Synchronic explanation

Paul de Lacy
Rutgers, The State University of New Jersey

John Kingston
University of Massachusetts Amherst

- This article was published in *Natural Language and Linguistic Theory* in April 2013
- The final publication is available at: http://www.springer.com/education+%26+language/linguistics/journal/11049
- Received: 7 March 2009 / Accepted: 7 December 2011 / Published: April 2013
- This version has been self-archived by the authors, following Springer's Self-Archiving Policy.

Abstract

The aim of this article is to show that synchronic cognitive constraints are responsible for some restrictions on human speech sound patterns; not all markedness asymmetries can be ascribed to Performance-based mechanisms of diachronic change. We identify evidence for synchronic constraints in sound patterns that are desirable from a Performance perspective yet are not attested. We also discuss recent experiments that provide evidence for psychologically and even neurophysiologically active restrictions; these patterns can be distinguished from statistical generalizations across the lexicon. We also argue that there is evidence that language learning and adult well-formedness judgments are determined by innate predispositions. Finally, we examine the methodology behind choosing a synchronic or diachronic account for a particular sound pattern when both potentially offer an explanation.

Keywords: Competence, deduction, diachronic, induction, markedness, Performance, synchronic

1 Introduction

In an extreme view, all human speech sound patterns are due to restrictions on diachronic actuation and transmission. In such an approach, the phonological component is able to output any structure respecting the formal properties of its objects and relations. The only reason that languages show systematic similarities is because there are particular grammars that are not learnable or are unstable in diachronic transmission. This approach to explanation relies on restrictions on learnability and language acquisition, particularly language transmission between individuals and generations that reflect how sounds are perceived and subsequently articulated; hence, it can be called 'diachronic explanation'.

The most extensive and sustained advocacy of diachronic explanation can be found in John Ohala's work (see e.g., Ohala 1974, 1981, 1983, 1987, 1993, 1995, 2005). Barnes

(2002) and Blevins (2004) present more recent expositions. We acknowledge the contributions of Ohala's and Barnes' work, but in order to sufficiently focus this article we will concentrate on Blevins' (2004) 'Evolutionary Phonology', as it advances the most sustained argument for replacing synchronic explanations of sound patterns with diachronic ones (also see Blevins 2006a, 2006b, 2008). For example, Blevins (2004:27) states, "There is no need to encode the primary content of phonological representations and constraint systems in the human mind, since these properties can be shown to emerge from natural processes of language change inherent in the transmission of language from one generation to the next." In any case, we intend the central points we make about Evolutionary Phonology to apply equally to other advocates of a diachronic perspective that seeks to eliminate synchronic grammatical constraints on languages' phonologies.

The alternative to theories that advocate diachronic explanation is that there are non-trivial cognitive restrictions in the phonological component of speakers' and listeners' grammars. In this view, the lack of certain sound patterns is due to the phonological component's inability to generate them; we will call this 'synchronic explanation'.

We argue for synchronic explanation: i.e., that there are synchronically active restrictions on the phonological component. We examine two types of evidence. One is grammars (or parts of grammars) that are learnable, but never generated. Section 2 focuses on place of articulation in neutralization and epenthesis. We argue that synchronic neutralization to and epenthesis of [k] (and dorsals generally) is unattested, yet is desirable for Performance reasons and thus expected if synchronic sound patterns have exclusively diachronic explanations.

The other type of evidence, discussed in Section 3, involves demonstrations of active synchronic restrictions. This evidence comes in four forms. First, the phonetic motivations of sound changes persist long after those changes have been phonologized. Second, sound changes can be optimizing in ways that suggest that they interact with an active synchronic grammar. Third, synchronic constraints actively regulate linguistic behavior and do so in ways that indicate they are a thing apart from generalizations across the lexicon or from any other aspect of the speaker's or listener's experience of their native language. Finally, language acquisition itself is apparently constrained by innate constraints on possible grammars.

To make one thing clear, we do not advocate an extreme 'synchronic explanation' position whereby every facet of every sound pattern is due to restrictions in the phonological component. The role for diachronic explanation is explored in Section 4, and argued to account for the typological frequency of grammars and sound patterns. For example, the fact that of the voiced stops the dorsal [g] is the least frequently attested cross-linguistically is not something that the phonological component necessarily should account for: it must be able to generate grammars both with and without [g]. Instead, Performance issues such as confusability in perception and difficulty in articulation are responsible for [g]'s relative low frequency. We then address the long-standing issue of whether a grammar's lack of attestation is due to synchronic restrictions or diachronic pressures; we argue that in many cases both are responsible.

2 Learnable grammars that cannot be generated

There are mismatches between diachronically desirable changes and synchronically attested sound patterns. Section 2.1 focuses on types of neutralization and epenthesis that are attested diachronically but never synchronically. Section 2.2 argues that a series of natural diachronic changes can easily lead to synchronically unattested grammars, a point recently made by Kiparsky (2006, 2008). In line with Kiparsky's work, we argue that synchronically active phonological restrictions are needed to account for such cases, as attested sound changes would otherwise produce them. In short, these restrictions show that languages' synchronic grammars are not merely the residue of the sound changes they have undergone, but actively limit the results of diachronic change.

2.1 Absolute restrictions

The following discussion will focus on two generalizations: (1) No language has epenthetic [k], and (2) No language has place neutralization in stops that yields [k]. These generalizations have been argued to extend to all dorsal segments (de Lacy 2006a); we focus on [k] for expositional purposes.

We use 'epenthesis' to refer to a situation where an output segment has no corresponding input segment. Specifically, in rule-based theories an epenthetic segment *x* is one that is the result of a rule $\varnothing \rightarrow x$; in Optimality Theory with Correspondence Theory (Prince and Smolensky 2004; McCarthy and Prince 1999), an epenthetic segment *x* is one that is not in a Correspondence relation with any input segment.

We use 'place neutralization' here to refer to a situation where a segment's input place of articulation is altered; specifically, no language has place neutralization in stops that yields [k]. In input→output terms, there is no input segment that differs in major Place of Articulation (PoA) from [k] (i.e., is not dorsal) and surfaces unfaithfully as [k], unless some incidental process such as assimilation or dissimilation interferes.[1] Restricting attention to syllable-final PoA neutralization, in output-only terms there are no wordforms with the shapes […Vk-CV…] and […Vt-V…] where the first morpheme is the same in both forms.

Neither of the assertions about [k] is new (e.g., Trubetzkoy 1939; Jakobson 1941; Lombardi 2002; de Lacy 2002, and many others). The focus in this section is on whether they are predicted to occur through mechanisms of diachronic change. Section 2.3 examines the validity of the generalizations as some recent work has cast doubt on them.

[1] Examples that do not count as 'neutralization' here are those involving contextual assimilation or dissimilation. For example, Bradley (2006) reports that the An^{ʔ}di^{ʔ}êm dialect of Katu allows only dorsals before coronal laterals (e.g., [klâm] 'urinate', [gluh] 'go out') while the Phúhòa dialect has coronals in these locations (cf. [tlâm], [tluh]); Mong Njua is cited as having free variation between dorsals and coronals in this environment (e.g., [k^{h}la]~[t^{h}la] 'to run; jump') (also see Rice 1996 and de Lacy 2006:370 for a similar case in South Vietnamese). These are cases where the OCP bans adjacent identical place specifications; the influence of context puts them outside of the scope of 'neutralization' as used here.

2.1.1 Synchronic explanation

There are a number of synchronic theories that can account for a lack of epenthetic [k]. For example, an Optimality Theoretic constraint system with an output markedness constraint *[dorsal] and no constraint that favors dorsals over other places of articulation prevents epenthetic [k]. Putting aside the effects of assimilation and dissimilation, tableau (1) demonstrates this impossibility, using the theory of markedness constraints presented in de Lacy (2002, 2004, 2006a).[2] The motivation for epenthesis is the constraint ONSET, which requires a syllable-initial consonant; it interacts with faithfulness constraints in the ranking ONSET, MAX » DEP. No candidate violates any IDENT constraint as faithfulness to input features is irrelevant in epenthesis; output constraints will fully determine output feature specification. Consequently, the only constraints that matter here are those that distinguish between different places of articulation.

(1)

/a/	*{dorsal}	*{dorsal,labial}	*{dorsal,labial,coronal}
(a) k̲a	*!	*	*
(b) p̲a		*!	*
☞ (c) t̲a			*

The candidate with [k] epenthesis (a) is harmonically bounded by the candidate with [t] epenthesis (c): there is no ranking of the constraints that produces epenthetic [k]; all rankings of these constraints favor the candidate with [t] epenthesis (Samek-Lodovici and Prince 1999). The crucial condition that makes this result valid is that there is no constraint that favors dorsals over all other places of articulation. While labials are not our focus here, the observations about [k] can also be seen for [p]: epenthetic [p] is harmonically bounded by [t], meaning that it cannot be epenthetic either; we discuss labial epenthesis and neutralization further below.

The added complexity with Place of Articulation is that glottals incur even fewer violations than coronals: i.e., [ʔ] does not violate any of the constraints in (1). Consequently, some languages have epenthetic [ʔ]. However, [ʔ] can be eliminated by other constraints (e.g., a ban on high sonority syllable margins; de Lacy 2002, 2006a), effectively promoting coronals to least marked status as the epenthetic segment in some languages.

The same point can be made for neutralization. With an input /ap/ and a constraint that forces /p/ to surface unfaithfully, there is no ranking of the constraints above that will force /p/ to become [k]. Again, /p/→[t] is a harmonic bound for /p/→[k].

[2] We use de Lacy's theory of markedness constraints here for convenience. The same point can be made with equal force using a Fixed Ranking theory (Prince and Smolensky 2004) or using constraints that relate structural complexity to markedness. All of these theories are able to produce systems in which [k] cannot be epenthetic.

(2)

/ap/	*{dorsal}	*{dorsal,labial}	*{dorsal,labial,coronal}
(a) ap		*!	*
(b) ak	*!	*	*
☞ (c) at			*

So, given a choice of epenthetic [k] or [t] and neutralization to [k] or [t], [t] will always win.[3] It is uncontroversial that synchronic theories are capable of implementing such a restriction. Other proposals include using fixed rankings of constraints (Lombardi 2002) and lack of specification of coronal place in an autosegmental approach (see Paradis and Prunet 1991 and references cited therein).

2.1.2 Diachronic explanation

Are phonological constraints like those just presented necessary? If languages with [k] epenthesis or neutralization to [k] are unlearnable or very unlikely to survive diachronic transmission intact, epenthetic [k] and neutralization to [k] would be unattested for purely Performance reasons. Exactly this alternative is advocated by Blevins (2004):

> "If we can demonstrate that principled diachronic explanations exist for particular sound patterns, considerations of simplicity would seem to dictate that explanations for the same phenomena should not be imported into, or otherwise duplicated within, synchronic accounts. In all cases where clear diachronic patterns exist for a particular synchronic pattern, this diachronic explanation makes a synchronic account redundant, since the optimal description should not account for the same pattern twice…A central premise of Evolutionary Phonology, then, is that principled diachronic explanations for sound patterns replace, rather than complement, synchronic explanations, unless independent evidence demonstrates, beyond reasonable doubt, that a separate synchronic account is warranted." (Blevins 2004:5)

In this alternative, the phonological component could generate epenthetic [k] and neutralization to [k] (or to any other PoA, see Hume and Tserdanelis 2002, Hume 2003)—i.e., there could be freely rankable constraints such as *dorsal, *labial, *coronal. The lack of attestation of epenthetic [k] and neutralization to [k] must instead be ascribed to mechanisms involved in diachronic change.

However, there is evidence that languages with [k] epenthesis and neutralization to [k] are desirable from a Performance point of view.

Epenthesis of [k] could develop in language change through misperception (the mechanism of sound change that Blevins refers to as CHANGE, see also Ohala 1981).

---

[3] To guarantee that /p/ will not neutralize to [k], restrictions on other markedness constraints are necessary (specifically that there is no markedness constraint that favors [k] over [t]). To ensure that /p/ maps to [t] and not to some other element (e.g., [f]), certain rankings of faithfulness constraints are also necessary. The role of faithfulness constraints in determining the output of neutralization is discussed in de Lacy (2006a: 253ff, 783ff).

Suppose a learner misperceives a vowel-vowel transition as having an intervening consonantal constriction. Exactly how this misperception happens is not of interest here—the fact that consonant epenthesis occurs means that (in a diachronic explanation) there must be some Performance factor that motivates the learner to store the form with an inserted segment. What is of interest is how the learner decides which consonant to insert.

Some languages have stop epenthesis. Famously, Axininca Campa has epenthetic [t] (Payne 1981; Spring 1990). A [t] is inserted at a variety of vowel-vowel junctures, exemplified in (3). Epenthetic consonants are underlined; reduplicants are double-underlined.

(3) Axininca Campa [t] epenthesis (Payne 1981)

(a) Root+suffix juncture

/i-N-√koma-i/ → [iŋkomati] ‘he will paddle’

cf. /i-N-√tʃ^{h}ik-i/ →[iɲtʃ^{h}iki] ‘he will cut’

(b) Suffix+suffix juncture

/i-N-√tʃ^{h}ik-aː-i/ → [iɲtʃ^{h}ikaːti] ‘he will cut again’

/√na-RED-wai-ak-i/ → [nata-nata-waitaki] ‘I will continue to carry it’

(c) Minimal word augmentation

/√t^{h}o/ → [t^{h}ota] ‘kiss, suck’

cf. /non-t^{h}o-RED/ →[nont^{h}onont^{h}o]

Why does Axininca Campa insert [t] and not [p], [tʃ], or [k]? From a transmissibility point of view, the answer must be that [t] has some perceptual or articulatory property that makes it more ‘desirable’ than these other options. Concretely, the property is possibly acoustic: coronals coarticulate less with surrounding vowels than labials and dorsals (Sussman, McCaffrey, and Matthews 1991; Sussman, Hoemeke, and Ahmed 1993; Fruchter and Sussman 1997; Sussman, Fruchter, Hilbert, and Sirosh 1998), and are therefore more readily separated perceptually from the flanking vowels. As such, they provide a clear break between vowels (assuming that the motivation for inter-vocalic epenthesis is to minimize the perceptual overlap between successive vowels).

Whatever the reason for [t] epenthesis, the issue is why [k] is not chosen in Axininca Campa, and in fact is *never* chosen in any language. There is a good deal of evidence that [k] has properties that can make it more desirable than [t] from a Performance point of view. In diachronic change, there are many languages in which *t has become [k]: Lynch et al. (2002:54) note that “across the languages of the world the sound change *t* to *k* is hugely more common than *k* to *t*”. This diachronic change is found in Hawai’ian, Luangiua, colloquial Samoan and several other Oceanic languages (Blust 1990; Lynch et al. 2002:Ch.4), as well as Fort Chipewyan Chipewyan (Haas 1968). The *t > k change happened in all phonological environments in the languages cited. For example, the Proto-Eastern Polynesian word for ‘man, people’ is *taŋata; in Hawai’ian, it is [kanaka] (Clark 1976; Pukui and Elbert 1986).

Blevins (2004: Section 5.4) discusses motivations for the *t > k change, observing that [k] has the longest VOT of all stops, and so proposes that “velars will sound, to the child, like good tokens of Category 1 stops [i.e., stops with long VOT and high amplitude]". There is an added subtlety here: Blevins observes that cases of *t > k have occurred when the parent language has no contrast between [t] and [k] (i.e., the change is diachronically non-neutralizing). This is probably irrelevant to their choice in epenthesis environments, especially because even languages without a [t]~[k] contrast do not epenthesize [k] (see discussion below).

An alternative motivation for [k] epenthesis lies in the fact that dorsals' articulation—and consequently their acoustics—vary with the backness of adjacent vowels, while for labials and especially coronals, the vowels' articulations and acoustics vary with the consonants' place instead (Stevens and Blumstein 1978; Blumsten and Stevens 1979, 1980; Sussman et al. 1991, 1993, 1998). These facts show why a dorsal would be preferred over a coronal or a labial in a language where place contrasts are deteriorating: the dorsal is phonetically more placeless than the others. Learners would opt for it instead of the other places if what they perceived were simply the presence of an oral closure without strong place cues. In other words, failing [ʔ], [k] is the most placeless consonant available; for a learner wanting an epenthetic stop but wishing to deviate least from the perceived speech signal, [k] is an excellent option.[4,5] It is important to point out that the fact that [ʔ] is 'better' than [k] in placelessness does not mean that [k] can never be epenthesized. Many languages do not have [ʔ] (see discussion in de Lacy 2006a), so in these languages [k] would be the best available placeless consonant.

In any case, the diachronic change of *t > k indicates that there is some Performance aspect of [k] that favours it over [t]. Therefore, [k] could be a better choice than [t] in epenthesis. As Blevins (2004: 128) states, "Coronal segments may have unique properties, but so do labials and dorsals." Therefore, there is no phonetic reason why epenthesis should consistently discriminate against [k].

A similar argument can be made for neutralization to [k]. Suppose a learner heard a morpheme in two different environments: [at#] and [a.t-o]. The learner might confuse [t] with [k] in the word-final environment with the consequence that there is now a pair of related word forms [ak] and [a.t-o] in his/her grammar. If the misperception was general enough so that every [t] in an onset corresponded to a [k] in a coda (while some onset [k]s corresponded to coda [k]s), in input→output terms the grammar would have to generate /t/→[k] neutralization in codas.

For neutralization, then, can [t] be misperceived as [k] solely in a limited environment such as the coda (or at least word-finally)? This is evidently the case in Peruvian Spanish. Pre-consonantal stops in loanwords are realized as [k]: e.g., [peksi] < 'Pepsi', [xikler] < 'Hitler'.[6] José Elías Ulloa (p.c.), a native speaker of the dialect, reports that a diachronic

---

[4] The motivations for [t] epenthesis and [k]/glottal epenthesis must be different. We suggested that the benefit in [t] epenthesis is that it minimizes vocalic overlap by having a consonant that is coarticulated least with the surrounding vowels; the motivation therefore means that the learner's desire to make the vowels as perceptually separate as possible overrides the desire to accurately reproduce the perceived speech signal. In glottal and potentially [k] epenthesis, the opposite holds. Glottals (particularly [h]) are significantly affected by their environment, and so is [k], but do not alter its acoustics. So, an epenthetic glottal or [k] would separate vowels perceptually, yet allow minimal deviation from the perceived speech signal by allowing the vowel articulations to continue through them.

[5] For the idea that the learner favors alternants that differ minimally in their perceptual properties, see Steriade (2001).

[6] Cristófaro-Silva (2003) shows that [k] replaces [t] in [tl] clusters in on-going sound changes in Brazilian Portuguese, e.g., *atleta* 'athlete' can be pronounced [akleta]. She argues that this and other changes in

change of pre-consonantal [p] to [k] is also partially in place: many people do not distinguish between [akto] 'act' and [apto] 'apt', with both realized as [akto]. Similarly, words like *abstracto* 'abstract' and *optional* 'optional' are usually pronounced as [akstrakto] and [oksjonal]. Finally, Elías Ulloa observes that when speakers of his dialect learn English as a second language, words like 'street' tend to be pronounced as [estri], but word-final [t] can be realized as [k]: i.e., [estrik]. All of these cases can be ascribed to misperception of [t] as [k] in pre-consonantal and word-final environments, showing that *t > k is possible in just these positions. However, Peruvian Spanish does not have alternations that show /t/→[k]; there are no wordform pairs where a morpheme has [t] in an onset and [k] in a coda.

The same point can be made for phonological change in Chinese dialects. Chen (1973) reports that oral and nasal coronal stops developed into dorsals in several daughter languages, as in the change from Middle Chinese to Classical Fuzhou. However, there is no evidence to show that there was any synchronic system in which underlying /t/ mapped to surface [k].

If misperception is responsible for *t > k in codas, then one would expect a situation in which a child learned a morpheme [at] as [ak], but retained the [t] in pre-vocalic environments [at-o], resulting in a synchronic /t/→[k] coda neutralization. However, no such system is attested.

In contrast, synchronic alternations involving coda neutralization to [t] do occur (e.g., Basque codas have /ogi-ʔapur/→[ot.ʔapur] 'bread-crumb'; Hualde 1991). Neutralization to [t] occurs in restricted environments such as reduplicants, in Taiwanese (e.g., /k-k'ak-RED/→[lak-k'i̱t] 'crack open' cf. /be-RED/→[le-bi], *[le-bik]; Li 1985) and Cantonese (/l-ç̧ap-RED/→[lap-ç̧i̱t]; Yip 1982). For Korean, Kim-Renaud (1986) reports that even /h/ surfaces as [t] in codas: e.g., /tʃoːh-ko/→[tʃoːt-kʰo][7] 'good and', cf. [tʃoːh-ʉni] 'as (it's) good'.

In short, the argument here is that diachronic evidence shows that [k] has at least some properties that are more desirable than [t] in terms of Performance; evidence from loanword adaptation and second language acquisition agrees. If Performance properties are what influences sound patterns, then one would expect epenthesis of [k] and neutralization to [k]. However, neither phenomenon is attested. This, then, is a Competence-Performance mismatch: a sound pattern is favoured by Performance factors, but the phonological component is unable to generate a grammar that reflects those Performance pressures.

An important issue lurks in the preceding discussion: the majority of epenthetic elements are glides and glottals and most neutralization involves debuccalization (i.e., neutralization to glottals). Is the lack of [k] epenthesis therefore accidental, due to the infrequency of stop epenthesis generally? Languages with [t] epenthesis are few—Axininca Campa is a clear case; other languages include Māori (de Lacy 2003 and references cited therein) and Odawa Ojibwa (Piggott 1993; Lombardi 2002); some other languages have coronal epenthesis (of [n] and rhotics).[8] In short, [t] epenthesis is not the

---

tautosyllabic clusters are motivated by a Performance factor, namely, the greater token frequency of [kl] over [tl].

[7] The [spread glottis] specification of the /h/ survives as aspiration on a following stop.

[8] Cases of [r] and [l] epenthesis in English dialects resemble glide epenthesis in that the liquid that's inserted is articulatorily similar to the vowel that precedes it (Gick 1999, 2002a,b).

most common type, but still attested. However, the lack of [k] epenthesis is still telling: there are situations where languages seem poised to have [k] epenthesis, but shy away from doing so, discussed in Section 2.2.

2.2 Many diachronic rights never make a synchronic wrong

Synchronically active phonological restrictions are needed to curb diachronic change when it would produce a synchronically impossible grammar. After all, it is an easy matter to identify a series of natural diachronic changes that would produce an unattested and undesirable language. The following argument builds on Kiparsky's (2006, 2008) work (also see de Lacy 2006b).

Kiparsky's (2008) argument refers to sonority-driven stress in Gujarati, described and analyzed in de Lacy (2002, 2006a). Stress is usually penultimate (e.g., [apwána] 'to give', [ekótеɾ] 'seventy-one'). However, it will fall on an [a] elsewhere if the penult is not [a] (e.g., [tádʒetəɾ], *[tadʒétəɾ] 'recently'). De Lacy (2002) proposes synchronic constraints that favour more sonorous stressed vowels over less sonorous ones, where the vowel sonority hierarchy is | a 〉 e,o 〉 i,u 〉 ə |. A diachronically-based account would not appeal to such a markedness hierarchy, and would have no related constraints. Instead, Kiparsky suggests that in a diachronic account the "intrinsic acoustic prominence of sonorous vowels may be reinterpreted as stress in sound change" (Kiparsky 2008:50). Suppose a language has stress on the word-final syllable. If a listener heard a word such as [pakí], [a]'s greater intensity and duration might mislead the listener into thinking that it bore stress: i.e., [páki]. In contrast, in a word such as [piká], [i] has less intensity and duration than [a], so there is no motivation for the listener to misperceive stress on [i]; the same applies to [pikí]. The resulting language would have 'sonority-driven stress' by having final stress in words like [pikí] and [piká], but non-final stress in words like [páki].

However, a natural sound change could easily change the stress facts. For example, Kiparsky observes that *a could become [ə] in all environments—stressed and unstressed—as it did in a sound change from Sanskrit. If this change happens in a daughter language of Gujarati, *[tádʒetəɾ] would become [tə́dʒetəɾ]. The problem with this form is that in synchronic terms stress has retracted to an antepenultimate schwa even though there is a more sonorous vowel—i.e., [e] —in penultimate position. In short, a natural sound change could result in a language that reverses the markedness relation between [ə] and mid vowels on the sonority hierarchy.

Notice also that such a language would be expected if grammars are simply the consequences of the sound changes the language has undergone: stress was on the antepenult before [a] was replaced by schwa and remains there after that sound change occurs. Descriptively, the distribution of stress is now lexical rather than determined by vowel sonority, but that fact is of no consequence in a theory of sound structure like Evolutionary Phonology where synchronic patterns are entirely a product of the sound changes the language has undergone. This sort of outcome is apparently even predicted by Evolutionary Phonology:

> "Natural sound patterns are precisely those with well-documented phonetic origins, and are transparent reflections of the phonologization of phonetically based sound

> change. Unnatural sound patterns may be the result of direct inheritance, or have one of the sources in (1) [analogy, rule inversion, rule telescoping, accidental convergence: deL&K]. Synchronic constraints incorporating naturalness or markedness are misguided, since whether a sound pattern is natural, crazy, or somewhere in between, is wholly a function of its history." (Blevins 2004:71)

Even though such a stress system could easily be inherited, it does not occur (see de Lacy 2004 for a typology).

Diachronic mechanisms therefore cannot account for the resistance to changes that would affect the sonority of the vowel without also affecting its ability to bear stress. That is, the diachronic explanation could allow the *a > ə change without necessarily requiring stress to shift back to the penult, so losing the sonority condition.

In contrast, the Competence theories presented in Kenstowicz (1997) and de Lacy (2004) make it impossible to construct a grammar in which schwa attracts stress away from a more sonorous vowel. These theories therefore predict that a sound change like *a > ə in a language with sonority-driven stress will necessarily alter the stress in words that have undergone the change: i.e., the *a > ə change must be simultaneous with the change in stress position to [tədʒétəɾ]—there is no stage in the language's history which would have [tə́dʒetəɾ].[9]

The sound changes discussed above—where *t > k—create a similar opportunity for a language to develop [k] epenthesis or neutralization to [k] in Eastern Polynesian languages. In the Eastern Polynesian language Māori, [t] is epenthesized to avoid a prosodic word (PrWd)-initial onsetless syllable in suffixation. An example is given in (4a). Evidence that the passive is /ia/ and gerund /aŋa/ is found in environments where the underlying form of the root has a final consonant (b) and where the suffixes can syllabify with the preceding vowel (c) or can form a foot (d).[10]

(4) [t] epenthesis in Māori, a conservative Central-Eastern Polynesian language

	UR	Isolation Form	Passive	Gerund
(a)	/mahue/ 'leave'	[mahue]	[mahue-t̠ia]	[mahue-t̠aŋa]
(b)	/hopuk/ 'catch'	[hopu]	[hopuk-ia]	[hopuk-aŋa]
(c)	/hika/ 'kindle'	[hika]	[hi.kai.a]	[hika-ŋa]
(d)	/toː/ 'drag'	[toː]	[toː-ia]	[toː-aŋa]

The passive and gerund undoubtedly existed in Proto-Polynesian: the morpheme has reflexes in all the major subgroups—Tongic, Samoic, and Eastern Polynesian (Pawley 1966; for extensive discussion supporting the proposed diachrony, see also Clark 1976: Section 3.2ff). (In some languages it marks ergative rather than passive.) All closely studied languages in the Polynesian family show allomorphy that is very similar to Māori's, so it is likely that epenthesis occurs in all these cases.

