Phonetically-based sound patterns:
Typological tendencies or phonological universals?

Juliette Blevins

Phonetic naturalness plays a major role in defining common types of sound change and the sound patterns resulting from them. However, synchronic typologies are the end result of a complex interaction of factors, only one of which is phonetic naturalness. Pre-existing sound patterns may make listeners more sensitive to certain contrasts than others or prime their categorization, phonetic properties may be interdependent, and probabilities of a given sound pattern may depend on the probabilities of particular initial states. Even when phonological patterns appear exceptionless, there is mounting evidence that these patterns reflect typological tendencies, which instantiate the complex interaction of articulatory variability and invariability, perceptual variability and invariability, and independent language-specific structural priming factors which may enhance or inhibit categorization in the course of language acquisition.

1. Variation and the phonetic basis of recurrent sound change

In a range of influential papers, Ohala (1981, 1983, 1989, 1993) gives examples of recurrent sound changes which are drawn from "a pool of synchronic variation". Ohala (1989) identifies a range of variation types, separating them broadly into those due to phonetic variation on the part of the speaker and those due to transforms on the part of the listener. These variation types are then associated with universal phonetic tendencies which lead to sound change.

Universal phonetic trajectories on the speaker's side can result from aerodynamic constraints, elasto-inertial constraints, constraints on gestural coordination, and, as suggested by Maddieson in his paper, from the still mysterious complex of transforms which relate clear speech to casual speech. Aerodynamic constraints are implicated in common patterns of obstruent devoicing, with devoicing more likely in oral stops with longer closure durations, and those farther back in the mouth. Elasto-inertial constraints include relationships between the amplitude of articulatory movements (e.g. jaw opening), and rate of articulation (fast vs. slow). As rate increases, amplitude decreases, meaning that certain properties of fast speech, including vowel reduction and consonant lenition, will be recurrent. Some universal phonetic trajectories, like utterance-final devoicing, may involve a confluence of these factors: voicing decay has been attributed to anticipation of a non-speech breathing vocal fold configuration, where both aerodynamics and laryngeal inertia are involved (Klatt and Klatt 1990:821; Myers and Hansen 2007).

Though Ohala (1989) refers to transforms on the part of the listener as part of the pool of "synchronic variation" from which sound change is drawn, his use of the term 'variation' is somewhat different from its standard usage in linguistic phonetics. It includes not only multiple distinct phonetic forms that a listener evaluates, but also biases in the perceptual system, or biases due to a speaker's linguistic experience, that effect interpretation of individual phonetic tokens. These biases can result in recurrent patterns of misperception, hypocorrection of predictable phonetic effects, or hypercorrection of the same. An alternative model distinguishes classical notions of variation across phonetic tokens from listener-based transforms as shown in (1). This typology, from Evolutionary Phonology (Blevins 2004, 2006a, 2006b, 2008), distinguishes simple confusion of acoustically similar sounds (CHANGE) from ambiguous localization of percepts (CHANCE), and both of these from the speaker-induced variation discussed above (CHOICE).

(1) General typology of sound change, S = speaker, L = listener (Blevins 2004; 2009a)

a. CHANGE: A phonetic signal is *misheard* by the listener due to perceptual similarities of the actual utterance with the perceived utterance.
Example: S says [aθa] L hears [afa]

b. CHANCE: A phonetic signal is accurately perceived by the listener but is intrinsically phonologically ambiguous, and the listener associates a phonological form with the utterance which differs from the phonological form in the speaker's grammar.
Example: S says [ʔa̲ʔ] for /aʔ/ L hears [ʔa̲ʔ] and assumes /ʔa/

c. CHOICE: Multiple phonetic signals representing variants of a single phonological form are accurately perceived by the listener, but due to variation across tokens, and differing frequency of tokens, the listener (a) acquires a proto-type or best exemplar of a phonetic category which differs from that of the speaker; and/or (b) associates a phonological form with the set of variants which differs from the phonological form in the speaker's grammar.

Type i: S says [tuʔəlaŋ], [tuʔᵊlaŋ], [tuʔlaŋ] for /tuʔəlaŋ/
L hears [tuʔlaŋ], [tuʔᵊlaŋ], [tuʔəlaŋ] and assumes / tuʔlaŋ /

Type ii: S says [si], [sʲi], [ʃi] for /si/
L hears [si], [sʲi], [ʃi] and says [ʃi] for /si/ (surface change)

All phonetically based recurrent sound changes likely involve a complex interaction of these factors at some level, but for many, one source appears to be primary.

The separation of speaker-induced variation (CHOICE,1c) from listener-directed change in this model is motivated by the hypothesis that token frequency of variants plays a central role in sound change with sources in CHOICE (see, e.g., Bybee 2001). This contrasts with common cases of directional misperception (CHANGE, 1a). When Finnish listeners are exposed to partially devoiced final long vowels, they will consistently judge these as shorter than non-devoiced vowels, providing a simple phonetic explanation for recurrent final vowel shortenings in the world's languages (Myers and Hanson 2007). While it could be that final vowel devoicing is variable at some early stage, we have evidence that in stages preceding vowel shortening or loss with this phonetic source, final vowel devoicing can be regular and invariable. In Gilbertese, word-final long high vowels shortened, and short high vowels were lost. Closely related Woleaian and Pulo Anna, however, show consistent voiceless reflexes of the same Proto-Micronesian word-final short vowels (Blevins 1997). A similar comparison might be made between Rotuman, where final unstressed vowels were coarticulated in preceding syllables, and presumably devoiced preceding loss, and those in Kwara'ae, where the same sound changes are in evidence, but with loss of final voiceless vowels incomplete (Blevins and Garrett, 1998). It could be, then, that in many sound changes where final vowel devoicing plays a role in final long vowel shortening or short vowel loss, devoicing is invariable prior to loss. For speaker-based variation, frequency of variants is relevant to direction and speed of change, while in typical cases of CHANGE (1a), where X is misperceived as Y, the primary variable is the probability of confusing X with Y, not the ambient frequency of X as opposed to other phonetic variants.

Among listener-based transforms, consistent universal directions of misperception (CHANGE, 1a) are distinguished from changes with sources in categorization difficulties (CHANCE, 1b). This distinction is motivated by the hypothesis that difficulties in categorization will show strong 'structure-preservation' effects (cf. de Chene and Anderson 1979; Kiparsky 1995; Blevins to 2009a), while perceptual biases will not. This hypothesis, stated in (2), proposes, where feature or segment localization, or general categorization is ambiguous (CHANCE), sound change will be structure preserving with greater-than-chance frequency.

