Reply to Commentaries

Juliette Blevins
Max Planck Institute for Evolutionary Anthropology

1 Evolutionary Phonology.

The primary goal of Evolutionary Phonology is to explain recurrent sound patterns in spoken languages across time and space. One domain of explanation is diachronic, and much of my own work has focused on diachronic explanations (e.g. Blevins 2004, 2006a, To appear a, To appear b, To appear c, this volume). Another domain of explanation is synchronic. A range of synchronic factors, from the mechanics of speech perception and production (e.g. Myers and Hansen 2005, To appear), to aspects of languages use (e.g. Blevins 2005a, 2005b), to general cognitive strategies in learning, categorization, and pattern-matching (e.g. Wedel 2004, 2005, To appear a, To appear b), are likely to play a major role in defining recurrent aspects of sound patterns.

In the context of modern phonological theories, various proposals have been made concerning the form and content of phonology-specific universal synchronic knowledge. These proposals range from the distinctive feature system of SPE to the markedness constraint inventory of OT. Two central questions arise with respect to putative phonological universals of this type. First, as Blevins (2004) and Kiparsky (this volume) emphasize, we must ask if these aspects of knowledge are truly innate, grounded in language use, or combine facets of both. Second, as stressed in Blevins (2004), we must ask if the knowledge in question is truly 'phonological', i.e. not reducible to phonetic knowledge, and not a special instance of cognitive processes that apply more generally to human learning and pattern extraction in other domains.

In this context, it seems that two of the commentaries in this volume misconstrue the central explanatory goal of Evolutionary Phonology (EP). Kiparsky mistakes the goal of EP as an attempt 'to eliminate UG', while De Lacy erroneously equates 'extra-phonological' explanations with 'extra-cognitive' ones. Though EP challenges the existence of phonological universals, it does not deny, and, indeed, it embraces work in phonetics and cognitive sciences more generally which demonstrates synchronic effects of innate knowledge or processing effects in these two domains. This means, in effect, that there is no substantive difference between Evolutionary Phonology, as presently conceived and developed, and the 'amphichronic' program that Kiparsky outlines, except, perhaps, in what one hypothesizes is the primary domain for explanation. These are two different names for the same general enterprise. With this as background, two issues are conflated in De Lacy's contribution: (i) the degree to which phonologies have diachronic vs. synchronic explanations; and (ii) whether the content of the purported 'phonological component' (PC), could emerge from phonetic substance (as suggested by Lindblom), or alternatively, whether it could reflect much more general (non-phonological) aspects of cognition.

A very strong claim made by both Kiparsky and De Lacy is that phonological systems show non-random recurrent patterns which are functionally arbitrary, in the midst of functionally motivated patterns. These effects are referred to by De Lacy as 'straightjacket effects' or 'functional ignorance' effects, and are claimed to provide evidence for a synchronic phonological component. (For further discussion of 'functional ignorance', and general 'underphonologization effects' see Blevins (to appear d)). Kiparsky attributes the absence of word-final obstruct voicing to an effect of this sort, while De Lacy identifies final voicing in Somali as an instance of sonority increase, and takes the absence of selective devoicing as evidence of a highly specific non-functionally motivated phonological constraint. Kiparsky adopts a general universal OT constraint prohibiting marked features in weak positions: voiced obstruents are marked, final position is weak, and therefore, final obstruct voicing should not occur in any spoken language. De Lacy attributes Somali final voicing to a general universal OT constraint hierarchy where high-sonority moraic consonants are marked, and low-sonority moraic consonants are unmarked. Facts regarding final voicing may be disputed (see below), but what of other 'marked' features, like obstruct aspiration? As Iverson and Salmons (this volume) highlight, final aspiration or final fortition is widespread in the world's languages,
and well documented for many languages, including German. It runs counter to the straightjacket effects suggested by Kiparsky and De Lacy, but is entirely consistent with two phonetic factors noted in the target article: (i) prosodic boundaries may be marked by laryngeal (spreading or closure) gestures; and (ii) in phrase-final position, the vocal folds may spread, anticipating the laryngeal configuration for relaxed non-speech breathing. If Kiparsky and De Lacy are correct, 'straightjacket effects' should be visible, and final aspiration should be absent from the world's languages. However, it is present, and robustly attested in many languages (Vaux and Samuels 2005). Here again, EP makes what appears to be the right predictions, while synchronic approaches incorporating synchronic phonological markedness constraints do not.