---

[9] A possibly more likely but uninteresting alternative to this development is that all [a]s become schwa, except those that are stressed. This alternative is uninteresting because the typical phonetic correlates of stress, greater duration and intensity, should prevent the shortening and quieting that would lead to a stressed [a] losing its sonority and becoming schwa.

[10] For the conditions surrounding passive formation in Māori and evidence for epenthesis, see de Lacy (2003), which builds on a great deal of previous work cited therein.

Proto-Central Eastern Polynesian (PCE)—the ancestor of Maori, Hawai'ian, and Tahitian—had *p, *t, and *k, and no *ʔ.[11] Consequently, PCE had a Māori-like situation: it had the passive and gerund, and very likely epenthesis of [t]. Subsequent developments of *k > ʔ occurred in Tahitian and Hawai'ian, and *t>k in Hawai'ian.

If PCE had [t] epenthesis, and *t became [k] in Hawai'ian, one would expect the epenthetic consonant to be [k] in Hawai'ian: i.e., PCE **alofat̲ia* should become Hawai'ian [alohak̲ia]. However, it in fact appears with a glottal stop: [alohaʔ̲ia] (Pukui and Elbert 1986; Elbert and Pukui 2001: Section 6.6.3). The conundrum is therefore why learners converted surface [t]s that had underlying /t/s to [k]s, while surface epenthetic [t]s were replaced by [ʔ]. The development of Hawai'ian presents a prime situation where [k] could be epenthetic, and should be, but is not.[12]

There is a straightforward synchronic explanation for the Hawai'ian situation: the phonological component is not capable of generating a grammar with epenthetic [k], as ruled out by the analysis in Section 2.1.1. The only option for the Hawai'ian speakers is to have epenthetic [ʔ].

In short, the development from Proto-Central Eastern Polynesian to Hawai'ian presented a situation where [k] epenthesis is expected. The fact that Hawai'ian has epenthetic [ʔ] instead of [k], whereas all non-epenthetic PCE [t]s became [k], indicates that the result of language change can be restricted by the synchronic grammar: if the phonological component is unable to create epenthetic [k], the fact that [ʔ] is the epenthetic consonant in Hawai'ian is explainable.

To be clear about the points made in the last two sections, we are not claiming that there is no role for diachronic explanation. Instead, we are asserting that no synchronic grammar can map underlying /t/ to surface [k], even though sound changes such as that which occurred in PCE can readily turn a *t into [k]. This then is a case of the kind Blevins (2004) argues does not and should not arise if a language's synchronic patterns are entirely a product of its history,

> "There is no need to encode the primary content of phonological representations and constraint systems in the human mind, since these properties can be shown to emerge from natural processes of language change inherent in the transmission of language from one generation to the next." (Blevins 2004:27)

This case also demonstrates the essential independence of sound changes from the regulation of synchronic grammars. Suppose that synchronic grammars rule out replacing or otherwise realizing /t/ as [k]; that is, no constraint ranking permits a dorsal output from a coronal input (or produces a dorsal output from scratch). Even so, sound change is not limited to only those diachronic changes that are synchronic input→output mappings. So,

---

[11] All daughters of PCE contrast three places of articulation for stops; the most common are [p t k]. PCE most likely had epenthesis because the suffix can be reconstructed back to Proto-Polynesian, and epenthesis in the passive and gerund is found in languages outside PCE.

[12] Our thanks to 'Oiwi Parker-Jones for his help with Hawai'ian. As in Māori, underlying final consonants surface: /malah-ia/→[malahia] 'ache', /paul-ia/→[paulia] 'finish', /inum-ia/→[inumia] 'drink'. When /ia/ can be incorporated into the same PrWd (i.e., with short bases), the [i] deletes post-vocalically: /ale-ia/→[aléa] 'swallow' (Elbert and Pukui 2001:84). Bases that end in /a/ and long bases force the passive into its own PrWd, so epenthesis occurs: /wela-ia/→[{wéla}{ʔ̲ía}] 'burn', /aloha-ia/→[{alóha}{ʔ̲ía}] 'love'.

even if /t/→[k] is impossible, this does not mean that *t > k is likewise impossible. Sound change motivations can encompass the kinds of misperceptions and reanalyses documented by Blevins and others. That is to say, a comprehensive account of how languages come to acquire the sound structures they do must encompass both a principled account of sound change of the kind embodied in Evolutionary Phonology and related work, and an equally principled and independent account of how the phonological component of the grammar is constituted synchronically. Despite the inherent appeal of the simplicity achieved by reducing one of these accounts to the other, the facts preclude such reduction, in either direction. We return to the division of labor between diachronic and synchronic explanation further in Section 4.

2.3 Evidence in epenthesis and neutralization

The preceding discussion is based on the proposal that [k] is never epenthetic and never the output of place neutralization. A broader claim is that epenthesis and neutralization never produce dorsals. This claim is not new, and has been supported by typologies in Lombardi (2002) and de Lacy (2002, 2006a). While epenthesis of and neutralization to glottal and coronal place of articulation are both well attested, no cases involving dorsals survive scrutiny. The aim here is not to refute every putative counter-example but to identify general themes that lead to misidentification of dorsals as either epenthetic or the output of neutralization.[13] In some cases the putative epenthetic segment is a morpheme; others involve suppletion, and in many cases of nasal neutralization the output is phonologically glottal but must be realized phonetically with velar or post-velar constriction because a nasal cannot be articulated any farther back.

We use ‘epenthetic’ here to refer to an output segment that does not correspond to any input segment. How can one tell if a segment is not present underlyingly? Canonical evidence is found in alternations. For example, suppose a root [pomo] and a suffix [a] form [pomota] when concatenated; it is possible that [t] is epenthetic here. An essential property of epenthetic segments is that they appear in a phonotactically well-defined environment. The [pomota] case fits the bill: the environment would be the onset.

However, it is essential to show that the non-appearance of the segment in the complementary environments is not due to deletion. For example, the underlying form corresponding to [pomota] could be /pomot-a/, with /pomot/ undergoing coda deletion under the pressure of NoCoda to form [pomo] in isolation. Alternatively, the /t/ could belong to a completely separate mono-segmental morpheme: i.e., /pomo-t-a/. The majority of such cases can be eliminated by examining morpho-syntactic restrictions on the putative epenthetic segment. If it is part of a morpheme, it will only ever appear in specific, well-defined morpho-syntactic environments; this follows because morphemes must satisfy the specific requirements of those environments. Consequently, if the [t] in [pomo-t-a] only ever occurs, say, before plural suffixes, it must belong to a morpheme

[13] Doing so would amount to taking up the notoriously difficult task of proving a negative. Instead, what we do here is demonstrate that many putative cases of dorsal epenthesis fail to meet the criteria for counting as epenthetic. These criteria are neither complex nor controversial, but rather surprisingly they have been ignored by those who have argued that dorsals are epenthetic just so long as they happen to occur in contexts that resemble those where genuine epenthetic segments do.

that has morpho-syntactic properties that require it to appear in the root+plural environment. In short, if a segment's distribution is severely morpho-syntactically restricted, it cannot be epenthetic. The exception to necessarily treating the segment as a morpheme would be if the particular morpho-syntactic environment presents the only *phonological* environment in which epenthesis could take place.

Severely restricted morpho-syntactic environments contrast with broad morpho-syntactic environments. Severely restricted morpho-syntactic environments involve a single morpheme, or a very small group of morphemes, often semantically defined. In contrast, phonological restrictions are known to be able to refer to morphological domains such as 'root' and 'stem', and to classes such as Class I and Class II affixes. Some have argued that such morphological domains have phonological counterparts. In any case, what we would call epenthesis could be restricted to such broadly defined morphological domains.

Another common diagnostic is unconditioned variability in the quality of the epenthetic segment. If a putative epenthetic [f] appears in the same phonological and morphological environments as a putative epenthetic [t], then at least one of the segments is not epenthetic. Many theories hold that epenthetic segments have features determined entirely by the constraint or rule system; consequently, epenthetic segments must have a consistent form, holding phonological and morphological environments constant.

Our claim is that there are no cases of epenthetic [k] in contexts where a dorsal articulation could not arise as the result of assimilation or dissimilation to its environment. Of course, epenthetic segments can be influenced by their surrounding environment: witness Lardil's /ɻil-a/→[ɻil-t̲-a] 'neck (nominative)' cf. /kaŋ-a/→[kaŋ-k̲-a] 'speech (nominative)', where the epenthetic stop agrees in place of articulation with the preceding segment (Hale 1973). In Dakota, epenthetic consonants are [j] before front (i.e., coronal) vowels [i e] and [w] before back (i.e., dorsal) vowels [a o u] (Shaw 1980:90). Similarly, the epenthetic consonant in Brahui assimilates to low back vowels in dorsality, voice, and continuancy, resulting in [ɣ] (e.g., [lumːa-ɣ̲-aːk] 'mother (masc.pl)') (Elfenbein 1997). A dissimilatory example is found in Seri. Marlett (1981, 2010) reports epenthetic [k]; however, it occurs only between a coronal and labial: e.g., /i-si-m-aai/ → [isk̲w̃ããĩ] '3:3-Independent Irrealis-Negative-make', /ma-t-m-aʔo/ → [matk̲w̃ãʔo] '2sg Direct Object-Dependent Realis-Negative-see'. Here, avoidance of homorganicity forces dorsality.

Finally, we reject non-alternating phonotactic evidence for epenthesis. If a language is observed to only have [k] in syllable codas, it could be assumed that underlying /t/ must neutralize to [k]. However, there are many other possible ways that languages can eliminate coda [t]s: e.g., deletion, coalescence, and lenition. Without alternations, it is extremely difficult to prove an input→output mapping /t/→[k]. The same goes for epenthesis: the observation that a language does not have surface vowel hiatus does not mean that any underlying hiatus situations are resolved by epenthesis. Even if it can be shown that a language has intervocalic [k] where its ancestor language had vowel hiatus, this is no indication that an earlier synchronic grammar permitted epenthetic [k] to repair hiatus. These points are developed at length in de Lacy (2006a).

Axininca Campa exemplifies epenthesis clearly. Its epenthetic [t] is found intervocalically in junctures between root+suffix and suffix+suffix, and to avoid sub-minimal words (see (3)). In short, the epenthetic [t] appears wherever it is needed phonologically, regardless of the morpho-syntax of its surroundings. A valid case of epenthetic [k] would have to meet the same standard: it must occur in a well-defined phonological environment, and not be restricted to a limited morpho-syntactic environment.

All putative examples of epenthetic [k] and neutralization to [k] that are known to us fail to meet the standards outlined above. In no case is it possible to demonstrate that the putative epenthesis is both phonotactically well defined and not severely morpho-syntactically restricted.

2.3.1 Mono-segmental affixes

An example of a mono-segmental affix that has the appearance of an epenthetic element is found in Koɖava, a Dravidian language. Ebert (1996:9) reports that “euphonic [k] is inserted between roots ending in a vowel or [n] and a following [a].” Examples are given in (5); [ɯ] is epenthesized after root-final consonants. Voiceless stops become voiced after nasals, so accounting for the ‘euphonic’ [g] in (d) and (e).

(5) Koɖava euphonic [k] (Ebert 1996)

(a) C-final stems		
/ʌɭɯd/ ‘write’	[ʌɭɯdɯ] ‘write’	[ʌɭɯdate] ‘don’t write’
(b) V-final stems		
/aɭa/ ‘sit’	[aɭa] ‘sit’	[aɭakate] ‘without sitting down’
/koɖɯ/ ‘give’		[koɖɯkate] ‘do not give!’
/kuɖi/ ‘drink’	[kuɖi] ‘drink’	[kuɖika] ‘let’s drink’
(c) /n/-final stems		
/tin/ ‘eat’		[tingadɯ] ‘let him eat’
/kan/ ‘greeting’		[kanga] ‘see you!’

The reason for treating [k] as epenthetic in (b) [aɭakate] is presumably that the bare root is realised as [aɭa], not *[aɭak], and the suffix is [ate] (not *[kate]), as can be seen in (a) [ʌɭɯd-ate]. It is clear that the [k] is not part of the root in [aɭakate] because underlying C-final roots undergo vowel epenthesis as in (a), while ‘sit’ in (b) is not *[aɭakɯ].

It is also possible to identify a phonotactically well-defined environment for [k] insertion: [k] is inserted to fill an onset. When a consonant is already available, [k] does not appear: [ʌɭɯd-ate], *[ʌɭɯd-kate], *[ʌɭɯdɯkate]. Roots ending in /n/ do not pose an insurmountable problem for this proposal: as in Lardil, there could be a condition that prefers the right edge of morphemes to align with the right edge of syllables, thus /kan-a/→[kan.ga], with [g] inserted to fill an otherwise empty onset. This condition is ignored in the case of /ʌɭɯd-ate/→[ʌɭɯdate] because *[ʌɭɯdkate] produces an unacceptable heterosyllabic cluster—only NC clusters are permitted.

However, Koɖava euphonic [k] cannot be epenthetic for two reasons. One is that it is morpho-syntactically restricted in its distribution: it can only appear between a verb root and specific suffixes. For example, /√kond-ɯn-avʌ̃/→[kondɯnavʌ̃], *[kondɯngavʌ̃] 'one who killed'; /√ʌɭɯd-ɯn-ʌ/→[ʌɭɯdɯnʌ], *[ʌɭɯdɯngʌ] 'I wrote'. This contrasts with Axininca Campa, for example, where epenthesis happens at any morphological boundary (including suffix+suffix), except at the prefix+root juncture, which is arguably due to a phonological sensitivity (i.e., alignment of PrWd with syllable edges) rather than a morphological one. The other reason is that in other hiatus situations glides are epenthesized: e.g., [elːi-j-uː] 'wherever', [boɳɖu-ʋ-aː] 'is it necessary?' (Ebert 1996:9).

If [k] is a morpheme, a good deal of sense can be made of its distribution. Like other morphemes, /k/'s distribution can be morpho-syntactically restricted to the verb root + suffix juncture. With [k] as a (perhaps semantically empty) morpheme, the only challenge is to explain why it does not appear after *every* verb stem. The answer is that [k] is deleted when its presence would violate the language's phonotactics. If realising a verb root segment and the [k] conflict, the verb root segment wins. For example, /ʌɭɯd-k-ate/ does not surface as [ʌɭɯdkate] because the cluster [dk] is banned in Koɖava (as is [dg]). In contrast, /k/ does surface in /tin-k-ad/→[tingadɯ] because homorganic NÇ clusters are permitted on the surface. Treating /k/ as a morpheme has ample precedent: many languages have semantically contentless 'thematic' morphemes (e.g., Attic Greek, Lupas 1972).

Chuukese (Trukese) is another case where mono-segmental morphemes give the appearance of being epenthetic [k]s. The causative prefix is a non-high vowel. It harmonizes in backness and lowness with the root vowel under complex conditions that are irrelevant here (Goodenough and Sugita 1980, 1990): see (6a). The preferential repair for vowel hiatus in Chuukese is merger to form a long vowel, as seen when a non-high vowel precedes the causative (6b), and when a non-high vowel follows the causative (6c). However, the causative cannot merge with a preceding *high* vowel, so a [k] appears (6d). Despite seemingly coming from nowhere, this [k] is not epenthetic because it is morpho-syntactically restricted: it appears only before the causative (and after the stative, discussed further below). In addition, if a high vowel or long vowel appears *after* the causative, a *glide*—not [k]—is epenthesized (6e).

(6) <u>The Chuukese causative (from Goodenough and Sugita 1980, 1990)</u>

(a) Before C(C)

[a-ssʌːr] 'cause to slide' [a-karara] 'cause to make a noise'

[æ-weːsi] 'cause to be finished' [o-kkufu] 'cause to loose'

(b) After a short non-high vowel → merger

/ki-ma-a-nʌːnʌ/ → [kimaːnʌːnʌ], *[kima<u>k</u>anʌːnʌ] 'guardian spirits'

/ma-a-wun/ → [mɔːwun] 'fight a war'

/sa-a-iːm^w/ → [saːjiːm^w] 'spontaneously-cause-pandanus leaf'

(c) Before a short non-high vowel → merger

/a-apa/ → [aːpa] 'move to one side'

/a-ʌtɨ/ → [aːtɨ] 'cause to be smoked'

/a-ɔnnut-a/ → [ɔːnnuta] 'put to sleep' (cf. [ɔnnut] 'sleep')

(d) After a short high vowel → [k]

[ni-k̲-a-ssʌːr] ‘game of sliding down a slope’

[ini-k̲-a-saf] ‘shoot-cause-vitiligo’

[ni-k̲-ɔ-p^{w}oːp^{w}o] ‘balloon’

(e) Before a high vowel or a long vowel → glide epenthesis

/a-it/ → [ajit] ‘make demonstration’

/a-eːw-in/ → [æjeːwin] ‘cause to enumerate’

/a-aː-j/ → [ajaːj] ‘cause to use/possess’

/a-aː-ŋeni/ → [ajaːŋeni] ‘cause use or possession to’

/a-oː-w/ → [ɔjoːw] ‘be caught’

The ‘reality’ aspect marker /a/ provides a near minimal pair with the causative (Goodenough and Sugita 1980:4). However, it behaves somewhat differently. After a short non-high front vowel, the aspect marker merges: e.g., /ke-a/→[kaː] (2sg), /e-a/→[aː] (3sg). In all other situations, there is glide epenthesis: /si-a/→[sija] (1incl), /wɨ-a/→[wɨwa] (1sg), /wo-a/→[wowa] (2pl). Unlike the causative, [k] does not make an appearance after a short high vowel: /si-a/→[sija], *[sika].

For other morphemes, in hiatus situations that cannot be resolved by merger, glide epenthesis occurs. For example, the honorific /i/ merges with a preceding high vowel (e.g., /æti-i-sɔm^w/→[ætiːsɔm^w] ‘man of chiefly rank’), but a homorganic glide is inserted after a non-high vowel: [nije-ji-sɔm^w] ‘woman of chiefly clan’, [fɛwɨ-wi-tʃo] ‘bicep’ (Goodenough and Sugita 1980:77). The same applies to the locative prefix /i/ (e.g., /i-e/→[ije] ‘here it is! (locative, emphatic)’). Finally, when a suffixing C_1VC_2 reduplicant has no base C_1 to copy, a glide is inserted: /eji-RED/→[eji-j-e̲j̲] ‘one hand of bananas’ (cf. /nɨkɨ-RED/→[nɨkɨ-nɨk] ‘be convinced’).

If the [k] in (6d) is not epenthetic, it must either be part of a suppletive allomorph of the causative (i.e., the causative has two morphs /a/ and /ka/, selected by the need to have an onset filled); or it could be the realization of a mono-segmental morpheme that subcategorizes for the causative’s environment.

The latter may be correct. We propose that the [k] that appears in (6d) is the realization of a morpheme /q/ that carries the meaning of ‘object focus’ (after Goodenough and Sugita 1980:xliv–xlv). The /q/ morpheme appears obligatorily with the two ‘object focus’ prefixes: the causative and the stative. Uvulars are banned in the output, so /q/ usually deletes: e.g., /q-a-ssʌʌr/ ‘*q*-causative-slide’→[assʌʌr], *[qassʌʌr]. However, when it is needed to avoid vowel hiatus it neutralizes to [k]: /ni-q-a-ssʌʌr/→[nikassʌʌr], *[niassʌʌr]. (The *q* morpheme cannot be underlyingly /k/ as [k]’s are perfectly acceptable in Chuukese outputs.)

A putative epenthetic [k] also occurs with the stative. However, we consider the stative prefix to consist of a moraic consonant /C^μ/ that fully assimilates to the root consonant: e.g., [posu] ‘stabbed’, cf. [p^μpos] ([pːos]) ‘be stabbed’ (Goodenough and Sugita 1980:xxiii; Davis 1999). The *q* morpheme follows it; when a root consonant is available the /q/ deletes as usual: /C^μ-q-posu/→[p^μ.pos]. However, when there is no root-initial consonant, the /q/ neutralizes to [k] and the stative assimilates to it as usual: /C^{μ}-q-æːpi/→[k^{μ}-k-æːp] 'transport it'.[14]

The restriction of [k] to occurring with object focus prefixes—with glides epenthesized elsewhere—indicates that it is a mono-segmental morpheme and not an epenthetic segment as Blevins (2008:105–106) proposes. Other languages cited as having epenthetic dorsals can also be reanalyzed as morphemes; the telltale sign is once again their restricted morpho-syntactic distribution.

2.3.2 Fixed segmentism in reduplication

Alderete et al. (1999) argue that reduplicants can pair with fully specified morphemes; the result is 'fixed segmentism' in reduplication: part of the reduplicant is constant. At first glance, some cases of fixed segmentism may appear to involve epenthesis. However, the apparent 'epenthetic' segments are always severely morpho-syntactically restricted, only occurring with one particular reduplicant.

For example, Howe (2004) identifies Sm'algyax as having a CVk-reduplicant: e.g., [ɬpun] 'whale' cf. [ɬʌk-ɬpun]. Alderete et al.'s (1999) approach would consider the fixed reduplicant to be a morpheme that subcategorizes for this particular morpho-syntactic environment: i.e., /RED-k-ɬpun/→[ɬʌ-k-ɬpun]. The [k] is not epenthetic here—it is present in the input.

Similarly, Murut has a [CVɣ] reduplicant (Prentice 1971:121): e.g., /RED-ɣ-aŋkup/→[gaɣaŋkup] (*no gloss*), /RED-ɣ-insilot/→[giɣinsilot] 'toothpick'. Notice that the /ɣ/ itself is reduplicated (it appears as [g] because [ɣ] is banned word-initially). Following Murut's phonotactic restrictions, the underlying /ɣ/ is deleted before consonants: e.g., /RED-ɣ-bulud/→[buβulud] 'ridges in which tuberous crops are planted'. In contrast to the fixed segmentism analysis, Trigo (1988:59ff) proposes that the [ɣ] is epenthetic—it appears to thwart vowel hiatus. The problem with the epenthesis proposal is that [ɣ] only shows up in *some* vowel hiatus situations, even with reduplicants: /RED+ulampoj/→[ulalampoj] {*no gloss*}, *[guɣulampoj]; /RED+indimo/→[indidimo] 'about five times', *[guɣindimo]. The fixed segmentism approach has a ready explanation for such 'exceptions': in these cases, the underlying /ɣ/ is not present.

Blevins (2008:105) provides a recent example, suggesting that Southern Oromo has an epenthetic [m] found in the reduplication of frequentative verbs. Stroomer (1987:54) reports that the pattern occurs in the Orma and Waata dialects: e.g., [eːm-eːge] 'he waited long', [tam-tataːniː] 'they stayed and stayed', [fuːm-fuːgite] 'she raised some children'. In contrast, we contend that this [m] is underlyingly present as a 'prespecified' part of the frequentative reduplicant: i.e., /RED-m-eːge/→[emeːge]; /RED-m-tataːni]→[tamtataːni].

[14] A (C)V reduplicant with a habitual/distributive meaning can prefix to these morphemes with unsurprising results: /RED_{σ}-C^{μ}-q-fætæn/ → [fæ-f^{μ}-fætæn] 'walk habitually'; /RED_{σ}-C^{μ}-q-aːtɨ/ → [a-k^{μ}-k-aːtɨ] 'smoke habitually'. (The reduplicant can only copy root material, thus *[ka-k^{μ}-k-aːtɨ].) Because geminate glides are not allowed, the prefixes infix after a glide: /RED_{σ}-C^{μ}-q-won/ → [w-o-k^{μ}-k-on], *[ow^{μ}-won]. This reduplicant can appear with the causative, with expected results: /RED_{σ}-C^{μ}-q-a-eːw-in/ → [æ-k^{μ}-k-æ-j-eːw-in] 'first ones, the first'. See Goodenough and Sugita (1980: Section 4.3) for further discussion.

The epenthetic and fixed segmentism theories make different predictions for other situations in which the same phonological environment occurs. For example, the Southern Oromo distributive plural is also marked by a prefixing reduplicant, but one that lacks any [m]: e.g., [aadiː] 'white (dist. pl.)', [didiːmaː] 'red (dist. pl.)', [gugurdaː] 'many/much (dist. pl.)', [babajːaː] 'distant houses (scattered)' (Stroomer 1987:100–101). In our analysis, the distributive plural morpheme is a plain reduplicant without any fixed segmentism. [gugurdaː] is not *[gumgurdaː] because there is no /m/ underlyingly. In contrast, the epenthetic analysis does not explain why the [m] fails to show up in the distributive plural: i.e., *[amadiː], *[bambajːaː].

Apart from the [m]'s severely restricted morpho-syntactic distribution, it's difficult to see the phonological rationale for epenthesizing it. The [m] appears before roots beginning with a consonant, e.g., [tamtataːni], where there is no hiatus to resolve. Alternatively, [m] would be epenthesized to fill a coda. While coda epenthesis can happen (e.g., to make a stressed syllable bimoraic), [m] appears even when a consonant from the base is available for copying, e.g., [bam-barbaːdani] 'they searched continuously' (Stroomer 1987:165), *[bar-barbaːdani]. ([rb] clusters are otherwise permitted: e.g., [arba] 'elephant'.)

In summary, the [m] in Southern Oromo has all the hallmarks of a morpheme—it is morpho-syntactically restricted and appears in phonologically unmotivated environments. As Alderete et al. (1999) argue for similar cases of reduplicant prespecification like Sm'algyax, the Southern Oromo [m] is instead better analyzed as fixed segmentism in a reduplicant.

2.3.3 Suppletion

Morphological suppletion can also occasionally look like epenthesis. Suppletive morphemes have more than one underlying form; the underlying form that surfaces is selected for its output wellformedness in different environments (Mascaró 1996).

An example is found in Buriat, which has been cited as having an epenthetic voiced dorsal (Poppe 1960; Rice 2004). The exact featural content of the 'epenthetic' consonant varies depending on its environment: it is velar [g] before front vowels, uvular [ʁ] between back vowels, and uvular [ɢ] after front and before back vowels (Poppe 1960). Poppe (1960:20) states that a [g]/[ʁ]/[ɢ] is epenthesized at stem-suffix junctures in certain vowel hiatus situations. For example, [dyː] 'younger brother' is realized as [dyːgeː] in the reflexive possessive form; compare [ɢar] 'hand' ~ [ɢaraː], *[ɢarɢaː].