(2) Structural Analogy (Blevins 2004:154)

In the course of language acquisition, the existence of a (non-ambiguous) phonological contrast between A and B will result in more instances of sound change involving shifts of ambiguous (hard to categorize) elements to A or B than if no contrast between A and B existed.

Support for this hypothesis is found in a range of cross-linguistic studies of phonetically motivated sound change including: compensatory lengthening; perceptual metathesis; dissimilation; and syncope (Blevins to 2009a). It may also be in evidence in some of the recurrent historical vowel lengthening patterns attributed to universal perceptual factors by Myers and Hansen (2005, 2007). In the following section, I suggest that post-vocoidal lengthening (in contrast to, e.g., pre-sonorant lengthening) shows strong historic structure-preservation effects: categorization of vowels as 'long' in these phonetic contexts appears to be 'primed' by the existence of unambiguous long vowels in other contexts. By incorporating knowledge of ambient sound patterns into a general model of sound change, we come closer to predicting the probability of these phonetically based sound changes for a given language type.

2. Post-vocoidal vowel lengthening and the role of ambiguity

Kavitskaya (2002) argues that compensatory lengthening is a consequence of phonologization of pre-existing differences in phonetic vowel duration. Phonetic factors leading to longer vowel durations include longer V-C transitions, longer vowels before particular consonants, and open-syllable lengthening. By the typology in (1), compensatory lengthening is a case of CHANCE (1b), , where long transitions can render the *localization of duration* ambiguous, and where vowels which are intermediate between 'long' and 'short' offer further ambiguities in classification. By (2), compensatory lengthening sound changes should occur more often when a language has a pre-existing length contrast than otherwise, since the existence of 'long' and 'short' categories has a priming effect on classification of vowels with intermediate durations. Is this true? Of the 80 languages documented by Kavitskaya (2002), 72 or 90%, have a pre-existing vowel length contrast, while 8, or 10%, do not.[1]

Perceptual metathesis is illustrated in (1b). A long-domain feature has an ambiguous localization, and this is resolved with a feature/segment in a non-historical position (Blevins & Garrett 1998; 2004). Of the 24 cases of perceptual metathesis surveyed in Blevins & Garrett (1998, 2004), 23 are structure-preserving. The only exception is Cayuga, an apparent sound change in-progress with "little evidence that originally post-vocalic laryngeals have been phonologically reinterpreted as prevocalic." (Blevins and Garrett 2004:512).

While similarly broad cross-linguistic surveys of dissimilation have not been made, Ohala's (1993) view of dissimilation as 'hypercorrection' makes similar predictions. Under Ohala's view, sound change caused by 'hypercorrection' does not create new segments. This is because dissimilation is "the result of the listener applying normalization processes to the speech signal: normalization requires recovering a (presumed) standard sound from a signal that differs in some way from the standard." In the model proposed in (1), long-distance dissimilations are also instances of CHANCE (1b), involving ambiguity in percept localization. The prediction of (2) is that long-distance dissimilations should result in sound patterns which are pre-existing. Is this the case? In the few well documented cases I am aware of, the answer is yes: Grassman's Law in Sanskrit, $*C^hVC^h > CVC^h$, occurs where CVC^h is pre-existing; Javanese rhotic dissimilation, *rVrV > lVrV, occurs where lVrV is pre-existing; and Iban sibilant dissimilation , *sVsV > tVsV, occurs where tVsV is pre-existing (Blevins, 2009b).

A recurrent phonetically-based sound change which appears to fall squarely into this category is postvocoidal vowel lengthening. In the case of postvocoidal lengthening, a long formant transition between vocoid and following vowel results in categorization ambiguity: is the significant transition affiliated primarily with the first segment or the second? If naïve learners associate this duration with the second segment, historical vowel lengthening will occur. Myers and Hansen (2005) list a number of languages with synchronic sound patterns of this type, including: a range of Bantu languages (Chiluba, Holoholo, Rimi, Lumasaaba, Shi, Sukuma, Nkore-Kiga, Luganda, Kinyambo, Kihehe, Bemba, Makua, Jita, and Kimatuumbi); Anufo, a Kwa language (Casali 1998:151); and Modern Japanese (Poser 1986). They also note two cases of similar sound changes: in Old Icelandic (Hock 1986:442), and in 16th Century English (Dobson 1968: 705). A prediction of (2) is that there will be a strong tendency for sound changes of GV > GV: to occur when an unambiguous vowel length contrast pre-exists in the language, since in the course of acquisition, long vowels prime learners to categorize vowels of ambiguous length as long. In languages with no pre-existing vowel length contrasts, there is no priming effect and the ambiguous vowel is categorized with all others, as short. Do these sound changes correlate with earlier vowel length contrasts, or, can vowel length contrasts arise from post-vocoidal lengthening alone?

TABLE 1. POST-VOCOID V-LENGTHENING SOUND CHANGES
AND PRE-EXISTING VOWEL LENGTH CONTRASTS

LANGUAGE	SOUND CHANGE	PRE-EXISTING VOWEL LENGTH CONTRAST?
Bantu (with reflexes in, e.g., Chiluba, Holoholo, Rimi, Lumasaaba, Shi, Sukuma, Nkore-Kiga, Luganda, Kinyambo, Kihehe, Bemba, Jita, Kimatuumbi, Makua)	*(C)GV > (C)GV:	**yes** Proto-Bantu
Old Icelandic (Old West Norse)	iV > jV: , V = [+back]	**yes** Proto-Norse, Old Norse
Middle Japanese	(j)iu > ju:	no Old Japanese **yes** Early Middle Japanese
16th Century English	ju > ju:	**yes** Middle English **yes** Early Mod. English (ca. 1600) /u:/, /o:/, /i:/, /e:/

Table 1 lists known independent instances of *GV > GV: sound changes, based on Myers and Hansen's survey.[2]

