar	'throat'	
	sapık	sapıklar	'pervert'	

A few complications arise, however. First, as shown in (9a), not all polysyllabic roots conform to Backness Harmony. In this case, suffix vowels generally exhibit harmony with the final root vowel. There is also a very small class of nouns, shown in (9b), which take *-ler*, although their final root vowel is [+Back]. Interestingly, the roots themselves may be harmonic.

(9) Exceptional Turkish nominatives:

	nom.sg.	*nom.pl.*		
a.	mezar	mezarlar	'grave'	(Inkelas et al. 2000)
	model	modeller	'model'	
	silah	silahlar	'weapon'	
	memur	memurlar	'official'	
	sabun	sabunlar	'soap'	
b.	etol	etoller	'fur stole'	(Göksel and Kerslake 2005)
	saat	saatler	'hour, clock'	
	kahabat	kahabatler	'fault'	

Anderson (1974:212) and Iverson and Ringen (1978) argue that suffix harmony in disharmonic roots found in (9a) requires the rule governing suffix harmony alternations to be distinguished from a sequence structure constraint governing root harmony. Since the rule and sequence structure constraint are otherwise identical, this constitutes a "duplication" (in the sense of Kenstowicz and Kisseberth 1977:136f.): harmony is stated at two points in the grammar.

Another possiblity, however, is to understand root disharmony as a type of lexical exceptionality. Under such an analysis, suffix harmony in disharmonic roots is entirely consistent with the theory of exceptionality proposed in *SPE* (Zonneveld 1978:197f.).[6] Chomsky and Halle assume that the specification of the target (i.e., the segment or segments to be changed) of a rule *R* must be marked [+*R*] by convention. A root or affix which fails to undergo *R* despite otherwise matching the structural description is simply said to be marked [−*R*]. In other words, no representation is ever truly an exception to a rule; rather, some underlying representations have non-default features which do not match the extended structural description of *R*, which requires that the target be [+*R*]. If disharmonic roots are marked [−Backness Harmony], then the final vowel of disharmonic roots will still trigger Backness Harmony, since the [−Backness Harmony] root is no longer the target but rather the trigger, which is not subject to the [+Backness Harmony] requirement. Under this account, root and suffix harmony are both derived from Backness Harmony, but they do not have the same ontology: suffix harmony is direct result of phonological rule application, whereas the tendency for harmonic roots arises from a dispreference for lexical exceptionality. Under this analysis, there is no need to state a sequence structure constraint producing root harmony.

Anderson also notes that a small number of roots which fail to undergo suffix harmony, like (9b), may themselves be harmonic. He takes this to be evidence for the necessity of duplication:

> …there are words which are exceptions to harmony across boundaries (e.g., *kabahat* 'fault', *kabahatti* 'his fault') but which are perfectly regular internally. Since the morpheme structure condition and the phonological rule in this case have distinct classes of exceptions, it is clear that they cannot be identified. (Anderson 1974:289)

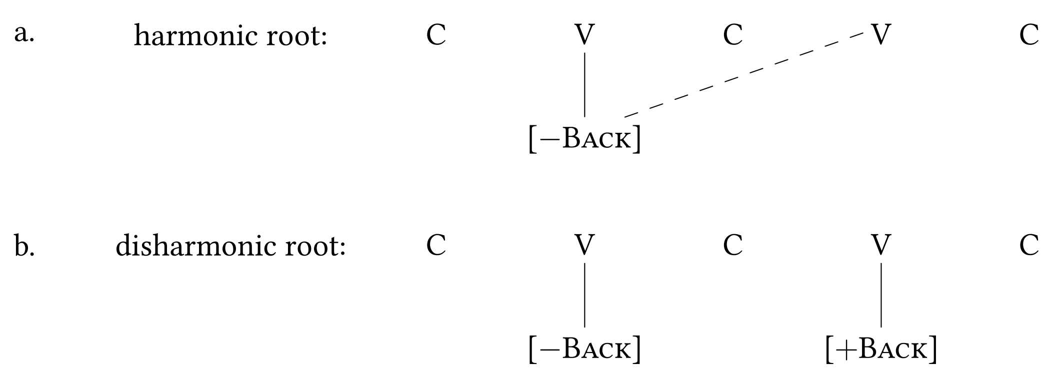
Under the assumptions so far, however, there is no reason to assume that *kabahat* (or any

[6] Kiparsky (1968:29f.) discusses a parallel case in Finnish, and proposes a similar separation between the exception-filled sequence structure constraint and an exceptionless suffix harmony process. Howard (1972:171f.) suggests an analysis of Finnish that is quite similar to the one proposed for Turkish below.

other root which fails to trigger suffix harmony) is [+Backness Harmony]. The tendency for root harmony holds of the Turkish lexicon as a whole, but there is no evidence that it holds of those few roots which select exceptional suffixes; indeed, there are disharmonic roots which select disharmonic suffixes (e.g., *etol* ‘fur stole’, nom.pl. *etoller*).

This analysis compares favorably to an alternative proposed by Clements and Sezer (1982) and Inkelas et al. (1997) which has other desirable properties but which makes no predictions about root harmony. Root vowels exhibit a robust contrast for backness (e.g., *kül* ‘ash’ vs. *kul* ‘servant’, *kepek* ‘bran’ vs. *kapak* ‘lid’), whereas backness of vowels in non-initial syllables is predictable in harmonic roots (note that there are no prefixes in Turkish). Clements and Sezer and Inkelas et al. propose that these vowels, as well as harmonizing suffix vowels, are underspecified for backness, whereas the non-initial vowels of disharmonic roots and of certain exceptional suffixes are fully specified. This is schematized below.

(10) Autosegmental underspecification in harmonic roots (after Clements and Sezer 1982):



One crucial detail is missing from this analysis: Backness Harmony needs to be prevented from overwriting the [+Back] specification of disharmonic roots, perhaps by structure preservation condition (Kiparsky 1985). However, such a condition makes it impossible to relate Backness Harmony to the tendency for harmonic roots, despite considerable evidence (reviewed below) that they are marked in Turkish.[7]

While harmony in non-derived environments can be inferred from the aforementioned suffix alternations, no evidence has yet been presented to show that Turkish speakers internalize the tendency for roots to conform to backness harmony. If Turkish speakers do not attend to this generalization, there is no need for the grammar to account for it. Several other "external" facts suggest that this is not the case. The discussion here is not intended to imply uncritical acceptance of evidence from loanword adaptation, language games, or particular psycholinguistic tasks as evidence for phonological grammar, but rather to illustrate additional evidence that is pertinent if the linking hypothesis is correct.[8]

The production of non-native word-initial onset clusters, discussed by Clements and Sezer (1982) and Kaun (1999), suggests that loanword adaptation respects BACKNESS HARMONY. Some speakers pronounce these non-native clusters, but in fast speech the cluster is split by anaptyxis. In the majority of cases, this vowel matches the following root vowel for backness.

(11) Variable non-native cluster adaptation (Clements and Sezer 1982:247):

a.	spiker	∼	sipiker	'announcer'
	fren	∼	firen	'brake'
b.	trablus	∼	tırablus	'Tripoli'
	kral	∼	kıral	'king'
c.	brom	∼	burom	'bromide'
	prusya	∼	purusya	'Prussia'

[7] On the other hand, it is possible to interpret the presence of a single backness specification per root as a sort of default. A precedent for this is the surface-oriented interpretation of the tonal Obligatory Contour Principle proposed by Goldsmith (1976:134) and Odden (1986), under which adjacent identical tones are automatically attributed to a single underlying tone. However, this is merely a notational variant of the rule exceptionality account in which [+BACKNESS HARMONY] is the default.

[8] Thanks to Bert Vaux and Kie Zuraw for bringing these studies to my attention.

It is unclear whether the cluster-splitting vowel is deleted in the non-native variant or epenthesized in the fast speech variant. Under either analysis, there is no ready explanation for the tendency of the cluster-splitting vowel to have the backness features of following vowels; if anything, one might expect it to determine the backness features of following vowels. All that can be said with certainty is that the adaptation of non-native onset clusters appears to proceed in such a fashion so that the lexical items in question are [+BACKNESS HARMONY].

Similar evidence comes from a language game discussed by Harrison and Kaun (2001). The game is native to the related language Tuvan, where it is used to convey a sense of "vagueness or jocularity"; it is not indigenous to Turkish, but can be taught quickly to children or adults. In this game, the base is reduplicated and the first vowel of the reduplicant replaced with a [+BACK] vowel. In (12a), the second [−BACK] vowel of the base is, in the reduplicant, "reharmonized" with the inserted [+BACK] vowel. The disharmonic roots of (12b) do not reharmonize.[9]

(12) Turkish reduplication game (Harrison and Kaun 2001:231):

a.	kibrit	kibrit-{kabrıt}	'match'
	bütün	bütün-{batın}	'whole'
b.	mali	mali-{muli}	'Mali'
	butik	butik-{batik}	'boutique'

Harrison and Kaun propose that those roots which fail to reharmonize are prespecified for backness throughout; this also requires that reharmonization is subject to a non-derived environment condition. In the full specification analysis adopted here, reharmonization is the result of BACKNESS HARMONY applying within the reduplicant; the lack of rehar-

[9]A similar contrast between harmonic and disharmonic roots is found in Tuvan (Harrison and Kaun 2001) and in an unrelated language game in Finnish (Campbell 1986).

monization in the reduplicants of disharmonic roots requires that the [−Backness Harmony] exception feature is copied under reduplication.[10]

A number of studies have investigated the role of harmony in word-spotting tasks, thought to mimic auditory word recognition and segmentation in natural settings. Many of these studies have been carried out in Finnish, which has a vowel harmony system similar to that of Turkish. Suomi et al. (1997) and Vroomen et al. (1998) task Finnish speakers with identifying harmonic disyllables in an auditory stream. When the syllable preceding the disyllabic target has a different backness specification than the target, recognition of the target is facilitated. Presumably, disharmony facilitates the recognition of word boundaries. Kabak et al. (2010) find that Turkish Backness Harmony has a similar effect: Turkish speakers are quicker and more accurate at the task of spotting the nonce target word *pavo* when preceded by a disharmonic juncture (e.g., *gölü-PAVO*) than when preceded by a harmonic juncture (e.g., *golu-PAVO*). Kabak et al. report that this effect does not obtain for speakers of French, a language which lacks vowel harmony. This implies that Turkish speakers have internalized the tendency of harmonic sequences to be root-internal and of disharmonic transitions to cross word boundaries.

The Turkish word-spotting experiment is adapted for infants by van Kampen et al. (2008). In this study, 9-month-old infants are familiarized with recordings of harmonic, disyllabic nonce words presented in isolation and tested with the head turn preference paradigm. Infants acquiring Turkish listen longer to nonce words preceded by a disharmonic juncture during familiarization (e.g., *lo-NETIS*), whereas infants acquiring German, a language which lacks vowel harmony, do not exhibit this preference. Similarly, van

[10]There are other cases which suggest that lexical diacritics are copied in reduplication. In Kinande, verbal reduplication requires a bisyllabic reduplicant (Downing 2000, Mutaka and Hyman 1990). This requirement has synchronic force in the grammar, since "reduplicated" monosyllabic roots in fact contain three copies of the root (so that the reduplicant is bisyllabic); furthermore, reduplicated forms of many trisyllabic verbs are ineffable. However, a few trisyllabic verbs exceptionally show full reduplication. This is a lexical property, but presence or absence of the exceptionalty feature that permits non-bisyllabic reduplicants only surfaces on the reduplicant itself.

Kampen et al. report that Turkish 6-month-old infants prefer to listen to harmonic nonce words such as *paroz* over disharmonic nonce words like *nelok*, but German 6-month-old infants show no such preference. However, there are some caveats for identifying these effects with grammatical computations: the domain for segmentation effects is considerably larger than the domain for harmony, so the two cannot be easily identified. All that can be said is that there is an obvious similarity between computing the contexts for vowel harmony and for the word segmentation heuristic.

3.1.2 Roundness harmony

Roundness Harmony is quite similar to Backness Harmony, but imposes an additional restriction, that targets be [+High].

(13) Roundness Harmony (condition: rightward application):

$$\begin{bmatrix} -\textsc{Cons} \\ +\textsc{High} \end{bmatrix} \longrightarrow \begin{bmatrix} =\textsc{Rnd} \end{bmatrix} \;/\; \begin{bmatrix} =\textsc{Rnd} \end{bmatrix} \; \mathrm{C}_0 \,\_\_$$

A [+High] vowel becomes [+Rnd] after a [+Rnd] vowel, and [−Rnd] after a [−Rnd] vowel, ignoring any intervening consonants, and applying from left to right.

If permitted to apply in non-derived environments, this rule accounts for the tendency of polysyllabic roots to contain only round or unround high vowels. In concert with Backness Harmony, Roundness Harmony also triggers alternations which account for the shape of the dative singular (dat.sg.) and genitive singular (gen.sg.) suffixes, among others.suffixes, among others. As was the case for Backness Harmony, disharmonic roots occur, but generally trigger suffix harmony.

(14) Turkish nominal suffix allomorphy:

	nom.sg.	*dat.sg.*	*gen.sg.*		
a.	ip	ipi	ipin	‘rope’	(Clements and Sezer 1982:216)
	kız	kızı	kızın	‘girl’	
	sap	sapı	sapın	‘stalk’	
	köy	köyü	köyün	‘village’	
	son	sonu	sonun	‘end’	
b.	boğaz	boğazı	boğazın	‘throat’	(Inkelas et al. 2000)
	pelür	pelürü	pelürün	‘onionskin’	
	döviz	dövizi	dövizin	‘currency’	
	yamuk	yamuğu	yamuğun	‘trapezoid’	
	ümit	ümiti	ümitin	‘hope’	

Few studies have directly investigated whether speakers are aware of the tendency for roots to conform to ROUNDNESS HARMONY. However, two of the external sources of evidence for BACKNESS HARMONY also bear on this question. First, the cluster-splitting vowel found in non-native onset clusters, discussed above, tends to agree in roundness with following high vowels (e.g., *prusya-purusya* ‘Prussia’). Secondly, ROUNDNESS HARMONY participates in reharmonization in the language game described by Harrison and Kaun: the second *ü* in *bütün* ‘whole’ reharmonizes to the [−RND] vowel *ı* in reduplicated *bütün-batın*.

3.1.3 Labial attraction

Lees (1966a:35) notes the tendency of Turkish high back vowels to be round after *a*-labial consonant sequences, and formalizes this as a phonological process.

(15) LABIAL ATTRACTION:

$$\begin{bmatrix} -\textsc{Cons} \\ +\textsc{Back} \\ +\textsc{High} \end{bmatrix} \longrightarrow \begin{bmatrix} +\textsc{Rnd} \end{bmatrix} \;/\; \text{ɑ}\, C_0 \begin{bmatrix} +\textsc{Cons} \\ +\textsc{Labial} \end{bmatrix} C_0 \,\rule{1em}{0.4pt}$$

This rule is significantly more complex than the harmony rules, and this may obscure the fact that it produces exceptions to ROUNDNESS HARMONY, producing *a*C_0*u* (e.g., *çapul* 'raid', *sabur* 'patient', *şaful* 'wooden honey tub', *avuç* 'palm of hand', *samur* 'sable'; Lees 1966b:285) rather than the expected *a*C_0*ı*. However, LABIAL ATTRACTION does not apply in derived environments: the gen.sg. of *sap* 'stalk is *sapın* rather than **sapun* that would be predicted if LABIAL ATTRACTION triggered alternations. Lees and Zimmer (1969) cite roots which do not conform to LABIAL ATTRACTION (e.g., *tavır* 'mode') but agree that they are surprisingly rare; so as to dispute the claim that such exceptions are rare, Clements and Sezer (1982) presents many additional examples of exceptions.

3.2 Evaluation

Zimmer (1969:311) administers two paired wordlikeness tasks designed to evaluate native speakers' knowledge of BACKNESS HARMONY, ROUNDNESS HARMONY, and LABIAL ATTRACTION in roots. Speakers are presented with a pair of nonce words, differing only in whether they obey or violate one of these three constraints, and then indicate the nonce word that is more Turkish-like.[11] Zimmer concludes that the former two rules are reflected

[11]Compared to the unpaired ratings tasks commonly used in wordlikeness research, paired rating tasks have considerably more statistical power (e.g. Gigerenzer et al. 2004), since there is little chance that any contrast between the phonotactically licit and illicit members of an otherwise-identical nonce words pair is caused by an omitted variable. Consequently, paired rating tasks are ideal for collecting wordlikeness judgements. The use of paired stimuli also makes more overt the purpose of the experiment from speakers, as opposed to the standard practice of concealing the purpose, the latter thought to increase response

in wordlikeness judgements, whereas Labial Attraction is not. Below, both lexical statistics and Zimmer's wordlikeness results are analyzed statistically; Labial Attraction is shown to be a statistically robust generalization over the Turkish lexicon, but no variant of Labial Attraction is reflected in Zimmer's wordlikeness study. In contrast, the two harmony processes have robust effects both on the lexicon and on wordlikeness. This dissociation between statistical tendencies generalizations and wordlikeness results provides further evidence against the assumption that phonotactic knowledge can be inferred directly from lexical statistics.

3.2.1 Lexical statistics

Counts are computed by regular expression matching on a 9,601-root subset of the TELL database, consisting of roots which show no surface variation in any inflected form.

To test for associations between the process (more specifically, the constraint that it imposes on roots) and type frequency in this database, each root was sorted into a 2×2 contingency table; the contents of each cell are specific to the process in question. The counts in this table are not expected to add up to 9,601, since many roots neither exemplify nor violate the process in question; for instance, monosyllabic words are irrelevant to root harmony. The Fisher exact test is used to compute a *p*-value representing the probability of the observed patterns arising under the null hypothesis that there is no association between the constraint and type frequency.

Backness harmony

Backness Harmony is exemplified in the lexicon insofar as there is a positive association between the backness of vowels in all adjacent syllables. Any root which contains vowels in adjacent syllables which disagree in backness is counted as disharmonic, even

variability (see Hertwig and Ortmann 2001:398f. for discussion).

	$[+\textsc{Back}]_1$	$[-\textsc{Back}]_1$	p-value
$[+\textsc{Back}]_{2\ldots n}$	3,089	1,704	1.19E−89
$[-\textsc{Back}]_{2\ldots n}$	1,698	2,250	

Table 3.2: TELL roots sorted according to BACKNESS HARMONY

	$[+\textsc{Rnd}]_i$	$[-\textsc{Rnd}]_i$	p-value
$[+\textsc{High}, +\textsc{Rnd}]_{i+1}$	613	261	1.02E−36
$[+\textsc{High}, -\textsc{Rnd}]_{i+1}$	581	2,841	

Table 3.3: TELL roots sorted according to ROUNDNESS HARMONY

if other vowel transitions in the root are harmonic; further assumptions are necessary to determine whether a root which has one disharmonic transition may in any sense "obey" harmony elsewhere. To construct the contingency table, roots are binned first according to the backness specification of the first nucleus, and then according to the backness of all following nuclei. For example, the first two syllables of *adalet* 'justice' are harmonic, but it is coded as disharmonic because there is a *a...e* transition later in the word. The resulting counts are shown in Table 3.2. 61% of roots conform to BACKNESS HARMONY, and the interaction between the backness of the first and of the subsequent vowels is significant.

Roundness harmony

ROUNDNESS HARMONY predicts correlation between the roundness of a vowel and the roundness of high vowels in the next syllable. Any root for which a vowel does not agree in roundness with a high vowel in the following syllable (e.g., *ümit*) is considered to be an exception. Roots are binned first according to the roundness of a vowel followed by a high vowel in the next syllable, and then according to the roundness of all following high vowels. The resulting counts are shown in Table 3.3. 83% of the roots conform to ROUNDNESS HARMONY, and the interaction between the roundness of the ith vowel and the roundness of the $(i+1)$th high vowel is significant.

The counts in the bottom row of Table 3.3 contain a number of roots which are apparent exceptions to ROUNDNESS HARMONY but conform to LABIAL ATTRACTION (bottom left), and which conform to ROUNDNESS HARMONY at the expense of LABIAL ATTRACTION (bottom right). Excluding these types of roots would have the effect of slightly increasing the overall rate of ROUNDNESS HARMONY, since the former is more common.

Labial attraction

Clements and Sezer (1982) also object to LABIAL ATTRACTION and argue that it is not "systematic".

> Even more decisive evidence against a rule of Labial Attraction is the existence of a further, much larger set of roots containing /...aCu.../ sequences in which the intervening consonant or consonant cluster does not contain a labial...We conclude that there is no systematic restriction on the set of consonants that may occur medially in roots of the form /...aCu.../. (Clements and Sezer 1982:225)

This claim can be evaluated using the Fisher exact test. Let P denote a sequence of one or more consonants, one of which is labial, and let T denote a sequence of one or more consonants none of which is labial. The null hypothesis is that *a*P*u* sequences, which conform to LABIAL ATTRACTION, are no more likely than would be expected from other *a*T*u* sequences violating ROUNDNESS HARMONY. The resulting counts are shown in Table 3.4. Whereas the sequence *a*P*u* is more than twice as likely as *a*P*ı*, the sequence *a*T*u* is 5 times less likely than *a*T*ı*. This interaction is significant, as predicted by LABIAL ATTRACTION, but contrary to Clements and Sezer's claim. In fact, *a...u* sequences are less, not more, common than *a...ı* sequences, presumably a consequence of ROUNDNESS HARMONY.

	a...u	a...ı	p-value
aP...	124	57	1.02E−36
aT...	136	590	

Table 3.4: TELL roots sorted according to LABIAL ATTRACTION

3.2.2 Wordlikeness ratings

Zimmer (1969) administers two variants of the paired nonce word rating task. The first used 23 native adult speakers who were permitted to select either nonce word as more like Turkish, or to indicate 'no preference'. For the purposes below, 'no preference' results are ignored. The second experiment used 32 native adults, none of whom appeared in the preceding study, and used a forced binary choice.

Each response is coded as *concordant* if the nonce word conforming to the process is preferred, and *discordant* if the disharmonic word is selected. To test for an association between the constraints, a non-parametric statistic, the Goodman-Kruskal (1954) γ is used. The γ statistic ranges between -1 (which would indicate that non-conforming nonce words are always preferred to conforming nonce words) and 1 (which indicates that conforming nonce words are always preferred).[12]

Backness harmony

Both of the Zimmer (1969) experiments include 5 pairs which differ in whether or not the nonce words conform to, or violate, BACKNESS HARMONY. As can be seen from Table 3.5, harmonic pairs are preferred approximately 6-to-1, and aggregating over speakers, no disharmonic member of a pair is favored. Speakers have a highly reliable preference for nonce words which exhibit BACKNESS HARMONY ($\gamma = 0.694$, $p = 1.7\text{E}{-}59$). It is interesting to note that the disharmonic nonce word which has the highest rating is found

[12]It also is possible to perform statistical tests aggregating over items, but for small number of items, such tests have very poor power.