There are several problems with the proposal that there is an epenthetic [g]. One is that it is severely morpho-syntactically restricted: [g] appears before three morphemes (the instrumental, the genitive, and the reflexive possessive). It does not appear with other morphemes, even when vowel hiatus occurs. For example, the initial vowel in the 'pure relational noun' suffix /iːji/ undergoes deletion after a vowel, not epenthesis: /ʃereː-iːji/→'table + pure relational noun' [ʃereːjiː], *[ʃereːgiːji] (Poppe 1960: 37). In fact, deletion—not epenthesis—is clearly the default resolution for resolving vowel hiatus in the language: /xana-iːji/ 'wall + pure relational noun'→[xaniːji]; /aχa-iːN/ 'elder brother + possessive'→[aχiːN].

Another problem is that the environments in which [g] appears are different for each of the three morphemes. [g] appears with the instrumental after all long vowels and diphthongs; with the reflexive possessive it appears after long vowels, diphthongs, and suffixes ending in /n/ (but not roots ending in /n/); with the possessive it appears after long vowels excluding /iː/ and diphthongs.

The problems discussed above do not arise if the [g] is seen as part of suppletive allomorphs of each of the possessive, instrumental, and reflexive possessive. As an example, the forms of the possessive are given in (7):

(7) Buriat Possessive allomorphy (Poppe 1960:36)

(a) After consonants = [ai]

[mal-ai] 'cattle', [ger-ai] 'yurt', [bulag-ai] 'well'

(b) After long vowels = [gai]

[ʃereː-ɢai] 'table', [bøː-gei] 'shaman'

(c) After stem-final [iː] and diphthongs = [n]

[zygiː] 'bee' ~ [zygiːn], [dalai-n] 'sea'

(d) After a short vowel = [iːn]

[aχa] 'elder brother' ~ [aχiːn], [esege] 'father' ~ [esegiːn]

We propose that the possessive consists of two suppletive allomorphs: /iːn/ and /gai/. Phonological restrictions determine which allomorph appears in the output (for similar cases see Mascaró 1996). Specifically, the allomorph that results in the smallest prosodic size is realized. For example, /aχa/ surfaces as disyllabic [aχiːn], not as trisyllabic *[aχagai].

Independent phonological restrictions (such as ONSET) force adjacent vowels to fuse: hence [aχiːn], *[a.χa.iːn]. Similarly, the final long vowel in /zygiː/ fuses with /iːn/ to create [zygiːn], which is shorter than the trisyllabic alternative *[zygiː-gai]. Adjacent consonants can also fuse. So, the /g/ in /gai/ fuses with root-final consonants: e.g., [ma<u>l</u>ai]; this output form is more desirable than *[mal-gai], which has two heavy syllables (versus a light+heavy in [ma.lai]). The form *[mal-iːn] fares worse than [mal-ai] because *[ma.liːn] has a final super-heavy (trimoraic) syllable.

Interestingly, fusion is blocked by one consideration: the features of underlying long vowels must be preserved. Consequently, /zygiː/ can fuse with /iːn/ to form [zygiːn] because both underlying /iː/s have surface correspondents that preserve their features. However, /ʃereː/ cannot fuse with /iːn/ because one of the long vowels would necessarily lose its features: *[ʃereːn], *[ʃeriːn]. However, failing to fuse would produce vowel hiatus, which is banned in Buriat: *[ʃe.reː.iːn]. The only remaining option is to use the [gai] allomorph, [ʃereː-ɢai]; even though this form does not minimize word size, it is the only way to both preserve long vowels and avoid hiatus.

Finally, we observe that suppletion is rife in Buriat's morphology. Many cases are even more obviously suppletive than the possessive. For example, the causative is realized as [uːl] after stems with short vowels, [ga] after stem-final liquids, [χa] after stem-final [d], and [lga] after long vowels and diphthongs (Poppe 1960:99).

2.3.4 Summary

To summarize, mono-segmental morphemes and suppletion can both be responsible for segments in the phonological output that look like they are epenthetic. However, on close inspection these segments behave like morphemes; they tellingly occur in severely morpho-syntactically restricted environments rather than phonologically definable ones. Genuine epenthetic segments are free with respect to the morphological environments in which they appear. In contrast to all the cases just reviewed, Axininca Campa's epenthetic [t] appears between any root and any following suffix, between any suffixes, and after subminimal roots, and it does so to prevent vowel hiatus or to increase a morpheme's weight (see Paradis and Prunet 1994 for relevant discussion).

2.3.5 Outside the synchronic realm

We use the term 'epenthetic segment' here to refer to a segment that (a) appears in a phonological output and (b) has no input correspondent. The term has also been used in broader ways. We discuss its use in diachronic change, cross-dialect comparison, and loanwords.

In regard to diachronic change, Vaux (2003) proposes that epenthetic [b], [ʃ], and [ʒ] occur in various dialects of Basque, citing Hualde and Gaminde (1998). The data derive from a comparison of how different dialects respond to a vowel hiatus involving a stem and the singular determiner [a]. However, Hualde and Gaminde (1998:42) state that their data show "the output for each of the historical (or, if one wishes, 'underlying') sequences".[15] In other words, 'epenthetic' is used here on the diachronic—not synchronic—dimension: modern [b]/[ʃ]/[ʒ] are 'epenthetic' in the sense that they did not appear in a particular word in an earlier stage of the language but now do. However, for such 'diachronic epenthesis' to be relevant to the concerns of this article, it must be demonstrated that an earlier synchronic grammar generated a particular output [b] (and [ʃ], and [ʒ]), which had no input correspondent at that stage in the language's history. Segments can arise in diachronic change through misperception and morphological misanalysis; consequently, the appearance of a [b] in a word that historically had no [b] does not necessarily mean that there was a grammar that generated a synchronically epenthetic [b].

The same point extends to cross-dialect comparison. Guitart (1976) discusses colloquial Cuban Spanish coda obstruent 'velarization': whereas Castilian Spanish distinguishes [akto] 'act' and [apto] 'fit', Cuban Spanish has [akto] for both. Does this mean that Cuban Spanish neutralizes /p/ to [k]? By no means. It is quite possible that there was a diachronic change of *apto to [akto] in the development of Cuban Spanish. However, as above, this diachronic change does not mean that there was a synchronic grammar in which /p/→[k]. There are no synchronic alternations showing that synchronic

[15] In all the Basque cases, the apparent epenthesis was influenced by its environment: [b] only appears between /u+a/ sequences in four of the dialects, while [ʒ] appears between /i+a/, and there is no epenthetic consonant for the other hiatus situations (/a+a/, /e+a/, /o+a/).

underlying /p/ and/or /t/ become [k] in Cuban Spanish codas. Similarly, Proto-Eastern Polynesian *t became Hawai'ian [k], but there are no synchronic alternations to show that /t/→[k]. Certainly, Richness of the Base requires consideration of what *could* happen to Cuban Spanish /p/ in codas or Hawai'ian /t/. However, without synchronic alternations there is no more reason to think that they map to [k] than to think that they delete, or become [n], or are resolved in any other way.

Some analyses of loanword adaptation have been argued to involve epenthesis of or neutralization to dorsals and labials (e.g., Adler 2006). It is common to refer to a segment in a loanword as 'epenthetic' if it does not occur in the source language's word. However, for a loanword segment to be truly epenthetic, in the sense relevant here, it must be demonstrated that the loanword segment has no input correspondent in its *underlying form*. For example, Māori [kɪˈɾihiˌmɛtɛ] is borrowed from English [ˈkɹɪsməs]. To argue that all—or even some—of the Māori word's [ɪ], [i], [ɛ] are epenthetic, it is necessary to demonstrate that there is a phonological component in which there is an underlying form /ˈkɹɪsməs/ which surfaces as [kɪ̠ˈɾihi̠ˌmɛtɛ]. To our knowledge, such a demonstration has not been made for Māori loanword adaptation—there is no evidence that the underlying form of [kɪˈɾihiˌmɛtɛ] is anything but /kɪɾihimɛtɛ/. More broadly, work on loanword adaptation rarely proves that loanword epenthesis involves actual input-output disparity (for the practical reason that it is extremely difficult to uncover input-output disparities in non-alternating forms). While there is no doubt that the synchronic grammar heavily influences the misperception of non-native forms, the idea that loanword segments that have no source word correspondent also have no *underlying* correspondent at any point in the loanword adaptation process has not been demonstrated.

2.3.6 Neutralization to 'ŋ'

A fairly large number of cases of neutralization to [ŋ] (or an [ŋ]-like sound) have been reported: e.g., Huallaga Quechua (Weber 1989), Genovese (Ghini 2001:173), Kagoshima Japanese (Haraguchi 1984; Trigo 1988), Seri (Marlett 1981:20), Yamphu (Rutgers 1998), Makassarese (Aronoff et al. 1987), Canadian French (Howe 2004) and the San Marcos dialect of Misantla Totonac (MacKay 1994:380), also see Rice (1996) and Howe (2004) for further recent discussion (also see Section 2.2 above regarding diachronic 'neutralization' in Chinese dialects). Neutralization to [ŋ] would certainly falsify the claim that there is no neutralization to dorsals, and shakes the claim that there is no neutralization to [k].

However, none of these cases have been demonstrated to be neutralization to a phonologically dorsal segment. Trigo (1988) argues that nasals in these cases are phonologically placeless, not [dorsal] (for at least part of the derivation) (also see Trigo 1991; Yip 1989, 1991; Piggott 1991). In a similar vein, de Lacy (2002, 2006a) has argued that the 'ŋ' is phonologically glottal and that its velar contact is due to phonetic implementation (for similar proposals, see Yip 1996; Baković 2001); in fact there are two distinct segments that have been misreported as 'ŋ' in this view: a glottal nasal glide (i.e., [ɦ̃]) and a glottal nasal stop [N].

In most cases, it is difficult to find phonological processes that reveal the place specifications of ‘ŋ’, and very few studies attempt to identify phonological processes that would do so. However, the few languages that do provide such evidence show that the apparent ‘ŋ’ is demonstrably not phonologically dorsal.

For example, the Nepalese language Yamphu is reported to have neutralization to ‘ŋ’ in codas: e.g., /hæn/→[hæŋ] ‘you (sg.)’, cf. /hæn-æʔ/→[hænæʔ] ‘you (sg.) + ergative’ ([ŋ] is allowed in onsets: e.g., [i.poː.ŋi.kːo] ‘ten + one’) (Rutgers 1998). However, assimilation reveals this ‘ŋ’ to not be phonologically dorsal. Oral stops assimilate in place to a following [ʔ], becoming glottal: e.g., /mo-dok-ha/→[modoʔha] ‘like those’, *[modokha]; /læːt-he-ma/→[læːʔhema] ‘to be able to do’ (Rutgers 1998:48). In contrast, nasal stops assimilate to following glottals to form ‘ŋ’: /pen-ʔi/→[peŋʔi] ‘he is sitting’; /hen-heː-nd-u-æn-de/→[heŋheːndwende] ‘can you open it?’ (Rutgers 1998:44). If this assimilated ‘ŋ’ is really *dorsal* [ŋ], the outcome of stop assimilation is inexplicable—one would expect stop assimilation to result in [k] instead of [ʔ]: i.e., /læːt-he-ma/→*[læːk.he.ma]. From a broader perspective, assimilation of PoA always results in *agreement* of PoA features. Therefore, the nasal that appears before [h] in Yamphu must be phonologically [glottal] (or placeless, if glottals are considered placeless segments).

Many putative ‘ŋ’s do not behave like dorsals in terms of their phonotactic restrictions; they instead behave like glottals. For example, it is common for glottals to be banned from onset position (e.g., Chamicuro and Macushi Carib do not allow [h] in onsets; Parker 1994). Buriat has been argued to have neutralization to ‘ŋ’ (Poppe 1960), but Buriat’s ‘ŋ’ is restricted to appearing in codas, just like glottals. If Buriat ‘ŋ’ is phonologically dorsal, it is curious that it alone is banned from onsets while the other dorsals [k], [g], and [x] are not.

In fact, there is no language that bans velars like [k g x ɣ] in *onsets* but allows them in *codas*. More generally, with the exception of glottals every PoA that is allowed in codas is also allowed in onsets (Goldsmith 1990; Beckman 1998; de Lacy 2006a: Section 3.2.3). So, the fact that Buriat’s ‘ŋ’ is only allowed in codas indicates that it is not dorsal, but rather glottal/placeless. The favoring of ‘ŋ’ in coda position is also seen in some cases of coda epenthesis, as in Buginese augmentation of sub-minimal words (Mills 1975; Lombardi 2002), Kaingáng in augmentation of stressed syllables (Wiesemann 1972:95–97), and to satisfy a requirement that (word/phrase-)final syllables be heavy as in Uradhi (Hale 1976).

On occasion, when putative ‘ŋ’s are seen to alternate, they alternate with glottals. For example, the glottal [h̃] appears in Aguaruna onsets, but is realized as ‘ŋ’ in codas: [suŋkuŋ] ‘influenza’ cf. [suŋ.ku.h̃-ãn] ‘influenza + accusative’ (Payne 1990:162). If ‘ŋ’ is glottal, this process is merely one of fortition: i.e., a glide becomes a nasal stop in coda position but otherwise retains all other features (i.e., nasality and glottal place of articulation).[16] If this ‘ŋ’ is really velar [ŋ], the motivation for the alternation is unclear; alternations involving [h] in other languages produce other glottals or coronals (e.g., Korean; Kim-Renaud 1986), not velars.

Finally, Howe (2004) has recently discussed velar epenthesis and neutralization, particularly of ‘ŋ’. He argues that putative ‘ŋ’s are truly dorsal because they have the same phonetic realization as a demonstrably dorsal [ŋ]. However, de Lacy (2002, 2006a) (also see Baković 2001) has argued that a dorsal phonetic realization cannot be used as a diagnostic for the phonological specification of ‘ŋ’ as both phonologically glottal/placeless [N] and dorsal [ŋ] are realized as phonetic [ŋ]. Evidence for whether a particular ‘ŋ’ is placeless/glottal or dorsal can *only* be found by examining its behaviour in phonological processes.[17] A placeless nasal will necessarily be pronounced as a dorsal nasal (i.e., as velar [ŋ] or uvular [ɴ]) if it has a complete oral closure because it is not possible to produce nasal air flow if the oral cavity is completely closed behind the uvula.

The one case that still perplexes us is found in some dialects of Uradhi (Hale 1976; Crowley 1983; Paradis and Prunet 1993, 1994; Howe 2004). In many dialects, phrase-final vowels are avoided by optionally adding an ‘ŋ’: e.g., [ama]~[amaŋ] ‘person’. There is no reported phonological process that allows us to determine whether the final ‘ŋ’ is velar or glottal/placeless here. Underlying final consonants are realized faithfully: e.g., /aŋen/→[aŋen] ‘dig (past)’.

However, in some dialects the added final consonant is realized optionally as [ʔ] if the preceding consonant is a stop. In other dialects, though, it is realized optionally as [k]: e.g., [ípiŋ]~[ípik̠̚] ‘water’ (cf. [amaŋ̠] ‘person’). The added consonant can undergo phonological processes such as post-[i] palatalization: e.g., [ipikʲ]~[ipiŋʲ] ‘water’.

Is the phrase-final [k] epenthetic? Unfortunately, it is not a clear-cut case. Aspects of this final consonant are remarkable, particularly that its orality/nasality is determined by long-distance agreement with another consonant. Gafos (1998) argues that such long-distance consonant agreement is only seen in reduplication (with the exception of anteriority and perhaps voicing, though Rose and Walker (2004) argue that even these feature harmonies are instantiated through correspondence relations as in reduplication).

Given the consonant’s behaviour in long-distance agreement, it is possible that it is a mono-segmental reduplicant that is prespecified as a dorsal stop but copies the preceding consonant’s [nasal] value. A problem might be that this morpheme is not morpho-syntactically restricted—it might appear in any morpho-syntactic environment when there

---

[16] Howe (2004) observes that Aguaruna poses a problem for Trigo’s (1988) view that ‘ŋ’ is a placeless glide. However, Aguaruna’s allophony follows straightforwardly if ‘ŋ’ is a glottal nasal stop (de Lacy 2002).

[17] Howe (2004) also argues against a placeless/glottal ‘ŋ’ analysis for the Chukchi nasal. In Chukchi, only /ŋ/ undergoes assimilation; /n/ and /m/ do not. Trigo (1988) argues that this type of selective assimilation can only be explained if the ‘ŋ’ is actually placeless. However, de Lacy (2002, 2006a) and Howe (2004) provide evidence that Chukchi ‘ŋ’ is truly dorsal. This interpretation is beside the point, however, as de Lacy (2002, 2006a) also shows that segments need not be placeless in order to undergo assimilation, and in fact any single place of articulation can assimilate in place while the others do not.

is a phrase-final vowel. Unfortunately, given the sources available to us, we can only suggest that further investigation is warranted.[18]

To summarize, a clear case of an epenthetic dorsal or neutralization to dorsal PoA has yet to be identified. Putative examples can be shown to involve morphemes, suppletion, glottal nasals, or assimilation.

Finally, we note that we have not discussed whether labials—or in fact coronals and glottals—can be epenthetic or the output of neutralization. It is not our aim to do so here as our focus is solely on whether dorsals *could* be epenthetic or the target of neutralization. For consideration of other places of articulation, see the extensive discussion in de Lacy (2002, 2006a) and the numerous sources cited therein.

3 Active synchronic restrictions

In an approach that relies exclusively on diachronic explanation, as does Blevins (2004), there is potentially no role for active synchronic restrictions:

> "The purpose of this book is to demonstrate that many of the similarities in sound shape across languages are best explained in terms of parallel evolution or direct genetic inheritance…With this foundation, synchronic grammars are liberated from the burden of explaining most cross-linguistic similarities in sound patterns, and can be modeled to best describe attested patterns and alternations [within each language: deL&K]. Synchronic constraint systems are minimal in form. They specify phonological categories – featural, segmental, and prosodic – and they specify possible relationships between these categories. Synchronic constraint systems do not express the occurrence of similar sound patterns across languages, when those sound patterns can be shown to have their source in direct inheritance, convergent evolution, or parallel evolution." (Blevins 2004:52)

The synchronic grammar that results from a language's history consists of essentially arbitrary statements of sound distributions that are encoded in the representations of and paradigmatic relations between morphemes and words. In an exclusively diachronic account of this kind, the sound changes that have brought about a particular sound pattern are phonetically motivated but the resulting sound pattern itself no longer is. In short, phonetics governs the past but not the present. A sound pattern's detachment from its original phonetic motivation permits it to enter subsequently into new patterns, which are themselves not phonetically motivated, leading eventually to idiosyncratic 'crazy rules' (Bach and Harms 1972):

> "Many languages have alternations that appear to be phonetically motivated because these alternations reflect sound changes which are phonetically natural. However,

[18] The only other approximately relevant case is neutralization of codas to uvular [χ] before obstruents in Surinam Carib (e.g., /enaːpɨ/→[enaːχ-potɨ] 'eat repeatedly'; Gildea 1995). It is unclear how relevant this process is to the issue at hand as uvulars may be a variant of gutturals, which can be grouped with glottals as a particular class of segments (McCarthy 1994). For further discussion, see de Lacy (2006a:134–135).

> other unnatural alternations are attested, suggesting that naturalness plays no role in constraining synchronic systems." (Blevins 2004:70)

It is certainly the case that subsequent sound changes can obscure the phonetic motivations for a particular sound pattern, by altering it or its environment, but the possibility of subsequent sound changes does not itself require that a sound pattern lose its phonetic motivation once the sound change that produced it is complete. On the contrary, this phonetic motivation can persist into the synchronic grammar of the daughter language long after the sound change is complete. Such persistence would be potent evidence against the 'been-there, done-that' character of the exclusively diachronic account advocated by Blevins; we present evidence that phonetic motivations persist in Section 3.1.

Blevins also argues that despite their phonetic origins, sound changes do not optimize phonological systems:

> "Like genetic mutations, the three phonetic sources of sound change defined in 2.2 [CHANGE, CHOICE, and CHANCE: dL&K] are in no way goal directed. Sound change happens, but it does not occur in order to make speech easier to articulate, easier to perceive or to transmit; it does not necessarily result in a more symmetrical, more stable or generally improved phonological system; for every case where it happens, there is a parallel case where it does not happen." (Blevins 2004:45)

We counter by describing a sound change that is clearly optimizing. Sound changes of this kind have been overlooked or perhaps undervalued in Evolutionary Phonology because its account relies so exclusively on the listener, particularly during the course of language learning, to initiate sound change. The case we describe in Section 3.2 is instead speaker-oriented, in the sense that it represents the speaker's active manipulation of the phonetic content of strings, for the purposes of enhancing a contrast. The evidentiary value of this example does not rest on its being optimizing; after all, in the quote above Blevins acknowledges that some sound changes have this effect. This example is instead important because the optimization is an intended consequence of the sound change rather than an accidental byproduct.

In Evolutionary Phonology, a mature language user knows the distributions of phonological categories (features, segments, and prosodic constituents) and the contents of the lexicon. Such a user may generalize inductively from this knowledge, but has no knowledge independent of these sources from which any further generalizations can be deduced. In Section 3.3, we show that the user's synchronic knowledge is not so limited and that restrictions on sounds' distributions are psychologically and even neuro-physiologically active. They are, moreover, distinguishable from statistical generalizations across the lexicon.

Evidence that language learning is, at least in part, determined by innate predispositions also challenges the argument that listeners know no more than what they can induce about sound patterns from observation. In Section 3.4, we present evidence of such predispositions' influence on language learning at a stage before the learner shows evidence of having learnt phonotactic patterns that are specific to the ambient language. In that section, we also present evidence that such predispositions also govern mature language users' assessment of the relative wellformedness of strings that do not occur in their language, and demonstrate that their assessments cannot be accounted for completely by extrapolating from their experience with patterns that do occur in that language.

3.1 Phonetic persistence

Tonogenesis in Athabaskan provides clear evidence that a sound change's phonetic motivation persists long after the sound change has been phonologized.[19] This persistence is unexpected if present-day sound patterns are to be explained entirely by their histories, as Evolutionary Phonology would have it. In its account of such patterns, there are phonetic motivations for language change, but there is no need for those motivations to persist *after* the change has been phonologized. Indeed, they cannot persist if the sounds participating in those patterns are to undergo subsequent, possibly conflicting sound changes as they often do. Nonetheless, the facts of tonogenesis in Athabaskan indicate that phonetic motivations may indeed persist long after phonologization.

In Proto-Athabaskan, glottalic stops, affricates, nasals, liquids, and glides contrasted with their non-glottalic counterparts at the ends as well as the beginnings of stems (Krauss 2005). Conservative languages spoken in Alaska and along the Pacific Coast in northern California and Oregon maintain these contrasts at the ends of stems, but in most of the rest of the family this contrast has been replaced by a tone contrast on the preceding vowel in stems ending in stops and affricates—henceforth just 'stops'—and the glottalic stops have merged with their non-glottalic counterparts.[20]

The development of tone from an earlier laryngeal contrast in an adjacent consonant is an extremely common sound change (Hombert, Ohala, and Ewan 1979), particularly in the language families of East and Southeast Asia. It can occur because one of the phonetic correlates of a laryngeal contrast in consonants is a difference in the fundamental frequency (F0) of adjacent vowels. These F0 differences become tone contrasts in the vowels and replace the original laryngeal contrast when the other phonetic correlates of that contrast are lost from the consonants. Tonogenesis from laryngeal contrasts in consonants would therefore be an instance of CHANGE, although it may, in the Athabaskan case, also involve aspects of CHOICE.

To facilitate the following discussion, the essential steps in the development of present-day tonal Athabaskan languages are laid out in (8):[21]

---

[19] The account of Athabaskan tonogenesis presented here condenses and updates the one presented in far more detail in Kingston (2005); that account supersedes the one first presented in Kingston (1985). All these attempts to explain how tone developed in this family rely on the prior work of Krauss (1978, published in 2005) and Leer (1981, 1999, 2001).

[20] This contrast has also been lost in stem-final stops in a few peripheral Alaskan languages without being replaced by a tone contrast.

[21] Although this table is both schematic and incomplete in its representation of how tonogenesis occurred in Athabaskan, it portrays all that's needed to support the analysis developed here. For the complete story, with exemplification, see Krauss (2005), Leer (1991, 1999, 2001), and Kingston (2005).

	Early		Late	
PA stem	> (a) High	> (b) Low	(a) > (c) Low	(b) > (d) High
i. vK'	v́K	v̀K	v̀K	v́K
ii. vK	v̀K	v́K	v́K	v̀K
iii. vR'	v́R'	v̀R'	v̀R'	v́R'
iv. vR	v̀R	v́R	v́R	v̀R
v. vʔ	v́ʔ	v̀ʔ	v̀ʔ	v́ʔ
vi. vv	v̀v	v́v	v́v	v̀v

(8) A schematic account of tonogenesis from P(roto)-A(thabaskan) in four exemplary tonal Athabaskan languages. K = stops, R = sonorants, C' = glottalic consonants, v = vowel, vʔ = glottalic full or long vowel, and vv = non-glottalic full or long vowel, v́ = high tone, and v̀ = low tone.

Three properties of tonogenesis in Athabaskan are relevant to the argument that a sound change's phonetic motivation can persist after it has been phonologized.

First, some Athabaskan languages have developed a high tone in stems that originally ended in a glottalic stop and a low tone in stems that ended in a non-glottalic stop, while in others the opposite tones have developed in these two kinds of syllables—these are the developments labeled "(a) High" and "(b) Low" in rows (i) and (ii) in (8).

Second, some of the languages that differ in whether they have high versus low tone from the original glottalic stops are so distantly related that this difference between them likely arose at the outset of this sound change, when dialects of the protolanguage diverged from one another—these are languages (a) and (b) in the 'Early' columns in (8). There is also no evidence of a pre-tonal period of development for any of these languages earlier than Proto-Athabaskan itself. Leer (1999) locates the tonal daughters that developed high tone (8a) from the glottalic consonants on the east side of the Canadian Cordillera, and those that developed low tone (8b) from this source on its west side. On both sides can also be found closely related languages that have the opposite tones, low in stems that ended in glottalic consonants in the protolanguage in east-side languages (8c) and high in such stems in west-side languages (8d). That these languages are otherwise very closely related to languages with the opposite tones indicates that these developments occurred later in their histories; that a language that was earlier of type (8a) changed into one of type (8c) or one that was earlier of type (8b) changed into one of type (8d). While none of the facts rule out such reversals happening more than once in the history of this family, only one stage must be postulated to account for those languages that retain the high or low tones that originally developed from glottalic consonants on the east and west side of the Cordillera, PA > (8a) or (8b), and only two stages for those that later reversed those tones, PA > (8a) > (8c) or PA > (8b) > (8d). The very close relationships between languages that have undergone two stages and others that have undergone just one indicates that the former reversed their tones, (8a) > (8c) or (8b) > (8d), long after the sound change that first introduced contrastive tone (i.e., the merger of stem-final glottalic stops with non-glottalic stops) was complete.