For Bantu, I make a conservative estimate of only one phonetically-based sound change, with possible diffusion, though it is also possible that this sound change occurred multiple times.[3] With the exception of Makua, all Bantu languages undergoing the *GV > GV: sound change had a pre-existing vowel length contrast inherited from Proto-Bantu. Bantu languages which have lost the inherited vowel length contrast are: all Kilimanjaro-Taita (E.60, E.74); Kilombero (G.50 = Pogoro, Ndamba); Northern Mozambique (P.30 = Makua, Lomwe, Chuabo, Marendje); Tumbuka (N.20); Cewa, Nyanja (N.30); Nsenga, Kunda, Nyungwe, Senta, etc. (N.40); Zone S (Botswana, Lesotho, Mozambique, South Africa, Zimbabwe); Lenje, Tonga etc. (M.60, in Zambia); ,most of North-East Coast (see Nurse 1999:5-7), parts of Rufiji-Ruvuma (N.10, P.10, P.20) and SW Tanzania (M.10, M.20), and Manda (Nurse 1999:23). Northern Mozambique includes Makua (P.31), a language with synchronic post-glide vowel lengthening in derived environments (see Cheng and Kisseberth 1979 on the Ikorovere dialect, spoken in the Tunduru district of Tanzania.) However, all Makua word lists I have found, including Kisseberth's (1996) extensive on-line Makua-English lexicon, and Kroeger's (2005) dialect survey of Coastal Mozambique, show evidence of long vowels in non-post-glide positions from other sources, including: penultimate lengthening; vowel concatenation; monosyllabic lengthening; *VNC > V:C; and loans. Is there evidence that post-vocoid lengthening giving rise to synchronic Makua vowel length alternations occurred at a time when there were no length contrasts in the language, after the loss of Proto-Bantu vowel length contrasts, but before other vowel length contrasts developed? Lack of detailed historical work on Makua does not allow us to answer this question at present. For Bantu as a whole then, there is a strong correlation between post-vocoid vowel-lengthening, and a pre-existing length contrast.

A clear case of a similar independent sound change is the change of front + non-front vowel sequences in Old Icelandic to jV: sequences: **liugan* 'lie' > *lju:ga*, **keosan* 'choose' > *kjo:sa*, (*)*θri:ar* (fem.) '3', etc. noted by Hock (1986:442). As shown by inputs like (*)*θri:ar*, a vowel length contrast pre-exists in Old Norse, inherited from Proto-Norse, which also had a vowel length contrast.

The 16th century English change of *ju > ju:* noted by Dobson (1968:705) is similar, though more limited. Though this change may post-date the Great English vowel shift in some areas, recall that under the vowel shift, length contrasts ca. 1600 were maintained for long vowels u: (< o:), o: < (ɔ:), i: (< e:, ɛ:), and e: (< ɛ:).

Japanese has a synchronic alternation between /iu/ and [ju:], discussed by Poser (1986). These synchronic alternations appear to be phonologizations of the Middle Japanese **(j)iu* > *ju*: sound change proposed by Martin (1987:46). Several etymologies are shown in (3).

(3) Middle Japanese **(j)iu* > *ju*: (Martin 1987:46).

a.	ki[w]u > kjuu	'nine'
b.	zifu > zi[w]u > zjuu	'ten'
c.	*ofo-ki-ku > ooki-ku > ooki-u > ookjuu	'big'
d.	(j)if-u > (j)iw-u > juu	'say' (cf. juu 'tie' < ju[w]-w < juf-u

As noted in Table 1, Old Japanese did not have a vowel length contrast. However, as indicated by the etymology in (3c), other sound changes in Middle Japanese also gave rise to contrastive vowel length, and at least some of these preceded post-vocoidal lengthening (Martin 1987:43-47; 79, Table 4). Of special interest here are sound changes pre-dating or coextensive with post-vocoidal lengthening. These include intervocalic *f-weakening and loss in (3c) (*ufu > uwu > uu; ofo > owo > oo)*, which begins about 700 AD, and is complete by about 1000, as well as monophthongization of *ou* > [o:] and *au* > [ɔ:] attested as early as 1300. In Japanese then, there is no evidence that post-vocoidal lengthening gave rise to a novel vowel length contrast. On the contrary, evidence for long vowels resulting from crasis of two short vowels or monophthongization are attested in the earliest Middle Japanese texts.

The recurrence of similar sound changes across time and space is attributed to the fact that most regular sound changes are phonetically motivated, and can be modelled in terms of one or more of the mechanisms outlined in (1). However, if we want to get closer to predicting when sound change will occur, it seems useful to distinguish speaker-based variation (1c) from modulations introduced by the listener. Speaker-based variation will display frequency effects which may play a role in the rate and distribution of sound change. Listener-based modulations do not typically show these frequency effects, and can be separated into two basic types: one based on significant misperception (1a) ; and another on ambiguities in the acoustic signal giving rise to difficulties in categorization (1b). This second type of listener-transform shows robust structure preservation effects, consistent with Structural Analogy (2). It remains to be seen whether a sound change like post-vocoidal lengthening can give rise to a novel vowel-length contrast, without the priming effects of a 'long vowel' category. Though Finnish speakers are more likely to hear a vowel as 'longer' when it is preceded by a longer (vs. shorter) transition duration (Myers and Hanson 2005), whether it is categorized by early learners as 'long' during the acquisition process may depend on pre-established categories of long vs. short vowels.

3. Distinguishing typological tendencies from phonological universals

Ohala's general approach to sound change, as well as the more detailed typology proposed in Evolutionary Phonology are embedded in probabilistic models of language change. Certain phonetically motivated changes are more common than others, and lead to common sound patterns, while other phonetically motivated changes are less common, and lead to less common sound patterns.[4] Sound patterns resulting from specific sequences of sound changes, or phonetically based sound change combined with analogical change are less common than sound patterns which have simple rather than complex origins. Recurrent sound changes with apparent phonetic origins in articulatory or perceptual universal tendencies are subject to laboratory testing. Where results are positive, these sound changes, and resulting recurrent synchronic sound patterns are explained in terms of phonetically motivated sound change.

When this is the case, phonological universals seeking to explain the same range of sound patterns seem unnecessary, and lead one to ask: How can true *phonological* universals be distinguished from typological tendencies resulting from recurrent phonetically based sound change? First, we must have some means of distinguishing phonological from phonetic observations generally, as suggested in (4), in order to evaluate whether certain claimed phonological universals are truly phonological, or a consequence of innate phonetic dispositions.

(4) Phonological vs. phonetic observations

i. Phonological observations are statements referring to phonological entities (features, segments, prosodic units, alternations, phonological constraints), where phonological entities are typically discrete and invariant.

ii. Phonetic observations are statements referring to phonetic entities (measurable aspects of speech perception, speech production, the acoustic record), where phonetic entities may be continuous and variable.