Experiment 1				Experiment 2			
HARMONIC		DISHARMONIC		HARMONIC		DISHARMONIC	
temez	19	temaz	3	pemez	30	pemaz	2
teriz	23	terız	0	teriz	28	terız	3
tokaz	21	tokez	1	tokaz	26	tokez	6
tipez	21	tipaz	1	tipez	24	tipaz	8
terüz	20	teruz	1	terüz	19	teruz	13

Table 3.5: Effects of BACKNESS HARMONY on wordlikeness (from Zimmer 1969)

Experiment 1				Experiment 2			
HARMONIC		DISHARMONIC		HARMONIC		DISHARMONIC	
törüz	19	töriz	1	pörüz	32	pöriz	0
tüpüz	22	tüpiz	0	tüpüz	31	tüpiz	1
takız	15	takuz	3	takız	22	takuz	10
tatız	12	tatuz	6	tatız	20	tatuz	12

Table 3.6: Effects of ROUNDNESS HARMONY on wordlikeness (from Zimmer 1969)

in the pair *terüz-teruz*, both of which violate ROUNDNESS HARMONY. While this is little more than an anecdote, this may be indicative of a link between the two processes, and their exceptions, in the minds of native speakers.

Roundness harmony

Both experiments include 5 pairs which differ in the presence or absence of ROUNDNESS HARMONY. As shown in Table 3.6, there is an approximately 5-to-1 preference for harmonic nonce words, and as was the case above, no disharmonic member of any pair is preferred overall, across all speakers. Turkish speakers have a reliable preference for nonce words to conform to ROUNDNESS HARMONY ($\gamma = 0.680$, $p = 1.1\text{E}{-47}$).

Labial attraction

Both experiments include 5 pairs which either conform to LABIAL ATTRACTION and violate ROUNDNESS HARMONY, or vice versa; the preferences are shown in Table 3.7. There is

Experiment 1				Experiment 2			
aPu		aPı		aPu		aPı	
tamuz	3	tamız	16	pamuz	15	pamız	17
tafuz	3	tafız	17	tafuz	21	tafız	11
tavuz	9	tavız	4	mavuz	16	mavız	16
tapuz	7	tapız	9	tapuz	17	tapız	15
tabuz	5	tabız	12	tabuz	16	tabız	16

Table 3.7: Effects of LABIAL ATTRACTION on wordlikeness (from Zimmer 1969)

a small preference against LABIAL ATTRACTION (and therefore in favor of ROUNDNESS HARMONY, though this is non-significant ($\gamma = -0.043$, $p = 0.305$).

Speakers do not have the preferences predicted by LABIAL ATTRACTION. At the same time, they do not have a preference for ROUNDNESS HARMONY either, as one might expect. At most, it could be said that LABIAL ATTRACTION is sufficiently robust so as to suspend the effect of ROUNDNESS HARMONY, with which it is in competition.

3.2.3 Discussion

It has been shown that while LABIAL ATTRACTION is a highly reliable generalization about Turkish roots, it has at best a minimal effect on wordlikeness judgements. In contrast, harmony processes have a similar statistical profile, but have quite robust effects on wordlikeness. The most plausible explanation for this is that LABIAL ATTRACTION does not trigger alternations; indeed, it is counter-exemplified by the effects of ROUNDNESS HARMONY in suffix allomorphy. As is noted by Inkelas et al. (1997:412f.), the lexicon of Turkish will, under the assumptions here, remain as it is whether or not LABIAL ATTRACTION has a synchronic reality.[13] It might even be possible to suggest that it does not exist at all, though under such analysis the absence of a preference in the direction of ROUNDNESS

[13] Another possibility is that such a constraint could result in a lexical "trend" over time. Even if such a trend is observed, however, it is quite difficult to establish that it is the result of a synchronic constraint at every intermediate stage of the language.

Harmony is without ready explanation.

Itô and Mester (1995a,b) and Ní Chiosáin and Padgett (1993) claim that Labial Attraction holds only over the native vocabulary, and that the Turkish lexicon is "stratified". This would be a potential confound for Zimmer's experiment, since it is not implausible that speakers in Zimmer's study would treat nonce words much like loanwords; indeed, many wordlikeness studies include instructions to the participants to treat the stimuli much as if they were loanwords (e.g., Hay et al. 2004). However, Inkelas et al. (2001) find that foreign words are more—not less—likely to conform to Labial Attraction than native words. One possible explanation is that many of the languages in contact with Turkish (including English, Farsi, and French) lack the /ɯ/ (*ı*) phoneme needed to contribute exceptions to a hypothetical Labial Attraction.

It is not obvious that it is desirable to exclude Labial Attraction as a possible rule. Inkelas et al. (1997:394, fn. 2) suggest that Labial Attraction may have even induced alternations at one point in the history of Turkish. However, Becker et al. cite the apparent inertness of Labial Attraction as evidence that naturalness constrains phonotactic learning.

> This is clearly a complex and somewhat unnatural phonotactic, both in terms of the nonlocality of environment and the conjunction of features from two distinct triggers, and it is therefore a welcome result that not all speakers readily encoded it into a generalizable constraint. (Becker et al. 2011:118)

If this is correct, it should be possible to show that a more "natural" variant of Labial Attraction is in fact better reflected in wordlikeness judgements, assuming it too is statistically valid. Inkelas et al. make a similar observation.

> Vowel labialization following labials is not a synchronic alternation in Turkish, yet it (unlike Labial Attraction per se) *is* a statistically supported ten-

	...u	...ɪ	p-value
P...	371	71	6.98E−49
T...	811	922	

Table 3.8: TELL roots sorted according to LABIAL ATTRACTION′

dency worthy of further research. (Inkelas et al. 2001:196)

This proposal is formalized below as LABIAL ATTRACTION′.[14]

(16) LABIAL ATTRACTION′: $\begin{bmatrix} -\textsc{Cons} \\ +\textsc{Back} \\ +\textsc{High} \end{bmatrix} \longrightarrow \begin{bmatrix} +\textsc{Rnd} \end{bmatrix} \ / \ \begin{bmatrix} +\textsc{Lab} \\ +\textsc{Cons} \end{bmatrix}$ —

The environment is now strictly local. Rounding of high vowels after labial consonants is also acoustically natural, as both are distinguished by low first and second formants. Finally, the rounding of a high back vowel after a labial consonant is widely attested (e.g., Vaux 1993). Furthermore, as shown in Table 3.8, LABIAL ATTRACTION′ is even more statistically reliable than Lees's original formulation. This reformulation targets a superset of Lees's original rule, and therefore distinguishes the stimulus pairs in Table 3.7; i.e., any *aPı* sequence violates LABIAL ATTRACTION′ just as much as it violates the original formulation. Yet, it is not clear that either form of the generalization is strongly reflected in wordlikeness judgements.

[14] Zimmer (1969) proposes another variant of LABIAL ATTRACTION which ignores intermediate consonantal place but requires an *a* trigger in the preceding syllable. However, this is neither supported by lexical statistics or the results of his study.

3.3 Conclusions

Statistical reliability is neither necessary nor sufficient to predict what speakers know about possible and impossible words in their language. Further, phonotactic constraints may go unlearned whether or not they are "natural". In the case of Turkish vowel restrictions, at least, it is precisely those constraints which are derive from phonological processes—albeit processes with a considerable number of exceptions—which are most clearly reflected in psycholinguistic tasks. This should not be taken to imply that all phonotactic constraints inferred from the lexicon are illusory: for instance, Frisch and Zawaydeh (2001) present psycholinguistic support for co-occurrence restrictions in Arabic posited by Frisch et al. (2004 [1995]) on the basis of lexical data. As a general principle, though, it should be apparent that lexical statistics do not contribute reliable evidence for the theory of phonotactics; the linguist who has identified a statistical tendency in the lexicon has much more work to do before it can be identified with the synchronic grammar. This is even more serious when a strong relationship between phonotactic and phonological representations is assumed (as it is here): contrary to common practices (e.g., Mester 1988, Padgett 1991, 1992), lexical statistics can not be taken as provide principled evidence for the nature of phonological features or of other components of phonological grammar.

Chapter 4

Structural and accidental gaps in English syllable contact

It is indisputable that many phonotactic restrictions are easily described with reference to prosodic primitives. In some cases, prosodic factors constrain the system of contrast. For instance, Latin has contrastive vowel and consonant length (e.g., *os* 'bone' vs. *ōs* 'mouth', *anus* 'ring' vs. *annus* 'year'), but the latter contrast is suspended in codas preceded by a diphthong or long monophthong; a syllable may contain a long vowel, or be checked by the first half a geminate consonant, but not both. However, constraints on underlying representations may also involve references to non-contrastive prosodic structures such as the syllable (e.g., Hooper 1973, Kahn 1976).[1] For instance, as noted by Haugen (1956), numerous restrictions on word-medial consonant clusters have a unified statement in prosodic terms:

[1]See Blevins 1995 for the claim that syllable structure is universally non-contrastive, and Elfner 2006 for arguments that a putative counterexample derives from an underlying vowel length contrast.

(17) Medial Cluster Law:

A medial cluster can consist maximally of a well-formed medial coda and a well-formed medial onset

As an illustration of this tautology, consider languages like Yokuts, which forbid complex codas and complex onsets, and which enforce these restrictions in complex words via processes such as vowel epenthesis. Newman (1944:26f.) notes that this imposes an an upper bound on the size of medial clusters: no cluster of more than two consonants can be parsed into a simple coda and simple onset. In the case of Yokuts, it is certainly possible to state this restriction without reference to syllable structure, as *CCC, a constraint on trisyllabic clusters (e.g., Ettlinger 2008:92f., Zuraw 2003:820f.). However, the aforementioned constraints on complex onsets and codas find independent motivation from the total absence of initial and final clusters in Yokuts;[2] with these two restrictions in place, a further constraint on medial triconsonantal clusters is otiose.[3] While the medial cluster law is certainly consistent with the hypothesis that syllable structure may be present in underlying representations (e.g., Anderson 1974:255, Vaux 2003), this need not be the case under Stampean occultation. If an underlying medial consonant sequence appears on the surface, it satisfies the medial cluster law by definition. If, however, an underlying cluster is modified by consonant deletion, coalescence, or vowel epenthesis, then it need not consist of a licit medial onset and medial coda.

Pierrehumbert (1994) begins a study of English word-medial consonant clusters with a restatement of the medial cluster law:

[2] The tendency of consonants to pattern with word boundaries and morph junctures has been long been noted (e.g., Hill 1954, Lass 1971, Moulton 1947); Kahn (1976:24f.) takes this as evidence for the syllable.

[3] Côté (2000:31f.) alludes to another critique of constraints like *CCC, namely that this constraint requires “counting”: see Isac and Reiss 2008:64f. for further discussion of the comparative merits of “counting” and “grouping” analyses in phonology.

> That is, in the absence of additional provisos, any concatenation of a well-formed coda and a well-formed onset is predicted to be possible medially in a word. (Pierrehumbert 1994:168)

However, Pierrehumbert reports that the vast majority of the "possible" clusters (i.e., those which conform to the medial cluster law) are in fact unattested. So as to acount for this, Pierrehumbert presents "provisos" in the forms of static co-occurrence restrictions, unrelated to any phonological alternation in English.[4] This chapter will argue, however, the static constraints proposed by Pierrehumbert are unnecessary, and that the only restrictions on the inventory of medial clusters in English (beyond the medial cluster law) are those which derive from well-known phonological processes.

4.1 English syllable contact clusters

The aforementioned study by Pierrehumbert, as well as further investigations of this domain by Duanmu (2009: chap. 8) and Hammond (1999: chap. 3), argue for the necessity of admitting static phonotactic constraints, and illustrate proposals for the architecture of the phonotactic system. However, there are a number of reasons to reconsider the findings of these authors in light of the proposals made in previous chapters.

4.1.1 The role of phonological processes

Duanmu, Hammond, and Pierrehumbert do not generally take into account the effects of phonological processes which target medial clusters in English. As a consequence,

[4] Given the arguments presented in the previous chapter, that not all lexical tendencies have a synchronic basis, one may question the intuition that any gaps in the English cluster inventory (or those that go beyond the medial cluster law) must be accounted for by the synchronic grammar. In addition to some informal statistical evidence (of the sort problematized in the previous chapter), Pierrehumbert administers a word-likeness task to validate the static constraints she proposes. However, this experiment is of a quite informal nature and the results are given only a superficial analysis, so it is less than probative.

some of the static constraints these authors identify may be in fact the product of English morphophonemics and Stampean occultation. For instance, Pierrehumbert writes that "nasal-stop sequences agree in labiality" (175) and posits a static constraint to account for this fact. However, this generalization is merely a narrower form of a restriction deriving from a process of Nasal Place Assimilation (see §4.2.2). Other derived constraints are simply not mentioned; for instance, Pierrehumbert does not discuss the highly reliable tendency of obstruent-obstruent clusters to agree in voicing (see §4.2.2); while Hammond (1999) does allude to this restriction, it is dismissed in light of a few apparent counterexamples (though see §4.2.1 below). In contrast, this chapter attempts to evaluate derived and static constraints on an equal footing.

4.1.2 The role of sparsity

Pierrehumbert (1994) infers static constraints from near-exceptionless gaps in the lexicon, but little effort is made to show that the patterns of lexical underrepresentation are not due to chance. Consequently, it is possible to suggest that some of these gaps are accidental rather than structural in nature. This is made all the more likely given the tendency of segment and cluster frequency distributions to be highly skewed (e.g., Pande and Dhami 2010, Sigurd 1968, Tambovtsev and Martindale 2007, Weiss 1961) so that it is difficult to distinguish between structural and accidental gaps. Furthermore, Pierrehumbert considers only triconsonantal clusters, but medial clusters may be as short as two consonants, as in *a*[n.t]*ics*, or as long as four, as in *mi*[n.str]*el*, and no justification is given for ignoring clusters of other lengths. If there is any effect of this focus, it is presumably to produce further sparsity in the distribution observed.

4.1.3 The role of morphological segmentation

Many components of Pierrehumbert's study cannot be replicated. Pierrehumbert limits her study to words she judges to be "morphologically simple" and "reasonably familiar"; the author's sensations thereof are not replicable, nor are they available to other researchers in any form. It has been suggested (e.g., Labov 1975, Schütze 1996) that the sensations (as well as cognitive limitations) of concerned parties should not be granted evidential status in the first place, given the potential for implicit bias; Labov calls the *Experimenter Principle*.

It is not uncommon for analysts to propose otherwise-unmotivated morphological junctures simply to preserve phonological or phonotactic generalizations. This is done by Chomsky and Halle (1968), for instance, to simplify principles of English stress assignment. Similarly, Rice (2009:546) analyses many words in Slave as compounds simply because they contain consonant clusters that rarely occur in morph-internal contexts. Applied indiscriminately, however, this heuristic trivializes both morphological segmentation and phonotactic generalization. For these reasons, the wordlist used in this study is derived from a publicly available database, and no experimenter intuitions are used.

4.2 Evaluation

After constructing a sample of syllable contact clusters in English simplex words, this sample is used to evaluate the coverage of static and derived constraints.

4.2.1 Method

Materials and procedure

Following Duanmu (2009: chap. 8) and Hammond (1999: chap. 3), who also consider restrictions on English medial clusters, a wordlist is generated using the English portion of the CELEX database (Baayen et al. 1996). Only words marked in CELEX as "monomorphemic" are used, and all words labeled in CELEX as non-native are excluded.[5] These more-stringent criteria exclude many words labeled exceptions in the studies by Duanmu or Hammond in their studies. For instance, nearly all the exceptions to OBSTRUENT VOICE ASSIMILATION (see §4.2.2 below) noted by Hammond (1999:74) are excluded either as complex words (e.g., *jurisdiction*, *madcap*, *tadpole*, *scapegoat*, *magpie*) or non-native (e.g., *vodka*, *smorgasbord*).

In contrast to prior studies, these criteria also exclude words which consist of a Latinate prefix and a bound stem (e.g., *inspect*, *excrete*). While Pierrehumbert rejects this analysis as as unmotcontroversial, it in fact has extensive formal and experimental support. First, Latinate prefixes simplify the statement of many morphophonemic details in English. For instance, Aronoff (1976:11f.) observes that Latinate forms which share the same bound stem also share irregular allomorphs of that stem under derivation.

(18) Bound stem-specific allomorphy:

a.	adhere	adhesion
	cohere	cohesion
b.	conceive	conception
	perceive	perception

[5] The first criterion results in the exclusion of proper names, which have long been noted to push the bounds of native language phonotactics (e.g., Trubetzkoy 1958:254).

Aronoff takes this to be evidence that *adhere* and *cohere*, for instance, share a bound stem. There is also an interaction between Latinate prefixes on verbs and the complements they select. Latinate verbs do not generally allow ditransitive, verb participle, or adjectival resultative constructions, all of which are acceptable with similar Anglo-Saxon verbs (e.g., Gropen et al. 1989, Harley 2009).

(19) Restrictions on Latinate verbal complements:

a.	show him the painting	$\sim$	*exhibit him the painting
b.	drink himself stupid	$\sim$	*imbibe himself stupid
c.	break it off	$\sim$	*terminate it off

Lexical decision also provide evidence for the segmentation of Latinate prefixed forms. Taft and Forster (1975, 1976) and Taft et al. (1986) find that nonce words like **re-sert*, which appear to be composed of a prefix and a bound stem, take longer to reject that non-words which lack apparent morphological structure, such as **refant*. Bound stems also show frequency effects independent of whole word frequency (Taft 1979, Taft and Ardasinski 2006). Finally, Emmorey (1989) and Forster and Azuma (2000) report facilitative priming, thought to implicating morphological relatedness, between pairs like *permit-submit*, which appear to share a bound stem.

4.2.2 Results

Filtering the CELEX data according to the above criteria results in a list of 6,619 simplex words. The full set of clusters and their frequencies are listed in Appendix C. The CELEX transcriptions of these words are then syllabified and phonologized using a procedure described in Appendix B. In all, the sample contains 23 different medial coda and 40 different medial onsets. Of the 920 ($= 21 \times 40$) medial clusters that would result from free

	attested	unattested	saturation	p-value
conforming	25	91	22%	.106
violating	4	40	9%	

Table 4.1: Dorsal-labial cluster attestation in the lexical sample

combination of medial coda and medial onset, 174 (19%) are attested.

Static constraints

To account for the 81% of "possible" but unattested clusters, Pierrehumbert (1994) proposes three static constraints on English medial clusters.

Dorsal-labial clusters Pierrehumbert (1994:173) writes that "velar obstruents occurred only before coronals in the clusters studied, never before labials or other velars", while noting that absence of velar-velar clusters is expected due to a separate constraint on geminate clusters (see §4.2.2). However, biliteral velar-labial clusters are found in words such as *a*[k.m]*e*, *ru*[g.b]*y*, or *pi*[g.m]*ent*. Velar-labial clusters are somewhat less common than velar-coronal clusters (e.g., *ve*[k.t]*or*), but such underrepresentation is not unlikely to occur by chance according to the Fisher exact test (Table 4.1).

Coronal obstruent codas Pierrehumbert (1994:175) claims that "clusters with a coronal obstruent in the coda do not occur", but at the same time observes exceptions like *a*[nt.l]*er*, *ke*[s.tr]*el* and *oi*[nt.m]*ent*. In the CELEX sample (Table 4.2), coda coronal obstruent clusters are not significantly less likely to occur than non-coronal obstruent clusters (e.g., *re*[p.t]*ile*). While not shown in tabular form, the same is true if attention is restricted to triconsonantal clusters ($p = .129$).

ABA clusters Pierrehumbert (1994:176) observes a "lack of clusters with identical first and third elements", ignoring presence or absence of voicing. Despite the fact that there

	attested	unattested	saturation	*p*-value
conforming	56	304	15%	.430
violating	37	243	13%	

Table 4.2: Coda coronal obstruent cluster attestation in the lexical sample

	attested	unattested	saturation	*p*-value
conforming	47	512	8%	.250
violating	0	25	0%	

Table 4.3: ABA cluster attestation in the lexical sample

are no exceptions to this generalization, these ABA clusters are not significantly less common than any other triconsonantal and quadraconsonantal clusters (Table 4.3).

Summary There is no statistical support for any of Pierrehumbert's static constraints.

Derived constraints

In *SPE*, Chomsky and Halle (1968) describe three phonological processes which target medial consonant clusters. As will be shown, these three processes have a profound influence on the English cluster inventory.

Obstruent voice assimilation Voice assimilation alternations are evidenced by the nonsyllabic allomorphs of the regular past (e.g., *nap*[t]-*nab*[d]) and noun plural (e.g., *lap*[s]-*lab*[z]), which take the voicing specification of a preceding obstruent;[6] voice assimilation is also claimed to operate across prefix and compound junctures (Davidsen-Nielsen 1974).

[6]Underlying /-d, -z/ are assumed here (e.g., Anderson 1973, Baković 2005:284f., Basbøll 1972, Chomsky and Halle 1968:210, Hockett 1958:282, Pinker and Prince 1988:102, Shibatani 1972); alternative analyses are put forth by Bloomfield (1933:210f.), Borowsky (1986:135), Hoard and Sloat (1971), Kiparsky (1985), Lightner (1970), Luelsdorff (1969), Miner (1975), Nida (1948:426), and Zwicky (1975).

	attested	unattested	saturation	p-value
conforming	35	329	10%	.002
violating	11	305	3%	

Table 4.4: Obstruent voice assimilation cluster attestation in the lexical sample

(20) Obstruent Voice Assimilation:

$$\begin{bmatrix}-\textsc{Son}\end{bmatrix} \longrightarrow \begin{bmatrix}= \textsc{Voi}\end{bmatrix} \; / \; \_\_ \begin{bmatrix}= \textsc{Voi} \\ -\textsc{Son}\end{bmatrix}$$

Pierrehumbert (1994) does not discuss a constraint against adjacent obstruents disagreeing in voice. However, the vast majority of medial obstruents clusters in simplex words are either uniformly voiced, as in *hu*[z.b]*and*, or uniformly voiceless, as in or *rha*[p.s]*osdy* (Table 4.4). Hetero-voiced clusters, like those in *a*[b.s]*inth* and *a*[s.b]*estos*, are far rarer than would be expected from chance.

Nasal place assimilation Nasal Place Assimilation (e.g., Borowsky 1986:65f., *SPE*:85, Halle and Mohanan 1985:62) permits [ŋ] to be described as an allophone of /n/ (see §B.3), and furthermore accounts for allomorphy in certain Latinate prefixes.

(21) *im-*/*in-* allomorphy:

a.	polite	i[m.p]olite
	balance	i[m.b]alance
b.	tangible	i[n.t]angible
	decent	i[n.d]ecent

The rule is formalized below.

	attested	unattested	saturation	p-value
conforming	31	2	94%	3.4E−07
violating	11	22	33%	

Table 4.5: Nasal place assimilation cluster attestation in the lexical sample

(22) NASAL PLACE ASSIMILATION:

$$\begin{bmatrix} +\textsc{Nas} \end{bmatrix} \longrightarrow \begin{bmatrix} = \textsc{Lab} \\ = \textsc{Cor} \\ = \textsc{Dor} \end{bmatrix} \;/\; \rule{1em}{0.4pt} \begin{bmatrix} = \textsc{Lab} \\ = \textsc{Cor} \\ = \textsc{Dor} \\ -\textsc{Son} \end{bmatrix}$$

Virtually all clusters consisting of a nasal coda followed by a homorganic obstruent (e.g., *pi*[m.p]*le*, *sta*[n.z]*a*, *mo*[ŋ.k]*ey*) are attested (Table 4.5). As Pierrehumbert (1994:175) observes, heterorganic clusters, like those *pli*[m.s]*oll* or *scri*[m.ʃ]*aw*, do occur, but in this sample they are significantly more rare.