Third, glottalic sonorants (nasals, liquids, and glides) still contrast with non-glottalic ones stem-finally in the tonal daughter languages, both after the first and second stages (row (iii) versus (iv) in 8), as does glottal stop with its absence (row (v) versus (vi) in 8), and the same tones have always developed in these stems as in stems that ended in glottalic versus non-glottalic stops, respectively, in the protolanguage. The tones are identical in languages in which high tones developed in syllables ending in glottalic stops (a, d) and in those where low tone developed in such syllables (b, c), and they are identical regardless of whether this difference arose when protodialects first diverged (a, b) or more recently (c, d). Because the laryngeal contrast has not been lost in sonorants nor in syllables ending in a glottal stop versus its absence, the tones that develop on preceding vowels remain redundant rather than contrastive in such stems, unlike those ending in stops.

Kingston (2005) argues that both high and low tones can develop directly from the original glottalic stops because these consonants may be pronounced in two different ways. Glottalic consonants are distinguished from non-glottalic ones by a constriction of the glottis that is tight enough to curtail or even cut off airflow through it. The glottis is closed by contracting the interarytenoid and lateral cricoarytenoid muscles while relaxing the posterior cricoarytenoid muscle, and the constriction is tightened by the forceful contraction of the thyroarytenoid muscle, which stiffens the inner bodies of the vocal folds and causes them to press firmly against one another. If this is all the speaker does, the voice quality of adjacent vowels is creaky and its F0 is low because the folds' outer covers remain slack. However, if the speaker also contracts the cricothyroid muscle at the same time, the folds' outer covers are stretched and the voice quality in the adjacent vowel is tense and its F0 high instead. The two muscles can be contracted independently because they are innervated by separate nerves, the cricothyroid by the superior laryngeal nerve and the thyroarytenoid by the recurrent laryngeal nerve.

Thus, in languages where a low tone developed from an original glottalic stop, speakers contracted the thyroarytenoid alone when pronouncing glottalic stops, while in languages where a high tone developed instead, they also contracted the cricothyroid. Whether just the thyroarytenoid was contracted or the cricothyroid was too does not appear to depend on any other articulation, neither laryngeal nor oral, but rather on the speaker's choice to contract just one of these muscles or both of them. A large and growing body of evidence of other language-specific phonetic differences shows that speakers commonly exercise their freedom to make such choices (Kingston and Diehl 1994).

This sound change is an instance of CHANGE. Once predictable differences in vowels' voice quality and F0 were originally produced by coarticulation with following stops that contrasted in whether they were glottalic. They became contrasting tones once this laryngeal contrast between the stops merged. This sound change also appears to be an instance of CHOICE, but one in which the *speaker* exercises the choice as to how a target segment is to be pronounced, rather than the listener choosing one from a number of variants.

What is particularly interesting about these choices is their persistence in the history of the tonal Athabaskan languages. The choice speakers of a particular protodialect made about how to pronounce the glottalic stops before their merger with non-glottalic stops continues to determine how the glottalic sonorants and glottal stop are pronounced by those of their present-day descendents that belong to types (a) and (b). This persistence is particularly striking since nothing appears to stand in the way of speakers changing how they pronounce the glottalic sonorants and glottal stop sometime between the present day and the time when the contrast was lost in the stops and transferred to tone in the preceding vowel. If it were once possible to choose whether to contract the cricothyroid as well as the thyroarytenoid and thus whether to raise or lower F0 in a preceding vowel, then it should have remained possible to do so in glottalic sonorants and glottal stop. Therefore, at any time during any tonal language's subsequent history, its speakers could have adopted the pronunciation of glottal constriction that has the opposite effect on F0 and tone in the preceding vowel from that which the original glottalic stops had. The result would be that stems that end today in glottalic sonorants or glottal stop would have the opposite tone from those that originally ended in glottalic stops. Such developments are sketched in (9), whose point of departure is languages of type (a) or (b):

Early		Late	
(a) High	(b) Low	(a) > *(e) High /_*K', Low /_{R', ʔ}	(b) > *(f) Low /_*K', High /_{R', ʔ}
v́K	v̀K	v́K	v̀K
v̀K	v́K	v̀K	v́K
v́R'	v̀R'	v̀R'	v́R'
v̀R	v́R	v́R	v̀R
v́ʔ	v̀ʔ	v̀ʔ	v́ʔ
v̀v	v́v	v́v	v̀v

(9) Hypothetical but unobserved subsequent tone reversals in stems ending in glottal sonorants (R') or glottal stop (ʔ).

This has never happened. No language of type (9e) or (9f) has ever arisen because when the sound change was phonologized, the phonetics of the pronunciation of glottal constriction were fixed in the grammar too. The fixing of the glottalic constriction's phonetics must have constrained glottalic sonorants and the glottal stop to be pronounced in the same way throughout the subsequent history of each tonal Athabaskan language as its own glottalic stops were when the sound change was actuated.

What then of the languages in which high tone has replaced low tone, or vice versa, more recently in the history of this family? These languages, too, indicate that a phonetic property has been fixed in the grammar: no languages of type (e) or (f) have ever arisen from a language of type (c) or (d) either. These later reversals of tone value are also a byproduct of speakers exerting their freedom to choose how to pronounce a glottalic consonant, although at these later stages in the history of the family they could only exercise this freedom in the pronunciation of the glottalic sonorants and glottal stop, because glottalic stops would have long since merged with their non-glottalic counterparts at the ends of stems. Once speakers chose a different pronunciation of these sounds and thereby reversed the value of the redundant tone on the preceding vowel, they also invariably reversed the value of the contrastive tone on vowels in stems that ended in glottalic stops in the protolanguage. They would do so because prior to the reversal, all stems bearing high tones would have been produced with similar F0 values, as would all stems bearing low tones, regardless of whether the tone was contrastive, as in stems ending in stops, or predictable, as in stems ending in sonorants or in full vowels. Once speakers chose to pronounce the glottalic sonorants and glottal stop differently, e.g., once speakers who had once pronounced these sounds with thyroarytenoid contraction alone chose to pronounce them with the cricothyroid contracted, too, then the F0 values on preceding vowels would have flipped from low to high. This flip would have been extended to stems ending in stops where the low tone was contrastive rather than predictable because those stems had similar F0 values. Even though the tone preceding glottalic sonorants and glottal stops remained redundant rather than contrastive, phonetic pattern congruity forced the contrastive tone to reverse its value when the redundant one did.

Phonetic pattern congruity arises from the constraint ranking illustrated in (10), which will produce a language of type (a) where high tones occur on stems that ended in a glottalic stop in Proto-Athabaskan or that end in /R'/ or /ʔ/ synchronically. This tableau thus models the input-output relations at a time in the language's history after glottalic and non-glottalic stops have merged to non-glottalic stops in stem-final position, and the contrast has shifted to tone on the preceding vowel. Tone differences remain predictable in stems not ending in stops. This stage is important to model because it's the one that precedes a possible reversal in tone values (here, the transition from a language of type (a) to one of type (c)) that maps both contrastive and predictable tones onto the opposite values, i.e., a reversal that is phonetically congruent.

The six possible stems are listed as inputs, with various arrays of outputs considered in each row. The relevant grammar fragment consists of a single faithfulness constraint and five markedness constraints. The faithfulness constraint, IDENT[T], ensures that an input specification for tone is preserved. It is violated in the output arrays in (10c–g) because tones are introduced on stems that aren't specified for them in the input, including on stems ending in glottal sonorants and glottal stop in the optimal output array in (10c). These violations are, however, not fatal because this output array does not violate the higher-ranked markedness constraint that prohibits vowels from appearing toneless before glottal consonants, *0/_C'. This markedness constraint expresses an obligation to recognize the phonetic effects (a particular voice quality and F0) of coarticulation with a following glottalic consonant as the realization of a tone.[22] The complementary markedness constraint that prohibits a tone from occurring on any vowel that precedes a non-glottalic consonant is also ranked above IDENT[T] even though all output arrays violate it at least once, and it cannot therefore decide between (10a–d). Multiple violations of this constraint rule out (10g), where all outputs are realized with high tones. The need for this ranking is established in tableau (12) below. The remaining markedness constraints, *L and *H, determine the phonetic value with which the tone T is realized. The ranking of *L above *H in this tableau ensures that T is realized as a high tone (10c) rather than a low one (10d). A language of type (b) would be produced by the opposite ranking of these two constraints. A number of cells in tableaux (10–13) contain more than one ! marking a fatal violation, because each row evaluates more than one input-output pair.

[22] The purpose of *0/_X' is the same as one that prohibits a vowel from occurring without nasalization following a nasal or nasalized segment in a language with nasal harmony such as Madurese (McCarthy and Prince 1999). The complementary constraint, *T/_X, mimics one banning nasalization following an oral segment.

(10)	vK	vR	vv	v_TK	vR'	vʔ	*0/_X'	*T/_X	IDENT [T]	*L	*H
(a)	vK	vR	vv	$\acute{v}_T$K	vR'	vʔ	*!*!	*			*
(b)	vK	vR	vv	$\grave{v}_T$K	vR'	vʔ	*!*!	*		*!	
☞(c)	vK	vR	vv	$\acute{v}_T$K	$\acute{v}_T$R'	$\acute{v}_T$ʔ		*	**		***
(d)	vK	vR	vv	$\grave{v}_T$K	$\grave{v}_T$R'	$\grave{v}_T$ʔ		*	**	*!*!*!	
(e)	$\acute{v}_T$K	vR	vv	vK	vR'	vʔ	*!*!	*	**		*
(f)	$\acute{v}_T$K	$\acute{v}_T$R	$\acute{v}_T$v	vK	vR'	vʔ	*!*!	***	****		***
(g)	$\acute{v}_T$K	$\acute{v}_T$R	$\acute{v}_T$v	$\acute{v}_T$K	$\acute{v}_T$R'	$\acute{v}_T$ʔ		*!*! *!*!	*****		*** ***

The tableaux in (11) and (12) show that the same output array is selected as optimal when, under Richness of the Base, stems ending in /R'/ or /ʔ/ are specified for tone in the input (11) or that stems ending in R or v are specified for tone (12). Both alternatives are expected under Richness of the Base because tone specification is predictable in stems ending in these ways. Top-ranking of *0/_X' ensures that stems ending in /R'/ or /ʔ/ surface with tone in the optimal output, regardless of whether they are specified for tone in the input (11c) or not (10c). Similarly, ranking *T/_X above Ident[T] ensures that (12g) is not preferred over (12c).

(11)	vK	vR	vv	v_TK	v_TR'	v_Tʔ	*0/_ X'	*T/_ X	IDENT [T]	*L	*H
(a)	vK	vR	vv	$\acute{v}_T$K	vR'	vʔ	*!*!	*	**		*
(b)	vK	vR	vv	$\grave{v}_T$K	vR'	vʔ	*!*!	*	**	*!	
☞(c)	vK	vR	vv	$\acute{v}_T$K	$\acute{v}_T$R'	$\acute{v}_T$ʔ		*			***
(d)	vK	vR	vv	$\grave{v}_T$K	$\grave{v}_T$R'	$\grave{v}_T$ʔ		*		*!*!*!	
(e)	$\acute{v}_T$K	vR	vv	vK	vR'	vʔ	*!*!	*	****		*
(f)	$\acute{v}_T$K	$\acute{v}_T$R	$\acute{v}_T$v	vK	vR'	vʔ	*!*!	***	*** ***		***
(g)	$\acute{v}_T$K	$\acute{v}_T$R	$\acute{v}_T$v	$\acute{v}_T$K	$\acute{v}_T$R'	$\acute{v}_T$ʔ		*!*! *!*!	***		*** ***

(12)	vK	v_TR	v_Tv	v_TK	vR'	vʔ	*0/_ X'	*T/_ X	IDENT [T]	*L	*H
(a)	vK	vR	vv	$\acute{v}_T$K	vR'	vʔ	*!*!	*	**		*
(b)	vK	vR	vv	$\grave{v}_T$K	vR'	vʔ	*!*!	*	**	*!	
☞(c)	vK	vR	vv	$\acute{v}_T$K	$\acute{v}_T$R'	$\acute{v}_T$ʔ		*	****		***
(d)	vK	vR	vv	$\grave{v}_T$K	$\grave{v}_T$R'	$\grave{v}_T$ʔ		*	****	*!*!*!	
(e)	$\acute{v}_T$K	vR	vv	vK	vR'	vʔ	*!*!	*	****		*
(f)	$\acute{v}_T$K	$\acute{v}_T$R	$\acute{v}_T$v	vK	vR'	vʔ	*!*!	***	**		***
(g)	$\acute{v}_T$K	$\acute{v}_T$R	$\acute{v}_T$v	$\acute{v}_T$K	$\acute{v}_T$R'	$\acute{v}_T$ʔ		*!*! *!*!	***		*** ***

Finally, the tableau in (13) shows that switching the ranking of *L relative to *H and making (13d) optimal rather than (13c) is all that's required to reverse the F0 value of the tone that realized T, whether T is contrastive or predictable. The tone values in (13e–g) have been switched to low (cf. (10)) to show that it's not the tone level that a syllable bears that rules out an output array but whether it bears a tone.

(13)	vK	vR	vv	v_TK	vR'	vʔ	*0/_X'	*T/_X	IDENT [T]	*H	*L
(a)	vK	vR	vv	$\acute{v}_T$K	vR'	vʔ	*!*!	*		*!	
(b)	vK	vR	vv	$\grave{v}_T$K	vR'	vʔ	*!*!	*			*
(c)	vK	vR	vv	$\acute{v}_T$K	$\acute{v}_T$R'	$\acute{v}_T$ʔ		*	**	*!*!*!	
☞(d)	vK	vR	vv	$\grave{v}_T$K	$\grave{v}_T$R'	$\grave{v}_T$ʔ		*	**		***
(e)	$\grave{v}_T$K	vR	vv	vK	vR'	vʔ	*!*!	*	**		*
(f)	$\grave{v}_T$K	$\grave{v}_T$R	$\grave{v}_T$v	vK	vR'	vʔ	*!*!	***	****		***
(g)	$\grave{v}_T$K	$\grave{v}_T$R	$\grave{v}_T$v	$\grave{v}_T$K	$\grave{v}_T$R'	$\grave{v}_T$ʔ		*!*! *!*!	*****		*** ***

There are two decisive rankings in this analysis. One is between the markedness constraints that regulate the distribution of tones (*0/_X' and *T/_X) and the faithfulness constraint that preserves input specifications for tone IDENT[T]. The other decisive ranking is between *L and *H. The first ensures that predictable tone is realized in stems ending in glottalic elements and not on stems that do not end in such an element, while the second ensures that both contrastive and predictable tones are realized as high or low. The independence of the specification for tone from the choice of the tone's F0 level further ensures that contrastive and predictable tones are always realized at the same F0 levels at any time in the language's history, regardless of whether those levels are ever reversed.

This independence also recognizes a subtler analytic demand, namely, the need to determine *whether* a feature specification occurs in the optimal output independently from *how* it is realized phonetically in that output. A similar and familiar case is contextual allophonic variation. For example, stops contrast for [voice] in a number of contexts in English, and IDENT[voice] needs to be ranked high enough to ensure that the contrast isn't neutralized in any of these contexts, but not so high as to prevent neutralization tautosyllabically after [s] or via flapping for coronal stops before unstressed vowels. In both contrasting and neutralizing contexts, the particular pronunciation of the members of the contrast or the product of their neutralization needs to be determined, too. For example, /b, d, g/ are typically pronounced as voiceless unaspirated stops at the beginnings of words (Caisse 1982; Docherty 1992), where /p, t, k/ are pronounced as voiceless aspirated stops; intervocalically before an unstressed vowel, the two series of stops are instead pronounced as voiced versus voiceless unaspirated stops; only voiceless unaspirated stops can appear after tautosyllabic [s]; and /t, d/ are both pronounced as flaps before unstressed vowels.

Vowels in the tonal Athabaskan languages vary similarly in F0 values and voice qualities, both as a function of whether they contrast for tone and the laryngeal articulation of a following consonant. In their case, the variation in pronunciation is not across contexts, but instead across languages, in that individual tonal languages differ in which F0 values and voice qualities are pronounced in vowels specified for tone or coarticulated with a following glottalic consonant.

To summarize, the phonological restriction on tone and its phonetic motivation necessarily co-existed and continued to do so throughout the histories of the individual tonal languages in the Athabaskan family. The persistence of their covariation is expected in a theory where synchronic phonological constraints remain intimately tied to their motivating phonetic conditions. In an approach to sound change like Evolutionary Phonology that severs the product of sound change from its phonetic motivation once it has been phonologized, there is not only no need to maintain the link between the phonologized output and its phonetic motivation, but indeed an overt and unfortunately counterfactual denial that any link can be maintained.

3.2 An optimizing sound change

Proponents of diachronic explanations have also claimed that sound change is not optimizing, although their arguments are inconsistent on this point. On the one hand:

> "Many languages have alternations that appear to be phonetically motivated because these alternations reflect sound changes which are phonetically natural. However, other unnatural alternations are attested, suggesting that naturalness plays no role in constraining synchronic systems." (Blevins 2004:70)

as well as:

> "There is no need to use markedness to incorporate notions of phonetic complexity into synchronic phonological description. Articulatory ease and perceptual contrast may play an indirect role in the typology of sound change, but the end results of these sound changes are phonological systems which are independent of their phonetic origins, and may ultimately obscure them." (Blevins 2004:78)

These passages jointly assert that a sound change and the synchronic pattern that it produces can readily be detached from their phonetic motivation. On the other hand, Blevins also argues that:

> "Like genetic mutations, the three phonetic sources of sound change defined in 2.2 [CHANGE, CHOICE, and CHANCE: dL&K] are *in no way* [our emphasis] goal directed. Sound change happens, but it does not occur in order to make speech easier to articulate, easier to perceive or to transmit; it does not necessarily result in a more symmetrical, more stable or generally improved phonological system; for every case where it happens, there is a parallel case where it does not happen." (Blevins 2004:45)

This passage denies that the phonetically motivated sound changes have any communicative or grammatical purpose. Such purposes may appear to be served, but the quotes indicate that sound changes serve such purposes at best temporarily and fortuitously and perhaps more in the eye of the analyst than the speakers and listeners of the language.

Although these quotes leave us uncertain about the strength of Blevins's (2004) views, here we examine the stronger of the two—that sound change cannot take place in order to optimize some aspect of the affected language's phonology and its transmission. As the last quote indicates, such optimization might be a fortuitous byproduct of a sound change, but it is not why the sound change happened. We agree that many sound changes are not optimizing and they did not happen to achieve such purposes, but nonetheless maintain that some sound changes have this effect and motivation. The case discussed here—the ongoing split in the pronunciation of the diphthong /aɪ/ in southern American English and similar changes in other dialects—is particularly interesting because the optimization apparently conflicts with well-grounded phonetic expectations. To anticipate, we offer an alternative explanation for this sound change than that proposed in our sources, which both removes the conflict and at the same time shows both that the sound change is optimizing and that the optimization is not fortuitous.

As documented in Thomas (2000), Moreton (2004), and Moreton and Thomas (2007) (and earlier accounts cited in these papers), /aɪ/'s pronunciation in southern American English has split or is still splitting: before [−voice] obstruents, it is pronounced as a more extreme diphthong—F1 is lower and F2 is higher in its offglide—and as a less extreme diphthong or even a open front monophthong (approximately [a̟]) elsewhere. Thomas (2000) and Moreton (2004) also show that listeners are significantly more likely to identify a following obstruent as [−voice] when diphthongization is more extreme.

What is striking about this split is that its direction appears to be exactly opposite what is expected on phonetic grounds. In the transition to [−voice] obstruents, formants are frequently cut off early because the glottis opens and voicing ceases before the oral constriction is complete. Voicing's continuation into the oral constriction of [+voice] obstruents permits formants to reach more extreme values at the end of a vowel preceding such consonants. An important consequence is that F1 is typically lower at the end of a vowel before a [+voice] than a [−voice] obstruent because voicing continued to excite this resonance during more of the F1 lowering caused by the oral closing gesture. This difference is in fact so reliable that listeners use it as a cue to the obstruent's voicing, at least when the vowel is not close and its steady-state F1 is not too low (Parker 1974; Wolf 1978; Walsh and Parker 1981, 1983; Walsh, Parker, and Miller 1987; Hillenbrand, Ingrisano, Smith, and Flege 1984; Summers 1987, 1988; Fischer and Ohde 1990; Diehl and Kingston 1991; Crowther and Mann 1992, 1994; Kingston and Diehl 1994, 1995; Kingston, Diehl, Kirk, and Castleman 2008). The lower value that F1 typically reaches before a [+voice] obstruent should enhance diphthongization in falling-sonority diphthongs such as [aɪ] because the close articulation of the off-glide also lowers F1. How then can diphthongal offglides become acoustically more extreme precisely in the context where the vowel-[−voice] consonant transition leads us to expect they would instead become less extreme, and vice versa?

After considering and rejecting a number of alternatives, Moreton (2004) proposes that this split in the pronunciation of /aɪ/ phonologizes their coarticulation with the following obstruent. The offglide of the diphthong is hyper-articulated before [−voice] obstruents because [−voice] obstruents are themselves hyper-articulated; specifically, they are produced with more extreme and faster articulations than [+voice] obstruents (Fujimura and Miller 1979; Summers 1987, 1988). [−voice] obstruents may be hyper-articulated to counteract the higher intraoral air pressure that would build up behind the oral closure as a result of the high rate of airflow through the glottal opening that devoices them. Coarticulation with hyper-articulated [−voice] obstruents would make the articulation of the end of any preceding vowel, including the offglide of a preceding diphthong, closer than before [+voice] obstruents. In other words, it is the articulatory interaction between the obstruent and the preceding vowel rather than its acoustic consequences that drives the sound change.

Moreton reports that in the offglides of the diphthongs /aɪ, eɪ, ɔɪ, aʊ/, F1 is lower before [t] than [d], and F2 is higher for the three front-gliding diphthongs and lower for the back-gliding diphthong. All these differences indicate a closer articulation of the offglides before the [-voice] stop than the [+voice] one. These coarticulatory differences extend into the nucleus of the diphthong /aɪ/, where F1 is still lower and F2 higher before [t] than [d]. Although the nuclear differences are still significant for this diphthong, they are smaller than in the offglide, as would be expected if they were produced by coarticulation with the following stops.[23]

Moreton and Thomas (2007) use the finding that the diphthong's nucleus coarticulates with the following stop, but less than the offglide, to unify Moreton's (2004) account of the Southern split with an analysis of Canadian Raising and similar patterns in other dialects. In Canadian Raising, the nucleus of /aɪ/ is closer [ʌ] before tautosyllabic [−voice] obstruents and more open [a] elsewhere. Moreton and Thomas describe a variety of English dialects that exhibit similar allophonic variation in the pronunciation of this diphthong and show that in each of them the offglide and nucleus are closer before tautosyllabic [−voice] obstruents than elsewhere. The dialects differ in whether it is the difference in the offglide that has phonologized, as in the Southern split, the difference in the nucleus, as in Canadian Raising, or something in-between. Moreton and Thomas describe the splits in the diphthongs as a "tendency of diphthongs to be dominated by the offglide before [−voice] codas, and by nuclei elsewhere" (Moreton and Thomas 2007:55). The offglide dominates before [−voice] obstruents because they shorten the nucleus more than the offglide, while the nucleus dominates before [+voice] obstruents (and elsewhere) because it is relatively long and the offglide relatively short in those contexts. Moreton (2004) also describes closer allophones or reflexes of sound changes before [−voice] obstruents than elsewhere in languages other than English, and notes that the reverse is not observed (see also Kingston 2002 for earlier discussion of the type example, Polish).

Can these sound changes be optimizing? Specifically, do they preserve or enhance the [voice] contrast in a perceptually weak position, the syllable coda? Moreton (2004) gives three arguments against such an interpretation. First, optimization cannot explain why vowels are only hyper-articulated before [−voice] consonants and not all others whose perceptual correlates might be weak in syllable codas, particularly the marked member of this opposition, the [+voice] obstruents. Second, hyper-articulation has the opposite acoustic effect from another correlate of [−voice]: it lowers F1 at the end of the vowel rather than cutting it off at higher frequency like the devoicing gesture does. Despite this conflict, Moreton's perceptual experiments showed that lower F1 in a diphthong induced more [−voice] judgments of a following stop, as did Thomas's (2000)

[23] Moreton did not find that nuclear F1 and F2 differed significantly before [t] compared to [d] for the other three diphthongs /eɪ, ɔɪ, aʊ/.

experiment. The studies cited above showed the opposite effect in monophthongs, where a lower F1 offset instead induced more [+voice] judgments. Kingston and Diehl (1994) argued that it does so because a lower F1 concentrates energy at lower frequencies next to the stop, much like voicing does during the stop closure itself. Third, even if listeners have learned to associate these acoustic differences with the [voice] contrast, that does not explain why hyper-articulation occurs before [−voice] rather than [+voice] obstruents.

Each of these arguments can be rebutted. First, why not hyper-articulate vowels before other consonants to convey their feature values better, especially before [+voice] obstruents, instead of before [−voice] obstruents? Fischer and Ohde (1990) showed that F1 offset values contribute little to the perception of the [voice] contrast following close monophthongs, because F1 is already so low in the vowel's steady-state that a noticeably lower offset value cannot be produced before [+voice] than [−voice] stops. The closeness of a diphthong's offglide would limit the perceptual value of F1's offset frequency similarly.

Because the perceptual value of F1 offset frequency is reduced or lost following close articulations, the perceptual value of low F1 during the offglide itself need not depend on any other phonetic correlate of the [voice] contrast. In this respect, it resembles the differences in vowel duration before obstruents contrasting for [voice], which are also detached perceptually from the inversely covarying durations of the obstruents' constrictions. These duration differences are at least as robust and perceptually valuable before syllable-final stops, e.g., in *muck* versus *mug*, as intervocalic stops, e.g., in *mucky* and *muggy*, even though the syllable-final stops are frequently not released audibly and listeners cannot detect that the [−voice] stop closure is longer than the [+voice] one. The perceptual effects of closer offglide articulations before [−voice] compared to [+voice] obstruents could detach similarly from other phonetic correlates of the [voice] contrast. This rebuttal addresses Moreton's third argument as well as his first.

The acoustic effects of closer articulation in the offglides would also not enhance place contrasts because F1 does not differ noticeably between any place in front of the uvula, and F2 varies as function of place.