Next, we must have some criteria for deciding whether an aspect of grammar is a true universal, or a by-product of tendencies in language change. Kiparsky (2006, 2008) has suggested a set of criteria for distinguishing true linguistic universals from typological tendencies. These appear in (5) with slight modification.

(5) Prerequisites for true phonological universals (after Kiparsky 2008)

a. Phonological universals should have no exceptions, and in this way, contrast with typological phonological generalizations which are, in principle, tendencies. (Tendencies may be "accidentally" exceptionless, but they are still a tendencies).

b. Phonological universals should constrain any kind of change. In a general sense, they define phonological "attractors" on which a variety of different diachronic paths may converge.

c. Phonological universals should emerge spontaneously within grammars.

d. Knowledge of phonological universals should be revealed in the way grammar unfolds in first language acquisition. Learners may construct grammars that violate typological generalizations, but they cannot, by definition, construct grammars that violate universals.

e. Phonological universals are rules or constraints that are part of every grammar and interact with each other dynamically in grammars. Typological generalizations are descriptive generalizations about grammars and are not necessarily a part of them.

(Where Kiparsky refers generally to 'universals', I have replaced this with 'phonological universals', focusing on the kind of evidence which might support a phonological universal over a universal or universal tendency with a clear basis in phonetically-based sound change.) While these appear to be the same criteria that Ohala, Maddieson, Yu, and many others, have used to distinguish phonological universals from typological tendencies, difficulties arise when what appear to be well documented phonetic universals (or universal tendencies) are repackaged as phonological constraints.

Consider, for example, Myers and Hansen's (2007) finding that final shortening can result from the failure to include voiceless portions of vowels as part of their overall duration. In this case, we have experimental evidence of a measurable gradient aspect of speech perception which plays a role in a phonological categorization task, and which is argued to play a role in recurrent sound changes where final short voiceless vowels are lost, and final long devoiced vowels are shortened. A translation of this finding into phonological terms might say something like "final voiceless vowels are non-moraic".[5] These two approaches are compared in (6).

(6) Phonetic vs. phonological accounts of final (voiceless) vowel

a. <u>Phonetic</u>: Identification of final vowels as long vs. short depends strongly on the duration of the voiced portion of the vowel, but is not significantly affected by the duration of the final voiceless portion of the vowel. This is due to the difficulty of perceiving voiceless vowels, which have low intensity. Such vowels are easily misperceived as silence, and this misperception can result in loss of short final voiceless vowels, as well as shortening of final long partially devoiced vowels.

b. <u>Phonological universal</u>: (Final) voiceless vowels are non-moraic (e.g. they do not contribute length/weight to syllables).

If we understand sound patterns involving voiceless vowels in terms of (6a), why, one might ask, do we need anything like (6b) as a phonological universal? The answer suggested by Kiparsky (2008) in (5) is that, in some cases, there may be universal aspects of sound patterns which cannot be explained by phonetic universals or phonetic universal tendencies. Within the model he advocates, where proposed universal constraints are violable within individual grammars, (5a) is particularly difficult to evaluate. If we do find a language where voiceless vowels are moraic, this is because the constraint in (6b) is dominated by a higher-ranked (faithfulness) constraint. For the other criteria in (5), the only one which will clearly distinguish phonological universals from phonetic universals or tendencies is (5e): but here, the real contrast is between grammatical generalizations and typological generalizations, not between phonological universals and typological ones. The phonological grammar contains rules or constraints that may interact with each other dynamically, but this interaction does imply universality; they could just as well be learned and language-specific.

How then do we evaluate the phonetic explanation in (6a) against the seemingly derivative phonological statement in (6b)? A reasonable working assumption is that if we have evidence for a phonetic explanation (whether a universal or strong tendency), and no *independent* evidence for a phonological universal, then we are justified in taking the phonetic explanation to be primary, and the phonological facts as derivative. In other words, there seems to be no justification for positing a phonological universal on the basis of facts which can be fully explained in phonetic terms. The debate in Blevins (2006a,b) and Kiparsky (2006, 2008) concerns final obstruent devoicing, not (final) voiceless vowels. I argue that the high frequency of final devoicing sound patterns in the world's languages is a consequence of phonetic universals and phonetic tendencies, while Kiparsky argues that the same distribution is due to a universal phonological constraint enforcing a preference for voiceless over voiced obstruents in final position.

Close to this debate, are facts concerning the distribution of voicing in obstruents generally, including word-initial obstruent clusters, the subject of Kreitman's paper. In the following discussion, I suggest gross probabilities for each initial cluster type, based on things we know about the phonetics of obstruent voicing, obstruent inventories generally, and the evolution of initial clusters. I then show that phonetic expectations are confirmed by filling out the typology. The universal implicational relationships suggested by Kreitman are disconfirmed, with common patterns of voicing in initial obstruent clusters showing all the hallmarks of typological generalizations, as opposed to phonological universals.

In (7) I outline some factors which are expected to contribute to the gross probabilities of voicing in word-initial obstruent-obstruent clusters. I limit myself here to stops, and use 'T' for voiceless oral stops, and 'D' for voiced oral stops.

(7) Gross probabilities for voice distribution in word-initial obstruent clusters

a. #TT	Should be most common since many languages lack a voicing contrast and have only T.
b. #DD	Should be least common since: (i) many languages lack D; (ii) voicing is difficult to maintain across long closure durations; (iii) no languages have only D, no T.
c. #TD	Should be less common than TT, since many languages lack a voicing contrast. Should be more common than DD, since there is no difficulty maintaining voicing. Can evolve from *TT, through assimilation of voicing with following vowel #TTV > #TDV.
d. #DT	Should be less common than TT, since many languages lack a voicing contrast. Should be more common than DD, since there is no difficulty maintaining voicing. Can evolve from *#VTTV > #VDTV > #DTV, via post-vocalic voicing and initial V-loss.

The general implicational relationships between singleton and geminate voiced and voiceless obstruents are assumed to follow from phonetic aerodynamic properties, summarized in (8), following Ohala (1997). To this, we can add some general knowledge about the evolution of word-initial CC-clusters. The most common phonetic sources of these clusters are unstressed vowel deletion and initial vowel loss, as schematized in (9).