Degemination The final alternation found in English medial clusters is the simplification of geminates which is characteristic of "level I" morphology, and is found in the irregular /-t/ past tense (e.g., *bend*/*ben*[t], *build*/*buil*[t]), in *-ly* deadjectival derivatives (e.g., *norma*[l]*y*, cf. *calm*[l]*y*), and Latinate prefix allomorphy (Borowsky 1986:102, *SPE*:148). DEGEMINATION is formalized here as a rule deleting the first of two segments agreeing on all feature values (except for voice, possibly).

	attested	unattested	saturation	p-value
conforming	173	643	21%	1.2E−10
violating	0	104	0%	

Table 4.6: Degemination in the lexical sample

(23) Degemination:

$$\begin{bmatrix} =\textsc{Lab} \\ =\textsc{Cor} \\ =\textsc{Dor} \\ =\textsc{Son} \\ \ldots \end{bmatrix} \longrightarrow \emptyset \,/\, \_\_ \begin{bmatrix} =\textsc{Lab} \\ =\textsc{Cor} \\ =\textsc{Dor} \\ =\textsc{Son} \\ \ldots \end{bmatrix}$$

The sample contains no sequences of identical segments, or of identical segments differing only in voice, something highly unlikely to arise by chance (Table 4.6); Degemination and Stampean occultation provide a natural explanation for this gap. It is interesting to compare the absense of geminates to the constraint against "ABA" clusters proposed by Pierrehumbert in this regardl: both are exceptionless, but only the former imposes a lexical tendency unlikely to arise by chance.

Summary All three of the *SPE* rules targeting medial clusters have a robust effect in constraining the inventory of possible word-medial syllable clusters; possible clusters which are surface exceptions to these three rules are much less likely to be attested than those which conform to them.

Computational models

Current computational models of phonotactic knowledge can "rate" possible clusters, assigning a numerical wellformedness score to any input. These models can be applied to a task of predicting which clusters are and are not attested by transforming these numerical values into a categorical prediction of either attestation or non-attestation. This is accomplished here with a soft-margin support vector machine (Cortes and Vapnik 1995) with a linear kernel, which attempts to find a single optimal numerical value about which to split attested and unattested clusters. This classifier is not intended to correspond to any component of a cognitively plausible model of phonotactic learning: it simply represents an upper bound for predicting the cluster inventory from positive data.

The models are scored using a "leave-one-out" scheme, in which each observation (a cluster) is scored using a model trained on all other observations. Four metrics are used to evaluate model fit. *Accuracy* represents the probability that a cluster is correctly classified as attested or unattested. Two additional metrics break down accuracy into constituent parts; *precision* represents the probability that a cluster which is predicted to be unattested is in fact unattested, and *recall* is the probability that an unattested cluster is predicted as such. It is possible to increase precision at the expense of recall, by predicting non-attestation for a greater number of clusters, or to maximize recall at the expense of precision by predicting all clusters to be unattested. F_1, the harmonic mean of these two measures, is a standard metric for quantifying this tradeoff; any increase in either precision or recall will result in an increase in F_1. The results are summarized in Table 4.7.

Null baseline In a classification task, the simplest baseline is one which uniformly predicts the most common outcome. Since only 19% of clusters are attested, 81% accuracy can be achieved simply by predicting all clusters to be unattested.

	accuracy	precision	recall	F_1
Baseline	0.812	0.812	1.000	0.896
Expected frequency	0.835	0.837	0.960	0.894
Derived constraints	0.838	0.835	0.967	0.897
DC & EF	0.861	0.866	0.969	0.914
Hayes and Wilson (2008)	0.835	0.964	0.833	0.894

Table 4.7: Results for the cluster classification task

Expected frequency Pierrehumbert (1994) proposes that the well-formedness of a syllable contact cluster is proportional to the product of the independent probabilities of the coda and of the onset that make it up; this is the cluster's *expected frequency*. Pierrehumbert reports that this is an excellent predictor of which complex clusters occur and which do not. This model does not impose any constraints which span the syllable boundary; rather, it is a model of which clusters might be expected to represent accidental gaps in the sample. This produces a small but significant improvement in accuracy over the null baseline (sign test, $p = 4.5\text{E}{-}05$).

Derived constraints By hypothesis, Obstruent Voice Assimilation, Nasal Place Assimilation, and Degemination rule out a large number of possible clusters. In all, they target 316 out of 920 possible clusters (34%) for neutralization; of these clusters, only 11 (3%) occur in the sample. Together, these three processes define a simple classifier in which a cluster is predicted to be attested only if it would not be neutralized by one of these processes. This results in increased precision and a small but significant improvement in accuracy compared to the null baseline (sign test, $p = 4.0\text{E}{-}09$).

Expected frequency and derived constraints It is possible to combine into a single classifier the intuitions of the expected frequency and derived constraint models, the former accounting for accidental gaps and the latter for structural gaps imposed by neutralizing phonological processes. Simultaneously accounting for both sources of cluster inventory gaps, this model outperforms all others in accuracy and F_1.

Maximum entropy phonotactics Hayes and Wilson (2008) present a model using the principle of maximum entropy to weigh a large number of competing phonotactic constraints. In one sense, this is isomorphic to the expected frequency model in that the constraint discovery mechanism is sensitive to the expected frequency of clusters: it favors constraints which rule out clusters with high expected frequency (but which are unattested) over those which have low expected frequency. Also like the expected frequency model, alternations play no role and static constraints like those proposed by Pierrehumbert (1994) may be posited.

Since this model has numerous experimenter-defined parameters, a close replication of Hayes and Wilson's original study is attempted: both their implementation and phonological feature specifications are used here. Following Hayes and White (in press), dictionary entries are syllabified using the procedure described in Appendix B, and a novel feature [± Coda] is added to allow the model to distinguish coda and onset consonants. Also, following Hayes and Wilson, constraints are limited to those s panning as many as three segments and the suggested "accuracy schedule" is used. Since the maximum entropy model produces slightly different scores on each run, the worst-performing of 10 runs is reported here, following Hayes and Wilson. This model has the poorest recall of any model; compared to the derived constraints baseline, the constraints induced by the maximum entropy model are narrower. This is particularly clear regarding possible clusters with a nasal coda followed by a non-homorganic obstruent, like *[m.kl]: the vast majority of such clusters, which would be neutralized by Nasal Place Assimilation, are unattested, but many are erroneously assigned the highest possible score by the maximum entropy model.

Summary Expected frequency and derived constraints effectively account for accidental and structural gaps in the cluster inventory. The Hayes and Wilson (2008) computational model only provides an inferior approximation of the derived constraints.[7]

4.2.3 Discussion

What then is to be said of the 366 clusters which do not violate a derived constraint, but which are yet unattested, like [b.z] or [z.n]? Insofar as the "discovery procedure" used by linguists (e.g., Pierrehumbert) and computers (e.g., the Hayes and Wilson model) are based on principled phonological primitives, yet fail to find meaningful gaps, the absence of these clusters appears to be phonologically arbitrary. It appears that further gaps in the cluster inventory cannot be described in phonological, structural terms. The remainder of this chapter is concerned with the nature of these gaps.

The probability of accidental gaps

Good (1953) proposes a method for estimating the probability of accidental gaps in a sample distribution. This takes the form of an estimate of p_0, the probability of outcomes with have zero frequency in the sample.

$$p_0 = \frac{n_1}{N}$$

In prose, the value of p_0 is the ratio of clusters that occur exactly once in the sample (n_1) to the size of the sample (N). In the data here, $n_1 = 67$ and there are 997 clusters in all, so were it possible to extend the lexical sample, there is an approximately 7% chance that the next cluster would "fill in" what is now a gap. This alone indicates the non-trivial amount

[7] McGowan (in press) claims that the frequency of individual syllable contact clusters in English is proportional to the change of sonority over the syllable boundary. However, McGowan reports that the change in sonority in fact accounts for only a small portion of the variance in cluster frequency. A pilot study showed that sonority change was not useful as a predictor of cluster attestation.

of "missingness" and the high likelihood of accidental gaps in such a sample.

Simulating medial clusters

It can be shown that the large number of possible-but-unattested clusters is a logical necessity given the sparse distribution of codas and onsets. The rank and type frequency (i.e., frequency in the lexicon) of medial codas and onsets in the lexical sample described below are displayed in log-log space in Figure 4.1. Both codas and onsets show a linear relationship between log rank and log frequency characteristic of Zipfian distributions (see Appendix D). As a result, an enormous lexical sample would be needed to realize all clusters predicted by the medial cluster law, even if there were no constraints on the combination of medial codas and onsets in English. To illustrate this fact, a simulation is used to create new "samples" of the same size as the lexical sample used here. The following procedure is repeated so as to generate new "observations" for the simulated sample.

(24) Simulation procedure:

a. Sample a medial coda according to the observed probabilities
b. Sample a medial onset according to the observed probabilities
c. Apply the *SPE* rules to the cluster formed by their concatenation

This procedure corresponds to the assumption that the medial cluster law the derived constraints impose the only structural restrictions on the cluster inventory. Cluster frequencies in one simulated sample are shown in Figure 4.2; points represent simulated frequencies and the line actual cluster frequencies. As can be seen, the observed and simulated frequencies are quite similar (i.e., $R^2 = 0.712$; $p = 4.5\text{E}{-}05$). To summarize, the sparse cluster inventory, which Pierrehumbert takes as evidence for static constraints on syllable contact clusters, would result even if said static constraints do not exist.

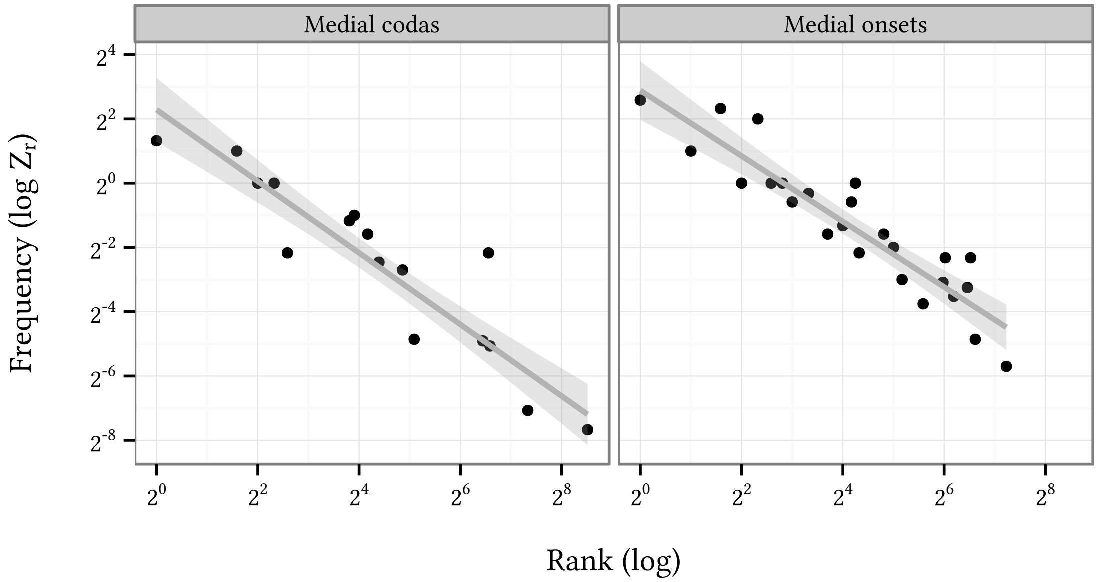


Figure 4.1: Medial coda and medial onset type frequencies in the lexical sample show a Zipfian distribution; frequencies have been smoothed using the Z_r transform (see Appendix D)

4.3 Conclusions

The foregoing results suggest that the only structural constraints on English syllable contact are derived from the phonological system: this study finds no evidence for static constraints. The many other unattested clusters can only be understood as accidental gaps which are a consequence of the finite nature of the English lexicon.

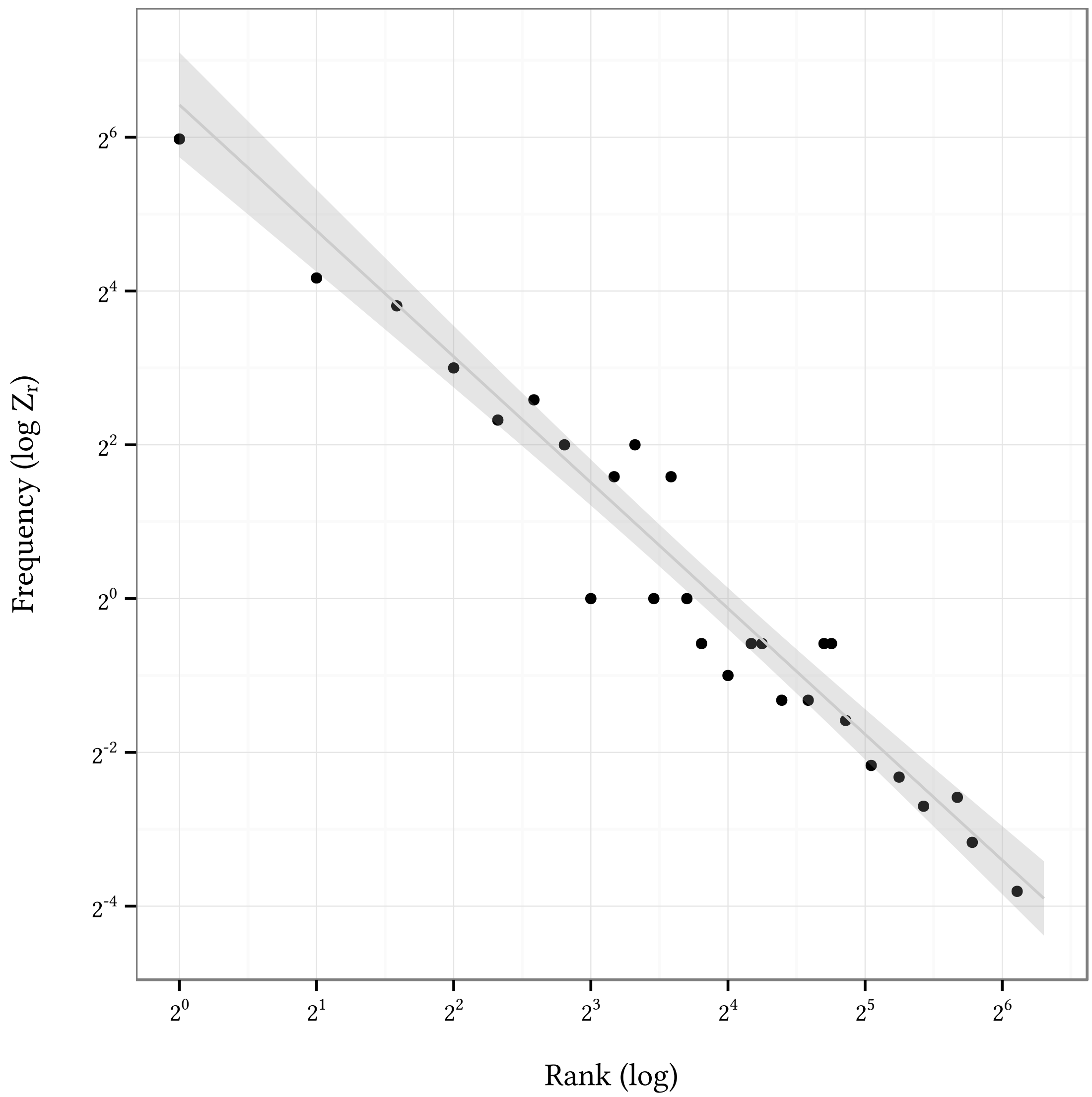


Figure 4.2: A simulated cluster inventory closely matches the observed cluster frequencies (represented by the smoothing line); frequencies have been smoothed using the Z_r transform (see Appendix D)

Chapter 5

Conclusions

5.1 Summary of dissertation and future directions

5.1.1 Chapter 1

Chapter 1 began with the claim that phonotactic theorists have not been sufficiently explicit about what they take as evidence for phonotactic knowledge. A core challenge in phonotactic theory is to determine which of the countless competing generalizations are internalized by speakers. Wordlikeness judgements, production and recognition experiments, loanword adaptation, alternative phonologies, and lexical statistics all are potential sources of phonotactic evidence, though there are important caveats associated with each source of evidence.

Some questions remain open regarding the relationship between loanword adaptation and phonotactic generalizations. A crucial question is what determines which phonotactic generalizations will be implicated in loanword adaptation. As indicated in the chapter, a naïve hypothesis that those generalizations which are phonologically "derived" cannot be maintained in light of evidence put forth by Peperkamp (2005). On the other hand, the model of loanword adapation advocated by Peperkamp is has little to say regarding how phonotactic knowledge and how it might influence non-native speech perception.

Shibatani's classic argument for the insufficiency of morpheme structure constraints was then presented. It was suggested that much of the distinction between *sequence structure constraints* and phonological processes is moot in light of *Stampean occultation.* Other sequence structure constraints may be identified with language-specific restrictions on the prosodic inventory. It therefore remains to be seen whether there are "static" constraints: those which are neither derived by phonological process or phonotactic parsing. According to one hypothesis, static constraints are inferred from statistical trends in the lexicon; this account makes predictions about gradient grammaticality and the independence of phonology and phonotactics which are taken up in in Chapters 2 and 3. Another account identifies phonotactic knowledge with inviolate markedness constraints and is considered in Chapter 4.

Finally, it was argued that whatever the merits of these accounts of static constraints, there is also some potential value to the null hypothesis that static constraints do not exist at all: such a principle would allow for a resolution of certain intransigent debates in phonology, such as the proper analysis of the Sanskrit "diaspirates". Other cases of this type, where one analysis derives a phonotactic constraint and another makes no such prediction, surely could be adduced and investigated experimentally.

5.1.2 Chapter 2

Chapter 2 is concerned with the claim that phonotactic wellformedness is not an "all-or-nothing" affair. This claim is first placed in a historical context. Then, three arguments are presented for *a priori* skepticism regarding the claims of gradience in wordlikeness judgements. First, a model of gradient grammatical judgements seems to require extraordinary abilities for which there is little evidence, whereas reporting categorical judgements calls only on well-established capacities. Secondly, intermediate ratings in wordlikeness tasks, long taken as evidence for gradient models, are in fact present in any rating task which permits intermediate ratings, even when this is nonsensical. Finally, there have been no serious attempts to compare current gradient models to simple categorical baselines.

The second half of this chapter is concerned with providing an evaluation which implements this comparison. A simple categorical baseline, as well as another baseline model measuring similarity to existing items, reliably outperform state-of-the-art models of gradient wellformedness. Once the categorical baseline effect is controlled for, gradient models are largely uncorrelated with the ratings. It is concluded that there is no evidence that intermediate ratings in wordlikeness tasks are the result of gradient grammaticality.

This chapter suggests many directions for further research. First, it throws down a gauntlet to computational modelers who are proponents of gradient grammaticality, and the evaluation therein must be kept up to date when the challenge is answered. Furthermore, the categorical baseline presented in this model is intentionally quite primitive, and some of the proposed improvements discussed in the chapter are worthy of implementation and evaluation. The evaluation in this chapter is based on a small amount of published wordlikeness data. It should be clear that the quality and quantity of this data is quite limited, but that it can be cheaply gathered. On analogy with undertakings like the English Lexicon Project (Balota et al. 2007), a publicly available database of lexical decision latencies, a carefully designed "English Wordlikeness Project" would be of great value to other phonotactic theorists, and would provide useful norming data for many other psycholinguistic tasks.

5.1.3 Chapter 3

The lexical/statistical model of phonotactic knowledge is the subject of Chapter 3. While it has long been speculated that statistical criteria could be used to determine which phonotactic generalizations are internalized and which are ignored, it is argued that lexical statistics are neither necessary nor sufficient to determine this fact. On the contray, there are many static phonotactic constraints which are statistically reliable but synchronically inert: case studies include Latin rhotacism and restrictions on the English vowel system. A longer case study considers constraints on Turkish vowels. A host of evidence, phonological, statistical, and "external", suggests that vowel harmony is active throughout the language, despite a significant amount of lexical exceptions. Considerably less clear is the status of the static constraint known as Labial Attraction; at best, it counteracts the effect of harmony.

Future work should further consider instances of "statistically reliable/synchronically inert" restrictions. Two obvious targets would be Mid-Atlantic short *a* and Turkish Labial Attraction, which would benefit from further careful psycholinguistic investigation.

5.1.4 Chapter 4

Chapter 4 begins with the observation that many phonotactic constraints can be described with reference to prosodic primitives, and that this is in fact consistent with Stampean occultation. Pierrehumbert, however, has argued that not all constraints on English syllable contact clusters can be attributed to prosodic restrictions, and that various static constraints are necessary. However, Pierrehumbert's study suffers from several methodological flaws. Principle among them is the failure to distinguish between derived and static constraints. A restudy shows that the former impose robust restrictions on the English syllable contact cluster inventory, whereas the static constraints proposed by Pierrehumbert lack statistical validity.

Many of the clusters which appear to be prosodically well-formed are not attested, but state-of-the-art computational models are not able to detect any systematicity in the patterns of missingness. Consequently, it is argued that many unattested clusters must be viewed as accidentap gaps, and that such is to be expected given the statistical properties of distributions of codas and onsets that make up medial clusters.

Further psycholinguistic research could be used to further evaluate the claim made in this chapter, that many unattested syllable contact clusters are accidental gaps. The analysis could profitably be extended to medial clusters in other languages.

5.2 A note on architectural matters

A major finding of this dissertation is that the "lexical/statistical" theory of phonotactics is not sufficiently in evidence. Consequently, it is premature to assume the existence of a grammatical module which performs the task of computing phonotactic wellformedness.

The alternative argued for here is that phonotactic knowledge is derived from familiar components of the ("narrow") phonological grammar: phonological processes and prosodic structures like syllables, feet, etc. These too must be compiled into a module capable of recognizing "possible" and "impossible' words, and work will need to be done to clarify how this is accomplished. But, unlike the "lexical/statistical" theory, this knowledge comes online only in response to phonological acquisition. As will be shown in what remains, this is in fact the view from research into phonological development.

5.3 Acquisition of phonotactic knowledge

As a sort of final summary of the theories under evaluation in this dissertation, consider the evidence provided by the order in which phonotactic, phonological, and lexical knowledge is acquired. Hayes (2004) argues that phonotactic learning occurs before lexical or phonological acquisition, and it therefore must be independent of other types of grammatical knowledge.

At 9 months of age, several studies imply that typically-develping infants have internalized langauge-specific preferences for native language phonotactics (Friederici and Wessels 1993, Jusczyk et al. 1994). A particularly telling study is performed by Jusczyk et al. (1993), who find that Dutch and English infants prefer to listen to lists of (likely unfamiliar) words in their native languages over those in English or Dutch, respectively. Crucially, these studies do not find the same preferences at earlier ages. Keeping in mind the usual caveats about a strong interpretation of negative results, this actually suggests that phonotactic acquisition occurs relatively late.

Some non-trivial amount of lexical learning occurs at an earlier age, for instance. Typically-developing infants know their names and the names of their caretakers as early as 4 months of age (Bortfeld et al. 2005, Mandel et al. 1995, Tincoff and Jusczyk 1999). As early as 6 months of age, infants know the visual referents of familiar words presented auditorily (Bergelson and Swingley 2012). At 7.5 months, phonological representations are sufficiently detailed to allow infants to discriminate between words like *cup* and mispronunciations like **tup* (Jusczyk and Aslin 1995). By 8 months, infants are able to locate both familiar and novel words in continuous speech (Jusczyk and Hohne 1997, Seidl and Johnson 2006). It is approximately at this time that the ability to discriminate non-native phonetic contrasts for vowels and consonants begins to decline (Best 1994, Polka and Werker 1994, Werker et al. 1981, Werker and Tees 1984, Werker and Lalonde 1988).

