

Prosodic-phonetic biases and the construction of phonological markedness

Elinor Payne
University of Oxford

This paper critically revisits the concept of phonological markedness through the scrutiny of two properties that are classically assumed to be of fundamental importance: universalism and mentally encoded norms of sound patterns. It argues that while both are individually relevant to the phonological markedness, or rather to what proponents of markedness seek to model or explain, these two properties relate to different things and should not be interwoven. Markedness is recast as the *language-specific* read-out of formal structural norms that, though mentally encoded, are not universal but constructed through learning. These norms have been differentially shaped by external factors that are themselves universally law-governed but are external to linguistics. A narrative of phonological markedness is thus constructed for each language. The iterative mutual influence of structures and speech behaviour is mediated by prosodic-phonetic biases that both promote and in time slowly re-fashion this narrative of markedness in the language concerned.

Markedness in language is characteristically a vague term, adopted by most phonological frameworks either implicitly or explicitly, but afforded different theoretical status and roles. At a purely descriptive level, markedness refers fairly uncontroversially to what is observed to be typologically frequent or infrequent. Beyond a simple statement of statistical distribution, however, other claims are theory-bound, that is, they presuppose assertions regarding the motivating forces behind such distribution, and are not axiomatic. For example, it is commonplace to talk of phonetic ‘simplicity’ or ‘naturalness’ in the unmarked, though such properties are not self-evident or clearly defined. Commonly proposed or assumed diagnostics of the unmarked include relative degree of articulatory ease, and early emergence in the trajectory of phonetic

acquisition (cf. Jakobson, 1962). For example, in infant speech acquisition, fricatives are typically noted to be acquired later than stops and this has been explained as reflecting the greater articulatory precision and control required in the production of fricatives and thus in the generation of a turbulent air-stream (as opposed to its 'simple' obstruction or a smooth airflow as in the case of stops). Factors of relative complexity are deemed not only to influence the order in which sounds are acquired but also to underlie typical adaptive strategies, for example fricative-stopping. The more 'marked' element (in this case, the fricative) constitutes a greater challenge to the developing child and is thus 'replaced', at least at early stages in development, by the less 'marked' element. During the babbling phase all manner of sounds are produced, including fricatives, although with lower frequency (MacNeilage & Davis, 2000: 154) and not necessarily with explicit 'segmental' control (i.e. as an intended target). Thus, fricatives are not so readily produced early on and are slower to be 'corralled' into the 'pen' of linguistic structure in a systematic way.

However, these challenges pertain to the developmental phase, and there is no external reason to suppose that they persist; once they have been acquired, for the typical adult speaker there is phonetically nothing strange or difficult about producing fricatives. In English, it could be argued that a residue of fricatives' 'otherness' can be seen in domain-initial fortition of fricatives (> stops), thereby decreasing the frequency of fricatives in speech; however, equally, stops may be lenited to fricatives in other contexts, thereby *increasing* their frequency. According to markedness theory, however, even markedness that is considered to be rooted in functionality is purported to stick around when the functional difficulties have long disappeared, and this raises some problems only resolvable through the proposition of a kind of abstract, 'natural' law, that pervades all languages and all instances of a language. A similar paradox can be found cross-linguistically. The typological relative rarity of clicks and ejectives is considered related in some way to their deemed phonetic 'complexity' or 'difficulty'. However, for the speakers of those languages in which clicks and ejectives occur, they are not infrequent, unnatural or difficult. On the contrary, for those speakers, clicks and ejectives form part of the regular and normal sound patterns in those languages.

The argument of complexity is also problematic, as the following example can illustrate. Aspirated stops are deemed to be more complex than unaspirated stops, but what form does this complexity take? Is it the case that 'aspiration' is seen as a quality that is 'added', as iconically reflected in the addition of a diacritic in IPA transcription, and if so to what extent is this merely an artifact of the way that these sounds are conventionally analysed and labelled? From

an articulatory point of view, aspiration results from a *different* co-ordination of oral and laryngeal parameters, but it is not clear in what sense this particular co-ordination is more complex than that required for unaspirated stops. Could one argue instead that aspirated stops are *acoustically* more complex, because they ‘add’ a separate and qualitatively distinct phonetic interval? Does that in turn presuppose greater auditory or psycho-acoustic complexity? Without external explanation or definition of complexity, the association of ‘markedness’ with ‘complexity’ is at risk of circularity: *aspirated stops are deemed marked because they are complex, and they are deemed complex because marked*. Moreover, as with the clicks and ejectives example above, in those languages where aspirated stops are prominent (either allophonically, as in most varieties of English and in Norwegian, or contrastively, such as Icelandic or Thai), there is no externally motivated (i.e. non-circular) basis for asserting that these are unusual or complex.