(8) Aerodynamic Voicing Constraint (Ohala 1997)

Voicing requires: (i) vocal cords adducted (lightly approximated at midlines) and (ii) air flowing through the vocal folds. Consequences of this are:

- a. Voicing is inhibited on obstruents
- b. Factors favoring obstruent voicing are:
  - i. shorter closure duration
  - ii. larger oral cavity
  - iii. active expansion of oral cavity
    (via larynx lowering, jaw lowering, augmenting velum elevation)
  - iv. velic leakage

(9) The most common phonetic sources of word-initial consonant clusters

i. unstressed vowel loss	$\#C_1VC_2V\ldots > \#C_1C_2V$
ii. initial vowel loss	$\#VC_1(C_2)V\ldots > \#C_1(C_2)V\ldots$

With unstressed vowel deletion (9i), we expect the distribution of voicing in C_1 to be balanced between voiced and voiceless, and likewise at C_2. The exception would be a language with regular intervocalic voicing. In this case, after a sound change like (9i), initial #TD and #DD clusters arise, but no #TT. Initial vowel loss as in (9ii) results in exposure of what were once intervocalic consonant clusters. If these clusters were subject to regular regressive voice assimilation, then the output of this sound change will be only #TT, #DD, but no mixed-voiced clusters.

Let us compare the general remarks above with the results of Kreitman's observations on initial obstruent clusters, based on a survey of over 60 languages. She finds no languages with only #DD, no languages with #DT that do not also show #TD, and no languages with any cluster including voicing which does not also have #TT. From these observations she proposes the implicational relations between initial obstruent clusters shown in (10).

(10) Voicing in word-initial obstruent clusters: proposed implicational relationships (Kreitman)

i.	#DD implies #TT
ii.	#DT implies #TD, #TT
iii.	#TD implies #TT

From the brief discussion above, it is clear why languages with only #DD do not occur, since they are expected to be very rare by (7b). However, there is no good phonetic explanation for the absence of languages with #DT that do not also show #TD, and such languages are expected. In addition, natural phonetic explanations lead one to expect #TD-only languages, just in case *#TTV > #TDV via regressive voice-assimilation with the following vowel (7c), or #DT-only languages (7d) where post-vocalic voicing and initial vowel loss occur in sequence.

As I now show, some of these expected language types do exist, even if they are comparatively rare. I begin with Gitksan, a Tsimshian language of the Skeena River valley of British Columbia. Gitksan does not have a phonemic voicing contrast. There are two series of stops: plain stops /p t c k kw q/ ; and ejective stops /p' t' c' k' k'w q'/. The plain stops undergo regular allophonic alternations, surfacing as voiced when they precede vowels, and with an aspirate release in word-final position, as illustrated by the data in (11a-e) from Rigsby and Ingram (1990:256). Their proposed rule is that "Plain voiceless stops become voiced stops when they precede a vowel" (cf. earlier statements of pre-vocalic voicing of Rigsby 1967, and Hoard 1978.) This rule is supported by their instrumental observations of the stops in pre-vocalic position. The Gitksan prevocalic voiced stops have VOTs that range from prevoiced to weakly voiced, to short VOTs, in contrast to aspirates which have consistently long positive VOTs. As a consequence of this allophony, all word-initial /#TT/ clusters surface as [#TD]. For example, /ptal/ 'rib' is pronounced [pdal] as in (11e).

(11) Gitksan plain stop allophony

i) plain voiceless stops are voiced preceding vowels
ii) plain voiceless stops have aspirate release finally

a. nɪbɪpʰ	/nə-pip/	'maternal uncle'
b. nɪbɪbiy'	/nə-pip-əy'/	'my maternal uncle'
c. nɪbɪptʰ	/nə-pip-t/	'his/her maternal uncle'
d. nɪbɪpdi:tʰ	/nə-pip-ti:t/	'their maternal uncle'
e. pdal	/ptal/	'rib'

Gitksan, then, looks like a language where #TD does not imply #TT, contra (10iii) above.[6]

The opposite pattern of a bias towards DT clusters (7d) is found in Begak (Ida'an), an Austronesian language of Sabah. Begak has a very limited number of initial obstruent stop-stop clusters (Goudswaard 2005:27-29), which correspond to intervocalic cluster of the same limited types. Apart from voiced and voiceless geminates, the most common cluster type consists of a voiced stop followed by a homorganic voiceless stop. There is one lexeme with attested /gb/, and several sound symbolic words with initial /kp/, but otherwise, all non-geminate initial obstruent clusters have the voice contour DT, as shown in (12b).[7]

(12) Initial obstruent clusters in Begak (Goudswaard 2005)

a. geminates

#TT		#DD	
ppa'	'thigh'	bbong	'skin disease'
ttas	'high'	dda'	'blood'
kkan	'cooked rice'	ggud	'edible soft part (of a coconut)'

b. homorganic obstruent clusters

#DT	
gkot	'work'
bpuk	'hair'
bpow	'a smell'
dtow	'sun, day'
dtu'	'far'

c. other

#TT		#DD	
kpar!	'sound of something falling'	gban	'forest, jungle'
kpil!	'sound of deer'		
kpis!	'sound of sneezing'		

Though one might imagine this situation is due to post-vocalic voicing, as outlined in (7d), the historical situation appears to be more unusual. Blust (1998) analyses initial and medial homorganic DT clusters as historic reflexes of a medial 'fortition' process applied to voiced stops in a number of North Sarawak languages, where fortition of *b, *d, and *g resulted in *bp*, *dt*, and *gk* respectively. As examples, he provides: PMP **tebu* 'sugarcane', Begak *təbpu*; PMP **qalejaw* 'day', Begak *dtow* (< **Vdow*), PMP **beRat* 'heavy', Begak *bəgkat*. The historical progression leading to initial DT clusters then appears to be that shown in (13).

(13) Unusual path to initial #DT preference: Begak, Northern Sarawak

VDV	>	VDTV	>	DTV
	D-fortition		initial vowel loss	

	Proto-Malayo-Polynesian	Begak	
a.	*tebu	təbpu	'sugarcane'
b.	*beRat	bəgkat	'heavy'
c.	*qalejaw	dtow (< *Vdow)	'day'

While the realisation of a long or fortis D as a voiced-voiceless contour is unusual cross-linguistically, it is not altogether unknown. Outside of Borneo, obstruent stop clusters, including geminate stops are realized similarly in Gooniyandi a language of the southern Kimberley of Western Australia: "the articulation is initially weak (giving a lenis syllable final consonant), becoming stronger finally (where the syllable initial articulation is fortis)" (McGregor 1990:78-80). In addition to the rarity of fortition processes of this sort cross-linguistically, the existence of contrasts with singleton T, D and geminate DD and TT, illustrated in (14), results in a situation where one cannot treat DT clusters as phonetic instantiations of long voiced or voiceless stops. Returning to Kreitman's survey, we see that Begak constitutes a counter-example to the proposed implicational relation in (10ii): initial #DT does not imply #TD. While this was our expectation, given the common phenomenon of post-vocalic voicing, the origin of the sound pattern in this case appears to be quite different, involving fortition/lengthening of D > DT, combined with initial vowel loss, as illustrated in (13).