Rebutting Moreton's second argument against treating the Southern split as preserving or enhancing the [voice] contrast not only bolsters the preservation/enhancement account, but suggests a quite different explanation for these sound changes. Recall that Moreton asked why F1 is lower in diphthongs' offglides before [−voice] than [+voice] obstruents, but lower at the offset of monophthongs before [+voice] than [−voice] obstruents. These apparently contradictory patterns can be resolved once it's recognized that the difference in the diphthongs' offglides is measured earlier than the difference in the monophthongs' offsets. The offglide difference is measured earlier because the articulatory target for the offglide can be (and is) reached well before the onset of the obstruent constriction. The formant frequencies could therefore change again after they reach whatever values they have during the diphthong's offglide, as the articulators move from that target to the following obstruent constriction. Whether any of those formant frequency changes are detectable and usable as cues to the obstruent's voicing depends on how different the articulation of the obstruent constriction is from the articulation of the diphthong offglide. In particular, whether F1 changes in a way that could contribute independently to the [voice] percept in the obstruent depends on whether the offglide's articulation is more open than that of the obstruent.

Moreton's measurements show that F1 is consistently higher during the offglide before [d] than [t], which means that F1 would fall more between the offglide and the vocalic offset before [d] than [t]. The listener could then detect more low frequency energy at the offset of the diphthong before [d] than [t], much as they would at the offset of more open monophthongs. The closer articulations of the diphthong's offglide before [−voice] than [+voice] obstruents do not therefore rule out F1's offset frequency contributing to the percept of low frequency energy near the stop in the same way at the offset of diphthongs as at the offset of monophthongs, quite the contrary. By not hyper-articulating diphthongs before [+voice] obstruents, speakers may be doing what they can to produce F1 transitions that fall enough to convey the presence of more low frequency energy next to those consonants than next to [−voice] obstruents.

The results of Moreton's two perceptual experiments can also be reinterpreted in this light. His listeners gave more [−voice] responses to stimuli with lower F1 values in the diphthong's offglide. F1's frequency in the offglide ranged across a nominal range of 300–450 Hz in his Experiment 2 and across a nominal range of 300–400 Hz in his Experiment 3 (ranges measured in the synthesized stimuli were 322–437 and 341–382 Hz, respectively). F1's offset frequency 25 ms after the end of the offglide was set at a nominal 200 Hz (Moreton reports no measured value), which is lower than even the lowest offglide value. Listeners may therefore have given more [−voice] responses when F1 was lower in the offglide because it fell too little to its offset value to produce as robust a low frequency energy percept as when F1 was higher in the offglide.

This second rebuttal also suggests that Moreton's attention to the hyper-articulation of diphthongs before [−voice] obstruents may have been misplaced, and that instead it is the *hypo*-articulation before [+voice] obstruents that needs to be explained. Diphthong offglides may be hypo-articulated in this context to raise F1 high enough that it can fall perceptibly in its transition from its offglide to its offset value.

The sound change itself in the Southern split, monophthongization of /aɪ/, phonologizes an exaggerated hypo-articulation. The resulting open monophthong with its high F1 steady-state followed by a low offset F1 differs more from the diphthong with its high F1-low F2 nucleus-to-low F1-high F2 offglide than the earlier somewhat hypo-articulated diphthong would have.

What then of Canadian Raising; have we lost the unification that Moreton and Thomas (2007) achieved between it and the Southern split? We don't think so. Canadian Raising has taken place in speech communities where the more open, hypo-articulated pronunciation of the diphthong's offglide before [+voice] obstruents has not been exaggerated or phonologized as an open monophthong. Members of these speech communities have chosen to exaggerate the other allophone instead by replacing the original more open nucleus [a] with the closer [ʌ]. Yet other speech communities appear to be caught between these extremes, neither completely monophthongizing the allophone before [+voice] obstruents nor choosing the categorically closer nucleus before [-voice] obstruents.[24]

---

[24] Some English-speaking communities do not pronounce /aɪ/ in noticeably different ways before obstruents differing in voicing (see Morton and Thomas 2007 for a list); they have chosen to exaggerate neither the hypo- nor the hyper-articulation.

One fact remains to be explained: /aɪ/ monophthongized everywhere except before [−voice] obstruents, where it remained a diphthong in the Southern split. Similarly, Canadian Raising has raised the nucleus of /aɪ/ only before tautosyllabic [−voice] obstruents. The difference between an identifiable environment and elsewhere tempts the analyst into seeking an explanation in which the identifiable environment has some identifiable positive effect that is absent elsewhere. For example, a following [−voice] obstruent could check a tendency to monophthongization that is given free rein elsewhere, or it could encourage a closer articulation of a nucleus that remains open elsewhere. The alternative analysis developed here for the Southern split requires the diphthong to monophthongize first before [+voice] obstruents, where this change makes a fall in F1 at the end of the close offglide more perceptible. This change would only later have generalized to syllables not ending in obstruents, including open syllables. It also predicts that monophthongization would not generalize or would generalize last to before [−voice] obstruents, precisely because its purpose is to preserve perceptible differences in low frequency energy next to the following stop. While we do not challenge the claim that a following [−voice] obstruent produced the hyper-articulated precursor that eventually acquired a categorically closer nucleus in /aɪ/ in dialects that have undergone Canadian Raising, the exaggeration that turned [a] into [ʌ] was, we think, prompted by the failure to monophthongize before [+voice] obstruents, i.e., by the desirability of enhancing the [voice] contrast in a perceptually weak context.

This sound change differs in another way from those discussed by Blevins: it is initiated by the speaker rather than the listener. The increase in distinctiveness in the realization of the [voice] contrast that results from either the Southern split or Canadian Raising may make listeners' perceptual task easier, but neither sound change was initiated by them. As speakers are likely to be motivated to convey the information content of their messages successfully, it should not be surprising that speaker-initiated changes like these would be optimizing. More generally, Blevins has artificially reduced the likelihood that a phonetically-motivated sound change would be optimizing by limiting their actuation to the listener's perceptual mistakes and choices. Once the other participant in the conversation and that participant's intentions are acknowledged, then sound changes that deliberately optimize information transmission become far more plausible.

3.3 Active synchronic constraints

In this section, we turn to the evidence that the constraints proposed to account for synchronic sound patterns are psychologically and even neuro-physiologically active. This evidence challenges the position Blevins takes:

> "…there is no clear role for markedness within synchronic phonology. Absolute universals and universal tendencies emerge from general pathways of language change, and have no independent status in the grammar." (Blevins 2004:20)

This evidence shows that such constraints are also not mere statistical generalizations across the lexicon that can be induced during word learning. These constraints influence the on-line categorization of sounds and the syllabification of segment strings.

Moreton (2002) presented listeners with two sets of stop-sonorant-vowel stimuli. The two sets differed in only the stop, which ranged incrementally from [d] to [b] in the first set and from [g] to [b] in the second. In both sets, the sonorant ranged incrementally between [l] and [w]. Listeners identified the members of the first set as beginning with ‘gl’, ‘gw’, ‘dl’, or ‘dw’ and the members of the second set as beginning with ‘gl’, ‘gw’, ‘bl’, or ‘bw’. Moreton showed that both [dl] and [bw] have a zero frequency of occurrence as onsets in the 18.5 million words in the London-Lund corpus of written and spoken British English. On statistical grounds alone, then, ‘dl’ and ‘bw’ responses are predicted to be equally disfavored.

The results were quite different. In responses to the [d-g][l-w] stimulus set, listeners were more than three times less likely to identify the sonorant as ‘l’ if they identified the stop as ‘d’ rather than ‘g’. In responses to the [b-g][l-w] stimulus set, they were actually more than one and a half times more likely to identify the sonorant as ‘w’ if they identified the stop as ‘b’ rather than ‘g’, contrary to what might be expected if statistical rarity inhibited ‘b’ responses.[25] These results show that it is possible to distinguish a zero that is the result of a phonotactic prohibition, i.e., the absence of /dl/ onsets, from one that is the consequence of an accidental gap, the absence of /bw/ onsets, contrary to what one would expect if listeners induced phonotactic constraints from the statistical properties of words during acquisition.[26]

What is the source of speakers’ knowledge of this distinction? In a theory with a non-trivial phonological component, it can be ascribed to an innate synchronic constraint that bans [dl], but not [bw]. In contrast, in a strictly diachronic explanation the zero frequencies of [dl] and [bw] have the same source, the process of diachronic transmission. Each occurs with zero frequency because in the course of the language’s history nothing has caused either to arise. There is no ongoing psychologically active constraint that restricts either structure, and nothing—not even lexical frequency—that can account for the difference in their preference.

In a follow-up experiment in which a vowel was inserted before the stop-sonorant-vowel string in the first stimulus set, the bias against ‘l’ when the stop was identified as ‘d’ disappeared. This result shows that the difference between [dl] and [bw] obtained in the first experiment is not merely a perceptual interaction between the two segments but a consequence of their syllabification. The sequence [d.l] is perfectly acceptable in English if the two segments are not both in the onset, as in such words as *bedlam*, *Hadley*, *Adler*.

Other phonotactic prohibitions have also been shown to influence on-line phoneme categorization (Moreton 1999; Moreton and Amano 1999; Coetzee 2004, 2005). We first take up the prohibition in Japanese against heterorganic consonant clusters that has been studied extensively by Dupoux and his colleagues (Dupoux, Kakehi, Hirose, Pallier, and Mehler 1999; Dehaene-Lambertz, Dupoux, and Gout 2000; Dupoux, Pallier, Kakehi, and

[25] Hallé, Segui, Frauenfelder, and Meunier (1998) show that French listeners find it difficult to distinguish [d, t] from [g, k] before [l], and Hallé, Best, and Bachrach (2003) and Hallé and Best (2007) show that Hebrew listeners, whose language permits [dl] and [tl] in onsets, are much better at this discrimination than French or English listeners. This difference shows that the constraint against onset [dl, tl] is low-ranked relative to the constraint preserving the contrast between coronals and dorsals in Hebrew but high-ranked in English and French.

[26] This generalization is supported by the fact that loanwords with [bw, pw] exist in some English dialects e.g., in *Bu̲enos Aires* and *Pu̲erto Rico*, but no dialect has loanwords with initial [dl, tl].

Mehler 2001), and then return to the English prohibition against /dl, tl/ onsets (Breen, Kingston, and Sanders 2013).

The Japanese example lays the foundation for a challenge to Evolutionary Phonology in three steps. First, Dupoux et al. (1999) show that the prohibition against heterorganic clusters can cause Japanese listeners to hallucinate an epenthetic [u] between the consonants in such clusters. Second, Dupoux et al. (2001) show that the repair cannot come from the lexicon because the lexicon would in particular instances provide other vowels than [u]. Taken together, these two findings undermine Evolutionary Phonology's claim that what speakers and listeners know about their language is derived from the surface forms (the pronunciations) they encounter in the speech of other speakers and generalizations that can be drawn from those surface forms. They do so because the form of the repair, [u] epenthesis, is not one exemplified by alternations between surface forms in Japanese. Although [u] epenthesis is the means by which prohibited clusters in loanwords are repaired, the hallucinations demonstrated by Dupoux et al.'s (1999) results show that Japanese listeners do not perceive the offending consonants as ever occurring side-by-side in the original pronunciations of the loans. They therefore could not have noticed the difference between the original pronunciation of the loan without an epenthetic [u] between any heterorganic consonants and its repair in which the epenthetic [u] separates them.

Third, Dehaene-Lambertz et al.'s (2000) event-related potential (ERP) study shows that Japanese speakers' brains do not distinguish a heterorganic cluster from one repaired by [u] epenthesis. Dehaene-Lambertz et al. interpret this finding as evidence that the phonotactics of Japanese filters the incoming signal and immediately supplies anything necessary to make its percept phonotactically legal. This finding suggests a different interpretation of the findings of the Dupoux et al. (1999, 2001) studies, one which does not challenge Evolutionary Phonology's claim that what speakers and listeners know about their language is limited to what they can glean from surface forms. If Japanese listeners' brains cannot distinguish between a string with a heterorganic cluster and one repaired by [u] epenthesis, as Dehaene-Lambertz et al.'s ERP results and the hallucinated [u]s documented by Dupoux et al. (1999, 2001) suggest, then perhaps no heterorganic clusters occur in the surface forms of these strings for these listeners.

Results of a new ERP study reported by Breen et al. (2013) and summarized below show that Evolutionary Phonology does not escape unscathed after all. The results of all these studies are discussed next in more detail to make their contribution to our critique of Evolutionary Phonology more explicit.

Dupoux et al. (1999) presented French and Japanese listeners with a continuum of stimuli from [ebuzu] to [ebzo], where the [u] was progressively truncated to nothing, and asked them respond whether the vowel [u] occurred between the two consonants. French listeners' 'yes' responses decreased monotonically from 100% down to 0% as the vowel was truncated, while Japanese listeners' 'yes' responses did not drop below 70% even to the stimulus from which the entire vowel had been removed. The phonotactic prohibition in Japanese against clusters such as [bz] creates a perceptual illusion or hallucination: where a vowel must occur, Japanese listeners hear one, even if there is actually no vowel there at all. Japanese listeners were also much poorer than French listeners at speeded discrimination of stimuli which differed in whether a vowel intervened between two consonants, e.g., [ebuzo] versus [ebzo].

Dupoux et al. (2001) showed that these effects reflect a phonotactic constraint and not just the occurrence of the vowel [u] between two consonants in many Japanese words, particularly in many loanwords. Listeners transcribed and made lexical decision judgments for non-word strings containing heterorganic clusters that either have a single lexical neighbor with [u] between the consonants (e.g., the string [sokdo] has the neighbor [sokudo] 'speed'), or a single lexical neighbor with some vowel other than [u] (e.g., the string [mikdo] has the neighbor [mikado] 'emperor'). If the illusion simply reflects the possibility that the string is a word once a vowel has been added to it, then listeners should transcribe [mikdo] with its missing [a] and identify it as a word as readily as they transcribe [sokdo] with its missing [u] and identify it as word. If they instead hallucinate that an [u] is present because the grammar supplies it to repair violations of the ban on heterorganic clusters such as [kd], then they should instead identify [sokdo] as a word more often than [mikdo]. They should also frequently transcribe [mikdo] with [u] between the two consonants, even though [mikudo] is not a word. In conformity with this alternative prediction, listeners inserted [u] into their transcription of [mikdo] strings nearly as often as into [sokdo], despite the absence of any corresponding word [mikudo]. They also they identified [mikdo] strings far less often as words than [sokdo] strings, despite the existence of the word *mikado*. Finally, response times in lexical decision to [mikdo] strings were as slow as those to the corresponding non-word strings [mikudo] while those to [sokdo] strings were as fast as those to the corresponding word strings [sokudo]. All these results follow if the phonotactics supplies the missing vowel rather than the lexicon. The phonotactic constraint must introduce the illusory vowel before any lexical item is activated because [u] was inserted into the transcriptions of [mikdo] strings nearly as often as into [sokdo] strings, despite the lexical competition from the [a] in [mikado], and [mikdo] was thus identified as a nonword as slowly as [mikudo].

As noted above, these two sets of results undermine Evolutionary Phonology's claim that what speakers and listeners know about their language is limited to what can be learned from surface forms and relationships between them. They could not learn that the language repairs heterorganic clusters by [u] epenthesis because there are no alternations between forms in which epenthesis has applied and those in which it has not. Of course, [u] epenthesis is robustly exemplified in the loan vocabulary, but to know that it has applied in those words, a Japanese speaker would have to know the original unrepaired pronunciation of the loan word in the source language. The evidence that they hallucinate vowels between prohibited heterorganic clusters suggests that even if they heard the loan's original pronunciation, they would not perceive any heterorganic consonants as occurring next to one another.

A more fundamental question lurks behind this discussion: how does a Japanese listener learn that such clusters are prohibited in the first place? Blevins (2004) proposes that a language's phonology is induced from positive evidence (see, for example, footnote 15, p. 237). In its strongest form, this proposal claims that Japanese listeners never learn that heterorganic clusters are actually prohibited; they would simply never encounter them in Japanese, either in native vocabulary or established loans. The evidence reviewed so far instead requires that Japanese actively prohibit heterorganic clusters, to explain why the person who first borrows a foreign word containing a heterorganic cluster immediately repairs it with [u] epenthesis. They do so because the language's phonotactics cause them to hallucinate a vowel between the offending consonants and never actually perceive the loan as having such a heterorganic cluster.

To account for this prohibition's activity, Japanese learners must have been able to learn that heterorganic clusters are prohibited from their absence in surface forms. Evolutionary Phonology permits language-specific constraints (Blevins 2004:244), so systematic absences or gaps in surface forms may be sufficient evidence to learn that the absent strings are actually prohibited.[27]

We turn finally to the behavioral and neuro-physiological data reported by Dehaene-Lambertz et al. (2000). In their experiment, French and Japanese listeners heard four repetitions of strings such as [igumo] or [igmo] that either did or did not have the vowel [u] between the two consonants. These four repetitions were followed either by a fifth repetition of the same string or by a string that differed in the presence or absence of the vowel [u] between the two consonants. Listeners labeled the fifth stimulus as the 'same' or 'different' from the first four as quickly as possible. French listeners responded correctly to different trials far more often than Japanese listeners, 95.1% versus only 8.9%. Responses were also significantly slower in different trials compared to same trials for the French but not the Japanese listeners. Like the accuracy data, the absence of any RT difference in the Japanese listeners' responses suggests they usually did not notice that the fifth stimulus was different in the different trials.

Simultaneous recordings of event-related potentials (ERPs) showed that the French listeners' brains also detected a difference that Japanese listeners' brains did not. An ERP obtained in an interval 139–283 ms following the moment when the fifth stimulus deviated from the four preceding stimuli on different trials was significantly more negative in voltage than that obtained on same trials for French but not Japanese listeners. Dehaene-Lambertz et al. interpret this ERP as arising when the brain detected the sensory mismatch between the different stimulus and the sensory memory of the preceding reference stimulus.[28] Just as the behavioral response showed that the Japanese listeners seldom consciously heard any difference between [igumo] and [igmo] strings, this early ERP shows that their brains did not notice any difference either. Dehaene-Lambertz et al. interpret these findings as evidence that the phonotactic restrictions of Japanese filter the incoming signal early in its processing and correct any phonotactic errors, here supplying a vowel between heterorganic consonants when the signal fails to do so. A constraint that is as bred in the bone as to prevent the brain from noticing the difference between [CC] and [CuC] would surely be evidence that a prohibition can be learned from absence of evidence for the prohibited surface pattern.

Breen et al. (2013) report an ERP experiment which evaluated the extent to which the phonotactic constraint against tautosyllabic [dl, tl] onset clusters in English filters signals containing these onset clusters and repairs them to [gl, kl]. English (and French) listeners frequently mistake [d] and [t] for [g] and [k] before [l] (Hallé, Segui, Frauenfelder, and Meunier 1998; Hallé, Best, and Bachrach 2003; Hallé and Best 2007), and as Moreton

[27] The difference in English listeners' responses to [dl] versus [bw] onsets in Moreton's (2002) results shows that the mere presence of a gap is not sufficient to learn that a string is prohibited, but we will set aside this difficulty for Evolutionary Phonology in the remainder of this discussion.

[28] This ERP's timing, polarity, and cortical topography closely resembles the mismatch negativity (MMN) obtained whenever the brain detects that the current stimulus is auditorily different from the immediately preceding one.

(2002) showed, English listeners are quite reluctant to respond ‘l’ after a stop that they identify as ‘d’ rather than ‘g’ (see also Massaro and Cohen 1983). Breen et al. presented listeners with a pair of CCV syllables on each trial; the first was the *prime* and the second the *target*. *Test* prime-target pairs differed in the ***status*** of the prime: *legal* [dwa gwa], *absent* [bwa gwa], or *illegal* [dla gla] (and their counterparts with voiceless stops).[29] For all three statuses, the initial consonant in the target differed by just one place feature from that in the prime. As illustrated by the examples, the onset clusters in the targets were always legal. Two other types of pairs were presented. In *identity* pairs, the prime was the same as the target: legal [gwa gwa], absent [gwa gwa], and illegal [gla gla], and in *control* pairs, the initial consonant in the target again differed from the prime in just one feature, voicing rather than place: legal [kwa gwa], absent [kwa gwa], and illegal [kla gla].[30]

Listeners rated the similarity of the target to the prime. In all three statuses, targets were rated as highly similar to their primes for identity pairs and highly dissimilar for control pairs. Dissimilarity ratings were also high for test pairs with legal and absent primes, e.g., [dwa gwa] and [bwa gwa], respectively, but not for illegal test pairs, e.g., [dla gla], which were rated as similar as the corresponding identity pair [gla gla]. As in the other behavioral studies mentioned above, English listeners strongly disprefer hearing [dl] (and [tl]).

The ERPs told quite a different story, at least early on. In an interval 200–350 ms after the target onset, the voltages recorded from right anterior electrodes were significantly more positive for both test and control trials than identity trials for all three kinds of primes, illegal as well as legal and absent, and test and control trials at these sites in this interval did not differ from one another for any prime status. The brain was apparently able to detect the place difference between target and prime as readily in illegal [dla gla] as legal [dwa gwa] and absent [bwa gwa] pairs. Later, in the interval 350–600 ms after the target's onset, at central posterior electrodes, the voltages were equally more positive for test and identity than control pairs for both illegal and absent primes, but for legal primes the voltages were no more positive for test than control pairs. During this interval at these scalp locations, the brain could no longer distinguish the target from the prime for illegal [dla gla] or absent [bwa gwa] pairs any better than the corresponding identity pairs [gla gla] or [gwa gwa]. In all respects but one, this later posterior potential exactly matches the similarity ratings—the exception is the absent test pairs [bwa gwa], which were rated as dissimilar as the legal test pairs [dwa gwa]

These results indicate that the brain can at first distinguish [d] from [g] before [l] as readily as it can before [w], and likewise [b] from [g] before [w]. Only later does [d] become indistinguishable from [g] before [l]. This result indicates that the phonotactics do not apply immediately, but only after some delay. Once they do apply, the original veridical percept that [d] is as dissimilar from [g] before [l] as before [w] is overwritten by correcting [d] to [g] before [l]. These findings do not indicate that the prohibition

[29] The terms for the statuses, ‘legal’, ‘absent’, and ‘illegal’, only describe the primes in the test pairs, as all the primes are legal for the other two types of pairs, identity and control pairs. These terms are nonetheless applied to pairs of these two types to make explicit which test pair each corresponds to.

[30] The identity trials are the same for legal and absent pairs because /gwa/ was the target for both the legal prime /dwa/ and the absent prime /bwa/.

against tautosyllabic /dl/ does not influence perception, but instead that it does not do so immediately. Phonotactics are late, in English if not Japanese.

More research is needed to determine whether the discrepancy between Dehaene-Lambertz et al.’s and Breen et al.’s results can be attributed to differences in methods, language, or the nature of the phonotactic constraints. Regardless of which of these differences is responsible for this discrepancy, Breen et al.’s results challenge Evolutionary Phonology's restriction on what can be learned to the properties of and relationships between surface forms. If it's possible to learn that a pattern is prohibited from its absence in surface forms, then English speakers should have learned that neither [dl] nor [tl] is a possible onset from their absence at the beginnings of syllables in English. That prohibition should also have made it impossible for English speakers to hear either [d] to [t] before [l] in an onset, a prediction that is apparently confirmed by English listeners' behavioral responses in experiments like those reported by Moreton (2002) and Hallé and Best (2007) as well as by Breen et al.’s similarity judgments. Yet English listeners’ early neural response indicates that [d, t] are at first veridically perceived as different from [g, k] in this context. If the brain can still detect the difference between these sounds as readily before [l] as before [w] long after the prohibition is learned, how is it that the prohibition against [dl, tl] onsets is not *unlearned* once the listener is exposed to syllables with such onsets in surface forms?

On the one hand, if the prohibition could be learned merely from the absence of a particular pattern in the learning data, i.e., from its absence in surface forms, then it should be just as easy to unlearn once the data change, as they do in the experiments in which listeners are exposed to [dl, tl] onsets. Unlearning should be especially easy given that the brain remains able to distinguish [d, t] from [g, k] before [l]. That the prohibition is not unlearned and still determines both the later neural response and the listener's overt behavioral response in Breen et al.’s experiment indicates that the listener has learned a prohibition against the co-occurrence of abstract categories rather than against the co-occurrence of superficial sounds.

On the other hand, the prohibition's pertinacity after [dl, tl] onsets are encountered is expected if markedness constraints outrank faithfulness constraints in the grammar at the onset of learning. If the learner is never exposed to [dl, tl] onsets, then the faithfulness constraint that preserves the contrast between [d, t] and [g, k] in other contexts will never be promoted above the markedness constraint that prohibits this contrast in this context. Because that markedness constraint prohibits particular patterns of abstract categories, it need not regulate the brain's early response to the incoming signal and prevent it from distinguishing [d, t] from [g, k] before [l] as easily as elsewhere. Instead, the constraint need only ensure that the categories onto which the signal is mapped conform to the language's phonotactics, as determined by the markedness constraints that continue to outrank the competing faithfulness constraints. Substantially more than the brief exposure to a previously unencountered pattern than that obtained in the laboratory during these experiments would be required to demote a well-established high-ranked constraint.

These results thus challenge Evolutionary Phonology's insistence that what is learned is the properties of surface forms and the relationships between such forms. Learning about surface forms would be much more labile in response to changes in those surface forms than listeners' behavioral responses in these experiments actually are. Their pertinacity is instead a product of the abstract encoding of phonotactic constraints on sound sequences. Despite their being applied to abstract categories rather than concrete instances of the affected sounds, these constraints are clearly psychologically active, in that they regulate perception and can even cause listeners to hallucinate sounds that aren't present in the signal or which belong to different categories altogether.

3.4 Emergence of constraints

A pattern or behavior can be identified as emergent if it regulates behavior in the absence of any external stimulus or model, as does vocal babble in deaf infants, or when the evidence for the pattern or behavior in the external stimulus is so incomplete or degenerate that the learner cannot determine the nature of the pattern or behavior from that stimulus. Blevins (2004) argues that there is little if any evidence that the stimulus is impoverished in crucial ways when it comes to acquisition of phonology:

> "[P]honological systems provide an abundance of stimuli, with productive phonological alternations robustly cued. … By the time an infant reaches the age of eight or nine months, when a first word might appear, the child will have heard hundreds of thousands of tokens of the sound patterns of the native language. The bulk of phonological learning then, for which there is overt evidence, points to data-driven learning." (Blevins 2004:223–224)

Here, we review two investigations of infants' knowledge of the ranking of markedness and faithfulness constraints reported by Jusczyk and colleagues (Jusczyk, Smolensky, and Allocco 2002; Jusczyk, Smolensky, Arnold, and Moreton 2003; see also Davidson, Jusczyk, and Smolensky 2004) and two investigations of English speakers' knowledge of markedness differences between onset clusters that don't occur in their native language reported by Berent and colleagues (Berent, Steriade, Lennertz, and Vaknin 2007; Berent, Lennertz, Smolensky, and Vaknin-Nusbaum 2009; cf. Peperkamp 2007).[31] These two sets of examples complement one another in showing the emergence of synchronic constraints in infants and adults, respectively.