As a theory attempting to explain recurrent sound patterns in spoken languages across time and space, Evolutionary Phonology might perhaps be viewed as having entirely disjoint goals from the generative tradition. This is evidently De Lacy's position, and might be attributed to Hamann as well. On this view, generative phonology is a theory of language competence, while EP would be a theory of transmissibility or performance. Carrying this separation further, one might be led to suggest that generative theories, including OT "should not account for frequency (typological, lexical, etc.) as frequency falls squarely in the domain of Performance factors" (De Lacy, p.14). (Hamann goes even further, dissociating EP from phonological theory generally by offering a narrow generative definition of the latter.) Yet De Lacy's insistence that generative phonology and OT do not, and should not, account for the typology of sound patterns, and phonological typology generally, is oddly incongruous with general practice over the past 40 years. Within the generative tradition, the goal of synchronic explanation has been taken to consist, at least in part, of the description and analysis of cross-linguistic sound patterns. This is true across a wide range of 'core' phenomena, including: feature/segment inventories (Chomsky and Halle 1968, chapter 7); vowel harmony systems (Archangeli and Pulleyblank 1994); stress patterns (Hayes 1995); syllable markedness constraints (Kager 1999:92-98); and even obstruent voicing constraints of precisely the type under investigation in the target article above (Hayes and Steriade 2004). In each case, typological patterns are assumed to provide evidence of implicit speaker knowledge. Irrespective of whether one regards this assumption as valid, it is plain that EP is an alternative theory that offers different types of explanations for precisely the same range of typological patterns.

2 Modelling sound change.

The general model of sound change outlined in Blevins (2004), and summarized in the target article is simplified along many dimensions. The primary focus is on phonetic sources of sound change, where sources can be roughly classified in terms of misperception, difficulty in percept localization, and articulatory variation. This focus is justified by the fact that perceptual asymmetries appear to influence patterns of sound change, independent of articulatory patterns, while difficulty in percept localization may result in strong structural analogy effects which are less visible for other types of sound change (Blevins to appear c). As Andersen's criticism points out, this restricted code does not suffice to account for other types of phenomena, including contact-induced change, the social dimensions of sound change, etc.

However, many of Andersen's criticisms stem from what may be regarded as artificial distinctions drawn between his earlier models and EP. (Use of the term 'alternations' in the target article refers only to synchronic alternations, and does not play a role in the model of sound change proposed.) Although the term 'allophonic change' is not applied to CHOICE, CHOICE is used as a cover-term for the evolution of conditioned allophones. The terms CHANGE, CHANCE and CHOICE (all in small caps) are perhaps not ideal, but they do refer to cross-classifications of the traditional, more descriptive terms that Andersen refers to. The central proposal is that this cross-classification is useful and illuminating, especially where relationships between sound change and pre-existing sound patterns are concerned. It also seems useful once one accepts that changes in pronunciation may or may not reflect changes in a phonological system, and likewise that the moment of 'phonologization' may occur long before two categories are in obvious contrast.

Another issue that Andersen raises is the usefulness of applying evolutionary concepts to language change (cf. Lass 1997), and cultural evolution more generally (cf. McElreath and Henrich, to appear). Similar criticisms have been made by a range of scholars, with useful recent summaries in Garrett (2006), Blevins
At the same time, there has been much progress in attempting to model grammatical competence and language change in terms of language as a complex adaptive system in which grammatical regularities are emergent probabilistic properties, resulting from the repeated interaction of innate biases, self-organizing properties of linguistic systems, and aspects of language use within a population. Numerous references were given in section 1.1 of the target article, including a range of scholars attempting to simulate language as a complex adaptive system. This is the sort of modeling that provides a formal context and validation for the application of the term 'evolutionary' to phonological systems, and, in particular to emergent sound patterns. I am fully in agreement with Lindblom (this volume) that EP should 'go computational', as must any realistic theory of the complex range of factors that interact in sound change, and linguistic change more generally. For recent work in this area with specific reference to EP, see Wedel (2004, 2005, to appear a, to appear b) and Mielke (2005).

As acknowledged above, the model of sound change in EP does not focus on the social factors and social contexts in which 'sorting' and adoption of variants takes place. Nevertheless, it is important to recognize that this process can, in principle, be distinguished from the existence of variants and their phonetic explanations. Consider, for instance, the many languages (French, German, Dutch) in which apical rhotics have shifted to uvular rhotics, or vice versa (Straka 1965, Howell 1987). A fundamental question for our understanding of this process is what apical and dorsal rhotics have in common, and whether shifts between these two sound types are based on articulation, perception, or both (cf. Engstrand et al., to appear). A very different question is how and why one variant has spread through a community (cf. Van de Velde and Van Hout 2002). The primary research focus of EP is on answering questions of the first type: what are the phonetic bases of sound change, and can we get closer to determining probabilities of sound change based on other aspects of the sound system and linguistic system more generally? Hence, in Labov's (1994) terms, Evolutionary Phonology concerns itself with 'internal factors', not 'social factors', that is, with 'initiation', not 'propagation' of sound change, a point that is made explicit in Blevins (2004).