(14) Singleton and cluster contrasts in Begak (Goudswaard 2005)

T	a**p**u'	'ancestor'	ba**t**u	'stone'
TT	a**pp**a'	'chew'	ba**tt**as	'bridge'
D	a**b**uk	'dust'	ba**d**as	'honeydew'
DD	ka**bb**un	'garden'	ba**dd**ok	'lotion'
DT	a**bp**ut	'bite'	ma**dt**uŋ	'repeatedly'

Are there universal patterns of voice-distribution in word-initial obstruent clusters? As with oral stop systems generally, voiceless implies voiced for singletons, geminates and clusters. The phonetic basis of this implicational relation was suggested early on in the study of the aerodynamics of voicing, and is summarized in (8) above. These tendencies, combined with our knowledge of the most common origins of initial clusters, suggest a wider range of attested initial clusters than in Kreitman's survey, and as shown above, some of these appear to exist. In general then, the cluster patterns observed by Kreitman are better viewed as typological tendencies with phonetic bases, rather than patterns resulting from the interaction of unmotivated phonological universals. Acoustic studies, like that presented by Kreitman, and future articulatory studies of glottal activity in the same cluster types, will allow us to better understand the range of gestures resulting in obstruent voicing in clusters, especially when this occurs initially, with no preceding voicing gesture.

4. Co-variation, universals, and underphonologization

The examination of sound changes above involving the evolution of vowel length and a range of voicing distributions in word-initial obstruent clusters illustrates that, while some sound patterns are more likely to be innovated than others, typologies are end results of a complex interaction of factors, with phonetic naturalness as per (1) playing a major role in defining common directions and patterns of change, but not the only role. Pre-existing sound patterns may make listeners more sensitive to certain contrasts than others (2), or prime their categorization of otherwise ambiguous tokens, as suggested for sound changes of post-vocoidal lengthening. In assessing probabilities for word-initial obstruent cluster voicing, one must consider not only the raw factors involved in the aerodynamics of obstruent voicing, but also the probabilities of voicing in C_1 and C_2 positions prior to the evolution of clusters, the cluster-evolution processes themselves, and the range of 'voicing' processes commonly occurring in pre- and post-vocalic contexts.

Though these factors are stated as independent ones, it is well known that certain phonetic properties co-vary. A range of these are noted by Maddieson (this volume), and Yu (this volume), and listed in (15).

(15) Some patterns of phonetic co-variation

i. Higher vowels have higher F0 than lower vowels. (Intrinsic F0 effect)
ii. Higher vowels are shorter than lower vowels.

iii. Bilabials have longer closure duration than velars.
iv. Bilabials have shorter VOT than velars.

v. Voiceless oral stops are shorter than voiced oral stops.
vi. Vowels preceding voiced oral stops are longer than vowels preceding voiceless oral stops.

Understanding the degree of variation and control in such co-variation is extremely important in assessing the probability and nature of phonologization resulting from certain phonetically motivated types of sound change. For example, consider the possibility of a final voicing sound pattern, resulting from a language with intervocalic VTV > VDV, followed by final V-loss. As pointed out by Blevins (2009a), though post-vocalic and intervocalic voicing/lenition are common cross-linguistically, the output of such processes are not necessarily voiced oral stops, but lax segments with shorter durations, where incomplete closure can also be involved (Lavoie 2001, Kirchner 2004). As a consequence, what might be expected from such a sequence is not a language with final voiced oral stops, but one with final voiced fricatives or approximants.

The nature of such co-variation becomes even more important when arguments for phonological universals, like that made recently by Moreton (2009), are based on 'underphonologization' of expected, phonetically based sound changes. 'Underphonologization' is a general terms for the unexpected low frequency of phonological systems which, based on current knowledge of phonetics, should have higher frequencies. Moreton (2009) argues that the existence of more (synchronic) tone-to-tone phonological processes cross-linguistically vs. tone-to-consonant voicing/aspiration processes is due to a universal phonological bias by which language learners prefer modular (all tone) markedness constraints to non-modular ones. Moreton's argument is based on a general measure of 'phonetic precursor robustness': where the effect of F_0 has been measured, Moreton (2009) pools results, and suggests that there is no evidence that phonetic interaction is greater between two tones than between a tone and obstruent laryngeal features.

(16) Underphonologization attributed to phonological universals (Moreton 2009)

> Phonological patterns relating tone to tone are more common than those relating tone to consonant voicing or aspiration (14 families vs. 5 families). This skewed distribution cannot be derived from differentials in the robustness of phonetic pre-cursor effects on F0, and is attributed instead, to universal learner biases towards "modular" (tone-tone vs. consonant-tone) phonological markedness constraints.

But in this schematic argument, two important factors, investigated by Yu and Maddieson respectively, are missing. The first is an assessment of how these co-varying differences in F_0 are perceived. As Yu demonstrates, not only are dynamic patterns perceived as longer than level F_0s, but high level F_0s are also perceived as having longer durations than mid and low level F_0s.[8] Until relevant perception studies are done, we cannot assume that acoustic 'phonetic precursor robustness' corresponds to perceptual robustness.[9]

The second missing ingredient is determination of the extent to which the measured interactions co-vary across different rates and styles of speech. In Maddieson's first experiment, 3 of 4 speakers showed increased F_0/vowel height differences in fast speech, but all showed reduced vowel duration-voicing ratios in fast speech. In the model of sound change outlined in (1), assessing the interaction of speech rate on co-varying phonetic properties is extremely important. Reduced vowel duration-voicing ratios in fast speech suggest that there should be underphonologization of vowel length arising from an earlier voicing contrast, since such co-variation is reduced in fast speech. In the same way, although the vowel length contrast in Lai appears to be robust in slow speech (Maddieson, Figure 1), the contrast as cued by covariation in coda C-length, is significantly reduced in fast speech. In this case, the expectation is that closed-syllable vowel shortening could occur in a language like Lai, despite co-variation in slow speech, since, in fast speech parallel cues are absent. Lack of co-variation, then, could be a source of 'overphonologization' of sound changes like closed syllable shortening, which are extremely common cross-linguistically.