While very few studies have investigated young infants' knowledge of phonological alternations, this is the subject of a fascinating study by White et al. (2008). Simplifying somewhat, the experimenters expose 8.5-month-old infants to an artificial language in which fricative voicing is contrastive, but voiced and voiceless variants of plosives are in complementary distribution, appearing only after vowels (*na-bevi*) and after voiceless consonants (*rot-pevi*), respectively. After familiarization, infants prefer to listen to nonce words preserving this complementary distribution of stops over nonce words which disrupt this distribution (e.g., *na-poli*, *rot-boli*), suggesting that the infants have extracted an allophonic generalization concerning plosive voicing. This too occurs before the earliest evidence for a more traditional type of phonotactic learning.

It is not until much later that phonotactic knowledge can be investigated using productive language skills; the seminal study by Smith (1973), for instance, begins when the target child, Amahl, is more than 2 years of age. Given the familiar assymetry between comprehension and production (the former leading), it seems unlikely that much can be learned about the chronology of acquisition from this “lagging indicator”.

While there are many gaps in current understanding which merit future research, there is no reason to think that phonotactic knowledge is acquired before a considerable amount of lexical acquisition has occurred, or before children are capable of extracting phonological alternations and applying them to novel words. This observation that infants’ phonotactic knowledge comes online only after highly specific phonological entries, subsyllabic representations, and alternation learning are available, is precisely what is predicted by the null hypothesis that phonotactic knowledge is derived from phonological processes and prosodic restrictions.

Appendix A

English wordlikeness ratings

A.1 Albright (2007)

	lexical density	−log p (bigram)	−log p (MaxEnt)	gross status	rating (7-point)
P L IY1 N	13	13.585	0.000	valid	5.32
B L AA1 D	13	17.609	0.000	valid	5.13
P L IY1 K	11	14.200	0.000	valid	5.06
P L EY1 K	14	15.576	0.000	valid	4.94
P R AH1 N JH	3	16.546	0.000	valid	4.94
B L UW1 T	14	15.692	0.000	valid	4.84
P L IH1 M	5	15.126	0.000	valid	4.71
B L EH1 M P	1	19.447	0.000	valid	4.69
B L AH1 S	14	13.806	0.000	valid	4.67
B L AE1 D	15	16.259	0.000	valid	4.65
B L IH1 G	4	16.347	0.000	valid	4.58
P R EH1 S P	4	17.214	0.000	invalid	4.50
B R EH1 N TH	4	21.255	0.000	valid	4.11
P R AH1 P T	4	18.487	0.000	valid	4.07
B R EH1 L TH	2	23.014	0.000	valid	3.14
P W IH1 S T	4	21.499	0.000	valid	2.94
B W AH1 D	2	20.596	0.000	valid	2.94
B W AA1 D	3	21.329	0.000	valid	2.94
P W AE1 D	2	24.103	0.000	valid	2.89
P W AH1 S	4	20.684	0.000	valid	2.61
P W EH1 T	6	20.998	0.000	valid	2.53
P T IY1 N	4	15.440	6.762	invalid	2.44
B W AE1 D	2	23.365	0.000	valid	2.41
B N IY1 N	2	21.180	7.296	invalid	2.39
P W AH1 D Z	0	25.210	0.000	valid	2.17
P N IY1 N	2	21.181	6.019	invalid	2.16
B N AH1 S	6	19.732	7.296	invalid	2.06
P N EH1 P	2	23.587	6.019	invalid	2.00
B N AA1 D	2	24.066	7.296	invalid	2.00
B Z IY1 N	1	16.896	19.097	invalid	2.00

P T AH1 S	3	14.169	6.762	invalid	1.94
P T EH1 P	3	17.228	6.762	invalid	1.86
B Z AH1 S	2	15.237	19.097	invalid	1.81
P N IY1 K	2	21.796	6.019	invalid	1.76
B D IY1 K	2	18.773	13.131	invalid	1.72
B D UW1 T	3	19.781	13.131	invalid	1.71
B D AH1 S	4	17.041	13.131	invalid	1.71
P T AE1 D	3	17.622	6.762	invalid	1.67
B Z AA1 D	1	20.151	19.097	invalid	1.63
B Z AY1 K	1	19.118	19.097	invalid	1.28

A.2 Albright and Hayes (2003), norming study

	lexical density	−log p (bigram)	−log p (MaxEnt)	gross status	rating (7-point)
S L EY1 M	15	17.469	0.000	valid	5.84
W IH1 S	34	11.208	0.000	valid	5.84
P IH1 N T	26	13.046	0.000	valid	5.67
P AE1 NG K	18	13.723	0.000	valid	5.63
S T IH1 P	18	12.599	0.000	valid	5.53
M IH1 P	33	12.345	0.000	valid	5.47
S T AY1 R	11	15.118	0.000	valid	5.47
M ER1 N	34	12.872	0.000	valid	5.42
P L EY1 K	14	15.576	0.000	valid	5.39
S N EH1 L	10	18.582	0.000	valid	5.32
S T IH1 N	18	10.899	0.000	valid	5.28
R AE1 S K	11	15.544	0.000	valid	5.21
T R IH1 S K	5	17.980	0.000	valid	5.21
S P AE1 K	17	14.205	0.000	valid	5.16
D EY1 P	22	14.193	0.000	valid	5.11
G EH1 R	25	13.044	0.000	valid	5.11
G L IH1 T	14	16.830	0.000	valid	5.11
S K EH1 L	16	16.356	0.000	valid	5.11
SH ER1 N	23	15.913	0.000	valid	5.11
T AA1 R K	18	17.702	0.000	valid	5.11
CH EY1 K	28	15.023	0.000	valid	5.05
G L IY1 D	14	16.118	0.000	valid	5.05
G R AY1 N T	4	17.626	0.000	valid	5.00
P R IY1 K	11	13.396	0.000	valid	5.00
SH IH1 L K	8	21.270	0.000	valid	4.89
D AY1 Z	39	12.730	0.000	valid	4.84
N EY1 S	23	14.952	0.000	valid	4.84
T AH1 NG K	18	15.046	0.000	valid	4.84
S K W IH1 L	6	18.210	0.000	valid	4.83
L AH1 M	35	11.569	0.000	valid	4.79
P AH1 M	30	11.121	0.000	valid	4.79
S P L IH1 NG	14	15.573	0.000	valid	4.72
G R EH1 L	3	14.624	0.000	valid	4.63
T EH1 SH	12	14.517	0.000	valid	4.63
T IY1 P	32	12.980	0.000	valid	4.63
B AY1 Z	35	12.821	0.000	valid	4.58
G L IH1 P	11	17.377	0.000	valid	4.53
CH AY1 N D	18	17.747	0.000	valid	4.37
P L IH1 M	5	15.126	0.000	valid	4.37
G UW1 D	29	15.448	0.000	valid	4.32

	B	L	EY1	F		6	19.485	0.000	valid	4.21
		G	EH1	Z		17	16.466	0.000	valid	4.21
	D	R	IH1	T		8	15.563	0.000	valid	4.16
	F	L	IY1	P		10	15.292	0.000	valid	4.16
		Z	EY1			23	15.208	0.000	valid	4.16
S	K	R	AY1	D		5	18.722	0.000	valid	4.11
		K	IH1	V		16	12.591	0.000	valid	4.05
	F	L	EH1	T		17	16.490	0.000	valid	4.00
		N	OW1	L	D	19	19.101	0.000	valid	4.00
	S	K	IH1	K		13	14.628	0.000	valid	4.00
	B	R	EH1	JH		7	17.318	0.000	valid	3.95
	K	W	IY1	D		10	16.039	0.000	valid	3.95
	S	K	OY1	L		9	19.350	0.000	valid	3.89
	D	R	AY1	S		12	17.758	0.000	valid	3.84
	F	L	IH1	JH		8	17.312	0.000	valid	3.79
	B	L	IH1	G		4	16.347	0.000	valid	3.53
		Z	EY1	P	S	7	24.825	0.000	valid	3.47
		CH	UW1	L		17	14.492	0.000	valid	3.42
		SH	AY1	N	T	8	18.503	0.000	valid	3.42
	SH	R	UH1	K	S	5	26.733	0.000	valid	3.32
	G	W	EH1	N	JH	0	22.722	0.000	valid	3.32
		N	AH1	NG		19	15.754	0.000	valid	3.28
S	K	W	AA1	L	K	1	25.752	0.000	invalid	3.26
	T	W	UW1			5	17.918	0.000	valid	3.17
	S	M	AH1	M		8	14.940	0.000	valid	3.05
	S	N	OY1	K	S	4	32.283	4.136	invalid	3.00
	S	F	UW1	N	D	1	23.241	3.507	valid	2.94
	P	W	IH1	P		4	20.928	0.000	valid	2.89
		R	AY1	N	T	8	14.412	0.000	valid	2.89
S	K	L	UW1	N	D	0	22.661	0.000	invalid	2.83
	S	M	IY1	R	G	0	27.601	0.000	invalid	2.79
	F	R	IH1	L	G	3	24.299	0.000	invalid	2.68
	SH	W	UW1	JH		0	28.270	0.000	invalid	2.68
	TH	R	OY1	K	S	0	32.485	4.136	invalid	2.68
	T	R	IH1	L	B	4	22.097	0.000	invalid	2.63
	K	R	IH1	L	G	1	23.719	0.000	invalid	2.58
	S	M	EH1	R	G	0	22.473	0.000	invalid	2.58
	TH	W	IY1	K	S	2	23.984	0.000	invalid	2.53
	S	M	EH1	R	F	0	23.136	0.000	invalid	2.47
	S	M	IY1	L	TH	0	26.377	0.000	invalid	2.47
	P	L	OW1	M	F	0	23.336	0.000	invalid	2.42
	P	L	OW1	N	TH	0	22.805	0.000	invalid	2.26
		TH	EY1	P	T	4	23.380	0.000	valid	2.26
	S	M	IY1	N	TH	0	25.043	0.000	valid	2.06
S	P	R	AA1	R	F	0	24.031	0.000	valid	2.05
	P	W	AH1	JH		0	23.205	0.000	valid	1.74

A.3 Scholes (1966), experiment 5

				lexical density	$-\log p$ (bigram)	$-\log p$ (MaxEnt)	gross status	rating (binary)
G	R	AH1	N	18	11.799	0.000	valid	33
K	R	AH1	N	21	11.597	0.000	valid	33
S	T	IH1	N	18	10.899	0.000	valid	33
S	M	AE1	T	13	16.654	0.000	valid	32

P	R	AH1	N	11	10.845	0.000	valid	32
S	L	ER1	K	12	17.846	0.000	valid	31
F	L	ER1	K	11	17.456	0.000	valid	31
B	L	AH1	NG	8	17.156	0.000	valid	31
D	R	AH1	NG	7	17.753	0.000	valid	31
T	R	AH1	N	12	11.975	0.000	valid	31
F	R	AH1	N	12	12.177	0.000	valid	29
S	P	EY1	L	16	15.851	0.000	valid	29
S	N	EH1	T	7	19.384	0.000	valid	28
P	L	AH1	NG	11	16.960	0.000	valid	28
SH	R	AH1	K	8	19.734	0.000	valid	27
G	L	AH1	NG	9	18.990	0.000	valid	27
M	R	AH1	NG	1	22.888	3.365	invalid	27
SH	L	ER1	K	4	23.711	0.000	invalid	22
S	K	IY1	P	15	16.845	0.000	valid	20
V	R	AH1	N	4	17.087	0.000	invalid	19
S	R	AH1	N	9	16.626	0.000	invalid	14
V	L	ER1	K	2	21.777	0.000	invalid	14
M	L	AH1	NG	4	21.300	10.164	invalid	13
SH	T	IH1	N	3	17.106	0.000	invalid	13
F	P	EY1	L	4	24.250	3.685	invalid	13
ZH	R	AH1	N	4	28.305	4.042	invalid	11
F	SH	IH1	P	2	22.640	10.198	invalid	11
SH	N	EH1	T	2	24.044	0.000	valid	10
F	T	IH1	N	2	14.767	3.685	invalid	10
Z	R	AH1	N	5	21.556	4.042	invalid	8
N	R	AH1	N	5	18.588	3.365	invalid	8
SH	M	AE1	T	1	20.389	0.000	valid	7
S	F	IY1	D	7	18.656	3.701	valid	7
Z	L	ER1	K	2	24.578	5.678	invalid	6
Z	T	IH1	N	1	23.600	5.678	invalid	6
F	S	EH1	T	4	19.079	10.198	invalid	6
V	Z	IH1	P	1	17.401	19.601	invalid	6
V	Z	AH1	T	1	15.806	19.601	invalid	6
ZH	L	ER1	K	2	33.442	5.678	invalid	5
SH	F	IY1	D	1	23.258	3.701	invalid	5
Z	N	AE1	T	1	25.541	5.678	invalid	4
F	N	EH1	T	2	23.969	3.315	invalid	3
F	K	IY1	P	1	23.905	3.685	invalid	3
V	T	IH1	N	2	22.639	3.685	invalid	3
Z	V	IY1	L	2	26.018	15.023	invalid	3
Z	M	AE1	T	1	21.983	5.678	invalid	2
ZH	M	AE1	T	1	26.800	5.678	invalid	2
F	M	AE1	T	4	21.800	3.315	invalid	2
SH	P	EY1	L	2	26.172	0.000	invalid	2
V	M	AE1	T	2	20.388	3.315	invalid	1
V	N	EH1	T	2	24.017	3.315	invalid	1
SH	K	IY1	P	2	26.976	0.000	invalid	1
Z	P	EY1	L	1	25.421	5.678	invalid	1
ZH	P	EY1	L	1	32.906	5.678	invalid	1
ZH	T	IH1	N	1	29.763	5.678	invalid	1
ZH	K	IY1	P	1	33.710	5.678	invalid	1
ZH	N	EH1	T	1	33.775	5.678	invalid	0
Z	K	IY1	P	1	27.547	5.678	invalid	0
V	P	EY1	L	2	25.782	3.685	invalid	0
V	K	IY1	P	1	26.586	3.685	invalid	0
ZH	V	IY1	L	1	32.181	15.023	invalid	0

Appendix B

English syllabification

For every entry in the CELEX database, there is a corresponding broad syllabified transcription of the word in a Received Pronunciation accent. This appendix describes an automated procedure used to process these transcripts and to separate medial clusters from their flanking nuclei, parsing the resulting sequences into coda and onset, and reversing allophonic processes targeting medial clusters.

While the segmental content of these transcriptions is precise, the CELEX syllabifications are unsystematic. Given the absence of contrastive syllabification in English (if not all languages: see Blevins 1995:221, Elfner 2006), any sequence of a medial consonant cluster and its flanking nuclei should receive the same syllabification in all words in which it occurs. This is not always the case with the CELEX transcriptions, however. For instance, the sequence [ɪstɹɪ] receives a different parse in *chemistry* [ˈkɛ.mɪ.stɹɪ] and *ministry* [ˈmɪ.nɪs.tɹɪ].[1] Consequently, these syllabifications are not used here.

[1]Note that word-final *y* is usually lax in Received Pronunciation (Wells 1982:II.294).

B.1 Ambiguous segments

The syllabification procedure begins by separating sequences of vocalic and consonantal segments. In English, *r* and onglides pattern with consonants or with vowels depending on the context in which they occur. The heuristic adopted here is that ambiguous segments which impose restrictions on adjacent vowels are themselves vocalic, and those which impose restrictions on adjacent consonants are consonantal.

Initially, between two vowels, or finally, *r* is consonantal. Before another consonant, however, *r* has been lost in Received Pronunciation. Even in *r*-ful dialects, though, post-vocalic non-onset *r* patterns with vowels, not coda consonants. Before non-onset *r* many vowel contrasts are suspended (e.g., Fudge 1969:269f., Harris 1994:255): compare American English *fern*/*fir*/*fur* to *pet*/*pit*/*putt*. In this position, *r* is the only consonant which permits variable glottalization of a following /t/ in *r*-ful British dialects (Harris 1994:258), and the only consonant which does not trigger variable deletion of a following word-final /t, d/ in American dialects (Guy 1980:8). This is shown in (25–26) below.

(25) /t/-Glottalization in *r*-ful British dialects:

a.	des[ɚt]	∼	des[ɚʔ]
	c[ɚt]ain	∼	c[ɚʔ]ain
b.	fi[st]	∼	*fi[sʔ]
	mi[st]er	∼	*mi[sʔ]er

(26) /t, d/-Deletion in American English:

a.	be[lt]	∼	be[l]
	me[nd]	∼	me[n]
b.	sk[ɚt]	∼	*sk[ɚ]
	th[ɚd]	∼	*th[ɚ]

Following Pierrehumbert (1994), pre-consonantal *r* is assigned to the preceding nucleus.

The front onglide is assigned to onset position when initial or preceded by a single consonant, as in [j]*arn* or *ju*[n.j]*or*. When the glide is preceded by two or more consonants, it is assigned to the nucleus. There is considerable evidence in support of this assumption. When [j] is assigned to the onset, it may be followed by any vowel (Borowsky 1986:276), but when it is nuclear, the following vowel is always [u], suggesting a nuclear affiliation (Harris 1994:61f., Hayes 1980:232). Clements and Keyser (1983:42) note that [j] is the only consonant which can follow onset /m/ and /v/: [mj]*use*, [vj]*iew*. Finally, [ju] sequences in words such as *spew* behave as a unit in language games (Davis and Hammond 1995, Nevins and Vaux 2003) and speech errors (Shattuck-Hufnagel 1986:130).[2]

The phonotactic properties of the back onglide [w] are quite different than those of the front onglide, and it is consequently assigned to the onset portion of medial clusters. Whereas [j] shows only limited selectivity for preceding tautosyllabic consonants (Kaye 1996), [w] only rarely occurs after onset consonants other than [k] (e.g., *tran*[kw]*il*), and never after tautosyllabic labials in the native vocabulary. Whereas [kj] is always followed by [u], [kw] may precede nearly any vowel (Davis and Hammond 1995:161).

B.2 Parsing medial consonant clusters

Medial consonant clusters are segmented into coda and onset using a heuristic version of the principle of onset maximization (e.g., Kahn 1976:42f., Kuryłowicz 1948, Pulgram 1970:75, Selkirk 1982:358f.) which favors parses of word-medial clusters in which as much of the cluster as possible is assigned to the onset. A medial onset is defined to be "possible" simply if it occurs word-initially (according to the rules defined above). As an example, the medial clusters in words such as *neu*[.tɹ]*on* or *bi*[.stɹ]*o* also occur in word-initial position (e.g., [tɹ]*ain*, [stɹ]*ike*), so the entire cluster is assigned ot the onset. In contrast, the cluster

[2] The glide is also assumed to be present in underlying representation (e.g., Anderson 1988, Borowsky 1986:278) rather than inserted by rule (e.g., *SPE*:196, Halle and Mohanan 1985:89, McMahon 1990:217) since presence or absense of the glide is contrastive (e.g., *booty*/*beauty*, *coot*/*cute*).

in *mi*[n.stɹ]*el* is not found word-initially; the maximal onset here is [stɹ] and the remaining [n] is assigned to the preceding coda.

In English, when a medial consonant cluster is preceded by a stressed lax vowel, as *wh*[ɪs.p]*er*, *v*[ɛs.t]*ige*, or *m*[ʌs.k]*et*, the first consonant of the cluster checks the lax vowel (Hammond 1997:3, Treiman and Zukowski 1990). As Harris (1994:55) notes, however, when the medial cluster is also a valid onset, as in *whi*[s.p]*er*, *ve*[s.ti]*ge*, and *mu*[s.k]*et*, onset maximization will incorrectly assign the entire cluster to the onset and leave the lax vowel unchecked. For this reason, onset maximization parses are modified to assign the first consonant of a complex medial consonant cluster to the coda before a stressed lax vowel (Pulgram 1970:48).

B.3 Phonologization

Following Pierrehumbert (1994), the traditional analysis of affricates as single segment (e.g., *SPE*:321f., Jakobson et al. 1961:24) rather than sequences of a stop and fricative (e.g., Hualde 1988, Lombardi 1990) is adopted here. In many languages, affricates pattern with simple onsets; for instance, Classical Nahua bans true onset clusters but permits the affricate series [ts, tʃ, tɬ] (Launey 2011:9). Other languages, such as Polish, distinguish affricates and stop-fricative sequences (Brooks 1965), providing further evidence that "true" affricates are represented as single segments (or single timing units), and in contrast with stop-fricative clusters (Clements and Keyser 1983:34f.).

In English, [ŋ] has been analyzed as a pure allophone of /n/ before underlying /k, g/ (with later deletion of /g/ in some contexts; Borowsky 1986:65f., *SPE*:85, Halle and Mohanan 1985:62), or as a phoneme in its own right (e.g., Jusczyk et al. 2002, Sapir 1925). Onset [ŋ] is totally absent in onset position, where it cannot be followed by a /k, g/ needed to derive the velar allophone, a fact predicted only by the former account, and English speakers have considerable difficulty producing initial [ŋ] (Rusaw and Cole 2009). Following Pierrehumbert (1994), the allophonic analysis is assumed here. When followed by /k, ɡ/, [ŋ] is mapped to /n/. When not followed by a velar stop (i.e., finally), [ŋ] is analyzed as underlying /nɡ/.