3.4.1 Emergence of constraints in pre-linguistic infants

Jusczyk et al. (2002) reports 10 studies using the head-turn preference procedure with 4.5-, 10-, 15-, and 20-month-old infants to determine whether markedness constraints start out ranked above faithfulness constraints, as hypothesized in accounts of language learning (Smolensky 1996; cf. the even stronger continual M » F bias proposed by Prince

[31] The need for gaps in that evidence to be interpreted as evidence of prohibitions against the missing structure has already been discussed in the preceding Section.

and Tesar 2004 and Hayes 2004).[32] The youngest infants in these studies, the 4.5-month-olds, are too young to have acquired any knowledge of the phonotactics of the ambient language (Jusczyk, Friederici, Wessels, Svenkerud, and Jusczyk 1993; Jusczyk, Luce, and Charles-Luce 1994), so evidence that markedness constraints are ranked above faithfulness constraints for infants this young would confirm the hypothesis that markedness constraints are innately ranked over faithfulness constraints in the initial state. This result would also demonstrate that the particular constraints observed to be ranked in this order are themselves innate.

Jusczyk et al.'s experiments examined the markedness constraint against heterorganic nasal-stop clusters, e.g., [nb], and the competing faithfulness constraint that requires the nasal's place of articulation in the input remain unchanged in the output. The markedness constraint can be referred to as AGREE, in acknowledgment that the typical repair is for the nasal to assimilate in place to the following stop in such clusters. The competing faithfulness constraint is then IDENT[PLACE]. If AGREE outranks IDENT[PLACE], the input /nb/ maps onto the output [mb]; the opposite ranking preserves the nasal's coronal place in the output [nb].

The stimuli in all the experiments had the same form: a syllable ending in a nasal, e.g., *un*, was presented first, then after a brief pause a syllable beginning with a stop, e.g., *ber*, was presented, and then following another pause two syllables would be presented as a single continuous utterance, e.g., *unber*. The first presentation of the two syllables in which they were separated by a pause is analogous to the input, because the pause makes AGREE irrelevant, while the following continuous presentation corresponds to a possible output. The example, *unber*, obeys IDENT[PLACE] but violates AGREE. Participants in the experiments were presented with lists of such input-output strings that conformed to one ranking from a source on one side in alternation with lists that conformed to another ranking from a source on the opposite side. Significantly longer looking times to the source producing one of the lists than at the other were interpreted as evidence that the list coming from that source corresponded better to the infants' expectations, where those expectations reflected their knowledge of the two constraints and their ranking.

Jusczyk et al.'s first three experiments established that 10-month-old infants' looking times differ in ways that would be expected if their perception is regulated by AGREE and IDENT[PLACE] individually and by the two together with AGREE ranked over IDENT[PLACE]. In conformity to AGREE, they looked longer to the source of lists in which the continuous presentation of the two syllables obeyed AGREE, e.g., to unmarked *um ber umber*, than to those that did not, i.e., marked *un ber unber*. IDENT[PLACE] is obeyed in both kinds of lists. In conformity to IDENT[PLACE], the infants also looked longer to the source of lists that obeyed this faithfulness constraint and which both obeyed AGREE, i.e., to faithful *um ber umber* than unfaithful *um ber iŋgu.* They also looked longer to the source of lists that obeyed AGREE and violated IDENT[PLACE], unmarked but unfaithful *un ber umber*, than to those that violated AGREE and obeyed IDENT[PLACE], marked but faithful *un ber unber*.

[32] Only the results from the youngest two age groups, 4.5 and 10 months, are discussed here, because only they are directly relevant to the question of whether markedness constraints outrank faithfulness constraints in the initial state.

Although these results conform to the hypothesis that markedness is ranked over faithfulness, infants have had sufficient exposure to the ambient language by 10 months that their looking times could reflect their having learned that it is English phonotactics that requires this ranking rather than its being the ranking in the initial state (see Jusczyk et al. 1993, 1994). Jusczyk et al. (2002) accordingly ran the same experiments with much younger 4.5-month-old infants. They produced the same pattern of results as the older 10-month-olds. As infants this young do not otherwise show any evidence of having learned the ambient language's phonotactics (Jusczyk, Friederici, Wessels, Svenkerud, and Jusczyk 1993; Jusczyk, Luce, and Charles-Luce 1994), this pattern can in their case instead be attributed to a ranking of markedness over faithfulness constraints in the initial state.

This simple story is complicated but not undermined by additional results obtained from 4.5-month-old infants reported by Jusczyk et al. (2003). They found that infants at this age did not look significantly longer to the source of unmarked and unfaithful *an bi andi* than to marked and faithful *an bi anbi*. The former stimulus obeys AGREE by assimilating the stop to the nasal's place rather than vice versa. If markedness were ranked over faithfulness, then infants should have looked longer to unmarked and unfaithful *an bi andi* than marked but faithful *an bi anbi*.

Jusczyk et al. (2003) suggest that perhaps infants showed no preference in this case because they could hear the difference in place in the stop between the input *an bi* and the output *andi*. This suggestion acknowledges that this repair is unattested in adult languages (except when the nasal is retroflex, Steriade 2001), where it is blocked by a higher-ranked positional faithfulness constraint requiring that place be preserved in onsets (Beckman 1997). The absence of this repair, the frequent resort to place assimilation by the nasal, and positional faithfulness to place of articulation in general can, of course, be explained by well-documented differences in the perceptibility of place cues in codas compared to onsets, or, if one prefers, before the close articulation of consonants compared to before the open articulation of vowels (Steriade 2001; Blevins 2004, and many earlier studies cited in these sources). Jusczyk et al. tested this hypothesis by comparing looking times to unmarked unfaithful *an bi ambi* with those to unmarked faithful *am bi ambi*. The infants looked significantly longer to the former than the latter. This result agrees with the earlier finding that infants at this age look longer to unmarked unfaithful *an bi ambi* than to marked but faithful *an bi anbi*, and thus shows that infants can detect the change in the coda nasal's place in the preferred stimuli. But it also shows that infants prefer to listen to an unmarked stimulus that violates a faithfulness constraint, *an bi ambi*, over an equally unmarked stimulus that does not, *am bi ambi*.

This result led Jusczyk et al. to an alternative interpretation of infants' preferences, in which they do not prefer stimuli that conform to their current ranking (markedness over faithfulness) over those that don't conform, but instead prefer stimuli that conform to their current ranking over those that conform to a different ranking. Unmarked unfaithful stimuli in which the stop has assimilated in place to the nasal, *an bi andi*, aren't preferred over unmarked faithful *am bi ambi* because no ranking can produce an output in which the stop has assimilated. Unmarked unfaithful *an bi ambi* is preferred over unmarked and faithful *am bi ambi* because the ranking of AGREE over IDENT[PLACE] produces the former while either ranking produces the latter. Their last experiment tested this hypothesis by comparing infants' looking times to unmarked unfaithful *an bi ambi* with those to unmarked unfaithful *an bi andi¸*which differ in whether obedience to AGREE and violation of IDENT[PLACE] are achieved by changing the place of the nasal in the coda or stop in the onset. Infants did not look significantly longer to the source of either stimulus type, as expected if their looking times reflect a preference for the current grammar over an alternative grammar.

Although these results have complicated the interpretation of Jusczyk et al.'s (2002) results, they do not undermine their most basic finding, that markedness is ranked over faithfulness for infants who are too young to have learned the ambient language's phonotactics. As noted at the beginning of this discussion, this finding both supports the hypothesis that this ranking characterizes the initial state and can therefore be interpreted as innate, and it supports the hypothesis that constraints with these properties are themselves innate.

3.4.2 Emergence of constraints in adults

Berent et al. (2007, 2009) compared English speakers' intolerance for onset clusters that differ in how much they deviate from the ideal rising sonority profile. Because none of the clusters examined occur as onsets in English, speakers of this language could treat them all as equally deviant. Alternatively, if they can access markedness constraints shared by speakers of all human languages, then they could distinguish onset clusters by how far they deviate from this ideal profile.

Berent et al. (2007) compared onset clusters consisting of smaller sonority rises than occur in English (stop-nasal, e.g., *bnif*), sonority plateaus (stop-stop, e.g., *bdif*), and sonority falls (liquid-stop, e.g., *lbif*), while Berent et al. (2009) compared onset clusters beginning with a nasal in which sonority rose (e.g., *mlif* or *nwat*) or fell (e.g., *mdif* or *nbat*).[33] Both studies used syllable-counting and discrimination tasks to compare English speakers' responses to these clusters. In the syllable-counting task, English speakers judged whether these monosyllables and disyllables created by epenthesizing a schwa between the two consonants in the onset consisted of one syllable or two, i.e., *bnif* versus *benif* (where 'e' stands for schwa). In the discrimination tasks, English speakers judged whether the monosyllables and the corresponding disyllable were different from one another. Berent et al. (2007) also tested the extent to which the disyllable served as an identity prime for the monosyllable, while Berent et al. (2009) replicated the auditory version of the discrimination task with spelled forms of the stimuli. If a cluster isn't tolerated, it can be repaired by perceptually epenthesizing a schwa between the two consonants, as in the disyllabic strings used in these experiments. If English speakers tolerate the more deviant of these clusters less well, they would be more likely to judge the monosyllable as disyllabic, to fail to discriminate it from the disyllable when presented either auditorily or visually,[34] and to be primed to recognize the monosyllable following the disyllable.

---

[33] Both studies also examined Russian speakers' responses in the various experiments. Their responses will not be discussed here because Russian permits clusters of all these types.

[34] They would fail to discriminate the CCVC from the CəCVC strings when presented visually if they encoded them phonologically by converting the graphemes to phonemes.

These expectations were all confirmed. First, English speakers' success at judging the monosyllables as consisting of just a single syllable was 63% for onset clusters with small sonority rises *bnif*, fell to 28% for those with plateaus *bdif*, and dropped again to 14% for those with falls *lbif*.[35] Monosyllables with rising sonority nasal-initial onsets *mlif* and *nwat* were judged to consist of just one syllable on 90% of trials, but falling sonority *mdif* and *nbat* were only correctly judged to be monosyllabic on 71% of trials.[36] All pair-wise differences were significant. Second, English speakers discriminated monosyllables with small sonority rises *bnif* from their epenthetic counterparts *bənif* on 66% of trials, but only discriminated monosyllables with plateaus or falls, *bdif* or *lbif*, from their epenthetic counterparts on 32% of trials. Monosyllabic rising sonority *mlif* and *nwat* were successfully discriminated from epenthetic *məlif* and *nəwat* on 81% of trials, while falling sonority *mdif* and *nbat* were only discriminated successfully from epenthetic *mədif* and *nəbat* on 73% of trials. The visual discrimination task replicated this difference: rising sonority 91% correct versus falling sonority 85% correct. Finally, epenthetic *bədif* failed to prime the corresponding monosyllable with a sonority plateau *bdif* compared to the actual identity prime *bdif*, but epenthetic *ləbif* primed *lbif* just as much as *lbif* itself did. To determine whether this last result is just a product of listeners' confusing epenthetic *ləbif* phonetically with *lbif*, Berent et al. (2007) manipulated the stimulus-response contingencies in a second version of the priming experiment so as to draw participants' attention to the epenthetic vowel's presence. If participants confused epenthetic *ləbif* phonetically with *lbif*, this manipulation should not have changed the extent to which epenthetic *ləbif* serves as an identity prime to *lbif*. Instead, it eliminated the priming of *lbif* by epenthetic *ləbif.* Berent et al. interpret this change as evidence that the grammar rather than the phonetics rendered *lbif* identical to *ləbif* by epenthesizing a schwa between the two consonants.

These results all show that English speakers can distinguish onset clusters by how much they deviate from the ideal rising sonority profile—falling sonority onsets are the worst, plateaus are better, and small rises are better still—and they can do so despite never having encountered any of these onset clusters. The ranking of markedness constraints that determines the deviance of these various onset clusters and the degree of English speakers' intolerance of them has emerged out of the fund of linguistic knowledge that they share with speakers of other languages.

Before closing this section, we must evaluate an alternative explanation of English speakers' responses in Berent et al.'s (2007, 2009) experiments, namely that they can be predicted from their experience with sound sequences that occur in English rather than universal sonority sequencing constraints on consonant clusters.[37] Blevins argues explicitly for such an inductive alternative to deduction from UG when she asserts that the stimulus for phonological learning is anything but impoverished (Blevins 2004:219–

[35] Percents are estimated by eye from the figures in Berent et al. (2007) to the nearest whole percent and rounded to the nearest whole percent from the values tabulated in Berent et al. (2009).

[36] The difference in the precision with which the results are presented here reflects the fact that means were only displayed in figures in Berent et al. (2007), while they were tabulated in Berent et al. (2009).

[37] There is perhaps no small risk of vicious circularity here, as the sound sequences an English speaker observes in that language are themselves likely to be governed by universal sonority sequencing constraints. The simulation presented here avoids that circularity only to the extent that it does not overtly try to use such constraints.

224). We demonstrate here that the predictions English speakers can generate from their observations of strings that occur in English fall short of accounting for English speakers' responses in Berent et al.'s (2007, 2009) syllable-counting experiments.[38]

The inductive alternative was simulated by:

1. Extracting the frequency of all word-initial CV, CCV, and CCCV strings that occur in English from version 2 of IPhOD (Vaden, Halpin, and Hickok 2009), and similarly the frequency of all occurring word-initial CəCV strings. These strings and their frequencies serve as the models from which an English speaker could learn whether a particular onset is possible in the language and how likely it is to occur, as well whether and how likely a CəCV string is. For brevity, these two sets of strings are referred to henceforth as 'onset' and 'epenthetic' strings.

2. The constraints comprising grammars that would produce the observed onset and epenthetic strings as optimal outputs were then generated using the UCLA Phonotactic Learner (Hayes and Wilson 2008). Table 1 shows how each segment was specified for distinctive features; these specifications are close but not identical to those in Hayes and Wilson (2008). Because onset strings differ fundamentally from epenthetic strings and there's no principled reason why constraints on one should be related to those on the other, the Learner was used to generate grammars for them separately. The final grammars for onset and epenthetic strings were the unions of the constraints generated by passing them each through the Learner 10 times.[39]

	cons	son	app	cont	nas	lab	cor	dor	ant	voi	strid	lat	syl
b	+	-	-	-	0	+	0	0	0	+	0	0	0
d	+	-	-	-	0	0	+	0	+	+	0	0	0
dʒ	+	-	-	-	0	0	+	0	0	+	+	0	0
g	+	-	-	-	0	0	0	+	0	+	0	0	0
p	+	-	-	-	0	+	0	0	0	-	0	0	0
t	+	-	-	-	0	0	+	0	+	-	0	0	0
tʃ	+	-	-	-	0	0	+	0	0	-	+	0	0
k	+	-	-	-	0	0	0	+	0	-	0	0	0
m	+	+	-	0	+	+	0	0	0	0	0	0	0
n	+	+	-	0	+	0	+	0	+	0	0	0	0
ŋ	+	+	-	0	+	0	0	+	0	0	0	0	0
v	+	-	-	+	0	+	0	0	0	+	0	0	0
ð	+	-	-	+	0	0	+	0	+	+	-	0	0
z	+	-	-	+	0	0	+	0	+	+	+	0	0
ʒ	+	-	-	+	0	0	+	0	-	+	+	0	0
f	+	-	-	+	0	+	0	0	0	-	0	0	0
θ	+	-	-	+	0	0	+	0	+	-	-	0	0

[38] After presenting our argument, we respond to Daland et al.'s (2011) case for the inductive alternative; specifically that the participants' responses in Berent et al.'s (2007, 2009) syllable-counting experiments can be 'projected' from the statistical properties of clusters in the English lexicon. Our argument was developed independently of theirs, at the same time as their paper became available.

[39] The Learner generates weights for each constraint; these were not used in the next step of the simulation because the strings they were next applied to are unattested in English, and were therefore likely to demand different constraint weights.

s	+	-	-	+	0	0	+	0	+	-	+	0	0
ʃ	+	-	-	+	0	0	+	0	-	-	+	0	0
l	+	+	+	0	0	0	+	0	+	0	0	+	0
r	+	+	+	0	0	0	+	0	-	0	0	0	0
w	+	+	+	0	0	+	0	+	0	0	0	0	0
j	+	+	+	0	0	0	+	0	0	0	0	0	0
h	-	-	+	+	0	0	0	0	0	-	0	0	0
ə	0	0	0	0	0	0	0	0	0	0	0	0	+

Table 1. Feature specifications used in representing word-initial C, CC, and CCC onsets and CəC epenthetic repairs in the inputs to the UCLA Phonotactic Learner (cf. Hayes and Wilson 2008). 0s are used for feature values that can reliably be predicted from other values. 'ə' represents both any V in the onset strings, CV, CCV, and CCCV, and the epenthetic strings, CəCV, and specifically the schwa in epenthetic strings.

3. The violations incurred by each CC sequence in Table 2 below and the corresponding CəC sequence for each of the constraints in these two grammars were tallied, as were those for a grammar that combined the two sets of constraints. These strings are a slightly expanded set of the CC onset clusters and their epenthetic repairs tested by Berent et al. (2007, 2009); the expansion tests a more complete set of the possible sonority rising, plateau, and falling onsets. (All inputs and candidates ended in a vowel.) All constraints that were not violated by any of the candidates were then purged, and the three resulting tableaux were submitted to the Maximum Entropy Grammar Tool (Capodieci, Hayes, and Wilson 2008). The frequencies of the onsets and their epenthetic repairs were the average percentages of one- versus two-syllable responses to monosyllabic stimuli with a particular onset type in Berent et al.'s syllable-counting experiments.[40] For our purposes here, the useful output of the Learner is how often a particular string is predicted by each grammar and how closely its predicted frequency matches the observed frequency with which the participants gave the corresponding one- versus two-syllable response.

Rise			Plateau		Fall		
pw	pn	tl	pt	bd	lp	rp	mt
bw	bn	dl	pk	bg	lt	rt	md
nw	tm	ml	tp	db	lk	rk	mk
	dm		tk	dg	lb	rn	mg
	km		kp	gb	ld	rd	np
	gm		kt	gd	lg	rg	nb
	kn						nk
	gn						ng

[40] Although English-speaking participants did not respond 'two syllables' to every disyllabic stimulus and the frequency of that response varied as a function of the sonority of the flanking consonants, they still responded 'two syllables' to all disyllabic stimuli on nearly 90% of trials or better. Because accuracy was so high overall on these trials and the range of variation as a function of the flanking consonants sonority is so slight, we made no attempt to model these responses.

Table 2: A modest superset of the onset clusters used in Berent et al.'s (2007, 2009) syllable-counting experiments and used here in simulating English speakers' generalization from observations of occurring word-initial CV, CCV, and CCCV onsets and CəCV strings in that language.

4. The extent to which the predicted frequencies matched the observed frequencies is a measure of how well participants' responses in the syllable-counting task can be predicted from their experiences with existing sound sequences in English. To get a more robust estimate of how well these predicted frequencies matched the observed frequencies, a new set of 'observed' frequencies was generated 1000 times by drawing from normal distributions defined by the means and confidence intervals of the observed frequencies of 'one-syllable' responses.[41] The predicted frequencies were generated from the means alone, so each new set of observed frequencies retests how well the predicted frequencies match the observed frequencies when the observed frequencies represent a new and more continuously variable sample from the same population that produced the original sample. This procedure may be thought of as a test of the generality of the predicted frequencies.

5. This match is most easily assessed by regressing the observed frequencies on the predicted ones. For this purpose, logistic regression models were fit to each set of observed frequencies of 'one-syllable' responses, using the frequencies predicted from the three maximum entropy models with the constraints generated by the Phonotactic Learner from:

a. Onset strings (CCC),
b. Epenthetic strings (CeC),
c. Onset and epenthetic strings (CCC+CeC).

To each of these models were then added additional parameters representing either the:[42]

---

[41] These values were determined by eye from Figure 1 in Berent et al. (2007) and from Figure 2 in Berent et al. (2009). For all sonority rises but /ml, nw/ they were 63±5%; for /ml, nw/ they were 90±3%; for sonority plateaus, they were 28±5%; for all sonority falls but /md, nb/, they were 14±5%; and for /md, nb/, they were 71±3%. (The exceptional onsets in each case are from Berent et al. (2009), all the others are from Berent et al. (2007).) These estimates are accurate to roughly 1%. The confidence intervals are confidence intervals for differences between means, which are √2 greater than the confidence intervals of the means themselves. The standard deviations of the normal distributions thus equal (confidence interval * √n)/(√2 * 1.96), where n is the number of participants in each experiment (16 for the syllable-counting experiment reported by Berent et al. (2007) and 26 for the one reported by Berent et al. (2009)). Because the values are percentages, they were limited to the range 0-100.

[42] Before doing so, it's vital to establish that these predictors are not collinear. Values of κ for the CCC and CeC predictors and the predictors for Profile, Difference, and Values range from 6.2–8.7, which are just above the value, 6, below which there is no appreciable collinearity (values close to 15 would indicate modest collinearity). These predictors may therefore be combined safely. The predictors in the Profile, Difference, and Values models are of course highly collinear, so they cannot be combined with one another.

d. Sonority profile of the onset, where a rise was coded as 1, a plateau as 0, and fall as -1 (Profile),
e. Sonority difference between the consonants in the onset, calculated from the following sonority values for the four manners of articulation (Difference):[43]

i. Stops	-3
ii. Nasals	-1
iii. Liquids	1
iv. Glides	3,

or
f–g. Sonority values of the consonants in the onset, using the values in (e) (Value). Two models were constructed using these values, one in which they were independent (f) and the other which included their interaction (g).

The parameters that the Profile, Difference, and Value models add to the models in (a–c) represent with various degrees of coarseness the sonority sequencing constraints of UG. If models including any of these parameters substantially improve the fit over models that only represent English speakers' experience of onset and epenthetic strings in the language, then that experience cannot fully explain how the differences between clusters influenced their syllable counts in Berent et al.'s (2007, 2009) stimuli.

Table 3 lists the mean residual deviances for the CCC, CeC, and CCC+CeC models without and then with these additional parameters. Whether a model with an additional parameter improves the fit over one without it can be assessed from the difference in residual deviance between the two models for the difference in their number of parameters. This difference is distributed as X^2 and its probability can be estimated using the difference in the number of parameters as the degrees of freedom in a X^2 test. The degrees of freedom for these tests in comparing models (d) and (e) to models (a–c) are 1 (44-43), for model (f) they are 2, and for model (g) they are 3. The cutoffs for significance at $p < 0.05$ and $p < 0.01$ for X^2 statistics for 1–3 degrees of freedom are 3.84, 5.99, 7.81 and 6.63, 9.21, 11.34, respectively. The smallest difference in residual deviance between models (d–g) and models (a–c) is 25.09, which indicates that all models with additional parameters significantly improve the fit over models (a–c). The only case where a difference in residual deviance falls below significance at $p < 0.01$ is for model (c+g) versus (c+f) where the difference, 2.93, is only marginally significant, $0.10 > p > 0.05$.

Models (*df*)	CCC	CeC	CCC+CeC
(a–c) (44)	436.70	646.47	345.66
(a–c)+(d) Profile (43)	345.49	621.38	246.90
(a–c)+(e) Difference (43)	390.80	589.65	260.21
(a–c)+(f) Value Independent (42)	333.55	579.81	226.69
(a–c)+(g) Value Interaction (41)	320.93	495.90	223.76

[43] These values were used rather than the 0–3 scale in Clements (1988) or Daland et al. (2011) so that the predictor would be centered. Centering constrains correlations between fixed effects. The values used for the clusters' sonority profiles in (b) were chosen for the same reason.

Table 3. Residual deviances for (a) CCC, (b), CeC, and (c) CCC+CeC models without additional parameters (top row), and with additional parameters representing the sonority Profile of the onset cluster (d), the sonority Difference between the consonants in the onset (e), the sonority Values of the consonants in the onset, either independently (f) or interacting (g). The numbers in parentheses are the degrees of freedom in each model.

To get an idea of when predicted and observed frequencies do not match and thus of the principal sources of residual variance, we have plotted in ascending order in Figure 1 how often the absolute value of the standardized residual corresponding to each cluster exceeds 5 for the 1000 iterations of the (a) CCC, (b) CeC, and (c) CCC+CeC models. The higher the bar for a particular cluster, the more often the predicted frequency mismatched its observed frequency to this extent. The three panels in the figure differ in how often the residual corresponding to a particular cluster exceeds the threshold, and likewise how often clusters with particular sonority profiles do so. The constraint sets from which the predicted frequencies are derived must therefore differ in which combinations of feature values they mark as ill-formed. The CCC model in (a), in which predicted frequencies are derived from constraints on onset strings, fails to predict the observed frequencies for a mixture of clusters with sonority falls and plateaus, while the CeC model in (b), where predicted frequencies are instead derived from constraints on epenthetic strings, fails to predict observed frequencies mostly for clusters with sonority falls, but also two with rises. The combined CCC+CeC model in (c) unsurprisingly produces fewer extreme residuals, and the clusters whose residuals are most often extreme are a mixture of sonority falls, plateaus, and rises.

Figure 2 is a similar plot for the models which include the clusters' sonority profile (Fig. 2A, a–c+d), the sonority difference between the consonants (Fig. 2B, a–c+e), or the sonority values of the two consonants and their interaction (Fig. 2C, a–c+g) (model (f), in which the sonority values of the individual consonants are independent, has been omitted because model (g) fits the data better). Clusters with sonority falls (black bars) still produce the greatest number of extreme residuals for the models that add any of these parameters to the CeC model, while those which add them the CCC or CCC+CeC models instead produce extreme residuals for a mixture of types of clusters. What is equally plain is that all of these models produce fewer extreme residuals for any type of cluster than the models that do not include any of these parameters, and that the model which produces the fewest is the one with predicted frequencies from the CCC and CeC constraint sets, the sonority values of C1 and C2, and their interaction.

