On formal universals in phonology

Andrew Nevins

Department of Linguistics, Harvard University, Cambridge, MA 02139.
nevins@fas.harvard.edu
www.fas.harvard.edu/~nevins

Abstract: Understanding the universal aspects of human language structure requires comparison at multiple levels of analysis. While Evans & Levinson (E&L) focus mostly on substantive variation in language, equally revealing insights can come from studying formal universals. I first discuss how Artificial Grammar Experiments can test universal preferences for certain types of abstract phonological generalizations over others. I then discuss moraic onsets in the language Arrernte, and how its apparent substantive variation ultimately rests on a formal universal regarding syllable-weight sensitivity.

The target article by Evans & Levinson (E&L), questioning not only the existence of universals but the methodology of pursuing them, raises many issues for discussion. In this commentary, I limit my remarks to the importance of formal (as opposed to substantive) universals – a distinction E&L do not draw with sufficient clarity. Formal universals in phonology are constituted by the analytic elements that human minds employ in constructing representations of sound structure. Put simply, formal universals refer to the set of available data structures (e.g., binary features, metrical grids, autosegmental tiers) and the possible operations on them that can be used in constructing a grammar of a language.
    Phonotactic dependencies of the form *Given segments A, B in the same word, if A has feature F, then B must have feature G* are

constrained by analytic restrictions on what can be referred to. These analytic restrictions are, by hypothesis, formal universals that are independent of the historical contingencies or cultural practices of any given language community. One of the best ways of studying formal universals of this kind is by constructing Artificial Grammar Learning experiments, using the methodology of cognitive science.

In one such recent study, Moreton (2008) conducted an experiment in which participants were taught a miniature artificial language containing phonotactic dependencies of the form outlined above. There were three conditions: in one, *F* and *G* were both vowel height; in a second, *F* and *G* were both obstruent voicing; and in a third, *F* was vowel height and *G* was obstruent voicing. Importantly, the rules of English phonotactics do not contain any of these three dependencies. The results, however, showed that the height-voice dependency was not learned by participants. Moreton's conclusion was that an analytic bias favors learning certain phonotactic dependencies over others; the resulting formal phonological universal is in (1):

> (1) Learning phonotactic dependencies of the form *Given segments A, B in the same word, if A has feature F, then B must have feature G* is universally easier when *F* and *G* are the same feature than when *F* and *G* are different features.

Formal universals like (1) lend themselves to eminently more possibilities for integration with the cognitive sciences than E&L's proposed research program based on "the dual role of biological and cultural-historical attractors" (target article, sect. 8, para. 6, E&L's thesis 5). Formal universals allow for experimental testing in laboratory conditions under which the historical-cultural factors are completely controlled for, and hence irrelevant to the outcomes.

It is worth considering how apparent exceptions to universals are analyzed in other fields. As an example, consider the case of the Jacana bird, one of nature's species exhibiting a "sex-role reversal," whereby it is the females that engage in polyandry and cuckolding of the males. At the right level of analysis, the sex-role reversal in these shorebirds is entirely unsurprising, because it is the males that perform the raising of the chicks. The correct asymmetry between sexes is not that males have multiple mates while females do not, but rather, that the sex that commits to what biologists call "parental investment" is the one who is stuck in the harem. When the universal is correctly formulated, the Jacana bird is actually an exception that proves the rule.

I argue that E&L err in concluding that there are no universals within human phonology based on apparent substantive exceptions:

> But in 1999, Breen and Pensalfini published a clear demonstration that Arrernte organizes its syllables around a VC(C) structure and does not permit consonantal onsets(Breen & Pensalfini 1999). With the addition of this one language to our sample, the CV syllable gets downgraded from absolute universal to a strong tendency, and the status of the CV assumption in any model of UG must be revised. (target article, sect. 2.2.2, para. 2)

Arrernte is not, as E&L would have it, but one language that recently "ruined the entire sample," so to speak. The question of VC syllabification in Australian languages was raised by Sommer (1970; 1981) on the language Oykangand, later insightfully analyzed in terms of onset-maximization by McCarthy and Prince (1986). There was, historically, a widespread loss of initial consonants throughout Australian languages, which Hale (1964) and Blevins (2001) attributed to stress shift and lenition processes. Although Arrernte was apparently no exception to this sweeping change, nonetheless, "25% of Arrernte words are pronounced in isolation with an initial consonant" (Breen & Pensalfini 1999, p. 2). To account for words such as *mp^w ar* and *taŋk*, Breen and Pensalfini (1999) have to propose that these words have an underlying hidden initial vowel, a red-flag for any "clear demonstration" that the language disallows consonantal onsets.

In general, the deduction of which syllabification pattern a word contains depends on particular phonological processes that refer to syllabic divisions. In this light, consider the following formal universal:

> (2) Stress assignment, weight-sensitive allomorphy, compensatory lengthening and prosodic morphology, when sensitive to distinctions among syllable types, refer exclusively to the representational unit of weight called the *mora*.

The phonological universal in (2), developed by Hyman (1985), McCarthy and Prince (1986), and Hayes (1989), is formal, not substantive in nature: it restricts the data structures that can be referred to by morphophonological processes, and is not about the substantive question of which segments can bear moras. In fact, Topintzi (2009) has gathered evidence from a wide range of languages demonstrating the ability of onset consonants to be moraic. The existence of metrical processes referring to onsets has been a topic of research for many years; see Davis (1988), Downing (1998), Goedemans (1998), and Gordon (2005), who discuss onset-sensitivity of stress in languages ranging from English and Italian to Pirahã and Iowa-Oto.

If vowels and onset consonants, but not coda consonants, are moraic in Arrernte, the statement of stress assignment and weight-sensitive allomorphy become quite straightforward in the light of (2). In Arrernte, stress is assigned within a word to the first vowel preceded by a consonant: *mp^w á.rem*, "is making," versus *i.k^w ént*, "policeman." Since onset consonants are moraic, the stress rule is simple: the left-most bimoraic syllable receives stress. Similarly, the statement of plural allomorphy in Arrernte is simple: bimoraic-or-greater forms like *i.ŋel* and *taŋk* take the suffix *-ewar*, while monomoraic forms like *ar* and *aŋk* take the suffix *-erir*. The reduplication patterns can receive a similar treatment in terms of moraic targets, within the prosodic morphology framework: for example, the copying of VC strings to a reduplicant is driven by the demand to fill a bimoraic template.

Like the Jacana bird's sex-role reversal, which has a mechanistic and principled explanation at a different level of primitives (partner with greater parental investment, instead of male and female), the patterning of weight-sensitive process in Arrernte exhibits a principled conformity to a formal universal at the level of which consonants are moraic, instead of in terms of syllabification. Rather than positing a silent initial vowel for 25% of Arrernte words, attention to the statement of formal universals enables a consistent representational property for syllables throughout the language. The universal in this case pertains to the set of data structures that learners use to encode sound patterns: moras, and only moras, are the formal unit that can be referred to by weight-sensitive properties.

E&L trumpet the slogan "A linguist who asks 'Why?' must be a historian" (sect. 7, epigram, quoting Haspelmath 1999, p. 205). Integration with the cognitive sciences, however, will come from mechanistic explanations, not from hand-waving at diachronic contingencies. Formal universals are restrictions on representational vocabulary, and they rear their heads even when history deals them an odd shuffle or, as in the case of artificial grammar experiments, no historical shuffle at all.