The idea of phonetic naturalness may have better success when it refers to the coincidence of certain properties, for example the coincidence of voicelessness with finality, for which a (physical, not abstract) law-governed explanation can be posited. Importantly, final voicing is possible, and is found – but external (i.e. physical) factors can conspire to make devoicing more likely in final position. Specifically, as discussed in Smith (1997) and Blevins (2004), there are quantifiable articulatory and aerodynamic factors leading to the transglottal pressure differential being too small for voicing to be maintained. It has been remarked that markedness is often a local phenomenon (cf. Greenberg, 1966), making it relative to a given context, hence there is greater theoretical purchase to be gained from making the more precise assertion that voicing is marked in obstruents in final position. The physical forces conspiring against voicing in that context are greater than in other contexts, and so the probabilities of devoicing occurring (and the chances in turn for this phonetic event to become grammaticalised over time) are greater. The loss of voicing leads to a loss of contrast and hence restricted distribution and neutralisation (unless the contrast is mapped onto some other feature, as in English with pre-fortis clipping). However, it is the loss of voicing that leads to restricted distribution and neutralization, rather than restricted distribution and neutralization dictating the loss of voicing. It is the probabilities that arise from physical properties and laws that shape structure, not the other way round. As Blevins (2004) argues, this kind of externally driven event requires no grammar-internal codification to explain it. The grammatical codification is the result of the processes, not the driver.

As critiqued by, among others, Blevins, in her *Evolutionary Phonology* (2004), theories that view markedness as an abstract ‘force’ effectively encode typologically statistical biases through linguistic conspiracy. So pervasive is this notion that, as Haspelmath (2006: 27) describes, the concept has become “an almost theory-neutral everyday term in linguistics”. However there are nuances to the abstract approach that are dependent on the specific strand of theory. In an extreme but nevertheless typical version, this encodification is made directly into or viewed as fundamentally embedded in the grammar. Also typical, though not inevitably associated with a grammar-internal approach, is the view that markedness is universal. For Prague School theorists (cf. Trubetskoy, 1931, 1939), markedness informs the notions of phonological complexity and contrast, providing a basis upon which to set up oppositional contrasts which pattern in an asymmetrical, non-equivalent way, hence features are either mark-bearing (e.g. voiced, nasalized, etc.) or non-mark-bearing (voiceless, non-nasalised, etc.). This asymmetry is understood to influence phonological ‘activity’, in a goal-orientated and hence teleological way and to reflect the representation of phonology in the mind. However, *early* Prague School accounts at least were not explicitly universalist, allowing markedness to be language-specific. In contrast, a universalist non-mentalist perspective can be seen later on in Greenberg’s *Language Universals* (1966), which identified cross-linguistic patterns but made no explicit claim about mental grammars. For Greenberg, markedness is a meta-grammatical concept *tout court*, describing what are merely observable patterns.

It is the thesis of this paper that while both patterns of mental representation *and* universalism, taken separately, are relevant to the formulation of markedness (in its descriptive sense), they are separate aspects of the conceptual apparatus and relate to different things. It will be argued that markedness relates to a descriptive mental read-out of *language-specific* regularities and norms, that are individually acquired and may shift, and that these patterns have been varyingly shaped by external factors (functional, phonetic, e.g. propensity for final devoicing) that are themselves universally law-governed. Thus universalism does influence mental structures, but only indirectly, and the nature of what is universal is not linguistic. Unifying the two – that is, either claiming that mental patterns are universal, or that what is universal has to be mentally represented – is flawed. In either version, it conjures a form of ‘natural law’ that is neither necessary or useful to understanding, despite being reinforced, either implicitly or explicitly, in much of phonological theory from the mid-20th century to the present day.

Although markedness was dealt with less explicitly in SPE (Chomsky & Halle, 1968) and subsequent work in rule-based generative phonology (e.g. Chen, 1973; Schachter, 1969), markedness values were nevertheless employed as a technical device to capture the relative ‘naturalness’ of phonological structures (where ‘naturalness’ was synonymous with cross-linguistic frequency). It was acknowledged that markedness was not an absolute quality – it could not explain all sound patterns and changes – but the notion of relative naturalness was nevertheless pervasive, and reached a more generalised expression in Natural Phonology theory (e.g. Hurch & Rhodes, 1996; Stampe & Donegan, 1980). The language-internalisation of phonetically grounded naturalness was an implicit imperative: for example, Kenstowicz and Kisseberth (1979: 251) state: “phonological theory must have some apparatus for expressing the fact that neutralization to a glottal stop [pre-consonantly or finally] is a natural rule as opposed to, say, neutralization to *l*”. The mechanism of internalization evolved in step with theory developments. Later, in his theory of Principles and Parameters (1981), Chomsky incorporates the idea of markedness by positing that parameter settings are not equal: “[the theory of markedness] imposes a preference structure on the parameters of UG [...] In the absence of evidence to the contrary, unmarked options are selected” (Chomsky, 1981: 8). Following in Jakobson’s footsteps, this predetermined markedness is in operation at acquisition, indeed it is viewed as the starting point for the acquisition process, and its main driver: “the unmarked case of any parameter represents the initial hypothesis that children make about the language to be acquired” (Kean, 1992). The universal nature of markedness in generative phonology is closely tied to the supposition of an innate code for language (UG), and this conceptual co-dependency makes it difficult to discern chicken from egg. UG predicts universal (natural) laws and properties such as markedness, and at the same time a theory of universal markedness provides justification for UG.