3 Final voicing, final devoicing, and related facts.

The target article illustrates the types of explanations that EP makes available, and contrasts the distinctive predictions it makes with those of other modern approaches. As an illustration, the article presents an overview of final obstruent devoicing processes, suggests diachronic phonetic explanations for these sound patterns, and then evaluates data which might suggest the existence of final obstruent voicing sound patterns as well. As two contributors take issue with some of the facts presented, it will be useful to address their concerns briefly, before turning to other, more general, empirical issues.

None of the commentaries in this volume dispute that word-final obstruent devoicing is a common sound pattern demanding explanation, and that this pattern is directly related to the aerodynamics of voicing, and a range of articulatory and perceptual factors that can result in the realization or interpretation of final voiced obstruents as voiceless. What is contested is: (i) Whether there are languages which show word-final voicing of obstruents (Kiparsky's section 2); and (ii) Whether final devoicing sound patterns show recurrent arbitrary restrictions which cannot be attributed to phonetic or other functional factors (De Lacy's 'straightjacket effect'). I will briefly discuss each of these points, and then turn to related facts brought up in the commentaries regarding obstruent laryngeal features and their distribution.

Kiparsky presents alternative analyses for the potential final-voicing sound patterns discussed in the target article. The reader can decide which analyses seem more plausible than others. However, certain facts should not be overlooked, nor the role of abstractness in phonological analyses generally. In the debate over final neutralization in Somali, a central question is what the correct phonological analysis of the opposition between /t k/ and /d g/ is in this language. Is it a tense/lax contrast, a fortis/lenis contrast, an aspirated/unaspirated contrast, a voiceless/voiced contrast, or something else? The debate takes place in the "conceptual prison" of the phonetics/phonology split from which Lindblom advises a final exodus: if the contrast in this language has multiple cues, and those cues are weighted differently in different contexts, why should one set of labels be favored over another?
For some Somali speakers, neutralization in final position can result in phones which are partially voiced. Armstrong (1964:4-8) is quite clear on this, describing final /b/, /d/, /dʒ/ and /g/ as either ‘voiceless (or with slight voicing) and with no release’ or as being produced with glottal closure and release, and Edmonson et al. find a similar range of allophones in word-final position. A related fact is that the geminates /bː, dː, ɖː, gː/ are also partially or fully voiced, so that generally, voicing is a phonetic feature associated only with the /b d ɖ g/ class. Kiparsky admits that there is a voiced variant in final position, but factors it out of the phonological analysis by fiat. The abstract phonological analysis he suggests represents final neutralization in terms of devoicing and deaspiration, with voicing attributed to "a separate phonetic implementation". But designation of voicing as "phonetic implementation" is purely arbitrary. Voicing could just as well be the rule, with glottalization and devoicing attributed to phonetic implementation, as suggested in the target article. Plainly, if surface voicing can simply be disregarded in phonological analyses of final neutralization processes, then there is no way of demonstrating that final-voicing sound patterns exist. For any case, the phonological analysis can be stated as 'final neutralisation', with voicing out-sourced to phonetic implementation. Indeed, counterexamples to any putative phonological universal can be overcome by disregarding inconvenient patterns or processes, and attributing them to 'phonetic implementation'. But recourse to this type of expedient classification should lead one to question the value of the original claim or universal.

Other facts disputed in the article relate to the 'straightjacket effects' suggested by De Lacy. De Lacy claims that selective devoicing (e.g. devoicing of only /g/ in a /b d g/ system) is unattested. As the original purpose of the target article was not to demonstrate selective final devoicing, but rather to illustrate aspects of the historical progression of final-devoicing supporting the phonetic aerodynamic account, no attempts were made to demonstrate that these patterns were robust and exceptionless. The evidence from Tonkawa, Frisian and Haisla is admittedly weak, because this evidence was selected to establish a different point. Frisian and Haisla may not be relevant, since Frisian lacks final /g/ (due to historical spirantization, resulting in modern [x]), while Haisla lacks final /b/. For Tonkawa, Hoijer's (1933) account is highly suggestive:

The voiceless media (b, d, g, gw) are pronounced in a manner about halfway between the corresponding English surds and sonants; somewhat as the b, d, and g of the central German dialects. They occur in all positions - initially, medially, and finally. In final position, g becomes the surd -k… dj varies with dz. It is approximated in pronunciation by the j of English 'judge'. In the final position it is entirely unvoiced, -tc… (Hoijer 1933:4)

But this description is absent from his later 1946 grammatical sketch, where the full series of stops, transcribed as /p t ʈ k kw/ are described as 'voiceless, unaspirated, and lenis' and often unreleased in final position (Hoijer 1946:290).