In addition to co-variation, we can return to (2), Structural Analogy, for another potential explanation for tone-tone vs. laryngeal-consonant-feature + tone interactions in recurrent sound change. If tone is more often contrastive on vowels than consonants due to longer vowel duration and greater overall energy (Gordon 1999, 2001), then perturbations of F0 which are ambiguous in source are more likely to be attributed particular pre-existing tone-bearing-units (in this instance, vowels with tonal specifications). Finally, one cannot rule out other more general cognitive factors relating to similar vs. dissimilar objects which could play a role.[10] These other factors which could result in under-phonologization are summarised in (17).

(17) Other factors which could result in under-(or over)-phonologization

a. <u>Perceptual biases</u>: Similar F0 differences can be perceived differently depending on context. This could result in under- or over-phonologization.

b. <u>Co-variation</u>: Lack of co-variation can result in underphonologization of certain associated properties (e.g. preceding vowel length/consonant voicing), but overphonologization of others (e.g. closed syllable shortening in a language like Lai)

c. <u>Structural analogy</u> (2): Tone is more often contrastive on vowels than consonants due to longer vowel duration and greater overall energy (Gordon 1999, 2001). By (2), perturbations of F0 which are ambiguous in source are more likely to be attributed to pre-existing unambiguous tone-TBU contrasts (e.g. vowels with tonal specifications), than to non-pre-existing contrasts.[11]

d. <u>General cognitive facilitation of observation under similarity</u>: There may be a general cognitive tendency to notice associations between similar things more easily than associations between dissimilar things. This result in overphonologization of tone-tone associations, which would be noticed before tone-consonant associations, and underphonologization of the latter. But this aspect of grammar would be a phonological instantiation of a much more general cognitive principle.

In sum, phonetic variation along the hypo-to-hyperarticulation continuum, along with misperception and perceptual ambiguity, define distinct sources of phonetically based sound change, all leading to recurrent sound patterns. Co-varying phonetic properties may not be consistent across different speech rates or styles, suggesting potential phonetic sources of under- and over-phonologization. Even when phonological patterns may appear exceptionless, there is mounting evidence that these recurrent patterns reflect typological tendencies, which instantiate the complex interaction of articulatory variability and invariability, perceptual variability and invariability, and independent language-specific structural priming factors which may enhance or inhibit categorization in the course of language acquisition.

References

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

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* 32:117-65.
Blevins, Juliette 2006b Reply to commentaries. *Theoretical Linguistics* 32: 245-56.
Blevins, Juliette 2008 Consonant epenthesis: natural and unnatural histories. In Jeff Good (ed.), *Language universals and language change*. Oxford: Oxford University Press. 79-107.
Blevins, Juliette 2009a Structure-preserving sound change: A look at unstressed vowel syncope in Austronesian. In Alexander Adelaar and Andrew Pawley (eds.). *Austronesian historical linguistics and culture history: a festschrift for Bob Blust*. Canberra: Pacific Linguistics. 33-49.
Blevins, Juliette 2009b Phonetic explanations for recurrent sound patterns: Diachronic or Synchronic? In Charles Cairns and Eric Raimy (eds.), *Contemporary Views on Architecture and Representations in Phonology*. Cambridge: MIT Press. 325-36.
Blevins, Juliette and Andrew Garrett 1998 The origins of consonant-vowel metathesis. *Language* 74: 508-56.
Blevins, Juliette and Andrew Garrett 2004 The evolution of metathesis. In Bruce Hayes, Robert Kirchner, and Donca Steriade (eds.). *Phonetically based phonology*. Cambridge: Cambridge University Press. 117-56.
Blust, Robert 1998 The position of the languages of Sabah. In Ma. Lourdes S. Bautista (ed.), *Pagtanaw, essays on language in honor of Teodoro A. Llamzon*. Manila: The Linguistic Society of the Philippines. 29-52.
Bybee, Joan 2001 *Phonology and language use*. Cambridge: Cambridge University Press.
Bybee, Joan 2008 Formal universals as emergent phenomena: The origins of structure preservation. In Jeff Good (ed.), *Language universals and language change*. Oxford: Oxford University Press. 108-21.
Casali, Roderick F. 1998 *Resolving hiatus*. New York: Garland.
Cheng, Chin-Chuan and Charles Kisseberth 1979 Ikorovere Makua tonology. *Studies in the Linguistic Sciences* 1:31-63.
De Chene, Brent and Stephen R. Anderson 1979 Compensatory lengthening. *Language* 55:505-35.
Dobson, E. J.. 1968 *English pronunciation 1500-1700*. 2nd edition. Oxford: Oxford University Press.
Fletcher, Harvey and Wilden A. Munson 1933 Loudness, its definition, measurement and calculation. *Journal of the Acoustic Society of America* 5:82-108.
Gordon, Matthew 1999 *Syllable weight: Phonetics, phonology, and typology.* PhD dissertation, UCLA.
Gordon, Matthew 2001 A typology of contour tone restrictions. *Studies in Language* 25:405-44.
Gess, Randall S. 1998 Compensatory lengthening and structure preservation revisited. *Phonology* 15:353-66.
Goudswaard, Nelleke 2005 *The Begak (Ida'an) language of Sabah*. Utrecht: LOT (Netherlands Graduate School of Linguistics).
Greenberg, Joseph H. 1969 Some methods of dynamic comparison in linguistics. In J. Puhvel (ed.) *Substance and structure of language*. Berkeley: University of California Press. 147-203.
Hoard, James E. 1978 Obstruent voicing in Gitksan: some implications for distinctive feature theory. In Eung-Do Cook and Jonathan Kaye (eds.), *Linguistic Studies of Native Canada*. Vancouver: University of British Columbia Press. 111-19.
Hock, Hans Heinrich 1986 Compensatory lengthening: in defense of the concept 'mora'. *Folia Linguistica* 20:431-60.
Kavitskaya, Darya 2002 *Compensatory lengthening: phonetics, phonology, diachrony*. London: Routledge.
Kiparsky, Paul 1995 The phonological basis of sound change. In John A. Goldsmith (ed.), *The Handbook of Phonological Theory*. Cambridge, MA: Blackwell. 640-70.
Kiparsky, Paul 2006 The amphichronic program vs. Evolutionary Phonology. *Theoretical Linguistics* 32: 217-236.
Kiparsky, Paul. 2008 Universals constrain change; change results in typological generalizations. In Jeff Good (ed.), *Language universals and language change*. Oxford: Oxford University Press. 23-53.
Kirchner, Robert 2004 Consonant lenition. In Bruce Hayes, Robert Kirchner, and Donca Steriade (eds.). *Phonetically based phonology*. Cambridge: Cambridge University Press. 313-45.