More recently, constraint-based accounts such as Optimality Theory take the formalization of markedness and place it at the very centre of their conceptual apparatus. In OT, outputs are the result of a conflict between markedness constraints (which exert a universal pull) and faithfulness constraints (which require adherence to the given forms of a language, thereby exerting a language-specific pull). How this conflict is resolved (i.e. the chosen output form) depends on the precise ranking of these constraints, with the likelihood that lower ranking constraints being dominated and therefore rendered inert. This novel reconfiguration of the mechanism by which (universal) markedness makes its presence felt in any given language ‘explains’

why the unmarked form may not always surface, or may only surface in a gradient manner. Thus, by way of an example, the universally default form is for final obstruents to be voiceless (as per the markedness constraint relating to this feature), but if the input to the system has final voiced obstruents, and if the faithfulness constraint is ranked more highly than the markedness constraint, then voiced final obstruents will surface. Notably, OT is silent on the nature of the input (the underlying representation) and why input forms abound with marked features in the first place (there are no restrictions on possible inputs – referred to as the ‘richness of the base’). Markedness is seen as an active force in the generation of output forms, rather than the fossilized products of diachronic processes that have washed up in the lexicon.

What both rule-based and constraint-based accounts of generative phonology have in common is to offer an *endogenous* account of markedness, that is, an abstract formalization of biases that are said to be grammar-internal (whether by a set of individual, disconnected rules or by constraints that are more globally encompassing). There have been several critiques of the formalization of markedness in this manner (e.g. Blevins, 2004; Hale & Reiss, 2000, 2001; Haspelmath, 2006; Hume, 2004). Blevins’s *Evolutionary Phonology* (2004) advocates instead for what can be described as an *exogenous* approach, whereby observed ‘markedness’ – typological biases in sound patterns and distributions – is to be viewed as nothing more than the consequences of external (i.e. functional, phonetic) factors. The theory draws on a large body of research in phonetics (e.g. Liljencrants & Lindblom, 1972; Lindblom, 1983; Ohala, 1974, 1978, 1981, 1983; Westbury & Keating, 1986) and Bybee’s theory of emergent phonology (2001), showing that functional attributes such as ease of articulation and perceptual distinctiveness, and the aerodynamic, acoustic and neurological principles that constrain these, can go a very long way to explain phonological structure and sound change.

For Blevins, the phonetic-diachronic evidence places explanation for phonological patterns truly outside of synchronic phonology and within (phonetically-driven) diachronic change:

[...] there is no clear role for markedness within synchronic phonology. Absolute universals and universal tendencies in sound patterns emerge from general pathways of language change, and have no independent status in the grammar [...] there is a great deal of empirical evidence against the direct incorporation of markedness into synchronic grammars. (Blevins, 2004: 20)

On this interpretation, the apparent fruits of markedness are emergent structures resulting from active forces external to phonology, not an active

synchronic force within phonology. Biases and patterns are not denied: Evolutionary Phonology is a “thesis not against markedness per se but against codification of markedness in grammar” (Blevins, 2004). Although coming from a very different theoretical viewpoint, Hale and Reiss reach a similar conclusion: “[M]any of the so-called phonological universals (often discussed under the rubric of markedness) are in fact epiphenomena deriving from the intersection of extragrammatical factors like acoustic salience and the nature of language change” (Hale & Reiss, 2000: 162).

The evidence for the ultimate phonetic grounding of sound patterns is vast and hard to refute. We may view this kind of phonetic explanation as a form of ‘phonetic determinism’, which in its extreme version would hold that phonological patterns are ultimately determined by phonetic facts. However, while phonetic determinism may have a critical role, and indeed the only role that is in any sense and without controversy governed by principles that are universal, it is arguably an incomplete theory insofar as it overlooks the potential influence of systemic factors. In its extreme form, phonetic determinism attempts to find perfect explanation of sound patterns from the way we hear and speak. However, while there may be a post hoc phonetic rationale for why, for example, final obstruent devoicing is so common cross-linguistically, unlike the aerodynamic principles underlying that tendency the pattern of final obstruent devoicing is not absolute. The physical consequences of a particular set of external, functional events and circumstances may be precisely correlated with those events and circumstances and thus predictable, but the *linguistic* consequences are not, for a linguistic system is not a naturally occurring object. Within the boundaries of what is physically possible, for a linguistic system there are many viable pathways. Hence there is no natural law – whether formal or functional – against final voiced obstruents, and indeed they are to be found (e.g. Norwegian *dag*, *labb*, *grad*). Phonetic determinism may be necessary to explain what happens when Y is in condition X, but it is not sufficient to explain or predict what this then means for a particular language system. For those who advocate a strict, impervious division between the abstract system and its implementation (cf. Hale & Reiss, 2000), and also for those contesting the universality of phonological structure (cf. Blevins, 2004), this may be a desirable, even intended, outcome. However, it has the unfortunate added effect of not being able to constrain external factors that, while universal in nature, are not universal in their application or influence. Neither abstract nor concrete universals can give us a complete explanation. Language-specific sound patterns are, as a result, caught between

the rock of phonological determinism and the hard place of phonetic determinism.