Selective devoicing is a topic, however, that has been treated by others. Hayes and Steriade (2004) invoke a place/voice/length markedness scale to account for selective devoicing. The scale is supported by a cross-linguistic survey of segment inventories, which illustrates "patterns of selective voicing neutralisation" (ibid p.11). The scale is also claimed to account for selective obligatory devoicing of velars and coronals over labials in a dialect of Sudanese Nubian (ibid p.11). Brown (this volume) notes that /g/-gaps in segment inventories have been attributed to the same selective devoicing processes, and careful allophonic descriptions of stop series show similar effects. For example, in Anejom (Lynch 2000:14), there is only a single series of stops /pʰ p t k/. Between vowels, the labial stops are fully voiced, while other stops are only partly voiced. Phonologization of this sort of pattern yields a language with /b t k/ (instead of /p t k/), and such inventories, though rare, appear to exist (e.g. Ket, as described by Vajda 2004). Unusual mixed final-neutralization patterns are found in Formosan languages. In Atayal (Rau 1992), with contrasting /p t k/ and /b r ɡ/; the voiced series are typically realized as voiced continuants, but word-final /b/ is neutralized to [p], /ɡ/ to [w], and /t/ to [j].

Furthermore, while De Lacy denies the existence of selective devoicing, he says nothing of the recurrent gaps in segment inventories that Hayes and Steriade (2004) and Donohue (2006) also attempt to explain. Many of these gaps can be attributed to the aerodynamics of obstructive voicing, though it is not clear that these phonetic explanations should be directly incorporated into synchronic phonologies. Brown's contribution shows that a range of different factors can play a role in defining recurrent gaps. For a /ɡ/ gap,
there is the selective devoicing already noted, but also selective lenition, which in many languages targets voiced velars before other voiced obstruents. More interestingly, Brown shows that some languages (e.g. Khalkha) may have only voiced velar stops as a consequence of historical velar lenition processes, and that other unexpected gaps are also attested. In addition to feature-based markedness constraints, synchronic principles of feature economy might also be invoked (Clements 2003, 2004). However, there are many stable consonant inventories which violate even the most permissive notions of feature economy (Blevins 2005d).

Another effect that De Lacy identifies as a consequence of arbitrary properties of the phonological component is the absence of k-epenthesis processes. He suggests that while /k/-epenthesis might be functionally motivated, the fact that /k/ is more marked than coronals or laryngeals entails that /k/-epenthesis does not occur (though compare Donohue 2006, on the unmarked status of velar stops). Yet Blevins (to appear b) demonstrates just such an 'unnatural' epenthesis process in Trukese, and velar stop/fricative epenthesis is also attested in Mongolian, Maru, and Land Dayak (Vaux 2003). In short, the putatively non-existent sound patterns on which arbitrary aspects of the synchronic phonological component are based, are indeed attested. There are few, if any, clear 'straightjacket effects', and one source of motivation for a (non-derivative, non-emergent) phonological component ultimately reflects simple misconceptions about attested sound patterns and sound inventories.

4 Concluding remarks.

This reply offers a welcome opportunity to clarify some misunderstood aspects of Evolutionary Phonology, and to integrate new facts and often stimulating commentary into this synthetic paradigm. Evolutionary Phonology suggests explanations for recurrent sound patterns in spoken languages across time and space. These explanations can be of many different types, ranging from historical phonetic ones, to functional usage-based ones, to general cognitive aspects of learning and generalization. A fundamental difference between this approach and generative models lies in the nature of the starting assumptions. Generative phonology assumes that recurrent sound patterns reflect speaker knowledge, and that much of this knowledge is innate and phonology-specific. Evolutionary Phonology does not pre-judge the issue, and is compatible with the evidence we have that phonological knowledge is learned and language-specific (e.g. Ernestus and Baayen 2003). Generative phonology also places a major explanatory burden on synchronic phonological grammars, while Evolutionary Phonology does not, and emphasizes the extent to which phonetically based sound change can explain recurrent sound patterns. As stressed at the outset, the competing models make different predictions about sound patterns of the world's languages. Evolutionary Phonology predicts that final obstruent voicing is possible, though it may be rare, or even unattested in the sample of languages for which we have reliable descriptions. It also predicts the widespread occurrence of final aspiration, and the existence, but rarity, of inventories like those discussed by Brown. Though the majority of work in Evolutionary Phonology to date has involved explanations for typological tendencies and has maintained a phonetics/phonology divide to engage the wider phonological community, the model invites increased computational implementation, and is particularly compatible with the general view of phonological categories and generalizations as emergent properties of complex adaptive systems.

References