Kisseberth, Charles 1996 Makhua-English lexicon. Available at: www.cbold.ddl.ish-lyon.cnrs.fr/CBOLD_Lexicons/Makhua.Kisseberth1996/Text/Emakhua.Kisseberth1996.txt

Klatt, Dennis and Laura Klatt 1990 Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustic Society of America* 87:820-57.

Kreitman, Rina (this volume) Mixed voicing word-initial onset clusters.

Kroeger, Oliver 2005 Report on a survey of coastal Makua dialects. SIL International.

Lavoie, Lisa 2001 *Consonant strength: Phonological patterns and phonetic manifestations*. New York: Garland.

Maddieson, Ian (this volume) Variation in co-variation: the search for explanatory principles.

Martin, Samuel E. 1987 *The Japanese language through time*. New Haven: Yale University Press.

McGregor, William 1990 *A functional grammar of Gooniyandi.* Amsterdam: John Benjamins.

Moreton, Elliott 2009 Underphonologization and modularity bias. In Steve Parker (ed.), *Phonological argumentation: Essays on evidence and motivation*. London: Equinox.

Myers, Scott and Benjamin Hansen 2005 The origin of vowel length neutralisation in vocoid sequences: evidence from Finnish speakers. *Phonology* 22:317-44.

Myers, Scott and Benjamin Hansen 2007 The origin of vowel length neutralization in final position: evidence from Finnish speakers. *Natural Language and Linguistic Theory* 25: 157-93.

Nurse, Derek 1999 Towards a historical classification of East African Bantu languages. In Jean-Marie Hombert and Larry M. Hyman (eds.), *Bantu Historical Linguistics*. Stanford: CSLI.

Ohala, John J. 1981 The listener as a source of sound change. In C. S. Masek, R. A. Hendrik, and M. F. Miller (eds.), *Papers from the parasession on language and behavior*. Chicago: Chicago Linguistic Society. 178-203.

Ohala, John J. 1989 Sound change is drawn from a pool of synchronic variation. In L. E. Breivik and E. H. Jahr (eds.), *Language Change: Contributions to the study of its causes*. [Series: Trends in Linguistics, Studies and Monographs No. 43]. Berlin: Mouton de Gruyter. 173-98.

Ohala, John J. 1990 The phonetics and phonology of aspects of assimilation. . In J. Kingston and M. Beckman eds., *Papers in Laboratory Phonology I*. Cambridge University Press, Cambridge. 258-75.

Ohala, John J. 1993 The phonetics of sound change. In Charles Jones (ed.), *Historical Linguistics: Problems and Perspectives*. London: Longman. 237-278.

Ohala, John J 1997 Aerodynamics of phonology. *Proceedings of the Seoul International Conference on Linguistics*. Seoul: Linguistic Society of Korea. 92-97.

Poser, William J. 1986 Japanese evidence bearing on the compensatory lengthening controversy. In E. Sezer and L. Wetzels (eds.), *Studies in Compensatory Lengthening*. Dordrecht: Foris. 167-87.

Rigsby, Bruce 1967 Tsimshian comparative vocabularies with notes on Nass-Gitksan systematic phonology. Paper presented at the Second International Conference on Salishan Languages.

Rigsby, Bruce and John Ingram 1990 Obstruent voicing and glottalic obstruents in Gitksan. *International Journal of American Linguistics* 56:251-63.

Tarpent, Marie-Lucie 1987 *A Grammar of the Nisgha Language*. PhD dissertation. Department of Linguistics, University of Victoria.

Yu, Alan (this volume) Tonal effects on perceived vowel duration.

Notes

[1] De Chene and Anderson (1979:517) suggested that all cases of compensatory lengthening were structure-preserving, developing only if vowel length contrasts pre-exist in a language. However, a range of exceptions have been noted (Gess 1998; Kavitskaya 2002).

[2] Anufo is not included. The historical phonology of Anufo, a Kwa language of north-eastern Ghana and northern Togo, is not well studied, but comparison with closely related Agni and Baoule as well as internal reconstruction suggest pre-existing long vowels in open and closed syllables, preceded by a range of consonants.

[3] If multiple parallel sound changes occurred, this will only increase the attested support for a principle like Structural Analogy (2).

[4] A similar, approach, with much less phonetic detail, is advocated by Greenberg (1969), and further developed by Bybee (2008). Campbell's (1998:295-98) general discussion of explanation and predictability in language change also advocates a probabilistic approach.

[5] An independent question is why there are so few languages with a contrast between modal voiced vowels and voiceless vowels. See Blevins (2004:199-201) for suggestions on why such a contrast might resist phonologization, and section 4 on general approaches to 'underphonologization'.

[6] Initial /TT/ clusters are relatively rare in Gitksan (Jason Brown, personal communication). Patterns similar to those in (11) are transcribed by Tarpent (1987) for Nisgha, a related Tsimshian language.

[7] Homorganic NC clusters also occur initially and medially (Goudswaard 2005:27-29). In medial position, NT vs. ND contrast (*kampu* 'out of breath' vs. *kambiŋ* 'goat'), but initially, only ND is attested, though the set of initial ND words is very small.

[8] It is not clear to me how the well known interactions of F_0 and intensity (equal loudness contours) documented by Fletcher and Munson (1933) are factored out in Yu's studies.

[9] Moreton (2009:16) admits to this possibility in his concluding discussion: "A still-viable alternative explanation … is that non-grammatical perceptual effects may skew the training data before it reaches the learner's pattern-finding mechanisms… The issue will only be settled in the lab."

[10] I am grateful to Andy Wedel for bringing general cognitive constraints of this sort to my attention.

[11] Compare Blevins (2004:199-201) on the possible underphonologization of voiceless vowels. There it is suggested that expected contrasts between voiced and voiceless vowels are rare due to (i) common loss of voiceless vowels due to perceptual factors; and (ii) the tendency to analyse voiceless vowels as /Vh/ or /hV/ sequences, where /h/ is pre-existing, as per (2).