‘Phonetically grounded’ versions of the OT model seek to extend the explanatory reach of the endogenous (formal) approach to sound patterns (cf. Archangeli & Pulleyblank, 1994; Bermudez-Otero & Börjars, 2006; Hayes & Steriade, 2004). More recent models of OT thus incorporate phonetic constraints, the assumption being that speakers have formalized knowledge of ‘phonetic difficulty’ and that this knowledge is part of the phonological system. To illustrate this, consider the rarity (and thus attributed markedness) of voiced geminate obstruents. No (known) language bans just voiceless geminates, or rather, the presence of a voiced obstruent geminate in a given language implies that of the corresponding voiceless geminate. As exogenous accounts like Evolutionary Phonology would highlight, there is a well-established phonetic factor conspiring against voiced obstruent geminates: the longer closure duration leads to an insufficient pressure difference between the sub-glottal and supra-glottal cavities, leading to an earlier cessation of phonation. Furthermore, the effect of this principle is observed more readily in stops articulated towards the back of the mouth (reducing the size of the supra-glottal cavity, and thereby leading to a quicker build-up of pressure), and this can ‘explain’ the greater rarity of voiced velar geminate stops over voiced bilabial geminate stops. For exogenous accounts, as we have seen, these principles are external to phonology and do not need to be encoded in the grammar. For the revised endogenous accounts (e.g. phonetically-grounded OT), however, they become enshrined as a property of grammar. Moreover, because of the mechanism of constraint ranking in OT, and the possibility of low-ranking constraints being rendered ‘inert’, this revised account can purportedly ‘explain’ the non-universal emergence of (universal) phonetic constraints.

As a *meta-linguistic* way of modelling how for an individual language a particular set of outputs can be generated, the revised (phonetically grounded) version of OT can arguably deliver. But is this indeed the desired goal? Modelling is not synonymous with explanation, and hijacking system-external pressures on phonological structure and porting them into the phonology itself is both unmotivated and unnecessary. The primary drive behind this model is to preserve ‘universalism’ in phonology, since functional constraints incorporate the true universalism of physical laws. But the need for universalism, and thus the placing of these physically-derived constraints, should be challenged. Does an individual need to have knowledge of cross-linguistic patterns and distributions, even when they run counter to her own language? The existence of final voiced obstruents is a ‘problem’ for cognitive (abstract phonological)

determinism because they violate an abstract universal of markedness. OT proposes a ‘neat fix’ to this by ranking constraints and then allowing low-ranking constraints to be violated (and the subsequent emergence of the marked). However, as Payne has argued elsewhere (2009), a low-ranked constraint is little more than a way of formalising the thought: “there is something that I, in my capacity as linguist, do not like about Property X because I know that it is uncommon and historically unstable, even if I know also that it occurs in Language Y with no evident communicative impairment”. Furthermore, for nativist versions of this theory, there is also an implicit claim that: “native speakers of Language Y know, at some unspecified level, that there is something ‘not quite right’ about Property X, though it does not in the least affect their speech behaviour (for the moment at least)”. This would seem to be a vast overstatement of what is required. A more reasonable starting point is that a speaker knows what is ‘normal’ (unmarked) for her own language: the burden of proof lies with the claim that she has to know more than this.

On the issue of innateness, OT in itself is theoretically neutral. Innateness is said not to be presupposed by the model: “OT only requires that CON be universal” (Tesar & Smolensky, 2000: 130). However, we shall see that this insistence on universality actually leads to the need for innateness. Optimality Theorists defend the universalism of markedness constraints by arguing that even when a given constraint appears not to be operative in a language (i.e. because violated by a higher-ranking conflicting constraint), its influence can emerge under certain conditions (i.e. where the dominating constraint is not relevant.) This is known as The Emergence of The Unmarked (TETU). “[TETU] lends support to view that there is a single universal constraint set for all languages” (McCarthy & Prince, 1994: 334).