Appendix C

English syllable contact clusters

cluster	cluster frequency	coda frequency	onset frequency	$-\log p$ (MaxEnt)	rule exceptions
N . D	79	365	88	0.000	
N . T	69	365	150	0.000	
M . P	60	161	73	0.000	
M . B	53	161	65	0.000	
N . G	44	365	48	0.000	
S . T	37	87	150	0.000	
K . S	34	94	92	0.000	
N . S	33	365	92	0.000	
N . K	30	365	63	0.000	
N . Y	23	365	98	0.000	
S . K	18	87	63	0.000	
K . T	17	94	150	0.000	
L . Y	16	96	98	0.000	
K . Y	14	94	98	0.000	
N . V	12	365	18	0.000	NPA
L . K	11	96	63	0.000	
L . T	11	96	150	0.000	
N . JH	11	365	16	0.000	
M . Y	9	161	98	0.000	
G . M	9	29	36	0.000	
T . S	9	34	92	0.000	
S . P	9	87	73	0.000	
T . R	8	34	32	0.000	
N . D R	8	365	10	0.000	
N . T R	8	365	19	0.000	
S . T R	8	87	19	0.000	
K . N	7	94	28	0.000	
Z . M	7	18	36	0.000	
T . Y	7	34	98	0.000	
L . M	7	96	36	0.000	
N . Z	7	365	13	0.000	
N . F	7	365	20	0.000	NPA
M . B R	6	161	6	0.000	
Z . L	6	18	28	0.000	
G . N	6	29	28	0.000	
L . B	6	96	65	0.000	

	L . D		6	96	88	0.000	
	L . S		6	96	92	0.000	
	G . Y		5	29	98	0.000	
	P . Y		5	21	98	0.000	
	L . V		5	96	18	0.000	
	L . F		5	96	20	0.000	
	B . Y		5	14	98	0.000	
	F . R		5	15	32	0.000	
	K . R		4	94	32	0.000	
	M . Z		4	161	13	0.000	NPA
	M . F		4	161	20	0.000	
	P . T		4	21	150	0.000	
	N . K	W	4	365	5	0.000	
	D . L		4	14	28	0.000	
	F . T		4	15	150	0.000	
	S . Y		4	87	98	0.000	
	K . S	T	3	94	10	0.000	
	K . SH		3	94	7	0.000	
	M . L		3	161	28	0.000	
	M . P	R	3	161	5	0.000	
	M . P	L	3	161	3	0.000	
	M . N		3	161	28	0.000	
	M . S		3	161	92	0.000	NPA
	G . L		3	29	28	0.000	
	P . R		3	21	32	0.000	
	P . S		3	21	92	0.000	
	T . L		3	34	28	0.000	
	V . R		3	6	32	0.000	
N	K . T		3	3	150	0.000	
	L . G		3	96	48	0.000	
	L . P		3	96	73	0.000	
	N . F	R	3	365	5	0.000	NPA
	N . K	R	3	365	4	0.000	
	N . S	T	3	365	10	0.000	
	N . TH		3	365	5	2.831	
	D . N		3	14	28	0.000	
	F . Y		3	15	98	0.000	
	S . M		3	87	36	0.000	
	K . W		2	94	8	0.000	
	K . M		2	94	36	0.000	
	K . L		2	94	28	0.000	
	K . D		2	94	88	3.007	VAssim
	Z . Y		2	18	98	0.000	
	Z . B		2	18	65	3.193	
M	P . T		2	4	150	0.000	
	G . Z		2	29	13	3.193	
	G . R		2	29	32	0.000	
	P . N		2	21	28	0.000	
	T . N		2	34	28	0.000	
	T . F		2	34	20	2.298	
	V . Y		2	6	98	0.000	
	L . W		2	96	8	0.000	
	L . JH		2	96	16	0.000	
	L . S	T	2	96	10	0.000	
	L . N		2	96	28	2.113	
	L . T	R	2	96	19	0.000	
	N . CH		2	365	3	0.000	
	N . S	T R	2	365	3	0.000	
	N . G	W	2	365	2	0.000	
	N . HH		2	365	3	2.368	
	N . G	R	2	365	3	0.000	
	N . K	L	2	365	2	0.000	
	N . SH		2	365	7	0.000	
	B . L		2	14	28	0.000	

	B . R		2	14	32	0.000	
	B . JH		2	14	16	3.193	
	B . S		2	14	92	3.193	VAssim
	D . M		2	14	36	0.000	
	D . Y		2	14	98	0.000	
	S . F		2	87	20	2.214	
	K . S	W	1	94	1	0.000	
	K . S	T R	1	94	3	0.000	
	K . B		1	94	65	3.007	VAssim
	K . T	R	1	94	19	0.000	
	M . D	R	1	161	10	0.000	NPA
	M . K		1	161	63	0.000	NPA
	M . F	R	1	161	5	0.000	
	M . K	W	1	161	5	0.000	NPA
	M . R		1	161	32	0.000	
	M . T		1	161	150	0.000	NPA
	M . B	L	1	161	1	0.000	
	M . SH		1	161	7	2.250	NPA
	M . F	L	1	161	1	0.000	
	M . D		1	161	88	0.000	NPA
	Z . JH		1	18	16	3.193	
N	G . HH		1	5	3	5.561	VAssim
N	G . R		1	5	32	0.000	
N	G . T		1	5	150	3.193	VAssim
N	G . S	T	1	5	10	3.193	VAssim
N	G . S		1	5	92	3.193	VAssim
M	P . K		1	4	63	0.000	
M	P . S		1	4	92	0.000	
	G . W		1	29	8	0.000	
	G . B		1	29	65	3.193	
	P . K		1	21	63	0.000	
	P . M		1	21	36	3.112	
	P . L		1	21	28	0.000	
	P . S	T	1	21	10	0.000	
	T . W		1	34	8	1.998	
	T . K		1	34	63	0.000	
	T . M		1	34	36	0.000	
N	S . K	R	1	1	4	0.000	
	V . L		1	6	28	0.000	
N	CH . B		1	1	65	5.841	VAssim
	L . CH		1	96	3	0.000	
	L . D	R	1	96	10	0.000	
	L . F	R	1	96	5	0.000	
	L . R		1	96	32	2.113	
	L . G	R	1	96	3	0.000	
	L . P	R	1	96	5	0.000	
	L . SH		1	96	7	0.000	
	SH . M		1	3	36	2.835	
	SH . R		1	3	32	2.835	
	SH . T		1	3	150	2.835	
	N . S	L	1	365	1	0.000	
	N . L		1	365	28	2.113	
	N . S	K	1	365	1	0.000	
	N . TH	R	1	365	1	2.831	
	TH . M		1	3	36	2.831	
	TH . L		1	3	28	2.831	
	TH . Y		1	3	98	2.831	
N	T . M		1	1	36	0.000	
	B . N		1	14	28	0.000	
	D . P		1	14	73	3.193	VAssim
	D . R		1	14	32	0.000	
	D . V		1	14	18	3.193	
	DH . M		1	1	36	2.831	

D	Z . W		1	1	8	3.193	
	F . G		1	15	48	3.007	VAssim
	F . N		1	15	28	2.182	
	F . TH		1	15	5	5.046	
	S . W		1	87	8	0.000	
	S . L		1	87	28	0.000	
	S . P	R	1	87	5	0.000	
	S . N		1	87	28	2.182	
	S . TH		1	87	5	5.046	
	S . B		1	87	65	3.007	VAssim
	K . CH		0	94	3	2.005	
	K . S	L	0	94	1	0.000	
	K . D	R	0	94	10	3.007	VAssim
	K . K		0	94	63	1.928	Degem
	K . Z		0	94	13	3.007	VAssim
	K . F	R	0	94	5	2.298	
	K . G		0	94	48	4.935	Degem, VAssim
	K . P		0	94	73	2.298	
	K . G	W	0	94	2	4.935	Degem, VAssim
	K . HH		0	94	3	2.368	
	K . K	W	0	94	5	1.928	Degem
	K . JH		0	94	16	5.011	VAssim
	K . G	R	0	94	3	4.935	Degem, VAssim
	K . P	R	0	94	5	2.298	
	K . B	L	0	94	1	3.007	VAssim
	K . V		0	94	18	3.007	VAssim
	K . K	L	0	94	2	1.928	Degem
	K . K	R	0	94	4	1.928	Degem
	K . B	R	0	94	6	3.007	VAssim
	K . P	L	0	94	3	2.298	
	K . S	K	0	94	1	0.000	
	K . TH		0	94	5	2.831	
	K . F	L	0	94	1	2.298	
	K . TH	R	0	94	1	2.831	
	K . F		0	94	20	2.298	
	M . S	W	0	161	1	0.000	NPA
	M . W		0	161	8	2.570	
	M . CH		0	161	3	2.250	NPA
	M . S	L	0	161	1	0.000	NPA
	M . M		0	161	36	2.570	Degem
	M . S	T R	0	161	3	0.000	NPA
	M . G		0	161	48	1.862	NPA
	M . G	W	0	161	2	1.862	NPA
	M . HH		0	161	3	2.368	
	M . JH		0	161	16	0.000	NPA
	M . G	R	0	161	3	1.862	NPA
	M . V		0	161	18	1.777	
	M . K	L	0	161	2	0.000	NPA
	M . K	R	0	161	4	0.000	NPA
	M . S	T	0	161	10	0.000	NPA
	M . S	K	0	161	1	0.000	NPA
	M . TH		0	161	5	2.831	NPA
	M . TH	R	0	161	1	2.831	NPA
	M . T	R	0	161	19	0.000	NPA
	Z . S	W	0	18	1	5.407	Degem, VAssim
	Z . W		0	18	8	0.000	
	Z . CH		0	18	3	3.193	VAssim
	Z . S	L	0	18	1	5.407	Degem, VAssim
	Z . D	R	0	18	10	3.193	
	Z . K		0	18	63	3.193	VAssim
	Z . Z		0	18	13	5.407	Degem
	Z . F	R	0	18	5	5.407	VAssim
	Z . S	T R	0	18	3	5.407	Degem, VAssim

	Z . G			0	18	48	3.193	
	Z . P			0	18	73	3.193	VAssim
	Z . G	W		0	18	2	3.193	
	Z . HH			0	18	3	7.775	VAssim
	Z . K	W		0	18	5	3.193	VAssim
	Z . R			0	18	32	2.142	
	Z . G	R		0	18	3	3.193	
	Z . T			0	18	150	3.193	VAssim
	Z . P	R		0	18	5	3.193	VAssim
	Z . B	L		0	18	1	3.193	
	Z . V			0	18	18	5.407	
	Z . K	L		0	18	2	3.193	VAssim
	Z . K	R		0	18	4	3.193	VAssim
	Z . B	R		0	18	6	3.193	
	Z . P	L		0	18	3	3.193	VAssim
	Z . S	T		0	18	10	5.407	Degem, VAssim
	Z . SH			0	18	7	5.407	VAssim
	Z . S	K		0	18	1	5.407	Degem, VAssim
	Z . N			0	18	28	2.182	
	Z . TH			0	18	5	8.238	VAssim
	Z . F	L		0	18	1	5.407	VAssim
	Z . TH	R		0	18	1	8.238	VAssim
	Z . D			0	18	88	3.193	
	Z . F			0	18	20	5.407	VAssim
	Z . S			0	18	92	5.407	Degem, VAssim
	Z . T	R		0	18	19	3.193	VAssim
N	G . S	W		0	5	1	3.193	VAssim
N	G . W			0	5	8	0.000	
N	G . CH			0	5	3	3.193	VAssim
N	G . S	L		0	5	1	3.193	VAssim
N	G . D	R		0	5	10	3.193	
N	G . K			0	5	63	5.121	Degem, VAssim
N	G . M			0	5	36	0.000	
N	G . L			0	5	28	0.000	
N	G . Z			0	5	13	3.193	
N	G . F	R		0	5	5	3.193	VAssim
N	G . S	T	R	0	5	3	3.193	VAssim
N	G . G			0	5	48	5.121	Degem
N	G . P			0	5	73	3.193	VAssim
N	G . G	W		0	5	2	5.121	Degem
N	G . K	W		0	5	5	5.121	Degem, VAssim
N	G . JH			0	5	16	3.193	
N	G . G	R		0	5	3	5.121	Degem
N	G . P	R		0	5	5	3.193	VAssim
N	G . B	L		0	5	1	3.193	
N	G . V			0	5	18	3.193	
N	G . K	L		0	5	2	5.121	Degem, VAssim
N	G . K	R		0	5	4	5.121	Degem, VAssim
N	G . B	R		0	5	6	3.193	
N	G . P	L		0	5	3	3.193	VAssim
N	G . Y			0	5	98	0.000	
N	G . SH			0	5	7	3.193	VAssim
N	G . S	K		0	5	1	3.193	VAssim
N	G . N			0	5	28	0.000	
N	G . TH			0	5	5	6.024	VAssim
N	G . F	L		0	5	1	3.193	VAssim
N	G . B			0	5	65	3.193	
N	G . TH	R		0	5	1	6.024	VAssim
N	G . D			0	5	88	3.193	
N	G . F			0	5	20	3.193	VAssim
N	G . T	R		0	5	19	3.193	VAssim
M	P . S	W		0	4	1	0.000	
M	P . W			0	4	8	3.112	

M	P . CH			0	4	3	4.255	
M	P . S	L		0	4	1	0.000	
M	P . D	R		0	4	10	3.007	VAssim
M	P . M			0	4	36	3.112	
M	P . L			0	4	28	0.000	
M	P . Z			0	4	13	3.007	VAssim
M	P . F	R		0	4	5	5.410	
M	P . S	T	R	0	4	3	0.000	
M	P . G			0	4	48	3.007	VAssim
M	P . P			0	4	73	5.410	Degem
M	P . G	W		0	4	2	3.007	VAssim
M	P . HH			0	4	3	2.368	
M	P . K	W		0	4	5	0.000	
M	P . R			0	4	32	0.000	
M	P . JH			0	4	16	5.011	VAssim
M	P . G	R		0	4	3	3.007	VAssim
M	P . P	R		0	4	5	5.410	Degem
M	P . B	L		0	4	1	6.118	Degem, VAssim
M	P . V			0	4	18	7.896	VAssim
M	P . K	L		0	4	2	0.000	
M	P . K	R		0	4	4	0.000	
M	P . B	R		0	4	6	6.118	Degem, VAssim
M	P . P	L		0	4	3	5.410	Degem
M	P . S	T		0	4	10	0.000	
M	P . Y			0	4	98	0.000	
M	P . SH			0	4	7	2.250	
M	P . S	K		0	4	1	0.000	
M	P . N			0	4	28	0.000	
M	P . TH			0	4	5	2.831	
M	P . F	L		0	4	1	5.410	
M	P . B			0	4	65	6.118	Degem, VAssim
M	P . TH	R		0	4	1	2.831	
M	P . D			0	4	88	3.007	VAssim
M	P . F			0	4	20	5.410	
M	P . T	R		0	4	19	0.000	
	G . S	W		0	29	1	3.193	VAssim
	G . CH			0	29	3	3.193	VAssim
	G . S	L		0	29	1	3.193	VAssim
	G . D	R		0	29	10	3.193	
	G . K			0	29	63	5.121	Degem, VAssim
	G . F	R		0	29	5	3.193	VAssim
	G . S	T	R	0	29	3	3.193	VAssim
	G . G			0	29	48	5.121	Degem
	G . P			0	29	73	3.193	VAssim
	G . G	W		0	29	2	5.121	Degem
	G . HH			0	29	3	5.561	VAssim
	G . K	W		0	29	5	5.121	Degem, VAssim
	G . JH			0	29	16	3.193	
	G . G	R		0	29	3	5.121	Degem
	G . T			0	29	150	3.193	VAssim
	G . P	R		0	29	5	3.193	VAssim
	G . B	L		0	29	1	3.193	
	G . V			0	29	18	3.193	
	G . K	L		0	29	2	5.121	Degem, VAssim
	G . K	R		0	29	4	5.121	Degem, VAssim
	G . B	R		0	29	6	3.193	
	G . P	L		0	29	3	3.193	VAssim
	G . S	T		0	29	10	3.193	VAssim
	G . SH			0	29	7	3.193	VAssim
	G . S	K		0	29	1	3.193	VAssim
	G . TH			0	29	5	6.024	VAssim
	G . F	L		0	29	1	3.193	VAssim
	G . TH	R		0	29	1	6.024	VAssim

G . D	0	29	88	3.193	
G . F	0	29	20	3.193	VAssim
G . S	0	29	92	3.193	VAssim
G . T R	0	29	19	3.193	VAssim
P . S W	0	21	1	0.000	
P . W	0	21	8	3.112	
P . CH	0	21	3	4.255	
P . S L	0	21	1	0.000	
P . D R	0	21	10	3.007	VAssim
P . Z	0	21	13	3.007	VAssim
P . F R	0	21	5	5.410	
P . S T R	0	21	3	0.000	
P . G	0	21	48	3.007	VAssim
P . P	0	21	73	5.410	Degem
P . G W	0	21	2	3.007	VAssim
P . HH	0	21	3	2.368	
P . K W	0	21	5	0.000	
P . JH	0	21	16	5.011	VAssim
P . G R	0	21	3	3.007	VAssim
P . P R	0	21	5	5.410	Degem
P . B L	0	21	1	6.118	Degem, VAssim
P . V	0	21	18	7.896	VAssim
P . K L	0	21	2	0.000	
P . K R	0	21	4	0.000	
P . B R	0	21	6	6.118	Degem, VAssim
P . P L	0	21	3	5.410	Degem
P . SH	0	21	7	2.250	
P . S K	0	21	1	0.000	
P . TH	0	21	5	2.831	
P . F L	0	21	1	5.410	
P . B	0	21	65	6.118	Degem, VAssim
P . TH R	0	21	1	2.831	
P . D	0	21	88	3.007	VAssim
P . F	0	21	20	5.410	
P . T R	0	21	19	0.000	
T . S W	0	34	1	0.000	
T . CH	0	34	3	5.818	
T . S L	0	34	1	0.000	
T . D R	0	34	10	4.913	Degem, VAssim
T . Z	0	34	13	3.007	VAssim
T . F R	0	34	5	2.298	
T . S T R	0	34	3	0.000	
T . G	0	34	48	3.007	VAssim
T . P	0	34	73	2.298	
T . G W	0	34	2	3.007	VAssim
T . HH	0	34	3	2.368	
T . K W	0	34	5	0.000	
T . JH	0	34	16	8.825	VAssim
T . G R	0	34	3	3.007	VAssim
T . T	0	34	150	1.906	Degem
T . P R	0	34	5	2.298	
T . B L	0	34	1	3.007	VAssim
T . V	0	34	18	3.007	VAssim
T . K L	0	34	2	0.000	
T . K R	0	34	4	0.000	
T . B R	0	34	6	3.007	VAssim
T . P L	0	34	3	2.298	
T . S T	0	34	10	0.000	
T . SH	0	34	7	1.907	
T . S K	0	34	1	0.000	
T . TH	0	34	5	2.831	
T . F L	0	34	1	2.298	
T . B	0	34	65	3.007	VAssim

	T . TH	R	0	34	1	2.831	
	T . D		0	34	88	4.913	Degem, VAssim
	T . T	R	0	34	19	1.906	Degem
N	S . S	W	0	1	1	2.214	Degem
N	S . W		0	1	8	0.000	
N	S . CH		0	1	3	3.912	
N	S . S	L	0	1	1	2.214	Degem
N	S . D	R	0	1	10	3.007	VAssim
N	S . K		0	1	63	0.000	
N	S . M		0	1	36	0.000	
N	S . L		0	1	28	0.000	
N	S . Z		0	1	13	5.221	Degem, VAssim
N	S . F	R	0	1	5	2.214	
N	S . S	T R	0	1	3	2.214	Degem
N	S . G		0	1	48	3.007	VAssim
N	S . P		0	1	73	0.000	
N	S . G	W	0	1	2	3.007	VAssim
N	S . HH		0	1	3	4.583	
N	S . K	W	0	1	5	0.000	
N	S . R		0	1	32	2.142	
N	S . JH		0	1	16	6.919	VAssim
N	S . G	R	0	1	3	3.007	VAssim
N	S . T		0	1	150	0.000	
N	S . P	R	0	1	5	0.000	
N	S . B	L	0	1	1	3.007	VAssim
N	S . V		0	1	18	5.221	VAssim
N	S . K	L	0	1	2	0.000	
N	S . B	R	0	1	6	3.007	VAssim
N	S . P	L	0	1	3	0.000	
N	S . S	T	0	1	10	2.214	Degem
N	S . Y		0	1	98	0.000	
N	S . SH		0	1	7	4.121	
N	S . S	K	0	1	1	2.214	Degem
N	S . N		0	1	28	2.182	
N	S . TH		0	1	5	5.046	
N	S . F	L	0	1	1	2.214	
N	S . B		0	1	65	3.007	VAssim
N	S . TH	R	0	1	1	5.046	
N	S . D		0	1	88	3.007	VAssim
N	S . F		0	1	20	2.214	
N	S . S		0	1	92	2.214	Degem
N	S . T	R	0	1	19	0.000	
	V . S	W	0	6	1	5.407	VAssim
	V . W		0	6	8	3.112	
	V . CH		0	6	3	5.443	VAssim
	V . S	L	0	6	1	5.407	VAssim
	V . D	R	0	6	10	3.193	
	V . K		0	6	63	5.076	VAssim
	V . M		0	6	36	3.112	
	V . Z		0	6	13	5.407	
	V . F	R	0	6	5	8.519	Degem, VAssim
	V . S	T R	0	6	3	5.407	VAssim
	V . G		0	6	48	3.193	
	V . P		0	6	73	6.304	VAssim
	V . G	W	0	6	2	3.193	
	V . HH		0	6	3	7.775	VAssim
	V . K	W	0	6	5	5.076	VAssim
	V . JH		0	6	16	3.193	
	V . G	R	0	6	3	3.193	
	V . T		0	6	150	3.193	VAssim
	V . P	R	0	6	5	6.304	VAssim
	V . B	L	0	6	1	6.304	
	V . V		0	6	18	10.296	Degem

V . K L	0	6	2	5.076	VAssim
V . K R	0	6	4	5.076	VAssim
V . B R	0	6	6	6.304	
V . P L	0	6	3	6.304	VAssim
V . S T	0	6	10	5.407	VAssim
V . SH	0	6	7	7.657	VAssim
V . S K	0	6	1	5.407	VAssim
V . N	0	6	28	2.182	
V . TH	0	6	5	8.238	VAssim
V . F L	0	6	1	8.519	Degem, VAssim
V . B	0	6	65	6.304	
V . TH R	0	6	1	8.238	VAssim
V . D	0	6	88	3.193	
V . F	0	6	20	8.519	Degem, VAssim
V . S	0	6	92	5.407	VAssim
V . T R	0	6	19	3.193	VAssim
N CH . S W	0	1	1	2.835	
N CH . W	0	1	8	4.832	
N CH . CH	0	1	3	8.652	Degem
N CH . S L	0	1	1	2.835	
N CH . D R	0	1	10	7.747	VAssim
N CH . K	0	1	63	2.835	
N CH . M	0	1	36	2.835	
N CH . L	0	1	28	2.835	
N CH . Z	0	1	13	5.841	VAssim
N CH . F R	0	1	5	5.133	
N CH . S T R	0	1	3	2.835	
N CH . G	0	1	48	5.841	VAssim
N CH . P	0	1	73	5.133	
N CH . G W	0	1	2	5.841	VAssim
N CH . HH	0	1	3	5.203	
N CH . K W	0	1	5	2.835	
N CH . R	0	1	32	2.835	
N CH . JH	0	1	16	11.659	Degem, VAssim
N CH . G R	0	1	3	5.841	VAssim
N CH . T	0	1	150	4.741	
N CH . P R	0	1	5	5.133	
N CH . B L	0	1	1	5.841	VAssim
N CH . V	0	1	18	5.841	VAssim
N CH . K L	0	1	2	2.835	
N CH . K R	0	1	4	2.835	
N CH . B R	0	1	6	5.841	VAssim
N CH . P L	0	1	3	5.133	
N CH . S T	0	1	10	2.835	
N CH . Y	0	1	98	2.835	
N CH . SH	0	1	7	4.742	
N CH . S K	0	1	1	2.835	
N CH . N	0	1	28	2.835	
N CH . TH	0	1	5	5.666	
N CH . F L	0	1	1	5.133	
N CH . TH R	0	1	1	5.666	
N CH . D	0	1	88	7.747	VAssim
N CH . F	0	1	20	5.133	
N CH . S	0	1	92	2.835	
N CH . T R	0	1	19	4.741	
N K . S W	0	3	1	0.000	
N K . W	0	3	8	0.000	
N K . CH	0	3	3	2.005	
N K . S L	0	3	1	0.000	
N K . D R	0	3	10	3.007	VAssim
N K . K	0	3	63	1.928	Degem
N K . M	0	3	36	0.000	
N K . L	0	3	28	0.000	