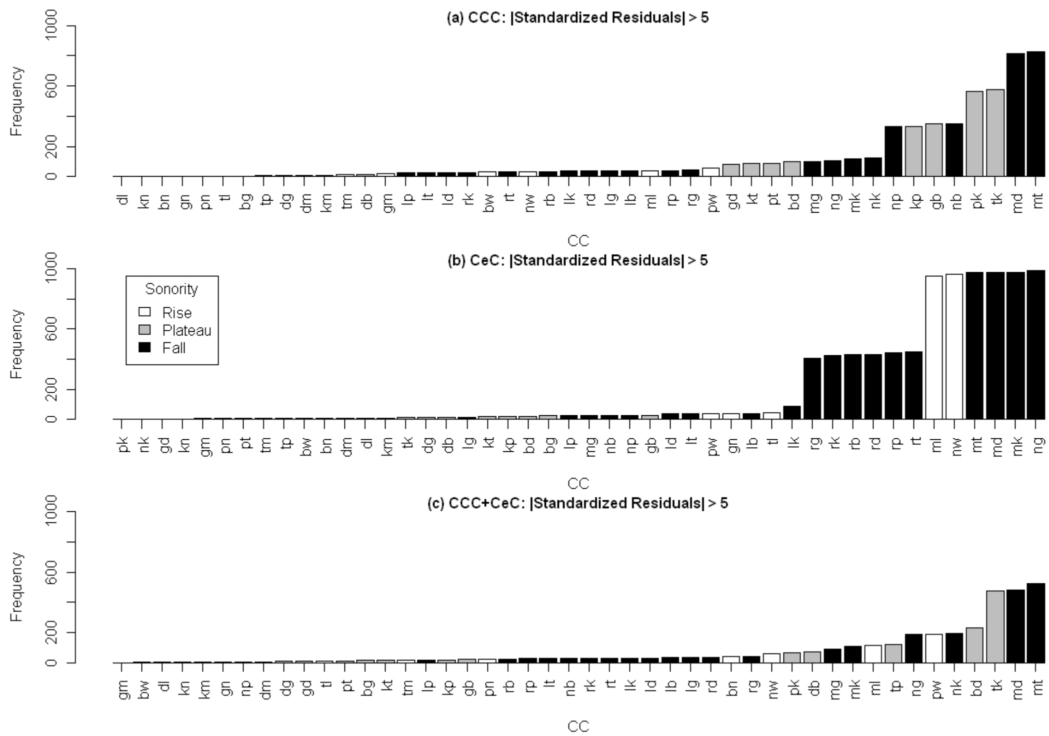

Figure 1. Ranked frequencies with which the absolute value of the standardized residual for a cluster exceeded 5, from 1000 iterations of the (a) CCC, (b) CeC, and (c) CCC+CeC models. White bars represent clusters in which sonority rises, gray bars those in which it plateaus, and black bars those in which it falls.

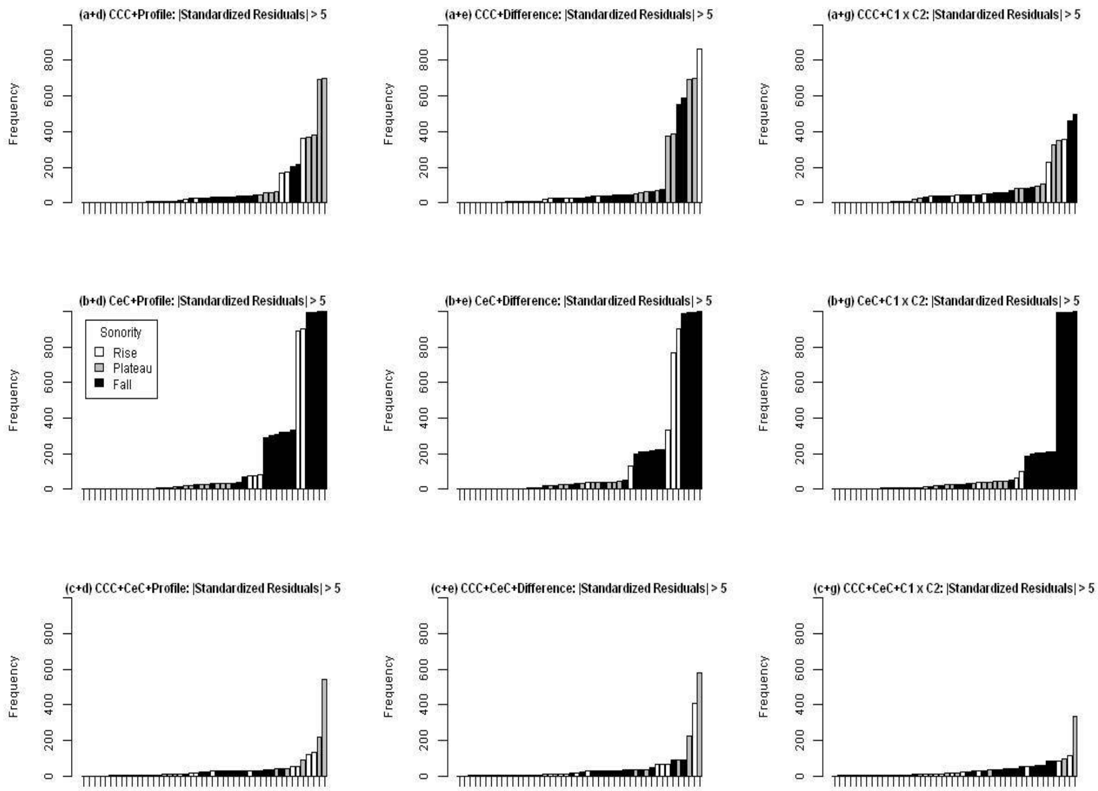

Figure 2. Ranked frequencies with which the absolute value of the standardized residual for a cluster exceeded 5, from 1000 iterations of the (left, a–c+d) Profile, (middle, a–c+e) Difference, and (right, a–c+g) interaction of Values models. White bars represent clusters in which sonority rises, gray bars those in which it plateaus, and black bars those in which it falls.

The noticeable improvement in fit obtained by adding the predictors from the Profile, Difference, or Values models shows that the Learner did not glean enough information from observations of the consonant clusters that occur in English or the corresponding epenthetic repairs to predict how often an English speaker would judge a particular non-occurring CCVC string as consisting of one syllable rather than two in Berent et al.'s (2007, 2009) syllable-counting experiments. These predictors, moreover, refer to the same properties or features of segments that universal constraints on sonority sequencing do and thus demonstrate that such constraints are useful in predicting participants' responses in these experiments. Finally, this improvement shows that the constraints generated by the Phonotactic Learner do not completely capture the differences in wellformedness between the various non-occurring consonant clusters examined in these experiments, despite the constraints' referring directly to the consonants' feature compositions.

We close by discussing Daland et al. (2011), who reject Berent et al.'s (2007) argument that English speakers' knowledge and experience of onset clusters that do occur in English predicts the distinctions they make between onset clusters that do not occur in the language. Daland et al. base their rejection on a demonstration that at least two models which represent the statistics of English speakers' experience, Hayes and Wilson's (2008) Phonotactic Learner and Albright's (2009) featural bigram model, produce wellformedness scores for the non-occurring clusters that correlate well with speakers' judgments of how English-like a variety of non-occurring clusters are.[44]

Daland et al. argue that these two models succeed because the inputs were syllabified, and the segments were represented by their features. Both criteria were met in our models, too, in that its statistics were based on word-initial onsets, which are necessarily preceded by a syllable boundary, and the segments were represented by their feature specifications (Table 1). The CCC, CeC, and CCC+CeC models show that English speakers' judgments of #CC strings that do not occur in English as CC or CəC also correlate with likelihood estimates based on the statistics of word-initial CV, CCV, CCCV onsets and CəC strings in English. However, the fit of both models was improved substantially by including additional predictors that directly represented the sonority sequence of the putative clusters. This improvement in fit shows that constraints derived from experience may fail to capture distinctions in the wellformedness between segment strings even when they refer directly to the segments' feature composition. The experience with English that is the source of these constraints must be supplemented if it is to account for English speakers' graded judgments of non-occurring clusters. This supplement could well be UG's sonority sequencing constraints.

Berent et al.'s (2007, 2009) results were used here to argue that UG's constraints emerge in adults' judgments of phonological strings that they do not encounter in their native language. An alternative account, which explains those judgments as extrapolations from constraints learned from the distributions of sound sequences they do encounter, was shown to fall short. To the extent that these arguments are well supported, UG not only guides infants' learning (Section 3.4.1) but continues to emerge throughout the lifespan whenever a speaker or listener encounters a novel form.

3.5 Summary

In this section, we have presented positive evidence in support of the claim that there are active synchronic restrictions: a sound change's phonetic motivation remains active after it has been phonologized (Section 3.1), sound changes can be optimizing in a way that reflects synchronic constraints (Section 3.2), the constraints attributed to synchronic phonologies are psychologically and neuro-physiologically active (Section 3.3), these constraints are not mere statistical generalizations across the lexicon (Sections 3.3 and 3.4.2), and the constraints of which synchronic phonologies are composed and their initial ranking are plausibly innate (Section 3.4.1). This evidence points to a rich, ongoing interchange between the phonetics and phonology and to a contentful synchronic phonology.

---

[44] These clusters were [*pw*, zr, mr, *tl*, dn, *km*, fn, *ml*, nl, *dg*, *pk*, lm, ln, rl, *lt*, rn, *rd*, *rg*]; the italicized ones were included in our modeling, too.

4 The roles of diachronic and synchronic explanation

We have argued that there are non-trivial phonological constraints that restrict the form of synchronic grammars; we therefore reject extreme versions of diachronic explanation that seek to ascribe every observed markedness asymmetry to transmissibility differences. However, we accept that diachronic change influences languages; in this section, we discuss the roles of both diachronic and synchronic change.

In Figure 3, C ∪ P contains all the grammars that an unfettered phonological component could generate; by 'unfettered' we mean that the phonological component could generate virtually any grammar, restricted only by the inherent limits on its formatives' and relations' formal properties. In contrast, C contains grammars that a restricted phonological component can generate—i.e., one that is limited by the kind of grammatical restrictions we have argued for in the previous sections. P contains all grammars that are transmissible, including those that are transmissible if there were no (significant) synchronic restrictions on the phonological component (N.B. Evolutionary Phonology considers P the attested set of languages). P contains subsets—the darker portions contain grammars that are more easily learnable (in an appropriate sense), and so are more likely to survive transmission intact than those grammars in lighter areas.

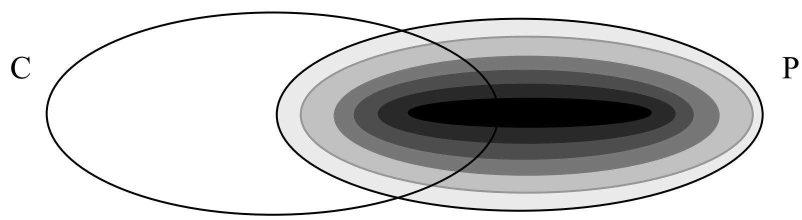

Figure 3. Competence, Performance, and attested grammars

In this article we have argued that P ∩ C' (i.e., what is in P but not in C) has content: there are grammars that are favored in diachronic actuation and transmission but cannot be generated by the phonological component. These include grammars with epenthetic [k] and neutralization to [k]. In contrast, an extreme diachronic explanation approach would hold that P ∩ C' is empty: the phonological component can generate any output, and the grammars that are diachronically favored are a proper subset of C.

It is uncontroversial that C ∩ P' is non-empty: there are grammars that can be generated by the phonological component but are unlearnable or cannot survive inter-generational transmission. For example, all extant phonological theories are capable of producing a grammar with just two segments and a CV syllable structure, but no such language exists. Such a language would be a resounding failure for functional reasons; given the limited distinct combinations, words would have to be very long—human memory limitations in parsing would result in an extremely limited vocabulary. Functional pressures rule such a language out, so there is no need for a theory of phonological competence to do so.

A more striking example involves onsetless syllables. Blevins (2008) observes that probably no language has CV syllables alone. Of course, Optimality Theory with a constraint ONSET ('Incur a violation for every onsetless syllable') predicts a language in which every syllable begins with a consonant. Such a grammar would have the ranking ONSET » MAX or DEP, and ONSET would not be blocked by any constraint that favored onsetless syllables (e.g., ALIGN-L(Root,σ); McCarthy and Prince 1993). Other constraints would rule out other syllable types (e.g., *COMPLEX, NOCODA, NOLONGVOWEL). Is this a problem for OT with constraints such as ONSET, NOCODA, and *COMPLEX, or more generally any formalist theory that can generate a grammar that bans onsetless syllables everywhere? No: while the phonological component defines a set of possible grammars, it does not guarantee that every definable grammar exists. In fact it has nothing to say about the popularity of individual grammars; the frequency of a grammar depends on Performance factors, so the lack of a language without onsetless syllables perhaps indicates the functional utility of using onsetless syllables as boundary markers for words (essentially adopting proposals by Trubetzkoy 1939:273ff).

We also think that it is uncontroversial that some grammars are easier to transmit than others—that is, they are more likely to survive the acquisition process intact. For example, the vowel inventory [i e a o u] is extremely stable (e.g., it has survived unchanged in Polynesia for over a thousand years). No doubt the stability is due to the inventory's perceptual and articulatory desirability (Liljencrants and Lindblom 1972; Lindblom 1986; Schwartz et al. 1997a,b; but see Kingston 2007 for reservations about these explanations). Grammars with a vowel system [i a o] are not robust in terms of these criteria, and so they are significantly fewer in number.

In terms of the frequency of vowel systems, what then are the roles of synchronic and diachronic explanation? Even if the learner comes to the task of learning a language with substantial innate knowledge of what a possible language can be, the variety of human languages at all levels of description, including the phonetic, is such that experience with speaking and listening will profoundly shape what is learned. This is the positive lesson from Berent et al.'s (2007, 2009) results (see Section 3.4.2): mature language learners' judgments are a joint product of universal predispositions and language-specific experience. Learning is, moreover, an occasion for mistransmission or misinterpretation of the current state of the language, which can lead to language change if it persists and spreads through the speech community. Vowel systems such as [i a o] lend themselves to misperception and misinterpretation, and so are liable to be altered in language change. In contrast, [i e a o u] vowel systems have communicative stability. Consequently, it is undeniable that transmissibility factors offer an account of relative typological frequency.

In contrast, it is not necessarily the case that the phonological component has any effect on relative typological frequency (though see Coetzee 2002; Moreton 2008, 2009). The sole necessary requirement on the phonological component is that it is capable of generating grammars with both [i e a o u] and [i a o] systems; the popularity of individual systems can be ascribed to Performance factors.

As another example, it is not necessarily a Competence theory's job to account for the fact that [g] is often absent while [b] and [d] are not. After all, every imaginable voiced stop inventory exists in terms of major place of articulation, as shown in (14).

(14) Voiced stop inventories

g	b	d	Nhanda (Blevins 2001), Catalan (Wheeler 2005)
g	b		Tigak (Beaumont 1979)
g		d	Wapishana (Tracy 1972), Ayutla Mixtec (Pankratz and Pike 1967)
	b	d	Sioux Valley (Santee) (Shaw 1980:17), Xavanté Macro-Je (Rodrigues 1999a)
g			Makurap (Rodrigues 1999b:112ff)
	b		Koasati (Kimball 1991)
		d	Diyari (Austin 1981), Nambiquara (Kroeker 1972)

It is the job of a theory of phonological Competence to produce grammars that generate all of the inventories in (14). There is nothing more it needs to explain. It is true that languages with a [g] and missing some other stop (like Tigak, Wapishana, Ayutla Mixtec, and Makurap) are rare, but there is no reason to ascribe their rarity to Competence mechanisms. In contrast, factors such as the difficulty in producing voicing in velar stops (i.e., [g]) (e.g., Ferguson 1975; Ohala 1983) mean that learners would be more likely to eliminate it from their inventories or reinterpret it as voiceless [k] or fricative [ɣ]. Therefore, an explanation of why [g] is rare relative to other voiced stops is not an explanation about markedness as a Competence concept—it is an account of Performance.

Kingston (2007) shows that [g] and other phonetically disfavoured stops are actually absent far less often than would be expected by chance. He proposes that these gaps are infrequent because languages are *efficient* in their use of distinctive features—if a language has a voiced stop at one major place of articulation, it will have them at the others as well.[45] Is efficiency part of Competence? We would argue that it is not. From the point of view of Performance, it is certainly preferable to use each means of contrast fully rather than partially, as otherwise either another contrast would have to be learned or words would have to be longer. But this preference for efficient use of distinctive features need not be incorporated into Competence.

In short, diachronic approaches and synchronic explanation theories are not incompatible in some areas. Formalist theories of phonology have no inherent interest in accounting for true universal tendencies. In the formalist conception, facts about typological frequency are likely to be accounted for by Performance mechanisms. In short, if both α and β are attested properties of language, the phonological component must be able to generate grammars with both α and β. If β is more common than α, the reason for their relative frequency can be sought in diachronic change and Performance. Of course, 'diachronic change' here means more than ease of transmissibility; war, disease, and migration could also influence the typological frequency of a particular sound pattern.

Of course, a methodological challenge arises with unattested grammars: is a grammar unattested because it cannot be generated, or because it has a very low chance of being

[45] Kingston's efficiency is equivalent to Clements's (2003) *economy* and likewise Ohala's (1979) maximal use of available features.

transmitted intact? For example, is lack of epenthetic [ʘ] indicative of a Competence restriction, or is it just due to the fact that [ʘ] is extremely rare in languages anyway?

We have argued above that one way to tell the two options apart is if there is a situation in which the language property is expected but fails to emerge. This was the case with epenthetic [k]. Its absence is particularly striking given the comparative ease with which it is produced by sound change from earlier *t. That is, a synchronic grammar could have [k] as an epenthetic segment or as the result of neutralization if sound change was not kept in check by a constraint ranking that harmonically bounds such an outcome. The evidence presented in Section 3 illustrates some of the other means of obtaining relevant evidence. In short, if it can be shown that a particular grammar is expected for Performance reasons, yet does not arise, it is highly likely there is an active phonological mechanism that prevents it.

If one's Performance theory predicts that an unattested sound pattern should be rare or non-existent, what can be determined about the role of synchronic phonological mechanisms? Blevins (2004) proposes a methodological principle:

> "A central premise of Evolutionary Phonology, then, is that principled diachronic explanations for sound patterns replace, rather than complement, synchronic explanations, unless independent evidence demonstrates, beyond reasonable doubt, that a separate synchronic account is warranted." (Blevins 2004:5)

We take this proposal to mean that if a grammar is unattested and there is a diachronic explanation for its lack of attestation, then there are no phonological mechanisms that prevent the grammar from being generated.

We question this methodological principle. If a sound pattern has a diachronic explanation, it does not necessarily follow that there must be no synchronic account. For example, while there may be good Performance pressures to avoid epenthetic [p], the grammar might also be incapable of generating epenthetic [p]. These points are independent from a synchronic point of view: establishing a Performance reason for a restriction in no way precludes a grammatical account.

At least some of the time it should be expected that synchronic and diachronic explanations agree: some synchronic constraints will ban the same grammars that some diachronic changes avoid. Why should some synchronic constraints have the same prejudices as diachronic change? As Performance pressures influenced the evolutionary development of innate phonological constraints, phonetically 'natural' constraints would be favored in survival, while 'unnatural' constraints may have been eliminated (see Chomsky and Lasnik 1977:487–9 for an example, and Pinker and Bloom 1990 for extended discussion of how universal characteristics of natural language could have evolved). Therefore, a number of innate constraints could militate against structures that have little perceptual, articulatory, or processing value.

We are certainly not denying the autonomy of synchronic constraints and diachronic change; after all, this article has been at pains to identify and delineate the distinct roles of synchronic and diachronic explanation. We are merely observing that *some* convergence in their effect (and therefore some overlap in roles) should be expected as the Performance pressures that have selected for particular constraints or constraint systems may also be those that influence synchronic perception, articulation, and processing. Therefore, the fact that a diachronic account predicts the presence or lack of a sound pattern does not rule out the possibility that there are also synchronic restrictions that have the same effect. Showing that there is a diachronic account for the lack of a sound pattern does not imply that there is no synchronic account.

However, the extent of the convergence is difficult to discern, as it is hard to tell whether a particular grammar cannot be generated because a Performance theory predicts that it will have low probability of survival. The only immediate options are to examine such issues using experiments of the variety described in Section 3.

In summary, we see diachronic explanation as crucial in accounting for some sound patterns. Diachronic approaches explain why some grammars that the phonological component can generate are unattested. They also account for the relative typological frequency of grammars (and no doubt partly for the popularity of particular sound patterns within languages). Moreover, if the Performance pressures that influence language actuation and transmission also influenced the survival of innate constraints, some convergence should be expected; there should be innate constraints that ban some of the grammars that are also avoided in diachronic change.

5 Conclusions

The aim of this paper has been to identify sources of evidence for synchronic cognitive phonological restrictions. We argued that there is a role for diachronic explanation, but it is confined to accounting for the non-attestation of untransmissible grammars and partially for the relative frequency of attested grammars. Performance factors certainly influence why [g] is typologically more rare than [d] and [b]; the phonological component may have no influence on typological frequency. Clear evidence for synchronic constraints is found when sound patterns that are well attested in diachronic change never result in synchronic alternations, such as [k] epenthesis and neutralization to [k]. We argued that such systems are avoided, indicating that the phonological component is incapable of generating grammars with [k] epenthesis and neutralization to [k].

We further argued that there is a necessary link between sound patterns and their phonetic motivations in Athabaskan tonogenesis, that sound change is synchronically optimizing, and that there is a variety of experimental evidence for psychologically active synchronic restrictions on sound patterns that cannot be induced from exposure to surface forms. We also identified evidence that infants have an innate initial ranking of markedness over faithfulness constraints and that universal constraints on sonority sequencing emerge in adults' judgments of the relative wellformedness of particular onset clusters.

In short, a strong version of diachronic explanation cannot hold—that there are no substantial cognitive restrictions on the phonological component. Active synchronic constraints are necessary.

Acknowledgements

We are particularly indebted to Elliott Moreton, six anonymous reviewers, and the associate editor for their many comments. We also wish to thank the following people: John McCarthy, Joe Pater, Kevin Mullin, Alexei Nazarov, Brian Smith, Presley Pizzo, Claire Moore-Cantwell, Robert Staubs, Haruka Fukazawa, Mafuyu Kitahara, Nina Topintzi, Clint Hartzell, Lisa Sanders, Minta Elsman, Wendell Kimper, Yelena Fainleib, Michael Key, Karen Jesney, Ellen Broselow, 'Oiwi Parker-Jones, and Stephen Marlett. We remain responsible for any errors.

REFERENCES

Adler, Allison. 2006. Faithfulness and perception in loanword adaptation: A case study from Hawaiian. *Lingua* 116: 1024-1045.

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

Alderete, John, Jill Beckman, Laura Benua, Amalia Gnanadesikan, John McCarthy, and Suzanne Urbanczyk. 1999. Reduplication with fixed segmentism. *Linguistic Inquiry* 30: 327–364.

Aronoff, Mark, Azhar Arsyad, Hasan Basri and Ellen Broselow. 1987. Tier configuration in Makassarese reduplication. In *CLS 23: Parasession on autosegmental and metrical phonology*, eds. Anna Bosch, Eric Schiller and Barbara Need. Vol. 2: 1–15. Chicago: Chicago Linguistic Society.

Austin, Peter. 1981. A grammar of Diyari, South Australia. Cambridge: Cambridge University Press.

Baković, Eric. 2001. Nasal place neutralization in Spanish. *University of Pennsylvania Working Papers in Linguistics* 7: 1–12.

Bach, Emmon, and Robert T. Harms. 1972. How do languages get crazy rules? In *Linguistic Change and Generative Theory: Essays from the UCLA conference on historical linguistics in the perspective of transformational theory*, eds. Robert Stockwell and Ronald Macaulay, 1–21. Bloomington: Indiana University Press.

Barnes, Jonathan Allen. 2002. Positional neutralization: A phonologization approach to typological patterns. PhD dissertation, University of California, Berkeley, CA.

Beaumont, Clive H. 1979. *The Tigak language of New Ireland. Pacific Linguistics Series B*, No. 58. Canberra: Australian National University.

Beckman, Jill N. 1997. Positional faithfulness, position neutralisation, and Shona vowel harmony. *Phonology* 14.1: 1–46.

Beckman, Jill N. 1998. Positional faithfulness. PhD dissertation, University of Massachusetts, Amherst, MA.

Berent, Iris, Tracy Lennery, Paul Smolensky, and Vered Vaknin-Nusbaum. 2009. Listeners' knowledge of phonological universals: Evidence from nasal clusters. *Phonology* 26: 75–108.

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

Blevins, Juliette. 2001. *Nhanda: An aboriginal language of Western Australia. Oceanic Linguistics Special Publication* 30. Honolulu: University of Hawai'i.

Blevins, Juliette. 2004. *Evolutionary Phonology: The emergence of sound patterns*. Cambridge: Cambridge University Press.

Blevins, Juliette. 2006a. A theoretical synopsis of Evolutionary Phonology. *Theoretical Linguistics* 32.2: 117–166.

Blevins, Juliette. 2006b. Reply to commentaries. *Theoretical Linguistics* 32.2: 245–256.

Blevins, Juliette. 2008. Consonant epenthesis: natural and unnatural histories. In *Language universals and language change*, ed. Jeff Good, 79–106. Oxford: Oxford University Press.

Blumstein, Sheila E., and Kenneth N. Stevens. 1979. Acoustic invariance in speech production: evidence from measurements of the spectral characteristics of stop consonants. *Journal of the Acoustical Society of America* 66: 1001–1017.

Blumstein, Sheila E., and Kenneth N. Stevens. 1980. Perceptual invariance and onset spectra for stop consonants in different vowel environments. *Journal of the Acoustical Society of America* 67: 648–662.

Blust, Robert. 1990. Three recurrent changes in Oceanic languages. In *Pacific Island languages: Essays in honour of G.B. Milner*, ed. J.H.C.S. Davidson, 7–28. London: University of London and School of Oriental and African Studies.

Bradley, Travic. 2006. Contrast and markedness in complex onset phonotactics. *Southwest Journal of Linguistics* 25.1: 29–58.

Breen, Mara, John Kingston, and Lisa Sanders. 2013. Perceptual representations of phonotactically illegal syllables. *Attention, Perception, and Psychophysics* 75:101-120.

Caisse, Michelle. 1982. Cross-linguistic differences in fundamental frequency perturbation induced by voiceless unaspirated stops. MA thesis, University of California, Berkeley, CA.

Capodieci, Frank, Bruce Hayes, and Colin Wilson. 2008. A maximum entropy model of phonotactics and phonotactic learning: Software. Version 1.1. http://www.linguistics.ucla.edu/people/hayes/Phonotactics/#software (Accessed 15 November 2011). UCLA.

Chen, Matthew. 1973. Cross-dialectal comparison: a case study and some theoretical considerations. *Journal of Chinese Linguistics* 1.1:38–63.

Chomsky, Noam, and Howard Lasnik. 1977. Filters and control. *Linguistic Inquiry* 8: 425–504.

Christófaro-Silva, Thaïs. 2003. Sound change in tautosyllabic consonantal clusters in Brazilian Portuguese. In *Proceedings of the 15th International Congress of Phonetic Sciences*, eds. Maria Josep Solé, Daniel Recasens, and Joaquín Romero, 1675–1678. Barcelona: Universitat Autònoma de Barcelona.

Clark, Ross. 1976. *Aspects of proto-Polynesian syntax*. Auckland: Linguistic Society of New Zealand.

Clements, G.N. 1988. Toward a substantive theory of feature specification. In *Papers from the 18th Annual Meeting of the North Eastern Linguistic Society (NELS 18)*, eds. Juliette Blevins and Juli Carter, 79–93. Amherst, MA: GLSA.

Clements, G.N. 2003. Feature economy in sound systems. *Phonology* 20.3: 287–333.