Some scholars (e.g. Becker & Potts, 2011; Flack, 2007; Hayes & Wilson, 2008) propose that certain types of constraint are induced by learners and thus not all constraints are innate. Flack (2007) distinguishes between functional and formal grounding of OT constraints. *Functional* constraints have to be available for all learners, of any language, to induce, and thus are more directly connected with phonetic (articulatory, perceptual, psycholinguistic) factors. *Formal* constraints may be *indirectly* connected – historically – to functional motivations, but they have since been grammaticalised to the point of opacity, and can no longer be induced from speech behaviour and thus, according to Flack, they must be induced via innate knowledge of markedness. Both kinds of constraint are said to be universal. One virtue of this approach at least is its acknowledgement of the redundancy in the innate specification of phonetically grounded markedness, and in itself, the universality of induced general

phonetic principles is not especially controversial (if we stop short of the theory requiring them to be encoded in the grammar). However, the assumption that *formal* constraints are also universal is problematical. Within the theory, formal constraints are those said not to be inducible from experience, and thus if they are universal they must by necessity be innate. Indeed, this requirement is only necessary if formal constraints are both universal and non-inducible. If formal constraints were language-specific, they could be induced through exposure to that language in the learning process, and hence there would be no need to resort to the innateness 'fix'.

I propose instead that formal markedness be construed as a set of language-specific properties, which is learnt through exposure, and can be explained diachronically through functional grounding *interacting with prior structure*. In the externalization of functional grounding, and in its non-universalism, this view is in line with exogenous theories such as Evolutionary Phonology. Where it departs is in the role of pre-existing (language-specific) structure in the selective permeation of functional grounding. Purely exogenous accounts acknowledge the rich source of phonetic structures that the variability and physical constraints inherent in speech give rise to, but stop short of examining how these may be differentially channelled by different languages. Without the role of language-specific structure, the non-universal and non-uniform influence of external factors cannot be explained other than as the result of unconstrained haphazardness, and without the acknowledgement of systemic influences, apparent directional stability in language change cannot be explained.

The *endogenous versus exogenous* polemic in the discussion on markedness can at one level be viewed as a tug-of-war over the primacy of phonology or phonetics in explanation. However, this polemic risks neglecting the importance of how the two interact. The phonological structures of any language are to a large degree constrained by the structures it has 'inherited'. Equally so, any 'live' processes emerging from speech behaviour (influenced by universal phonetic principles) can only act on what is there. So, while there may be general, universally-governed forces pushing in the direction of final devoicing, if there are no final voiced obstruents at a particular diachronic stage in a language, then this general force is irrelevant. Other 'forces' acting on voiceless obstruents may become, or appear, more prominent because of this distribution. If our language is tonal, like Norwegian for example, a reduced frequency of voiceless codas may impact on the alignment of tone within the syllable (through the action of other functional forces). And so on and so forth. What may seem a trivial point has important ramifications. Pre-

existing structure acts as a series of constraints on the specifics of speech behaviour and on the influence that speech behaviour has the *opportunity* to exert. In this way, a language's structural inheritance will make certain strategies more likely, and thus for certain patterns, and pathways to change, to emerge.

Through the iterative, constraining influence between existing structure and speech behaviour, apparent structural 'conspiracies' may arise, with a particular structure generating phonetic patterns that then reinforce that very same structure. In this way the 'unmarked' properties of that particular language emerge, are reinforced, and in turn exert an influence. Thus there is a high propensity for stressed syllables to lengthen in production. A reinforcement and codification of this lengthening may lead to phonologically long elements, which may be the nucleus vowel or a coda consonants, or, as in Norwegian, both of these in alternation. In terms of speech processing, a viable mechanism for this reinforcement is through differential frequencies and thus intensity of exemplar activation. The notion of system-conformity is not new, nor its role in apparent diachronic 'conspiracies'. It recalls in particular some aspects of Kiparsky's structure preservation thesis: "System-conforming variants have a selective advantage which causes them to be preferentially adopted. In this way, the language's internal structure can channel its own evolution, giving rise to long-term tendencies of sound change." (Kiparsky, 2006: 221). What is being newly proposed here is that system-conformity is in response to a set of learnt – and deeply and continually reinforced – biases in the phonetic implementation of structure – 'prosodic-phonetic biases' that, when combined, describe the unmarked properties of a language.

Speech is awash with these biases. Even phenomena that may be considered 'merely' the result of biomechanical constraints, such as coarticulation, show language-specific biases in the detail of their execution (e.g. gestural overlap in /kl/ clusters, see Gibbon, Hardcastle, & Nikolaidis (1993), or the varying propensity in Norwegian and English for open release in clusters, see Simonsen et al., 2016). The bias here is in temporal co-ordination, but biases may also be in the spatial dimension. Honikman (1964) introduces the notion of distinctive articulatory settings, conditioned in part by the overall structural properties of the language and resulting in minute differences in the implementation of what is ostensibly the same or similar phonological structure: "The internal articulatory setting of a language is determined, to a great extent, by the most frequently occurring sounds and sound combinations in that language". This setting gives rise to what is experienced by the speaker as the norm, that is, something natural, or accord-

ing to the thesis advanced here, ‘unmarked’. For example, she describes how “in English, the lateral tongue contraction [described for English coronals] gives to those unaccustomed to this setting the impression that the tongue is somewhat tensed, but the Englishman is not aware of any tension and feels the tongue to be relaxed” (Honikman, 1964).