N	K . Z			0	3	13	3.007	VAssim
N	K . F	R		0	3	5	2.298	
N	K . S	T	R	0	3	3	0.000	
N	K . G			0	3	48	4.935	Degem, VAssim
N	K . P			0	3	73	2.298	
N	K . G	W		0	3	2	4.935	Degem, VAssim
N	K . HH			0	3	3	2.368	
N	K . K	W		0	3	5	1.928	Degem
N	K . R			0	3	32	0.000	
N	K . JH			0	3	16	5.011	VAssim
N	K . G	R		0	3	3	4.935	Degem, VAssim
N	K . P	R		0	3	5	2.298	
N	K . B	L		0	3	1	3.007	VAssim
N	K . V			0	3	18	3.007	VAssim
N	K . K	L		0	3	2	1.928	Degem
N	K . K	R		0	3	4	1.928	Degem
N	K . B	R		0	3	6	3.007	VAssim
N	K . P	L		0	3	3	2.298	
N	K . S	T		0	3	10	0.000	
N	K . Y			0	3	98	0.000	
N	K . SH			0	3	7	0.000	
N	K . S	K		0	3	1	0.000	
N	K . N			0	3	28	0.000	
N	K . TH			0	3	5	2.831	
N	K . F	L		0	3	1	2.298	
N	K . B			0	3	65	3.007	VAssim
N	K . TH	R		0	3	1	2.831	
N	K . D			0	3	88	3.007	VAssim
N	K . F			0	3	20	2.298	
N	K . S			0	3	92	0.000	
N	K . T	R		0	3	19	0.000	
	L . S	W		0	96	1	0.000	
	L . S	L		0	96	1	0.000	
	L . L			0	96	28	2.113	Degem
	L . Z			0	96	13	0.000	
	L . S	T	R	0	96	3	0.000	
	L . G	W		0	96	2	0.000	
	L . HH			0	96	3	2.368	
	L . K	W		0	96	5	0.000	
	L . B	L		0	96	1	0.000	
	L . K	L		0	96	2	0.000	
	L . K	R		0	96	4	0.000	
	L . B	R		0	96	6	0.000	
	L . P	L		0	96	3	0.000	
	L . S	K		0	96	1	0.000	
	L . TH			0	96	5	2.831	
	L . F	L		0	96	1	0.000	
	L . TH	R		0	96	1	2.831	
	SH . S	W		0	3	1	5.049	
	SH . W			0	3	8	2.835	
	SH . CH			0	3	3	6.746	
	SH . S	L		0	3	1	5.049	
	SH . D	R		0	3	10	5.841	VAssim
	SH . K			0	3	63	2.835	
	SH . L			0	3	28	2.835	
	SH . Z			0	3	13	8.056	VAssim
	SH . F	R		0	3	5	5.049	
	SH . S	T	R	0	3	3	5.049	
	SH . G			0	3	48	5.841	VAssim
	SH . P			0	3	73	2.835	
	SH . G	W		0	3	2	5.841	VAssim
	SH . HH			0	3	3	7.417	
	SH . K	W		0	3	5	2.835	

SH . JH			0	3	16	9.753	VAssim
SH . G	R		0	3	3	5.841	VAssim
SH . P	R		0	3	5	2.835	
SH . B	L		0	3	1	5.841	VAssim
SH . V			0	3	18	8.056	VAssim
SH . K	L		0	3	2	2.835	
SH . K	R		0	3	4	2.835	
SH . B	R		0	3	6	5.841	VAssim
SH . P	L		0	3	3	2.835	
SH . S	T		0	3	10	5.049	
SH . Y			0	3	98	2.835	
SH . SH			0	3	7	6.956	Degem
SH . S	K		0	3	1	5.049	
SH . N			0	3	28	5.017	
SH . TH			0	3	5	7.880	
SH . F	L		0	3	1	5.049	
SH . B			0	3	65	5.841	VAssim
SH . TH	R		0	3	1	7.880	
SH . D			0	3	88	5.841	VAssim
SH . F			0	3	20	5.049	
SH . S			0	3	92	5.049	
SH . T	R		0	3	19	2.835	
N . S	W		0	365	1	0.000	
N . W			0	365	8	2.570	
N . M			0	365	36	2.570	
N . P			0	365	73	2.178	NPA
N . R			0	365	32	2.113	
N . P	R		0	365	5	2.178	NPA
N . B	L		0	365	1	2.178	NPA
N . B	R		0	365	6	2.178	NPA
N . P	L		0	365	3	2.178	NPA
N . N			0	365	28	2.113	Degem
N . F	L		0	365	1	0.000	NPA
N . B			0	365	65	2.178	NPA
TH . S	W		0	3	1	5.046	
TH . W			0	3	8	2.831	
TH . CH			0	3	3	6.743	
TH . S	L		0	3	1	5.046	
TH . D	R		0	3	10	5.838	VAssim
TH . K			0	3	63	2.831	
TH . Z			0	3	13	8.052	VAssim
TH . F	R		0	3	5	5.046	
TH . S	T	R	0	3	3	5.046	
TH . G			0	3	48	5.838	VAssim
TH . P			0	3	73	2.831	
TH . G	W		0	3	2	5.838	VAssim
TH . HH			0	3	3	7.414	
TH . K	W		0	3	5	2.831	
TH . R			0	3	32	4.973	
TH . JH			0	3	16	9.750	VAssim
TH . G	R		0	3	3	5.838	VAssim
TH . T			0	3	150	2.831	
TH . P	R		0	3	5	2.831	
TH . B	L		0	3	1	5.838	VAssim
TH . V			0	3	18	8.052	VAssim
TH . K	L		0	3	2	2.831	
TH . K	R		0	3	4	2.831	
TH . B	R		0	3	6	5.838	VAssim
TH . P	L		0	3	3	2.831	
TH . S	T		0	3	10	5.046	
TH . SH			0	3	7	6.953	
TH . S	K		0	3	1	5.046	
TH . N			0	3	28	5.014	

	TH . TH			0	3	5	7.877	Degem
	TH . F	L		0	3	1	5.046	
	TH . B			0	3	65	5.838	VAssim
	TH . TH	R		0	3	1	7.877	Degem
	TH . D			0	3	88	5.838	VAssim
	TH . F			0	3	20	5.046	
	TH . S			0	3	92	5.046	
	TH . T	R		0	3	19	2.831	
N	T . S	W		0	1	1	0.000	
N	T . W			0	1	8	1.998	
N	T . CH			0	1	3	5.818	
N	T . S	L		0	1	1	0.000	
N	T . D	R		0	1	10	4.913	Degem, VAssim
N	T . K			0	1	63	0.000	
N	T . L			0	1	28	0.000	
N	T . Z			0	1	13	3.007	VAssim
N	T . F	R		0	1	5	2.298	
N	T . S	T	R	0	1	3	0.000	
N	T . G			0	1	48	3.007	VAssim
N	T . P			0	1	73	2.298	
N	T . G	W		0	1	2	3.007	VAssim
N	T . HH			0	1	3	2.368	
N	T . K	W		0	1	5	0.000	
N	T . R			0	1	32	0.000	
N	T . JH			0	1	16	8.825	VAssim
N	T . G	R		0	1	3	3.007	VAssim
N	T . T			0	1	150	1.906	Degem
N	T . P	R		0	1	5	2.298	
N	T . B	L		0	1	1	3.007	VAssim
N	T . V			0	1	18	3.007	VAssim
N	T . K	L		0	1	2	0.000	
N	T . K	R		0	1	4	0.000	
N	T . B	R		0	1	6	3.007	VAssim
N	T . P	L		0	1	3	2.298	
N	T . S	T		0	1	10	0.000	
N	T . Y			0	1	98	0.000	
N	T . SH			0	1	7	1.907	
N	T . S	K		0	1	1	0.000	
N	T . N			0	1	28	0.000	
N	T . TH			0	1	5	2.831	
N	T . F	L		0	1	1	2.298	
N	T . B			0	1	65	3.007	VAssim
N	T . TH	R		0	1	1	2.831	
N	T . D			0	1	88	4.913	Degem, VAssim
N	T . F			0	1	20	2.298	
N	T . S			0	1	92	0.000	
N	T . T	R		0	1	19	1.906	Degem
	B . S	W		0	14	1	3.193	VAssim
	B . W			0	14	8	3.112	
	B . CH			0	14	3	5.443	VAssim
	B . S	L		0	14	1	3.193	VAssim
	B . D	R		0	14	10	3.193	
	B . K			0	14	63	3.193	VAssim
	B . M			0	14	36	3.112	
	B . Z			0	14	13	3.193	
	B . F	R		0	14	5	6.304	VAssim
	B . S	T	R	0	14	3	3.193	VAssim
	B . G			0	14	48	3.193	
	B . P			0	14	73	6.304	Degem, VAssim
	B . G	W		0	14	2	3.193	
	B . HH			0	14	3	5.561	VAssim
	B . K	W		0	14	5	3.193	VAssim
	B . G	R		0	14	3	3.193	

B . T	0	14	150	3.193	VAssim
B . P R	0	14	5	6.304	Degem, VAssim
B . B L	0	14	1	6.304	Degem
B . V	0	14	18	8.082	
B . K L	0	14	2	3.193	VAssim
B . K R	0	14	4	3.193	VAssim
B . B R	0	14	6	6.304	Degem
B . P L	0	14	3	6.304	Degem, VAssim
B . S T	0	14	10	3.193	VAssim
B . SH	0	14	7	5.443	VAssim
B . S K	0	14	1	3.193	VAssim
B . TH	0	14	5	6.024	VAssim
B . F L	0	14	1	6.304	VAssim
B . B	0	14	65	6.304	Degem
B . TH R	0	14	1	6.024	VAssim
B . D	0	14	88	3.193	
B . F	0	14	20	6.304	VAssim
B . T R	0	14	19	3.193	VAssim
D . S W	0	14	1	3.193	VAssim
D . W	0	14	8	1.998	
D . CH	0	14	3	5.099	VAssim
D . S L	0	14	1	3.193	VAssim
D . D R	0	14	10	5.099	Degem
D . K	0	14	63	3.193	VAssim
D . Z	0	14	13	3.193	
D . F R	0	14	5	3.193	VAssim
D . S T R	0	14	3	3.193	VAssim
D . G	0	14	48	3.193	
D . G W	0	14	2	3.193	
D . HH	0	14	3	5.561	VAssim
D . K W	0	14	5	3.193	VAssim
D . JH	0	14	16	5.099	
D . G R	0	14	3	3.193	
D . T	0	14	150	5.099	Degem, VAssim
D . P R	0	14	5	3.193	VAssim
D . B L	0	14	1	3.193	
D . K L	0	14	2	3.193	VAssim
D . K R	0	14	4	3.193	VAssim
D . B R	0	14	6	3.193	
D . P L	0	14	3	3.193	VAssim
D . S T	0	14	10	3.193	VAssim
D . SH	0	14	7	3.193	VAssim
D . S K	0	14	1	3.193	VAssim
D . TH	0	14	5	6.024	VAssim
D . F L	0	14	1	3.193	VAssim
D . B	0	14	65	3.193	
D . TH R	0	14	1	6.024	VAssim
D . D	0	14	88	5.099	Degem
D . F	0	14	20	3.193	VAssim
D . S	0	14	92	3.193	VAssim
D . T R	0	14	19	5.099	Degem, VAssim
DH . S W	0	1	1	8.238	VAssim
DH . W	0	1	8	2.831	
DH . CH	0	1	3	6.024	VAssim
DH . S L	0	1	1	8.238	VAssim
DH . D R	0	1	10	6.024	
DH . K	0	1	63	6.024	VAssim
DH . L	0	1	28	2.831	
DH . Z	0	1	13	8.238	
DH . F R	0	1	5	8.238	VAssim
DH . S T R	0	1	3	8.238	VAssim
DH . G	0	1	48	6.024	
DH . P	0	1	73	6.024	VAssim

	DH . G	W		0	1	2	6.024	
	DH . HH			0	1	3	10.606	VAssim
	DH . K	W		0	1	5	6.024	VAssim
	DH . R			0	1	32	4.973	
	DH . JH			0	1	16	6.024	
	DH . G	R		0	1	3	6.024	
	DH . T			0	1	150	6.024	VAssim
	DH . P	R		0	1	5	6.024	VAssim
	DH . B	L		0	1	1	6.024	
	DH . V			0	1	18	8.238	
	DH . K	L		0	1	2	6.024	VAssim
	DH . K	R		0	1	4	6.024	VAssim
	DH . B	R		0	1	6	6.024	
	DH . P	L		0	1	3	6.024	VAssim
	DH . S	T		0	1	10	8.238	VAssim
	DH . Y			0	1	98	2.831	
	DH . SH			0	1	7	8.238	VAssim
	DH . S	K		0	1	1	8.238	VAssim
	DH . N			0	1	28	5.014	
	DH . TH			0	1	5	11.069	Degem, VAssim
	DH . F	L		0	1	1	8.238	VAssim
	DH . B			0	1	65	6.024	
	DH . TH	R		0	1	1	11.069	Degem, VAssim
	DH . D			0	1	88	6.024	
	DH . F			0	1	20	8.238	VAssim
	DH . S			0	1	92	8.238	VAssim
	DH . T	R		0	1	19	6.024	VAssim
D	Z . S	W		0	1	1	8.600	Degem, VAssim
D	Z . CH			0	1	3	6.385	VAssim
D	Z . S	L		0	1	1	8.600	Degem, VAssim
D	Z . D	R		0	1	10	6.385	
D	Z . K			0	1	63	6.385	VAssim
D	Z . M			0	1	36	3.193	
D	Z . L			0	1	28	3.193	
D	Z . Z			0	1	13	8.600	Degem
D	Z . F	R		0	1	5	8.600	VAssim
D	Z . S	T	R	0	1	3	8.600	Degem, VAssim
D	Z . G			0	1	48	6.385	
D	Z . P			0	1	73	6.385	VAssim
D	Z . G	W		0	1	2	6.385	
D	Z . HH			0	1	3	10.968	VAssim
D	Z . K	W		0	1	5	6.385	VAssim
D	Z . R			0	1	32	5.334	
D	Z . JH			0	1	16	6.385	
D	Z . G	R		0	1	3	6.385	
D	Z . T			0	1	150	6.385	VAssim
D	Z . P	R		0	1	5	6.385	VAssim
D	Z . B	L		0	1	1	6.385	
D	Z . V			0	1	18	8.600	
D	Z . K	L		0	1	2	6.385	VAssim
D	Z . K	R		0	1	4	6.385	VAssim
D	Z . B	R		0	1	6	6.385	
D	Z . P	L		0	1	3	6.385	VAssim
D	Z . S	T		0	1	10	8.600	Degem, VAssim
D	Z . Y			0	1	98	3.193	
D	Z . SH			0	1	7	8.600	VAssim
D	Z . S	K		0	1	1	8.600	Degem, VAssim
D	Z . N			0	1	28	5.375	
D	Z . TH			0	1	5	11.431	VAssim
D	Z . F	L		0	1	1	8.600	VAssim
D	Z . B			0	1	65	6.385	
D	Z . TH	R		0	1	1	11.431	VAssim
D	Z . D			0	1	88	6.385	

D	Z . F			0	1	20	8.600	VAssim
D	Z . S			0	1	92	8.600	Degem, VAssim
D	Z . T	R		0	1	19	6.385	VAssim
	F . S	W		0	15	1	2.214	
	F . W			0	15	8	3.112	
	F . CH			0	15	3	4.255	
	F . S	L		0	15	1	2.214	
	F . D	R		0	15	10	3.007	VAssim
	F . K			0	15	63	1.884	
	F . M			0	15	36	3.112	
	F . L			0	15	28	0.000	
	F . Z			0	15	13	5.221	VAssim
	F . F	R		0	15	5	5.326	Degem
	F . S	T	R	0	15	3	2.214	
	F . P			0	15	73	3.112	
	F . G	W		0	15	2	3.007	VAssim
	F . HH			0	15	3	4.583	
	F . K	W		0	15	5	1.884	
	F . JH			0	15	16	5.011	VAssim
	F . G	R		0	15	3	3.007	VAssim
	F . P	R		0	15	5	3.112	
	F . B	L		0	15	1	6.118	VAssim
	F . V			0	15	18	10.11	Degem, VAssim
	F . K	L		0	15	2	1.884	
	F . K	R		0	15	4	1.884	
	F . B	R		0	15	6	6.118	VAssim
	F . P	L		0	15	3	3.112	
	F . S	T		0	15	10	2.214	
	F . SH			0	15	7	4.464	
	F . S	K		0	15	1	2.214	
	F . F	L		0	15	1	5.326	Degem
	F . B			0	15	65	6.118	VAssim
	F . TH	R		0	15	1	5.046	
	F . D			0	15	88	3.007	VAssim
	F . F			0	15	20	5.326	Degem
	F . S			0	15	92	2.214	
	F . T	R		0	15	19	0.000	
	S . S	W		0	87	1	2.214	Degem
	S . CH			0	87	3	3.912	
	S . S	L		0	87	1	2.214	Degem
	S . D	R		0	87	10	3.007	VAssim
	S . Z			0	87	13	5.221	Degem, VAssim
	S . F	R		0	87	5	2.214	
	S . S	T	R	0	87	3	2.214	Degem
	S . G			0	87	48	3.007	VAssim
	S . G	W		0	87	2	3.007	VAssim
	S . HH			0	87	3	4.583	
	S . K	W		0	87	5	0.000	
	S . R			0	87	32	2.142	
	S . JH			0	87	16	6.919	VAssim
	S . G	R		0	87	3	3.007	VAssim
	S . B	L		0	87	1	3.007	VAssim
	S . V			0	87	18	5.221	VAssim
	S . K	L		0	87	2	0.000	
	S . K	R		0	87	4	0.000	
	S . B	R		0	87	6	3.007	VAssim
	S . P	L		0	87	3	0.000	
	S . S	T		0	87	10	2.214	Degem
	S . SH			0	87	7	4.121	
	S . S	K		0	87	1	2.214	Degem
	S . F	L		0	87	1	2.214	
	S . TH	R		0	87	1	5.046	
	S . D			0	87	88	3.007	VAssim

S . S	0	87	92	2.214	Degem

Appendix D

Parameterizing Zipf's Law

Zipf (1949) famously observes the linear relationship between log rank r and log word frequency $f(r)$ in several linguistic samples. This is generalized below:

(27) $$f(C, \alpha) = \frac{C}{r^{\alpha}}$$

C is a constant, sensitive to sample size. Zipf assumes a 1-to-1 relationship, implying $\alpha = -1$, but it is possible to compute an optimal estimate for this parameter by taking the logarithm of both sides of the equation and solving for values of C and α that minimize the error term ε; this can be done efficiently with linear regression:

(28) $$\log f(r) \sim \log C + \alpha \log r + \varepsilon$$

Good (1953) notes that sparse distributions exhibit quantization at low frequencies, resulting in an artificially long flat right tail imposing an upward bias on estimates of α. Church and Gale (1991:29) propose a transform which eliminates this quantization. The vectors r, n are defined so such that n_i is the number of types which occur at frequency r_i (that is, n is a vector of frequencies of individual type frequencies). Z contains the elements of n by normalized by the points to the left and right.

(29) $Z_i = \dfrac{2n_i}{r_{i+1} - r_{i-1}}$

Church and Gale do not define this transform for the lowest and highest points (i.e., when $i = 1$ or N), but a natural extension of their definition is to scale the endpoints according to the next intermost point, as defined below.

(30) $Z_1 = \dfrac{n_1}{r_2 - r_1}$

(31) $Z_N = \dfrac{n_N}{r_N - r_{N-1}}$

The effect of applying this transform to sparse frequency data is shown in Figure D.1.

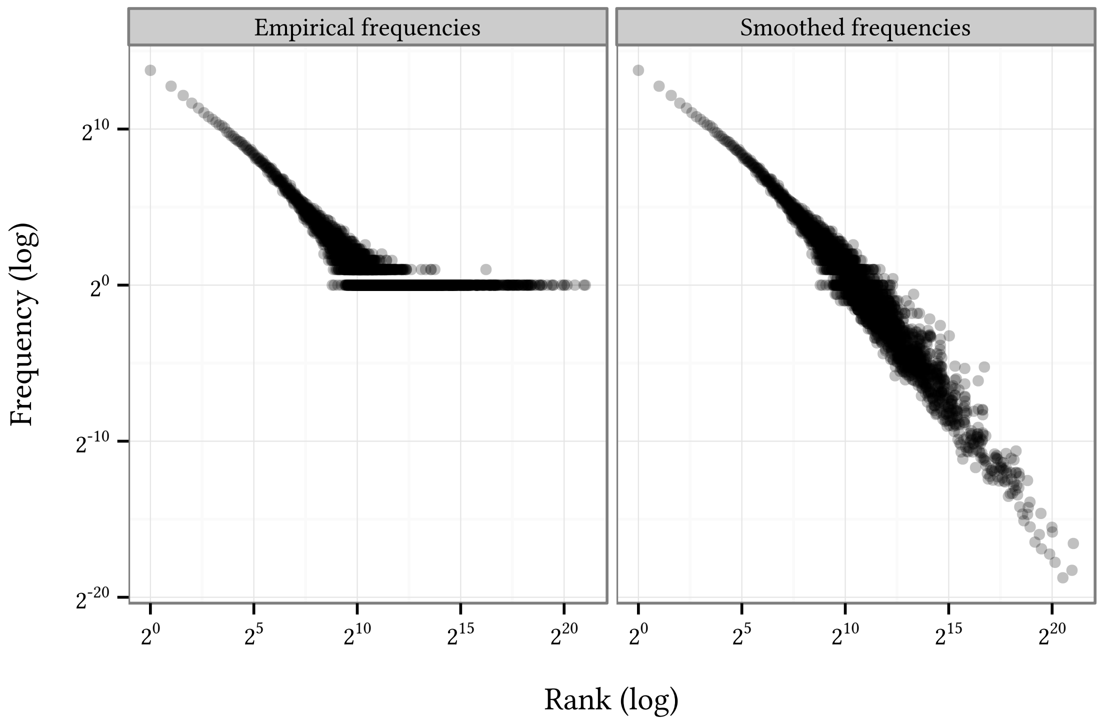


Figure D.1: Left panel: word frequencies from the SUBTLEX-US frequency norms (Brysbaert and New 2009); Right panel: the same data smoothed with the Z_r transform

Bibliography

Albright, Adam. 2007. Natural classes are not enough: Biased generalization in novel onset clusters. Ms., MIT.

Albright, Adam. 2009. Feature-based generalisation as a source of gradient acceptability. *Phonology* 26:9–41.

Albright, Adam, and Bruce Hayes. 2003. Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition* 90:119–161.

Albright, Adam, Giorgio Magri, and Jennifer Michaels. 2008. Modeling doubly marked lags with a split additive model. In *Proceedings of the 32nd annual Boston University Conference on Lanugage Development*, volume 1, 36–47. Somerville, MA: Cascadilla.

Alderete, John, and Mark Bradshaw. In press. Samoan root phonotactics: Digging deeper into the data. *Linguistic Discovery* to appear.

Anderson, John M. 1988. More on slips and syllable structure. *Phonology* 5:157–159.

Anderson, Stephen R. 1970. On Grassman's Law in Sanskrit. *Linguistic Inquiry* 1:387–396.

Anderson, Stephen R. 1973. Remarks on the phonology of English inflection. *Language and Literature* 1:33–52.

Anderson, Stephen R. 1974. *The organization of phonology*. New York: Academic Press.