Coetzee, Andries W. 2002. Between-language frequency effects in phonological theory. Ms., University of Massachusetts, Amherst, MA.

Coetzee, Andries W. 2004. What it means to be a loser: non-optimal candidates in Optimality Theory. PhD dissertation, University of Massachusetts, Amherst, MA.

Coetzee, Andries W. 2005. The OCP in the perception of English. In *Prosodies: Selected papers from the phonetics and phonology in Iberia conference, 2003*, eds. Sonia Frota, Marina Vigario, and Maria Joao Freitas, 223–245. New York: Mouton de Gruyter.

Crowley, Terry. 1983. Uradhi. In *The handbook of Australian languages*, eds. R.M.W. Dixon and Barry Blake, 307–428. Amsterdam: John Benjamins.

Crowther, Court S., and Virginia A. Mann. 1992. Native-language factors affecting use of vocalic cues to final-consonant voicing in English. *Journal of the Acoustical Society of America* 92: 711–722.

Crowther, Court S., and Virginia A. Mann. 1994. Use of vocalic cues to consonant voicing and native language background: the influence of experimental design. *Perception and Psychophysics* 55: 513–525.

Daland, Robert, Bruce Hayes, James White, Marc Garellek, Andrea Davis, and Ingrid Norrmann. 2011. Explaining sonority projection effects. *Phonology* 28: 197–234.

Davidson, Lisa, Peter Jusczyk, and Paul Smolensky. 2004. The initial and final states: Theoretical implications and experimental explorations of Richness of the Base. In *Constraints in Phonological Acquisition,* eds. René Kager, Joe Pater, and Wim Zonneveld, 321–368. Cambridge: Cambridge University Press.

Davis, Stuart. 1999. On the representation of initial geminates. *Phonology* 16: 93–104.

de Lacy, Paul. 2002. The formal expression of markedness. PhD dissertation, University of Massachusetts, Amherst, MA.

de Lacy, Paul. 2003. Maximal words and the Māori passive. In *Optimality Theory in phonology: A reader,* ed. John J. McCarthy, 495–512. Oxford: Blackwell.

de Lacy, Paul. 2004. Markedness conflation in Optimality Theory. *Phonology* 21.2: 1–55.

de Lacy, Paul. 2006a. *Markedness: Reduction and preservation in phonology.* Cambridge: Cambridge University Press.

de Lacy, Paul. 2006b. Transmissibility and the role of the phonological component. *Theoretical Linguistics* 32.2: 185–196.

Dehaene-Lambertz, Ghislaine, Emmanuel Dupoux, and Ariel Gout. 2000. Electrophysiological correlates of phonological processing: A cross-linguistic study. *Journal of Cognitive Neuroscience* 12: 635–647.

Diehl, Randy L., and John Kingston. 1991. Phonetic covariation as auditory enhancement: The case of the [+voice]/[-voice] distinction. In *Current phonetic research paradigms: Implications for speech motor control, PERILUS, XIV*, eds. Olle Engstrand and Catharina Kylander, 139–143. Stockholm: University of Stockholm.

Docherty, Gerard R. 1992. *The timing of voicing in British English obstruents.* Dordrecht: Foris Publications.

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.

Dupoux, Emmanuel, Christophe Pallier, Kazuhiko Kakehi, and Jacques Mehler. 2001. New evidence for prelexical phonological processing in word recognition. *Language and Cognitive Processes* 16: 491–505.

Ebert, Karen. 1996. *Koḍava*. *Languages of the World/Materials*, 104. Nürnburg: LINCOM Europa.
Elbert, Samuel H., and Mary K. Pukui. 2001. Hawaiian grammar. Honolulu: University of Hawai'i Press.
Elfenbein, Josef. 1997. Brahui phonology. In *Phonologies of Asia and Africa*, ed. Alan S. Kaye, 797–811. Winona Lake, Indiana: Eisenbrauns.
Ferguson, Charles A. 1975. Sound patterns in language acquisition. In *Developmental psycholinguistics: Theory and applications*, ed. Daniel P. Dato, 1–16. Washington, DC: Georgetown University Press.
Fischer, Rebecca M., and Ralph N. Ohde. 1990. Spectral and duration properties of front vowels as cues to final stop-consonant voicing. *Journal of the Acoustical Society of America* 88: 1250–1259.
Fruchter, David, and Harvey M. Sussman. 1997. The perceptual relevance of locus equations. *Journal of the Acoustical Society of America* 102: 2997–3008.
Fujimura, Osamu, and Joan E. Miller. 1979. Mandible height and syllable-final tenseness. *Phonetica* 36: 263–272.
Gafos, Adamantios. 1998. A-templatic reduplication. *Linguistic Inquiry* 29: 515–527.
Gick, Bryan. 1999. A gesture-based account of intrusive consonants in English. *Phonology* 16:1: 29–54.
Gick, Bryan. 2002a. The American intrusive l. *American Speech* 77: 167–183.
Gick, Bryan. 2002b. An X-ray investigation of pharyngeal constriction in American English schwa. *Phonetica* 59: 38–48.
Ghini, Mirco. 2001. *Asymmetries in the phonology of Miogliola*. New York: Mouton de Gruyter.
Gildea, Spike. 1995. Comparative Cariban syllable reduction. *International Journal of American Linguistics* 62.1: 62–102.
Goldsmith, John. 1990. *Autosegmental and metrical phonology*. Oxford and Cambridge, MA: Blackwell.
Goodenough, Ward H., and Hiroshi Sugita. 1980. *Trukese-English dictionary*. Philadelphia: American Philosophical Society.
Goodenough, Ward H., and Hiroshi Sugita. 1990. *Trukese-English dictionary/Pwpwuken Tettenin Fóós: Chuuk-Ingenes. Supplementary volume: English-Trukese and index of Trukese word roots.* Philadelphia: American Philosophical Society.
Guitart, Jorge M. 1976. *Markedness and a Cuban dialect of Spanish*. Washington, DC: Georgetown University Press.
Haas, Mary R. 1968. Notes on a Chipewyan dialect. *International Journal of American Linguistics* 34: 165–175.
Hale, Kenneth. 1973. Deep-surface canonical disparities in relation to analysis and change: An Australian example. *Current Trends in Linguistics* 11: 401–458.
Hale, Kenneth. 1976. Phonological developments in a Northern Paman language: Uradhi. In *Languages of Cape York*, ed. Peter Sutton, 41–49. Canberra: Australian Institute of Aboriginal Studies.
Hallé, Pierre A., Catherine T. Best, and Asaf Bachrach. 2003. Perception of /dl/ and /tl/ clusters: A cross-linguistic perceptual study with French and Israeli listeners. In *Proceedings of the 15th International Congress of Phonetic Sciences*, eds. Maria

Josep Solé, Daniel Recasens, and Joaquín Romero, 2893–2896. Barcelona: Universitat Autònoma de Barcelona.

Hallé, Pierre A., and Catherine T. Best. 2007. Dental-to-velar perceptual assimilation: A cross-linguistic study of the perception of dental stop+/l/ clusters. *Journal of the Acoustical Society of America* 121: 2899–2914.

Hallé, Pierre A., Juan Segui, Uli Frauenfelder, and Christine Meunier. 1998. Processing of illegal consonant clusters: A case of perceptual assimilation? *Journal of Experimental Psychology: Human Perception and Performance* 24: 592–608.

Haraguchi, Shosuke. 1984. Some tonal and segmental effects of vowel height in Japanese. In *Language sound structure: Studies in phonology presented to Morris Halle by his teacher and students*, eds. Mark Aronoff and Richard T. Oehrle, 145–156. Cambridge: MIT Press.

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

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

Hillenbrand, James, Dennis R. Ingrisano, Bruce L. Smith, and James E. Flege. 1984. Perception of the voiced-voiceless contrast in syllable-final stops. *Journal of the Acoustical Society of America* 76: 18–26.

Hombert, Jean-Marie, John J. Ohala, and William G. Ewan. 1979. Phonetic explanations for the development of tones. *Language* 55: 37–58.

Howe, Darin. 2004. Vocalic dorsality in revised articulator theory. http://www.ucalgary.ca/dflynn/files/dflynn/Howe2004.pdf. Accessed 14 November 2011.

Hualde, José Ignacio. 1991. *Basque phonology*. London: Routledge.

Hualde, José Ignacio, and Inaki Gaminde. 1998. Vowel interaction in Basque: A nearly exhaustive catalogue. *Studies in the Linguistic Sciences* 28.1: 41–77.

Hume, Elizabeth. 2003. Language specific markedness: The case of place of articulation. *Studies in Phonetics, Phonology, and Morphology* 9.2: 295–310.

Hume, Elizabeth, and Georgios Tserdanelis. 2002. Labial unmarkedness in Sri Lankan Portuguese Creole. *Phonology* 19.3: 441–458.

Jakobson, Roman. 1941. *Kindersprache, Aphasie, und allgemeine Lautgesetze*. Uppsala: Almqvist and Wiksell [tr. Child Language, Aphasia and Phonological Universals, Mouton].

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

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: 31–73.

Jusczyk, Peter W., Paul Smolensky, Karen Arnold, and Elliott Moreton. 2003. Acquisition of nasal place assimilation by 4.5-month-old infants. Ms., Johns Hopkins University, Baltimore, MD.

Kenstowicz, Michael. 1997. Quality-sensitive stress. *Rivista di Linguistica* 9.1: 157–188. Also Rutgers Optimality Archive #33.

Kimball, Geoffrey D. 1991. *Koasati grammar.* Lincoln: University of Nebraska Press.

Kim-Renaud, Young-Key. 1986. *Studies in Korean linguistics*. Seoul: Hanshin.

Kingston, John. 1985. The phonetics and phonology of the timing of oral and glottal events. PhD dissertation, University of California, Berkeley, CA.

Kingston, John. 2002. Keeping and losing contrasts. In *Proceedings of the 28th Annual Meeting of the Berkeley Linguistics Society*, eds. Julie Larson and Mary Paster, 155–176. Berkeley: Berkeley Linguistics Society.

Kingston, John. 2005. The phonetics of Athabaskan tonogenesis. In *Athabaskan prosody*, eds. Sharon Hargus and Keren Rice, 137–184. Amsterdam: John Benjamins.

Kingston, John. 2007. The phonetics-phonology interface. In *The Cambridge handbook of phonology*, ed. Paul de Lacy, 401–434. Cambridge: Cambridge University Press.

Kingston, John, and Randy L. Diehl. 1994. Phonetic knowledge. *Language* 70: 419–454.

Kingston, John and Randy L. Diehl. 1995. Intermediate properties in the perception of distinctive feature values. In *Phonology and Phonetics: Papers in Laboratory Phonology IV,* eds. Bruce Connell and Amalia Arvaniti, 7-27. Cambridge, UK: Cambridge University Press.

Kingston, John, Randy L. Diehl, Cecilia J. Kirk, and Wendy A. Castleman. 2008. On the internal perceptual structure of distinctive features: The [voice] contrast, *Journal of Phonetics* 36: 28-54.

Kiparsky, Paul. 2006. The amphichronic program vs. Evolutionary Phonology. *Theoretical Linguistics* 32.2: 217–236.

Kiparsky, Paul. 2008. Universals constrain change; change results in typological generalizations. In *Linguistic universals and language change*, ed. Jeff Good, 23–52. Oxford: Oxford University Press.

Krauss, Michael. 2005. Athabaskan tone. In *Athabaskan prosody*, eds. Sharon Hargus and Keren Rice, 55–136. Amsterdam: John Benjamins.

Kroeker, Barbara J. 1972. Morphophonemics of Nambiquara. *Anthropological Linguistics* 14: 19–22.

Leer, Jeff. (1979). Proto-Athabaskan Verb Stem Variation, Part One: Phonology. Alaska Native Language Center Research Papers 1. Fairbanks, University of Alaska.

Leer, Jeffrey. 1999. Tonogenesis in Athabaskan. In *Cross-linguistic studies of tonal phenomena: Tonogenesis, typology, and related topics*, ed. Shigeki Kaji, 37–66. Tokyo: Institute of the Study of the Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.

Leer, Jeffrey. 2001. Shift of tone markedness in Northern Tlingit and Southern Athabaskan. In *Cross-linguistic studies of tonal phenomena: Tonogenesis, Japanese accentology, and other topics*, ed. Shigeki Kaji, 61–86. Tokyo: Institute of the Study of the Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.

Li, Paul Jen-Kuei. 1985. A secret language in Taiwanese. *Journal of Chinese Linguistics* 13.1: 91–121.

Liljencrants, Johan, and Björn Lindblom. 1972. Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language* 48: 839–862.

Lindblom, Björn. 1986. Phonetic universals in vowel systems. In *Experimental phonology*, eds. John J. Ohala and Jeri J. Jaeger, 13–44. New York: Academic Press.

Lombardi, Linda. 2002. Coronal epenthesis and markedness. *Phonology* 19.2: 219–251.

Lupas, Liana. 1972. *Phonologie du Grec Attique*. The Hague: Mouton.

Lynch, John, Malcolm Ross, and Terry Crowley, eds. 2002. *The Oceanic languages*. Richmond: Curzon Press.

MacKay, Carolyn J. 1994. A sketch of Misantla Totonac phonology. *International Journal of American Linguistics* 60.4: 369–419.

Marlett, Stephen Alan. 1981. The structure of Seri. PhD dissertation, University of San Diego, San Diego, CA.

Marlett, Stephen Alan. 2010. Round vowel and dorsal consonant epenthesis in Seri. In *Work Papers of the Summer Institute of Linguistics, University of North Dakota Session*, ed. Joan Baart, http://arts-sciences.und.edu/summer-institute-of-linguistics/work-papers/2010.cfm. Accessed 3 February 2013. Grand Forks: University of North Dakota.

Mascaró, Joan. 1996. External allomorphy as emergence of the unmarked. In *Current trends in phonology: Models and methods*, eds. Jacques Durand and Bernard Laks, 473–483. Salford: European Studies Research Institute, University of Salford.

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

McCarthy, John J. 1994. The phonetics and phonology of Semitic pharyngeals. In *Phonological structure and phonetic form: Papers in laboratory phonology III*, ed. Patricia Keating, 191–233. Cambridge: Cambridge University Press.

McCarthy, John J., and Alan Prince. 1993. *Prosodic morphology I: Constraint interaction and satisfaction*. Rutgers Technical Report TR-3. New Brunswick: Rutgers University Center for Cognitive Science.

McCarthy, John J., and Alan Prince. 1999. Faithfulness and identity in prosodic morphology. In *The prosody-morphology interface*, eds. René Kager, Harry van der Hulst, and Wim Zonneveld, 218–309. Cambridge: Cambridge University Press.

Mills, Roger F. 1975. Proto South Sulawesi and Proto Austronesian phonology. PhD dissertation, University of Michigan, Ann Arbor, MI.

Moreton, Elliot. 1999. Evidence for phonological grammar in speech perception. In *Proceedings of the 14th International Congress of Phonetic Sciences*, eds. John J. Ohala, Yoko Hasegawa, Manjari Ohala, Daniel Granville, and Ashlee C. Bailey, 2215–2217. Berkeley: University of California, Berkeley.

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

Moreton, Elliott. 2004. Realization of the English postvocalic [voice] contrast in F1 and F2. *Journal of Phonetics* 32: 1–33.

Moreton, Elliott. 2008. Analytic bias and phonological typology. *Phonology* 25.1: 83–127.

Moreton, Elliott. 2009. Underphonologization and modularity bias. In *Phonological argumentation: Essays on evidence and motivation*, ed. Steve Parker, 79–101. Sheffield: Equinox Publishing.

Moreton, Elliott, and Shigeaki Amano. 1999. Phonotactics in the perception of Japanese vowel length: Evidence for long-distance dependencies. In *Proceedings of the 6th European Conference on Speech Communication and Technology (EUROSPEECH'99)*, 2679–2682 Budapest: Hungary. http://www.unc.edu/~moreton/Papers/Eurospeech1999.pdf. Accessed 1 March 2006.

Moreton, Elliott, and Erik R. Thomas. 2007. Origins of Canadian Raising in voiceless-coda effects: A case study of phonologization. In *Laboratory phonology 9*, eds. Jennifer Cole and Jose I. Hualde, 37–63. Berlin: Mouton de Gruyter.

Ohala, John J. 1974. Experimental historical phonology. In *Historical linguistics II. Theory and description in phonology*, eds. John M. Anderson and Charles. Jones, 353–389. Amsterdam: North Holland.

Ohala, John J. 1979. Moderator's introduction to symposium on phonetic universals in phonological systems and their explanation. In *Proceedings of the International Congress of Phonetic Sciences*, Vol. 3, 181–185. Copenhagen: Institute of Phonetics.

Ohala, John J. 1981. The listener as a source of sound change. In *Papers from the Parasession on Language and Behavior*, eds. Carrie S. Masek, Roberta A. Hendrick, and Mary F. Miller, 178–203. Chicago: Chicago Linguistic Society.

Ohala, John J. 1983. The origin of sound patterns in vocal tract constraints. In *The production of speech*, ed. Peter F. MacNeilage, 189–216. New York: Springer.

Ohala, John J. 1987. Explanation in phonology: Opinions and examples. In *Phonologica 1984*, eds. Wolfgang U. Dressler, Hans C. Luschützky, Oskar E. Pfeiffer, and John R. Rennison, 215–225. Cambridge: Cambridge University Press.

Ohala, John J. 1993. The phonetics of sound change. In *Historical linguistics: Problems and perspectives*, ed. Charles Jones, 237–278. London: Longman.

Ohala, John J. 1995. Phonetic explanations for sound patterns: implications for grammars of competence. In *Proceedings of the 13th International Congress of the Phonetic Sciences*, Vol. 2, eds. Kjell Elenius and Peter Branderud, 52–59. Stockholm: KTH and Stockholm University.

Ohala, John J. 2005. Phonetic explanations for sound patterns. Implications for grammars of competence. In *A figure of speech. A festschrift for John Laver*, eds. William J. Hardcastle and Janet M. Beck, 23–38. London: Erlbaum.

Pankratz, Leo and Eunice V. Pike. 1967. Phonology and morphotonemics of Ayutla Mixtec. *International Journal of American Linguistics* 33: 287–299.

Paradis, Carole and Jean-François Prunet, eds. 1991. *Phonetics and phonology: The special status of coronals: Internal and external evidence*. San Diego, Academic Press.

Paradis, Carole and Jean-François Prunet. 1993. A note on velar nasals: The case of Uradhi. *The Canadian Journal of Linguistics* 38.4: 425–439.

Paradis, Carole and Jean-François Prunet. 1994. A reanalysis of velar transparency cases. *The Linguistic Review* 11: 101–140.

Parker, Frank. 1974. The coarticulation of vowels and stop consonants. *Journal of Phonetics* 2: 211–221.

Parker, Stephen. 1994. Laryngeal codas in Chamicuro. *International Journal of American Linguistics* 60.3: 261–271.

Pawley, Andrew. 1966 Polynesian languages: A subgrouping based on shared innovations in morphology. *Journal of the Polynesian Society* 75: 39–64.

Payne, David L. 1981. *The phonology and morphology of Axininca Campa*. Summer Institute of Linguistics Publications in Linguistics 66. Dallas: SIL.

Payne, David L. 1990. Accent in Aguaruna. In *Amazonian linguistics: Studies in Lowland South American languages*, ed. Doris L. Payne, 161–184. Austin: University of Texas Press.

Peperkamp, Sharon. 2007. Do we have innate knowledge about phonological markedness? Comments on Berent, Steriade, Lennertz, and Vaknin. *Cognition* 104: 631–637.

Piggott, Glyne. 1991. Apocope and the licensing of empty-headed syllables. *The Linguistic Review* 8: 287–318.

Piggott, Glyne. 1993. Satisfying the minimal word. *McGill Working Papers in Linguistics* 8.2: 194–233.

Pinker, Steven, and Paul Bloom. 1990. Natural language and natural selection. *Behavioral and Brain Sciences* 13: 707–784.

Poppe, Nicholas N. 1960. *Buriat grammar. Uralic and Altaic Series Vol. 2*. Bloomington: Indiana University Press.

Prentice, David John 1971. *The Murut languages of Sabah. Pacific Linguistics Series C*, No.18. Canberra: Australian National University.

Prince, Alan and Paul Smolensky. 2004. *Optimality Theory: Constraint interaction in generative grammar*. Oxford: Blackwells.

Prince, Alan and Bruce Tesar. 2004. Learning phonotactic distributions. In *Constraints in phonological acquisition*, eds. René Kager, Joe Pater, and Wim Zonneveld, 245-291. Cambridge: Cambridge University Press.

Pukui, Mary K., and Samuel Elbert. 1986. *Hawaiian dictionary. Revised and enlarged edition.* Honolulu: University of Hawai'i Press.

Rice, Keren. 1996. Default variability: the coronal-velar relationship. *Natural Language and Linguistic Theory* 14.3: 493–543.

Rice, Keren. 2004. Sonorant relationships: The case of liquids. Handout from a talk presented at the Canadian Linguistics Association.

Rodrigues, Aryon D. 1999a. Macro-Jê. In *The Amazonian Languages*, eds. R.M.W. Dixon and Alexandra Y. Aikhenvald, 165–206. Cambridge: Cambridge University Press.

Rodrigues, Aryon D. 1999b. Tupí. *The Amazonian Languages*, eds. R.M.W. Dixon and Alexandra Y. Aikhenvald, 107–122. Cambridge: Cambridge University Press.

Rose, Sharon, and Rachel Walker. 2004. A typology of consonant agreement as correspondence. *Language* 80: 475–531

Rutgers, Roland. 1998. *Yamphu: Grammar, texts, and lexicon*. Leiden: Research School CNWS.

Samek-Lodovici, Vieri, and Alan Prince. 1999. Optima. Rutgers Optimality Archive #363.

Schwartz, Jean-Luc, Louis-Jean Boë, Nathalie Vallée, and Christian Abry. 1997a. Major trends in vowel system inventories. *Journal of Phonetics* 25: 233–253.

Schwartz, Jean-Luc, Louis-Jean Boë, Nathalie Vallée and Christian Abry. 1997b. The dispersion-focalization theory of vowel systems. *Journal of Phonetics* 25: 255–286.

Shaw, Patricia A. 1980. *Theoretical issues in Dakota phonology and morphology.* New York and London: Garland Publishing Inc.

Smolensky, Paul. 1996. *The initial state and 'Richness of the Base' in Optimality Theory.* Technical Report JHU-CogSci-96-4, Cognitive Science Department, Johns Hopkins University. Rutgers Optimality Archive #154.

Spring, Cari. 1990. Implications of Axininca Campa for prosodic morphology and reduplication. PhD dissertation, University of Arizona, Tucson, AZ.

Steriade, Donca. 2001. The phonology of perceptibility effects: The P-map and its consequences for constraint organization. Ms., MIT, Cambridge, MA. http://www.linguistics.ucla.edu/people/steriade/papers/P-map_for_phonology.doc. Accessed 9 Sept 2008.

Stevens, Kenneth N., and Sheila E. Blumstein. 1978. Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America* 64: 1358–1368.

Stroomer, Harry J. 1987. *A comparative study of three Southern Oromo dialects in Kenya*. Hamburg: Helmut Buske Verlag.

Summers, W. Van. 1987. Effects of stress and final-consonant voicing on vowel production: Articulatory and acoustic analysis. *Journal of the Acoustical Society of America* 82: 485–492.

Summers, W. Van. 1988. F1 structure provides information for final-consonant voicing. *Journal of the Acoustical Society of America* 84: 485–492.

Sussman, Harvey M., David Fruchter, Jon Hilbert, and Joseph Sirosh. 1998. Linear correlates in the speech signal: The orderly output constraint. *Behavioral and Brain Sciences* 21: 241–299.

Sussman, Harvey M., Katherine Hoemeke, and Farhan Ahmed. l993. A cross-linguistic investigation of locus equations as a phonetic descriptor for place of articulation. *Journal of the Acoustical Society of America* 94: 1256–1268.

Sussman, Harvey M., Helen A. McCaffrey, and Sandra A. Matthews. l991. An investigation of locus equations as a source of relational invariance for stop place categorization. *Journal of the Acoustical Society of America* 90: 1309–1325.

Thomas, Erik R. 2000. Spectral differences in /ai/ offsets conditioned by voicing of the following consonant. *Journal of Phonetics* 28: 1–25.

Tracy, Frances. 1972. Wapishana phonology. In *Languages of the Guianas*, ed. Joseph E. Grimes, 78–84. Norman: Summer Institute of Linguistics of the University of Oklahoma.

Trigo, Rosario L. 1988. On the phonological derivation and behavior of nasal glides. PhD dissertation, MIT, Cambridge, MA.

Trigo, Rosario L. 1991. On pharynx-larynx distinctions. *Phonology* 8: 113–136.

Trubetzkoy, Nikolai S. 1939. *Grundzüge der Phonologie*. Güttingen: Vandenhoeck and Ruprecht.

Vaden, K. I., H. R. Halpin, and G. S. Hickok. 2009. *Irvine Phonotactic Online Dictionary*, Version 2.0. www.iphod.com.

Vaux, Bert. 2003. Consonant epenthesis and hypercorrection. Ms., Harvard University, Cambridge, MA.

Walsh, Thomas, and Frank Parker. 1981. Vowel termination as a cue to voicing in post-vocalic stops. *Journal of Phonetics* 9: 105–108.

Walsh, Thomas, and Frank Parker. 1983. Vowel length and transition cues to [+/- voice] in post-vocalic stops. *Journal of Phonetics* 11: 407–412.

Walsh, Thomas, Frank Parker, and C.J. Miller. 1987. The contribution of F1 decline to the perception of [± voice]. *Journal of Phonetics* 15: 101–103.

Weber, David John. 1989. *A grammar of Huallaga (Huánaco) Quechua*. Berkeley: University of California Press.

Wheeler, Max. 2005. *The phonology of Catalan.* Oxford: Oxford University Press.

Wiesemann, Ursula. 1972. *Die phonologis che und grammatische Struktur der Kaingang Sprache.* The Hague: Mouton.

Wolf, Catherine G. 1978. Voicing cues in English final stops. *Journal of Phonetics* 6: 299–309.

Yip, Moira. 1982. Reduplication and CV skeleta in Chinese Secret languages. *Linguistic Inquiry* 13.4: 637–662.

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

Yip, Moira. 1991. Coronals, consonant clusters, and the coda condition. In *Phonetics and phonology: The special status of coronals: Internal and external evidence*, eds. Carole Paradis and Jean-François Prunet, 61–78. San Diego: Academic Press.

Yip, Moira. 1996. Phonological constraints, optimality, and phonetic realization in Cantonese. In *UCI working papers in linguistics*, ed. Sze-Wing Tang, 141–166. University of California, Irvine.

Paul de Lacy
Linguistics Department
Rutgers, The State University of New Jersey
http://www.pauldelacy.net

John Kingston
Linguistics Department
University of Massachusetts Amherst
http://people.umass.edu/jkingstn/