The speaker/listener experience is a critical element and the dimension through which phonological structure and phonetic processes can interact. Biases in the way a language user perceives, processes and produces sounds reinforce the very structures that also give rise to these biases. Biases are both the result of and the promoter of familiar routines, and therefore self-replicating, aiding ease and speed of both production and perception, and specifying the mapping between input and output. During the process of acquisition, the developing child not only learns mastery of phonetic ability and the abstract system of contrasts of her language, but critically also learns how to map the one onto the other. An infant will at an early age acquire the meaningful contrast between voiced and voiceless stops, for example – the cornerstone for recognizing words, and indeed for building her own lexicon. She will also learn that this contrast is made in a particular way in her language, with particular co-ordination and timing of articulators resulting in a particular set of acoustic properties (for example more or less voicing during the oral articulation, presence or absence of pre- or post-aspiration, duration of the preceding vowel). She will learn how and where this varies. As her perception becomes increasingly attuned to those properties, she learns the boundaries of what is ‘unmarked’ for her language. A voiced stop articulated in a non-native manner will be an outlier, will diminish perceptual coherence and ease, and can thus be considered marked. Similarly, as she begins to tune her own productions, she lays down routines and reinforces neurological pathways, and these become ‘the unmarked’ behaviours for her.

The biases specify the mapping between abstract structures and the phonetic forms, or devices, used to implement them. Mappings may not be one-to-one: just as a structure may be implemented using a range of phonetic devices, a given phonetic device may be a ‘multiple signifier’, mapping onto more than one structure. Thus, while there may be superficial similarities between languages in the phonological distinctions made, and in the phonetic devices employed to mark these, the ecology of available devices and the different demands placed on these will vary. Payne and colleagues (2015, 2016, forthcoming) examine relative VC timing in both English and Norwegian, and find the different degree of competition for this phonetic device (signalling coda obstruent voicing in both languages, but additionally

signalling syllable weight in Norwegian) results in a distinctly different set of mappings. These mappings represent prosodic-phonetic biases that are subtly but significantly different between the two languages that affect the acquisition pathways taken. It is argued here that these subtly differing biases are at the heart of how language-specific (un)markedness arises. With continual microvariation in performance, they also provide a mechanism through which the ‘unmarked’ may shift, as the precise centres of perceptual and production biases very gradually change.

The biases may also become a source of phonological innovation as reinforced behaviours are mapped to innovative abstract structures. Hence while the reinforcement of a pre-existing (phonetic) form may in itself appear conservative, its association with a new or different function (a ‘re-mapping’) is creative. Thus a localized bias in the implementation of a particular structure may become more pervasive in the language, and have consequences elsewhere in the system. In effect this leads to reification and spread of a particular phonetic device, a phenomenon Payne (2006) calls a phonetic *motif* and illustrates with the example of consonant gemination in Italian. In brief, multiple sources of geminate consonants can be identified in (standard) Italian, the combined presence of which raises the frequency (and therefore also the ‘unmarkedness’) of geminates. The three main lexical sources are: i) the inheritance of pre-existing structure in the form of Latin geminates (e.g. BUCCAM > *bocca*); ii) a historical process of cluster assimilation (e.g. SCRIPTUM > *scritto*); and iii) a historical process of prosodic (post-tonic) lengthening (e.g. FEMINA > *femmina*). Additionally, central and southern varieties of Italian also have a process of post-lexical gemination (*raddoppiamento sintattico*), the sources of which mirror the conditions for novel (non-Latin) lexical geminates (i.e. the assimilation of a (Latin) word-final consonant to a word-initial consonant across a word boundary, e.g. AD > a; a [n:]otte, ‘at night’; and the lengthening of a word-initial consonant after a final, open stressed syllable, e.g. menu [f:]isso, ‘fixed menu’).

One could ascribe a ‘structure preservation’ interpretation to this expansion to new lexical items, and even to the novel, post-lexical context. In other words, this could merely be a process of analogical extension within the abstract domain. However, elsewhere, Payne (2005) identifies a pervasive phonetic behaviour that appears to underpin the prosodically-conditioned geminates. She finds *all* consonants, singletons and geminates, to be longer in post-tonic position. The reinforced phonetic bias (post-tonic lengthening) and the conformity of the structural outcome (with pre-existing geminates) suggest a confluence of compatibility leading to a structural innovation. Furthermore, the