Anderson, Stephen R. 1985. *Phonology in the twentieth century: Theories of rules and theories of representations*. Chicago: University of Chicago Press.

Anderson, Stephen R. 2000. Reflections on "On the phonetic rules of Russian". *Folia Linguistica* 34:11–28.

Anttila, Arto. 1997. Deriving variation from grammar. In *Variation, Change and Phonoloy Theory*, ed. Frans Hinskens, Leo Wetzels, and Roeland van Hout, 35–68. Amsterdam: John Benjamins.

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

Armstrong, Sharon L., Lila R. Gleitman, and Henry Gleitman. 1983. What some concepts might not be. *Cognition* 13:263–308.

Aronoff, Mark. 1976. *Word formation in generative grammar*. Cambridge: MIT Press.

Baayen, R. Harald, Richard Piepenbrock, and Léon Gulikers. 1996. CELEX2. Linguistic Data Consortium: LDC96L14.

Bagemihl, Bruce. 1995. Language games and related areas. In *The handbook of phonological theory*, ed. John Goldsmith, 697–712. Cambridge: Blackwell.

Bailey, Todd M., and Ulrike Hahn. 2001. Determinants of wordlikeness: Phonotactics or lexical neighborhoods. *Journal of Memory and Language* 44:586–591.

Baković, Eric. 2005. Antigemination, assimilation and the determination of identity. *Phonology* 22:279–315.

Baldi, Philip. 1994. Some thoughts on Latin rhotacism. *General Linguistics* 34:209–215.

Balota, David A., Melvin J. Yap, Michael J. Cortese, Keith A. Hutchison, Brett Kessler, Bjorn Loftis, James H. Neely, Douglas L. Nelson, Greg B. Simpson, and Rebecca Treiman. 2007. The English Lexicon Project. *Behavior Research Methods* 39:445–459.

Basbøll, Hans. 1972. Remarks on the regular plural formation of English nouns. *Language and Literature* 3:39–42.

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

Berent, Iris. 2008. Are phonological representations of printed and spoken language isomorphic? Evidence from the restrictions on unattested onsets. *Journal of Experimental Psychology: Human Perception and Performance* 34:1288–1304.

Berent, Iris, and Joseph Shimron. 2003. Co-occurrence restrictions on identical consonants in the Hebrew lexicon: Are they due to similarity? *Journal of Linguistics* 39:31–55.

Berent, Iris, Joseph Shimron, and Vered Vaknin. 2001. Phonological constraints on reading: Evidence from the Obligatory Contour Principle. *Journal of Memory and Language* 44:644–665.

Berent, Iris, Donca Steriade, Tracy Lennertz, and Vered Vaknin. 2007. What we know about what we have never heard: Evidence from perceptual illusions. *Cognition* 104:591–630.

Bergelson, Elika, and Daniel Swingley. 2012. At 6 to 9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences* 109:3253–3258.

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. Chicago: Chicago Linguistic Society.

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

Best, Catherine T. 1994. The emergence of native-language phonological influences in infants: A perceptual assimilation model. In *The transition from speech sounds to spoken words: The development of speech perception*, ed. Judith Goodman and Howard C. Nusbaum, 167–224. Cambridge: MIT Press.

Blevins, Juliette. 1995. The syllable in phonological theory. In *The handbook of phonological theory*, ed. John Goldsmith, 206–244. Oxford: Blackwell.

Bloch, Bernard. 1941. Phonemic overlapping. *American Speech* 16:278–284.

Bloomfield, Leonard. 1926. A set of postulates for the science of language. *Language* 2:153–163.

Bloomfield, Leonard. 1933. *Language*. New York: Henry Holt & Co.

Bloomfield, Leonard. 1962. *The Menomini language*. New Haven: Yale University Press.

Borowsky, Toni. 1986. Topics in the lexical phonology of English. Doctoral dissertation, University of Massachusetts, Amherst. Published by Garland, New York, 1991.

Borowsky, Toni. 1989. Structure preservation and the syllable coda in English. *Linguistic Inquiry* 7:145–166.

Borowsky, Toni, and Armin Mester. 1983. Aspiration to roots: Remarks on the Sanskrit diaspirates. In *Papers from the 19th regional meeting of the Chicago Linguistic Society*, 52–63. Chicago: Chicago Linguistic Society.

Bortfeld, Heather, James L. Morgan, Roberta M. Golinkoff, and Karen Rathbun. 2005. Mommy and me: Familiar names help launch babies into speech-stream segmentation. *Psychological Science* 16:298–304.

Brooks, Maria Zagorska. 1965. On Polish affricates. *Word* 20:207–210.

Brown, Jason. 2010. *Gitksan phonotactics*. München: LINCOM.

Brown, Roger, and Donald Hildum. 1956. Expectancy and the perception of syllables. *Language* 32:411–419.

Brysbaert, Marc, and Boris New. 2009. Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods* 41:977–990.

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

Cairns, Charles E. 1972. Review of Scholes 1966. *Foundations of Language* 9:135–142.

Campbell, Lyle. 1986. Testing phonology in the field. In *Experimental phonology*, ed. John J. Ohala and Jeri J. Jaeger, 163–186. Orlando: Academic Press.

Chao, Yuen-Ren. 1934. The non-uniqueness of phonemic solutions of phonetic systems. *Bulletin of the Institute of History and Philology (Academia Sinica)* 4:363–397.

Chomsky, Noam. 1955. The logical structure of linguistic theory. Ms., Harvard University and MIT. Revised version published by Plenum, New York, 1975.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge: MIT Press.

Chomsky, Noam. 1986. *Barriers*. Linguistic Inquiry monographs. Cambridge: MIT Press.

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

Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English*. Cambridge: MIT Press.

Chomsky, Noam, and George A. Miller. 1963. Introduction to the formal analysis of natural languages. In *Handbook of mathematical psychology*, ed. R. Duncan Luce, Robert R. Bush, and Eugene Galanter, II.269–321. New York: Wiley.

Church, Kenneth W., and William A. Gale. 1991. A comparison of the enhanced Good-Turing and deleted estimation methods for estimating probabilities of English bigrams. *Computer Speech and Language* 5:19–54.

Clayton, Mary L. 1976. The redundance of underlying morpheme-structure conditions. *Language* 52:295–313.

Clements, George N., and Samuel Jay Keyser. 1983. *CV phonology: A generative theory of the syllable.* Cambridge: MIT Press.

Clements, George N., and Engin Sezer. 1982. Vowel and consonant disharmony in Turkish. In *The structure of phonological representations*, ed. Harry van der Hulst and Norval Smith, 213–254. Dordrecht: Foris.

Cleveland, William S., and Susan J. Devlin. 1988. Locally weighted regression: An approach to regression analysis by local fitting. *Journal of the American Statistical Association* 83:596–610.

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.

Cohen, Paul. 1970. The tensing and raising of short *a* in the metropolitan area of New York City. Master's thesis, Columbia University.

Coleman, John, and Janet Pierrehumbert. 1997. Stochastic phonological grammars and acceptability. In *Proceedings of the 3rd meeting of the ACL Special Interest Group in Computational Phonology*, ed. John Coleman, 49–56. Somerset, NJ: Association for Computational Linguistics.

Coltheart, Max, Eddy J. Davelaar, Jon T. Jonasson, and Derek Besner. 1977. Access to the internal lexicon. In *Attention and performance VI*, ed. Stanislav Dornic, 535–555. Hillsdale, NJ: Lawrence Erlbaum.

Cortes, Corinna, and Vladimir Vapnik. 1995. Support-vector networks. *Machine Learning* 20:273–297.

Côté, Marie-Hélène. 2000. Consonant cluster phonotactics: A perceptual approach. Doctoral dissertation, MIT.

Davidsen-Nielsen, Niels. 1974. Syllabification of English words with medial *sp*, *st*, *sk*. *Journal of Phonetics* 2:15–45.

Davidson, Lisa. 2005. Addressing phonological questions with ultrasound. *Clinical Linguistics and Phonetics* 19:619–633.

Davidson, Lisa. 2006a. Phonology, phonetics, or frequency: Influences on the production of non-native sequences. *Journal of Phonetics* 34:104–137.

Davidson, Lisa. 2006b. Phonotactics and articulatory coordination interact in phonology: Evidence from non-native production. *Cognitive Science* 30:837–862.

Davidson, Lisa. 2010. Phonetic bases of similarities in cross-language production: Evidence from English and Catalan. *Journal of Phonetics* 38:272–288.

Davidson, Lisa, and Rolf Noyer. 1997. Local phonology in Huave: Nativization and the ranking of faithfulness constraints. In *Proceedings of the 15th West Coast Conference on Formal Linguistics*, 65–80. Stanford: CSLI.

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

Davis, Stuart, and Michael Hammond. 1995. On the status of onglides in American English. *Phonology* 12:159–182.

Dell, François. 1973. *Les règles et les sons.* Paris: Hermann.

Dell, François, and Mohamed Elmedlaoui. 1985. Syllabic consonants and syllabification in Imdlawn Tashlhiyt Berber. *Journal of African Languages and Linguistics* 7:105–130.

Downing, Laura J. 2000. Morphological and prosodic constraints on Kinande verbal reduplication. *Phonology* 17:1–38.

Dresher, B. Elan. 2009. *The contrastive hierarchy in phonology*. Cambridge: Cambridge University Press.

Duanmu, San. 2009. *Syllable structure: The limits of variation.* Oxford: Oxford University Press.

Dupoux, Emmanuel, Kazuhiko Kakehi, Yuki Hirose, Christophe Pallier, and Jacques Mehler. 1999. Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance* 25:1568–1578.

Elfner, Emily. 2006. Contrastive syllabification in Blackfoot. In *Proceedings of the 25th West Coast Conference on Formal Linguistics*, 141–149. Somerville, MA: Cascadilla.

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.

Emmorey, Karen. 1989. Auditory morphological priming in the lexicon. *Language and Cognitive Processes* 4:73–92.

Ettlinger, Marc. 2008. Input-driven opacity. Doctoral dissertation, University of California, Berkeley.

Ferguson, Charles A. 1975. 'Short *a*' in Philadelphia English. In *Studies in linguistics in honor of George L. Trager*, ed. M. Estellie Smith, 259–274. The Hague: Mouton.

Fikkert, Paula. 1994. On the acquisition of prosodic structure. Doctoral dissertation, Leiden University.

Fischer-Jørgensen, Eli. 1952. On the definition of phoneme categories on a distributional basis. *Acta Linguistica* 7:8–39.

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

Forster, Kenneth I., and Tamiko Azuma. 2000. Masked priming for prefixed words with bound stems: Does *submit* prime *permit*? *Language and Cognitive Processes* 15:539–561.

Fowler, Carol A. 1987. Consonant-vowel cohesiveness in speech communication as revealed by initial and final consonant exchanges. *Speech Communication* 6:231–244.

Fowler, Carol A., Rebecca Treiman, and Jennifer Gross. 1993. The structure of English syllables and polysyllables. *Journal of Memory and Language* 32:115–140.

Frauenfelder, Ulrich H., R. Harald Baayen, Frauke M. Hellwig, and Robert Schreuder. 1993. Neighborhood density and frequency across languages and modalities. *Journal of Memory and Language* 32:781–804.

Friederici, Angela D., and Jeanine E. Wessels. 1993. Phonotactic knowledge and its use in infant speech perception. *Perception and Psychophysics* 54:287–295.

Fries, Charles C., and Kenneth L. Pike. 1949. Coexistent phonemic systems. *Language* 25:29–50.

Frisch, Stefan A., Nathan R. Large, and David B. Pisoni. 2000. Perception of wordlikeness: Effects of segment probability and length on the processing of nonwords. *Journal of Memory and Language* 42:481–496.

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

Frisch, Stefan A., and Bushra A. Zawaydeh. 2001. The psychological reality of OCP-Place in Arabic. *Language* 77:91–106.

Fudge, Erik C. 1969. Syllables. *Journal of Linguistics* 5:253–286.

Gallagher, Gillian. In press. Speaker awareness of non-local ejective phonotactics in Cochabamba Quechua. *Natural Language and Linguistic Theory* to appear.

Gamble, Geoffrey. 1989. Spanish loans in Wikchamni. In *General and Amerindian ethnolinguistics: In remembrance of Stanley Newman*, ed. Mary R. Key and Henry M. Hoenigswald, 123–128. Berlin: Mouton de Gruyter.

Gibson, Edward, and James Thomas. 1999. Memory limitations and structural forgetting: The perception of complex ungrammatical sentences as grammatical. *Language and Cognitive Processes* 14:225–248.

Gigerenzer, Gerd, Stefan Krauss, and Oliver Vitouch. 2004. The null ritual: What you always wanted to know about significance testing but were afraid to ask. In *Sage handbook of quantitative methodology for the social sciences*, ed. David Kaplan, 391–408. Thousand Oaks, CA: Sage.

Goldsmith, John. 1976. Autosegmental phonology. Doctoral dissertation, MIT.

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

Good, Irving J. 1953. The population frequencies of species and the estimation of population parameters. *Biometrika* 40:237–264.

Goodman, Leo A., and William H. Kruskal. 1954. Measures of association for cross-classification. *Journal of the American Statistical Association* 49:732–764.

Gorman, Kyle. In press. Exceptions to rhotacism. Paper presented at CLS 48 (to appear in the proceedings).

Gouskova, Maria. 2004. Relational hierarchies in Optimality Theory: The case of syllable contact. *Phonology* 21:201–250.

Graff, Peter, and T. Florian Jaeger. In press. Locality and feature specificity in OCP effects: Evidence from Aymara, Dutch, and Javanese. Paper presented at the CLS 45 (to appear in the proceedings).

Greenberg, Joseph H., and James J. Jenkins. 1964. Studies in the psychological correlates of the sound system of American English, I and II. *Word* 20:157–177.

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

Gropen, Jess, Steven Pinker, Michelle Hollander, Richard Goldberg, and Ronald Wilson. 1989. The learnability and acquisition of the dative alternation in English. *Language* 65:203–257.

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

Gruber, Robert P., and Richard A. Block. 2005. Effects of caffeine on prospective duration judgements of various intervals depend on task difficulty. *Human Psychopharmacology* 20:275–285.

Guy, Gregory R. 1980. Variation in the group and the individual: The case of final stop deletion. In *Locating language in time and space*, ed. William Labov, 1–35. New York: Academic Press.

Göksel, Aslı, and Celia Kerslake. 2005. *Turkish: A comprehensive grammar*. New York: Routledge.

Hale, Kenneth. 1965. Some preliminary observations on Papago morphophonemics. *International Journal of American Linguistics* 31:295–305.

Hale, Mark, and Charles Reiss. 2008. *The phonological enterprise*. Oxford: Oxford University Press.

Halle, Morris. 1959. *The sound pattern of Russian*. The Hague: Mouton.

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

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

Hammond, Michael. 1997. Vowel quantity and syllabification in English. *Language* 73:1–17.

Hammond, Michael. 1999. *The phonology of English*. Oxford: Oxford University Press.

Hamp, Eric P. 1953. Morphophonemes of the Keltic mutations. *Language* 27:230–247.

Hare, Richard M. 1955. Theology and falsification. In *New essays in philosophical theology*, ed. Anthony Flew and Alasdair MacIntyre, 99–103. New York: Macmillan.

Harley, Heidi. 2009. A morphosyntactic account of the 'Latinate' ban on dative shift in English. Paper presented at the University of California, Santa Cruz.

Harris, John. 1994. *English sound structure*. Cambridge: Blackwell.

Harrison, K. David, and Abigail Kaun. 2001. Patterns, pervasive patterns and feature specification. In *Distinctive feature theory*, ed. T. Alan Hall, 211–237. Berlin: Mouton de Gruyter.

Haugen, Einer. 1956. Syllabification in Kutenai. *International Journal of American Linguistics* 22:196–201.

Havlicek, Larry L., and Nancy L. Peterson. 1976. Robustness of the Pearson correlation against violations of assumptions. *Perceptual and Motor Skills* 43:1319–1334.

Hay, Jennifer, Janet Pierrehumbert, and Mary E. Beckman. 2004. Speech perception, well-formedness and the statistics of the lexicon. In *Phonetic interpretation: Papers in Laboratory Phonology VI*, ed. John Local, Richard Ogden, and Rosalind A.M. Temple, 58–74. Cambridge: Cambridge University Press.

Hayes, Bruce. 1980. A metrical theory of stress rules. Doctoral dissertation, MIT.

Hayes, Bruce. 2000. Gradient well-formedness in Optimality Theory. In *Optimality Theory: Phonology, syntax, and acquisition*, ed. Joost Dekkers, Frank van der Leeuw, and Jeroen van de Weijer, 88–120. Oxford: Oxford University Press.

Hayes, Bruce. 2004. Phonological acquisition in Optimality Theory: The early stages. In *Constraints in phonological acquisition*, ed. René Kager, Joe Pater, and Wim Zonneveld, 158–203. Cambridge: Cambridge University Press.

Hayes, Bruce, and James White. In press. Phonological naturalness and phonotactic learning. *Linguistic Inquiry* to appear.

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

Hertwig, Ralph, and Andreas Ortmann. 2001. Experimental practices in economics: A methodological challenge for psychologists? *Behavioral and Brain Sciences* 24:383–403.

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

Hill, Archibald A. 1954. Juncture and syllable division in Latin. *Language* 30:439–447.

Hoard, James E. 1975. On incorporating Grassmann's Law into Sanskrit phonology. In *Proceedings of the 1st annual meeting of the Berkeley Linguistics Society*, 207–220. Berkeley: Berkeley Linguistics Society.

Hoard, James E., and Clarence Sloat. 1971. The inflectional morphology of English. *Glossa* 5:47–56.

Hockett, Charles F. 1958. *A course in modern linguistics*. London: Macmillan.

Hoenigswald, Henry M. 1965. A property of "Grassmann's Law" in Indic. *Journal of the American Oriental Society* 85:59–60.

Holden, Kyril. 1976. Assimilation rates of borrowing and phonological productivity. *Language* 52:131–147.

Hooper, Joan. 1973. Aspects of natural generative phonology. Doctoral dissertation, University of California, Los Angeles.

Howard, Irwin. 1972. A directional theory of rule application in phonology. Doctoral dissertation, MIT.

Hualde, Jose Ignacio. 1988. Affricates are not contour segments. In *Proceedings of the 7th West Coast Conference on Formal Linguistics*, 143–157. Stanford: Stanford Linguistics Association.

Huang, James. 1982. Logical relations in Chinese and the theory of grammar. Doctoral dissertation, MIT.

Inkelas, Sharon, Gunnar Ó. Hansson, Aylin Küntay, and C. Orhan Orgun. 2001. Labial attraction in Turkish: An empirical perspective. *Turkic Languages* 5:169–197.

Inkelas, Sharon, Aylin Küntay, C. Orhan Orgun, and Ronald Sprouse. 2000. Turkish Electronic Living Lexicon (TELL). *Turkic Languages* 4:253–275.

Inkelas, Sharon, C. Orhan Orgun, and Cheryl Zoll. 1997. The implications of lexical exceptions for the nature of grammar. In *Derivations and constraints in phonology*, ed. Iggy Roca, 393–418. Oxford: Clarendon Press.

Isac, Daniela, and Charles Reiss. 2008. *I-language: An introduction to linguistics as cognitive science.* Oxford: Oxford University Press.

Itô, Junko. 1989. A prosodic theory of epenthesis. *Natural Language and Linguistic Theory* 7:217–259.

Itô, Junko, and Armin Mester. 1995a. The core-periphery structure of the lexicon. In *University of Massachusetts Occasional Papers 18: Papers in Optimality Theory*, ed. Jill N. Beckman, Laura W. Dickey, and Suzanne Urbanczyk, 181–209. Amherst, MA: GLSA.

Itô, Junko, and Armin Mester. 1995b. Japanese phonology. In *The handbook of phonological theory*, ed. John Goldsmith, 817–838. Oxford: Blackwell.

Iverson, Gregory K., and Catherine Ringen. 1978. Exception devices in phonology. In *Proceedings of the 12th International Congress of Linguists*, ed. Wolfgang U. Dressler and Wolfgang Meid, 780–785. Innsbruck: University of Innsbruck.

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

Jakobson, Roman, Gunnar Fant, and Morris Halle. 1961. *Preliminaries to speech analysis: The distinctive features and their correlates.* Cambridge: MIT Press.

Johnson, C. Douglas. 1972. *Formal aspects of phonological description.* The Hague: Mouton.

Jones, Daniel. 1964. *An English pronouncing dictionary.* Cambridge: Heffer, 9th edition.

Jusczyk, Peter W., and Richard N. Aslin. 1995. Infant's detection of the sound patterns of words in fluent speech. *Cognitive Psychology* 46:65–97.

Jusczyk, Peter W., Angela D. Friederici, Jeanine E. Wessels, Vidgis Svenkerud, and Ann Jusczyk. 1993. Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language* 32:402–420.

Jusczyk, Peter W., and Elizabeth A. Hohne. 1997. Infants' memory for spoken words. *Science* 277:1984–1986.

Jusczyk, Peter W., Paul A. Luce, and Jan Charles-Luce. 1994. Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language* 33:630–645.

Jusczyk, Peter W., Paul Smolensky, and Theresa Allocco. 2002. How English-learning infants respond to markedness and faithfulness constraints. *Language Acquisition* 10:31–37.

Jäger, Gerhard. 2007. Maximum entropy models and Stochastic Optimality Theory. In *Architectures, rules, and preferences: Variations on themes by Joan W. Bresnan*, ed. Annie Zaenen, Jane Simpson, Tracy H. King, Jane Grimshaw, Joan Maling, and Chris Manning, 467–479. Stanford: CSLI.

Kabak, Barış. 2007. Hiatus resolution in Turkish: An underspecification account. *Lingua* 117:1378–1411.

Kabak, Barış, and William J. Idsardi. 2007. Perceptual distortions in the adaptation of English consonant clusters: Syllable structure or consonantal contact constraints? *Language and Speech* 50:23–52.

Kabak, Barış, Kazumi Maniwa, and Nina Kazanina. 2010. Listeners use vowel harmony and word-final stress to spot nonsense words: A study of Turkish and French. *Journal of Laboratory Phonology* 1:207–224.

Kahn, Daniel. 1976. Syllable-based generalizations in English phonology. Doctoral dissertation, MIT. Published by Garland, New York, 1980.

van Kampen, Anja, Güliz Parmaksiz, Ruben van de Vijver, and Barbara Höhle. 2008. Metrical and statistical cues for word segmentation: Vowel harmony and word stress as cues to word boundaries by 6- and 9-month-old Turkish learners. In *Language acquisition and development: Proceedings of GALA 2007*, 313–324. Newcastle: Cambridge Scholars Publishing.

Kaplan, Ronald, and Martin Kay. 1994. Regular models of phonological rule systems. *Computational Linguistics* 20:331–378.

Kaun, Abigail. 1999. Epenthesis-driven harmony in Turkish. In *Proceedings of the 25th annual meeting of the Berkeley Linguistics Society*, 95–106. Berkeley: Berkeley Linguistics Society.

Kavitskaya, Darya, and Peter Staroverov. 2008. Opacity in Tundra Nenets. In *Proceedings of the 27th West Coast Conference on Formal Linguistics*, 274–282. Somerville, MA: Cascadilla.

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. Stanford: CSLI.

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

Kaye, Jonathan. 1982. Harmony processes in Vata. In *The structure of phonological representations*, ed. Harry van der Hulst and Norval Smith, 385–452. Dordrecht: Foris.