prosodic-phonetic bias is functionally grounded in that an increase in duration assigns a greater proportion of the acoustic signal, and thus greater acoustic prominence, to what is prosodically prominent. Crucially, the *detail* of this mapping is *language-specific*. Payne and Eftychiou (2006) find *pre*-tonic lengthening of consonants in Cypriot Greek, a language that also has lexical and post-lexical geminates. Interestingly, Cypriot Greek also differs from (standard) Italian in that it allows lexical *word-initial* geminates, such as [p:ɛfti] ‘falls’ and [pɛfti] ‘Thursday’. By default, word-initial *lexical* geminates cannot ever be post-tonic, but can be – and often are – pre-tonic. Thus there is a similar, *but different*, harmony between prosodic-phonetic biases and structure. The biases use a similar phonetic device, but the direction of its alignment with prosodic prominence is language-specific. A further curious typological difference in the implementation of gemination between Italian and Cypriot Greek is the relative position of laryngeal adjustments. Cypriot Greek voiceless geminate stops are post-aspirated (Armosti, 2011) whereas evidence for Sienese Italian (closely related to standard Italian) reveals voiceless geminate stops to be *pre*-aspirated (Stevens, 2011).

These patterns of behaviour show a degree of formal cohesion that is pervasive and language-specific, and that closely interacts with and may influence existing structures. It is through the iterative processes of reinforcement between structure and behaviour that non-universal narratives of markedness arise. In this framework, the notion of markedness is recast as the set of language-specific structures and prosodic-phonetic biases that are induced by learning, but are ultimately constrained both by the structures inherited and by external, universal phonetic constraints. This interpretation weaves a path between extreme phonological determinism on the one hand and extreme phonetic determinism on the other, by preserving the mental status of markedness knowledge but rejecting the universal nature of markedness and the encoding of universal external forces, and by elaborating how phonetic behaviour may permeate structure in a non-random way. Within this framework, the ‘value’ of particular (preferred) abstract phonetic forms in a particular language goes beyond just its statistical frequency. Biases in structure are seen to be reflected in behaviour (production and perception) but neither are these simply reducible to behaviour, or to be codified by abstract natural (and innate) laws.

References

- Archangeli, D.B., & Pulleyblank, D.G. (1994). *Grounded phonology*. Cambridge, MA: MIT Press.
- Armosti, S. (2011). *The phonetics of plosive and affricate gemination in Cypriot Greek* (PhD thesis). University of Cambridge.
- Becker, M., & K.F. Potts (2011). The emergence of the unmarked. In M. van Oostendorp, C. J. Ewen, E. Hume, & K. Rice (Eds.), *The Blackwell companion to phonology. Volume 3: Phonological processes* (pp. 1363–1379). Chichester: Wiley-Blackwell.
- Bermudez-Otero, R., & Börjars, K. (2006). Markedness in phonology and syntax: The problem of grounding. *Lingua*, 116, 710–756.
- Blevins, J. (2004). *Evolutionary phonology*. Cambridge: Cambridge University Press.
- Bybee, J. (2001). *Phonology and language use*. Cambridge: Cambridge University Press.
- Chen, M. (1973). On the formal expression of natural rules in phonology. *Journal of Linguistics*, 9, 223–249.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.
- Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York: Harper & Row.
- Flack, K. (2007). *The sources of phonological markedness* (PhD thesis). University of Massachusetts at Amherst.
- Gibbon, F., Hardcastle, W.J., & Nikolaidis, K. (1993). Temporal and spatial coordination in /kl/ sequences: A cross-linguistic investigation. *Language and Speech*, 36, 261–277.
- Greenberg, J. (1966). *Language universals, with special reference to feature hierarchies*. The Hague: Mouton.
- Hale, M., & Reiss, C. (2000). Substance abuse and dysfunctionality: Current trends in phonology. *Linguistic Inquiry*, 31, 157–169.
- Hale, M. & Reiss, C. (2001). Phonology as cognition. In N. Burton-Roberts, P. Carr, & G. Docherty (Eds.), *Phonological knowledge: Conceptual and empirical issues* (pp. 161–184). Oxford: Oxford University Press.
- Haspelmath, M. (2006). Against markedness (and what to replace it with). *Journal of Linguistics*, 42, 25–70.
- Hayes, B., & Steriade, D. (2004). Introduction: The phonetic bases of phonological markedness. In B. Hayes, R. Kirchner, & D. Steriade (Eds.), *Phonetically based phonology* (pp. 1–33). Cambridge: Cambridge University Press.

- Hayes, B., & Wilson, C. (2008). A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry*, 39, 379–440.
- Honikman, B. (1964). Articulatory settings. In D. Abercrombie, D.B. Fry, P.A.D. MacCarthy, N.C. Scott, & J.L.M. Trim (Eds.), *In honour of Daniel Jones* (pp. 73–84). London: Longman.
- Hume, E. (2004). Deconstructing markedness: A predictability-based approach. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*, 13, 182–198.
- Hurch, B., & Rhodes, R.A. (1996). *Natural Phonology: The state of the art*. Berlin: Mouton de Gruyter.
- Jakobson, R. (1962). *Selected writings. Volume I: Phonological studies*. The Hague: Mouton.
- Kean, M. (1992). Markedness: An overview. In W. Bright (Ed.) *International Encyclopedia of Linguistics* (vol. 2) (pp. 390–391). New York: Oxford University Press.
- Kenstowicz, M., & Kisseberth, C. (1979). *Generative phonology*. San Diego: Academic Press.
- Kiparsky, P. (2006). Amphichronic linguistics vs. Evolutionary Phonology. *Theoretical Linguistics*, 32, 217–236.
- Liljencrants, J., & Lindblom, B. (1972). Numerical simulation of vowel systems: The role of perceptual contrast. *Language*, 48, 839–862.
- Lindblom, B. (1983). Economy of speech gestures. In P.F. MacNeilage (Ed.), *The production of speech* (pp. 217–245). New York: Springer Verlag.
- MacNeilage, P.F., & Davis, B.L. (2000). On the origin of internal structure of word forms. *Science*, 288, 527–531.
- McCarthy, J.J., & Prince, A. (1994). The emergence of the unmarked: Optimality in prosodic morphology. *Proceedings of the North Eastern Linguistic Society*, 24, 333–379.
- Ohala, J.J. (1974). Phonetic explanation in phonology. *Papers from the parasession on Natural Phonology, Chicago Linguistic Society*, 10, 251–74.
- Ohala, J.J. (1978). Production of tone. In V. Fromkin (Ed.), *Tone: A linguistic survey* (pp. 5–39). New York: Academic Press.
- Ohala, J.J. (1981). The listener as a source of sound change. *Papers from the parasession on language and behavior, Chicago Linguistic Society*, 17, 178–203.
- Ohala, J.J. (1983). The origin of sound patterns in vocal tract constraints. In P.F. MacNeilage (Ed.), *The production of speech* (pp. 189–216). New York: Springer.

- Payne, E. (2005). Phonetic variation in Italian consonant gemination. *Journal of the International Phonetic Association*, 35(2), 153–189.
- Payne, E. (2006). Phonetic motifs and the formation of sound structure. *UCL Working Papers in Linguistics*, 18, 321–343.
- Payne, E. (2009). [Review of] J. Blevins (2004) “Evolutionary Phonology: The emergence of sound patterns”, CUP. *Journal of the International Phonetic Association*, 39(1), 87–94.
- Payne, E., & Eftychiou, E. (2006). Prosodic shaping of consonant gemination in Cypriot Greek. *Phonetica*, 63, 175–198.
- Payne, E., Post, B., Garmann, N.G., & Simonsen, H.G. (2015). VC timing acquisition: Integrating phonetics and phonology. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, UK: the University of Glasgow. Retrieved from <http://www.icphs2015.info/pdfs/Papers/ICPHS0764.pdf>
- Payne, E., Post, B., Garmann, N.G., & Simonsen, H.G. (2016, March). *Competition for phonetic devices: A cross-linguistic investigation of VC timing*. Poster presented at BAAP 2016, Lancaster.
- Payne, E., Post, B., Garmann, N.G., & Simonsen, H.G. (forthcoming). The acquisition of long consonants in Norwegian. In H. Kubozono (Ed.), *The phonetics and phonology of geminates*. Oxford: Oxford University Press.
- Schachter, P. (1969). Natural assimilation rules in Akan. *International Journal of American Linguistics*, 35, 342–355.
- Simonsen, H.G., Garmann, N.G., Payne, E., Post, B., Holm, E., & Hansen, P. (2016, February). *Cross-linguistic patterns in children’s productions of consonant clusters*. Paper presented at The XIV Nordic Symposium on Child Language, Oslo.
- Smith, C. (1997). The devoicing of /z/ in American English: Effects of local and prosodic context. *Journal of Phonetics*, 25, 471–500.
- Stampe, D., & Donegan, P. (1980). The study of natural phonology. In D.A. Dinnsen (Ed.), *Current approaches to phonological theory* (pp. 126–173). Bloomington: Indiana University Press.
- Stevens, M. (2011). Consonant length in Italian: Gemination, degemination and preaspiration. In S.M. Alvord (Ed.), *Selected proceedings of the 5th conference on laboratory approaches to Romance phonology* (pp. 21–32). Somerville, MA, USA: Cascadilla Proceedings Project.
- Tesar, B., & Smolensky, P. (2000). *Learnability in Optimality Theory*. Cambridge, MA: The MIT Press.
- Trubetzkoy, N. (1931). Die phonologischen Systeme. *Travaux du Cercle Linguistique de Prague*, 4, 96–116.

- Trubetzkoy, N. (1939). *Grundzüge der Phonologie*. Göttingen: Vandenhoeck & Ruprecht.
- Westbury, J., & Keating, P. (1986). On the naturalness of stop consonant voicing. *Journal of Linguistics*, 22, 14–166.