Kaye, Jonathan. 1996. Do you believe in magic? The story of s+C sequences. In *A festschrift for Edmund Gussmann*, ed. Henryk Kardela and Bogdan Szymanek, 155–176. Lublin: Lublin University Press.

Kaye, Jonathan, and Jean Lowenstamm. 1985. A non-linear treatment of Grassmann's Law. In *Proceedings of the 15th meeting of North East Linguistic Society*, 220–233. Amherst, MA: GLSA.

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

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. Salford, UK: European Studies Research Institute, University of Salford.

Kenstowicz, Michael, and Charles Kisseberth. 1977. *Topics in phonological theory*. New York: Academic Press.

Kenstowicz, Michael, and Charles Kisseberth. 1979. *Generative phonology: Description and theory*. San Diego: Academic Press.

Kenstowicz, Michael, and Hyang-Sook Sohn. 2001. Accentual adaptation in North Kyungsang Korean. In *Ken Hale: A life in language*, ed. Michael Kenstowicz, 239–270. MIT Press.

Kessler, Brett, and Rebecca Treiman. 1997. Syllable structure and the distribution of phonemes in English syllables. *Journal of Memory and Language* 37:295–311.

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

Kiparsky, Paul. 1965. Phonological change. Doctoral dissertation, MIT.

Kiparsky, Paul. 1968. *How abstract is phonology?* Bloomington: Indiana University Linguistics Club.

Kiparsky, Paul. 1985. Some consequences of lexical phonology. *Phonology Yearbook* 2:85–138.

Kirby, James P., and Alan C.L. Yu. 2007. Lexical and phonotactic effects on wordlikeness judgements in Cantonese. In *Proceedings of the International Congress of the Phonetic Sciences XVI*, 1389–1392.

Kisseberth, Charles. 1970. On the functional unity of phonological rules. *Linguistic Inquiry* 1:291–306.

Klausenburger, Jurgen. 1976. (De)morphologization in Latin. *Lingua* 40:305–320.

Kuryłowicz, Jerzy. 1948. Contribution à la théorie de la syllabe. *Bulletin de la Société Polonaise de Linguistique* 8:80–114.

Labov, William. 1975. *What is a linguistic fact?* Lisse: Peter de Ridder.

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

Lass, Roger. 1971. Boundaries as obstruents. *Journal of Linguistics* 7:15–30.

Launey, Michel. 2011. *An introduction to Classical Nahuatl.* New York: Cambridge University Press.

Lees, Robert B. 1966a. On the interpretation of a Turkish vowel alternation. *Anthropological Linguistics* 8:32–39.

Lees, Robert B. 1966b. Turkish harmony and the description of assimilation. *Türk Dili Araştırmaları Yıllığı Belletene* 1966:279–297.

Levelt, Clara, Niels O. Schiller, and Willem J.M. Levelt. 2000. The acquisition of syllable types. *Language Acquisition* 8:237–264.

Lightner, Theodore M. 1970. On regular inflectional endings in English. *Papers in Linguistics* 5:503–519.

Lignos, Constantine, and Kyle Gorman. In press. Revisiting frequency and storage in morphological processing. Paper presented at CLS 48 (to appear in the proceedings).

Lipinski, John, and Prahlad Gupta. 2005. Does neighborhood density influence repetition latency for nonwords? Separating the effects of density and duration. *Journal of Memory and Language* 52:171–192.

Lombardi, Linda. 1990. The nonlinear organization of the affricate. *Natural Language and Linguistic Theory* 8:375–425.

Luce, Paul A., and Nathan R. Large. 2001. Phonotactics, density, and entropy in spoken word recognition. *Language and Cognitive Processes* 16:565–581.

Luelsdorff, Philip A. 1969. On the phonology of English inflection. *Glossa* 3:39–48.

Luka, Barbara J., and Lawrence W. Barsalou. 2005. Structural facilitation: Mere exposure effects for grammatical acceptability as evidence for syntactic priming in comprehension. *Journal of Memory and Language* 52:436–459.

MacEachern, Margaret. 1999. *Laryngeal co-occurrence restrictions.* New York: Garland.

Mandel, Denise R., Peter W. Jusczyk, and David B. Pisoni. 1995. Infants' recognition of the sound patterns of their own names. *Psychological Science* 6:314–317.

Marr, David. 1982. *Vision.* New York: Freeman.

Marslen-Wilson, William D., and Alan Welsh. 1978. Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology* 10:29–63.

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.

Massaro, Dominic W., and Michael M. Cohen. 1983. Phonological constraints in speech perception. *Perception and Psychophysics* 34:338–348.

Matthews, Peter H. 1972. *Inflectional morphology: A theoretical study based on aspect of Latin conjugation.* Cambridge: Cambridge University Press.

Mattys, Sven L., and Peter W. Jusczyk. 2001. Phonotactic cues for the segmentation of fluent speech by infants. *Cognition* 78:91–121.

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

McCarthy, John J. 2003. Comparative markedness. *Theoretical Linguistics* 29:1–51.

McCawley, James. 1973. The role of notation in generative phonology. In *The formal analysis of natural languages*, ed. Maurice Gross, Morris Halle, and Marcel-Paul Schutzenberger, 51–62. The Hague: Mouton.

McGowan, Kevin B. In press. Gradient lexical reflexes of the syllable contact law. Paper presented at CLS 45 to appear in the proceedings.

McMahon, April. 1990. Vowel shifts, free rides and strict cyclicity. *Lingua* 80:197–225.

McQueen, James M. 1998. Segmentation of continuous speech using phonotactics. *Journal of Memory and Language* 39:21–46.

Mester, Armin. 1988. *Studies in tier structure.* New York: Garland.

Miner, Kenneth L. 1975. English inflectional endings and unordered rules. *Foundations of Language* 12:339–365.

Moreton, Elliott. 2002. Structural constraints in the perception of English stop-sonorant clusters. *Cognition* 84:55–71.

Moulton, William G. 1947. Juncture in Modern Standard German. *Language* 23:212–226.

Murray, Robert W., and Theo Vennemann. 1983. Sound change and syllable structure in Germanic phonology. *Language* 59:514–528.

Mutaka, Ngessimo, and Larry M. Hyman. 1990. Syllables and morpheme integrity in Kinande reduplication. *Phonology* 7:73–119.

Myers, Scott. 1987. Vowel shortening in English. *Natural Language and Linguistic Theory* 5:485–518.

Nevins, Andrew, and Bert Vaux. 2003. Metalinguistic, shmetalinguistic: The phonology of shm-reduplication. In *Papers from the 39th meeting of the Chicago Linguistic Society*, 702–721. Chicago: Chicago Linguistic Society.

Newman, Stanley. 1944. *Yokuts language of California.* New York: Viking Fund Publications in Anthropology.

Newmeyer, Frederick J. 2007. Commentary on Sam Featherston, 'Data in generative grammar: The stick and the carrot'. *Theoretical Linguistics* 33:395–399.

Nida, Eugene A. 1948. The identification of morphemes. *Language* 24:414–441.

Noether, Gottfried E. 1981. Why Kendall tau? *Teaching Statistics* 3:41–41.

Norris, Dennis, James M. McQueen, Anne Cutler, and Sally Butterfield. 1997. The possible-word constraint in the segmentation of continuous speech. *Cognitive Psychology* 34:191–243.

Ní Chiosáin, Máire, and Jaye Padgett. 1993. On the nature of consonant-vowel interaction. Technical Report LRC-93-09, Linguistics Research Center, University of California, Santa Cruz.

Odden, David. 1986. On the role of the Obligatory Contour Principle in phonological theory. *Language* 62:353–383.

Odden, David. 2012. Formal phonology. *Nordlyd* 39:24–43.

Ohala, John J., and Manjari Ohala. 1986. Testing hypotheses regarding the psychological manifestation of morpheme structure constraints. In *Experimental phonology*, ed. John J. Ohala and Jeri J. Jaeger, 239–252. Orlando: Academic Press.

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

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

Pan, Ning, and William Snyder. 2003. Setting the parameters of syllable structure in early child Dutch. In *Proceedings of the 27th annual Boston University Conference on Language Development*, 615–625. Somerville, MA: Cascadilla.

Pan, Ning, and William Snyder. 2004. Acquisition of /s/-initial clusters: A parametric approach. In *Proceedings of the 28th annual Boston University Conference on Language Development*, 436–446. Somerville, MA: Cascadilla.

Pande, Hemlata, and Hoshiyar S. Dhami. 2010. Mathematical modeling of occurrence of letters and word initials in texts in Hindi. *SKASE Journal of Theoretical Linguistics* 7:19–38.

Payne, Arvilla. 1980. Factors controlling the acquisition of the Philadelphia dialect by out-of-state children. In *Locating language in time and space*, ed. William Labov, 143–178. New York: Academic Press.

Peperkamp, Sharon. 2005. A psycholinguistic theory of loanword adaptation. In *Proceedings of the 30th meeting of the Berkeley Linguistics Society*, ed. Marc Ettlinger, Nicholas Flesicher, and Mischa Park-Doob, 341–352. Berkeley: Berkeley Linguistics Society.

Peperkamp, Sharon, and Emmanuel Dupoux. 2003. Reinterpreting loanword adapations: The role of perception. In *Proceedings of the 15th International Congress of Phonetic Sciences*, 367–370. Adelaide: Causal Productions.

Phelps, Elaine. 1975. Sanskrit diaspirates. *Linguistic Inquiry* 6:447–464.

Phelps, Elaine, and Michael Brame. 1973. On local ordering of rules in Sanskrit. *Linguistic Inquiry* 14:387–400.

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

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: Cambridge University Press.

Piggott, Glyne L. 1975. More on the application of phonological rules. *Recherces linguistiques à Montréal / Montreal Working Papers in Linguistics* 5:113–150.

Pinker, Steven, and Alan Prince. 1988. On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. In *Connections and Symbols*, ed. Steven Pinker and Jacques Mehler, 73–193. Cambridge: MIT Press.

Pitt, Mark A., and James M. McQueen. 1998. Is compensation for coarticulation mediated by the lexicon? *Journal of Memory and Language* 39:347–370.

Polka, Linda, and Janet F. Werker. 1994. Developmental changes in perception of non-native vowel contrasts. *Journal of Experimental Psychoology: Human Perception and Performance* 20:421–435.

Postal, Paul. 1968. *Aspects of phonological theory*. New York: Harper and Row.

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

Prince, Alan, and Paul Smolensky. 1993. Optimality Theory: Constraint interaction in generative grammar. Technical Report TR-2, Rutgers University Center for Cognitive Science. Published by MIT Press, Cambridge, 2004.

Pulgram, Ernst. 1970. *Syllable, word, nexus, cursus*. The Hague: Mouton.

Pylkkänen, Liina, Andrew Stringfellow, and Alec Marantz. 2002. Neuromagnetic evidence for the timing of lexical activation: An MEG component sensitive to phonotactic probability but not to neighborhood density. *Brain and Language* 81:666–678.

Pylyshyn, Zenon. 1984. *Computation and cognition: Towards a foundation for cognitive science*. Cambridge: MIT Press.

Rice, Keren. 2009. Athapaskan: Slave. In *The Oxford handbook of compounding*, ed. Rochelle Lieber and Pavol Stekauer, 542–573. Oxford: Oxford University Press.

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

Rosch, Eleanor. 1975. Cognitive representations of semantic categories. *Journal of Experimental Psychology: General* 104:192–233.

Rose, Sharon, and Lisa King. 2007. Speech error elicitation and cooccurrence restrictions in two Ethiopian Semitic languages. *Language and Speech* 50:451–504.

Rusaw, Erin, and Jennifer Cole. 2009. Learning constraints that oppose native phonotactics from brief experience. Paper presented at the Mid-Continental Workshop on Phonology.

Safarewicz, Jan. 1932. *Le rhotacisme latin.* Wilno: Towarzystwo Przyjaciół Nauk.

Sag, Ivan A. 1974. The Grassmann's Law ordering pseudoparadox. *Linguistic Inquiry* 5:591–607.

Sag, Ivan A. 1976. Pseudosolutions to the pseudoparadox: Sanskrit diaspirates revisited. *Linguistic Inquiry* 7:609–622.

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

Sapir, Edward. 1930. Southern Paiute, a Shoshonean language. *Proceedings of the American Academy of Arts and Sciences* 65:1–269.

Saporta, Sol. 1955. Frequency of consonant clusters. *Language* 31:25–30.

Saporta, Sol, and Donald Olson. 1958. Classification of intervocalic clusters. *Language* 34:261–266.

de Saussure, Ferdinand. 1916. *Cours de linguistique générale.* Paris: Éditions Payot & Rivages.

Schindler, Jochem. 1976. Diachronic and synchronic remarks on Bartholomae's and Grassmann's laws. *Linguistic Inquiry* 7:622–637.

Scholes, Robert J. 1966. *Phonotactic grammaticality.* Berlin: Mouton.

Schütze, Carson T. 1996. *The empirical base of linguistics: Grammaticality judgments and linguistic methodology.* Chicago: University of Chicago Press.

Schütze, Carson T. 2011. Linguistic evidence and grammatical theory. *Wiley Interdisciplinary Reviews: Cognitive Science* 2:206–221.

Seidl, Amanda, and Elizabeth K. Johnson. 2006. Infant word segmentation revisited: Edge alignment facilitates target extraction. *Developmental Science* 9:565–573.

Selkirk, Elisabeth O. 1982. The syllable. In *The structure of phonological representations*, ed. Harry van der Hulst and Norval Smith, 337–385. Dordrecht: Foris.

Shademan, Shabnam. 2006. Is phonotactic knowledge grammatical knowledge? In *Proceedings of the 25th West Coast Conference on Formal Linguistics*, 371–379. Somerville, MA: Cascadilla.

Shademan, Shabnam. 2007. Grammar and analogy in phonotactic well-formedness. Doctoral dissertation, University of California, Los Angeles.

Shattuck-Hufnagel, Stefanie. 1986. The representation of phonological information during speech production planning: Evidence from vowel errors in spontaneous speech. *Phonology Yearbook* 3:117–149.

Shibatani, Masayoshi. 1972. The phonological representations of English inflectional endings. *Glossa* 6:117–127.

Shibatani, Masayoshi. 1973. The role of surface phonetic constraints in generative phonology. *Language* 49:87–106.

Sigurd, Bengt. 1968. Rank-frequency distributions for phonemes. *Phonetica* 18:1–15.

Silverman, Daniel. 2000. Dynamic versus static constraints in English truncation. In *Proceedings of the 19th West Coast Conference on Formal Linguistics*, 466–478. Somerville, MA: Cascadilla.

Smith, Nelson V. 1973. *The acquisition of phonology: A case study*. Cambridge: Cambridge University Press.

Smolensky, Paul. 1996. On the comprehension/production dilemma in child language. *Linguistic Inquiry* 7:720–731.

Sohn, Ho-Min. 1971. *a*-raising in Woleaian. *University of Hawaii Working Papers in Linguistics* 3:15–36.

Sprouse, Jon. 2007. Continuous acceptability, categorical grammaticality, and experimental syntax. *Biolinguistics* 1:118–129.

Sprouse, Jon. 2011. A test of the cognitive assumptions of magnitude estimation: Commutativity does not hold for acceptability judgements. *Language* 87:274–288.

Sprouse, Jon, and Diogo Almeida. Submitted. Power in acceptability judgement experiments and the reliability of data in syntax. Ms., University of California, Irvine, and New York University, Abu Dhabi.

Stampe, David. 1973. *A dissertation on natural phonology*. New York: Garland.

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

Stemberger, Joseph P. 1980. Another look at Grassman's Law. *Glossa* 14:113–135.

Steriade, Donca. 1995. Underspecification and markedness. In *The handbook of phonological theory*, ed. John Goldsmith, 114–174. Cambridge: Blackwell.

Stevens, Stanley S. 1946. On the theory of scales of measurement. *Science* 103:677–680.

Suomi, Kari, James M. McQueen, and Anne Cutler. 1997. Vowel harmony and speech segmentation in Finnish. *Journal of Memory and Language* 36:422–444.

Suárez, Lidia, Seok Hui Tan, Melvin J. Yap, and Winston D. Goh. 2011. Observing neighborhood effects without neighbors. *Psychonomic Bulletin and Review* 18:605–611.

Taft, Marcus. 1979. Recognition of affixed words and the word frequency effect. *Memory and Cognition* 7:263–272.

Taft, Marcus, and Sam Ardasinski. 2006. Obligatory decomposition in reading prefixed words. *The Mental Lexicon* 1:183–199.

Taft, Marcus, and Kenneth I. Forster. 1975. Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior* 14:638–647.

Taft, Marcus, and Kenneth I. Forster. 1976. Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behavior* 15:607–620.

Taft, Marcus, Gail Hambly, and Sachiko Kinoshita. 1986. Visual and auditory recognition of prefixed words. *Quarterly Journal of Experimental Psychology* 38A:351–366.

Tambovtsev, Yuri, and Colin Martindale. 2007. Phoneme frequencies follow a Yule distribution. *SKASE Journal of Theoretical Linguistics* 4:1–11.

Tincoff, Ruth, and Peter W. Jusczyk. 1999. Some beginnings of word comprehension in 6-month-olds. *Psychological Science* 10:172–175.

Trager, George L. 1930. The pronunciation of 'short *a*' in American Standard English. *American Speech* 5:396–400.

Trager, George L. 1934. What conditions limit variants of a phoneme? *American Speech* 9:313–315.

Trager, George L. 1940. One phonemic entity becomes two: The case of 'short *a*'. *American Speech* 17:30–41.

Treiman, Rebecca. 1983. The structure of spoken syllables: Evidence from novel word games. *Cognition* 15:49–74.

Treiman, Rebecca. 1986. The division between onsets and rimes in English syllables. *Journal of Memory and Language* 25:476–491.

Treiman, Rebecca, Carol A. Fowler, Jennifer Gross, Denise Berch, and Sarah Weatherston. 1995. Syllable structure or word structure? Evidence for onset and rime units with disyllabic and trisyllabic stimuli. *Journal of Memory and Language* 34:132–155.

Treiman, Rebecca, Brett Kessler, Stephanie Knewasser, Ruth Tincoff, and Margo Bowman. 2000. English speakers' sensitivity to phonotactic patterns. In *Papers in Laboratory Phonology V: Acquisition and the lexicon*, ed. Michael Broe and Janet Pierrehumbert, 269–282. Cambridge: Cambridge University Press.

Treiman, Rebecca, and Andrea Zukowski. 1990. Towards an understanding of English syllabification. *Journal of Memory and Language* 29:66–85.

Trubetzkoy, Nikolai S. 1958. *Grundzüge der Phonologie*. Göttigen, Germany: Vandenhoeck & Ruprecht, 2nd edition.

Ussishkin, Adam, and Andrew Wedel. 2003. Gestural motor programs and the nature of phonotactic restrictions: Evidence from loanword phonology. In *Proceedings of the 22nd West Coast Conference on Formal Linguistics*, 505–518. Somerville, MA: Cascadilla.

Vaden, Kenneth, Harry R. Halpin, and Gregory S. Hickok. 2009. Irvine phonotactic online dictionary. URL `http://www.iphod.com`.

Vaux, Bert. 1993. The origins of Altaic labial attraction. *Harvard Working Papers in Linguistics* 2:228–237.

Vaux, Bert. 2003. Syllabification in Armenian, Universal Grammar, and the lexicon. *Linguistic Inquiry* 34:91–125.

Vaux, Bert. 2011. Language games. In *The handbook of phonological theory*, ed. John Goldsmith, Jason Riggle, and Alan C.L. Yu, 722–750. Oxford: Wiley-Blackwell, 2nd edition.

Vitevitch, Michael S., and Paul A. Luce. 1998. When words compete: Levels of processing in perception of spoken words. *Psychological Science* 9:325–329.

Vitevitch, Michael S., and Paul A. Luce. 1999. Probabilistic phonotactics and neighborhood density in spoken word recognition. *Journal of Memory and Language* 40:374–408.

Vitevitch, Michael S., and Paul A. Luce. 2005. Increases in phonotactic probability facilitate spoken nonword repetition. *Journal of Memory and Language* 52:193–204.

Vitevitch, Michael S., Paul A. Luce, Jan Charles-Luce, and David Kemmerer. 1997. Phonotactics and syllable stress: Implications for the processing of spoken nonsense words. *Language and Speech* 40:47–62.

Vogt, Hans. 1954. Phoneme classes and phoneme classification. *Word* 10:28–34.

Vroomen, Jean, Jyrki Tuomainen, and Beatrice de Gelder. 1998. The roles of word stress and vowel harmony in speech segmentation. *Journal of Memory and Language* 38:133–149.

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

Weiss, Mary. 1961. Über die relative Häufigkeit der Phoneme des Schwedischen. *Statistical Methods in Linguistics* 1:41–55.

Wells, John C. 1982. *Accents of English*. Cambridge: Cambridge University Press. 3 volumes.

Werker, Janet F., John H.V. Gilbert, Keith Humphrey, and Richard C. Tees. 1981. Developmental aspects of cross-language speech perception. *Child Development* 52:349–355.

Werker, Janet F., and Chris E. Lalonde. 1988. Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology* 24:672–683.

Werker, Janet F., and Richard C. Tees. 1984. Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7:49–63.

White, Katherine S., Sharon Peperkamp, Cecilia Kirk, and James L. Morgan. 2008. Rapid acquisition of phonological alternations by infants. *Cognition* 107:238–265.

Whitney, William D. 1889. *Sanskrit grammar*. Cambridge: Harvard University Press, 2nd edition.

Wolf, Matthew. 2011. Limits on global rules in Optimality Theory with Candidate Chains. *Phonology* 28:87–128.

Wolf, Matthew, and John J. McCarthy. 2009. Less than zero: Correspondence and the null output. In *Modeling ungrammaticality in Optimality Theory*, ed. Curt Rice and Sylvia Blaho, 17–66. London: Equinox.

Yip, Moira. 1989. Feature geometry and cooccurrence restriction. *Phonology* 6:349–374.

Zimmer, Karl E. 1969. Psychological correlates of some Turkish morpheme structure conditions. *Language* 45:309–321.

Zipf, George K. 1949. *Human behavior and the principle of least effort: An introduction to human ecology*. Cambridge: Addison-Wesley.

Zonneveld, Wim. 1978. *A formal theory of exceptions in generative phonology*. Lisse: Peter de Ridder.

Zuraw, Kie. 2003. Optimality Theory in linguistics. In *Handbook of brain theory and neural networks*, ed. Michael Arbib, 819–822. Cambridge: MIT Press, 2nd edition.

Zwicky, Arnold M. 1965. Topics in Sanskrit phonology. Doctoral dissertation, MIT.

Zwicky, Arnold M. 1970. The free-ride principle and two rules of complete assimilation in English. In *Proceedings of the 6th annual meeting of the Chicago Linguistic Society*, 579–588. Chicago: Chicago Linguistic Society.

Zwicky, Arnold M. 1975. Settling on an underlying form: The English inflectional endings. In *Testing linguistic hypotheses*, ed. David Cohen and Jessica R. Wirth, 129–185. New York: Wiley.

Zwicky, Arnold M., and Geoffrey K. Pullum. 1987. Plain morphology and expressive morphology. In *Proceedings of the 13th annual meeting of the Berkeley Linguistics Society*, 330–340. Berkeley: Berkeley Linguistics Society.