The construction of markedness: narratives of phonologisation

Elinor Payne
Faculty of Linguistics, Philology and Phonetics
University of Oxford
Some questions...

- What is ‘markedness’?
- Where does it come from?
- (Why) do we need it?
- Some views
- An alternative view
Caution: work in progress!

Unapologetically at the conceptual stage...

Part of a wider set of research questions:

• What is the nature of phonological knowledge?

• What determines phonological structure?

• Is phonology to be seen in any way as autonomous from phonetic substance?
A little background on ‘markedness’

- (Haspelmath, 2006): despite being embraced by virtually every different theory— and perhaps because of this—the concept has become “an almost theory-neutral, everyday term in linguistics”

- **Original concept (Prague School)**
  - Abstract notion of complexity and contrast. Specification for phonological distinction; oppositions, relations between pairs of phonemes;
  - Phonological features come in pairs, which are polar opposites, and pattern in a nonequivalent way: mark-bearing (voiced, nasalised, rounded) vs markless (voiceless, non-nasalised, unrounded)
  - Tells us about the intrinsic content of features and their representation in our minds. Influences phonological activity (thus, outcome of neutralisation is the unmarked form)

- **Universalist perspective** (Greenberg, *Language Universals* 1966)
  - Identification of cross-linguistic patterns (universals)

- **More recent and contemporary approaches:**
  - 2) not a concept to be developed in any interesting way from a phonological perspective (Hale & Reiss 2000, 2001, Hume 2003, Haspelmath, 2006; Blevins, 2004)
Some diagnostics of markedness

- Frequency of use; abnormality / rarity (in a language / in the world)

- “Complexity” (e.g. unaspirated stop / aspirated stop); “difficulty” (e.g. b < d < g < G) ???

- Early vs late child acquisition (related to difficulty; cf Jakobson)

- ‘Taken for granted, ordinary, unusual’

- Appearance in neutralised context (e.g. voiceless/voiced)

- Restricted distribution: e.g. in German voiced obstruents cannot appear in syllable coda (Dixon’s functional markedness)

- Stability (in language change)
‘Local markedness’

• Observation that markedness is not an absolute property, but often relative to a given context (cf Greenberg)

• E.g. ‘voicing’ is ‘marked’ for obstruents but unmarked for sonorants
What is markedness?

• Does not mean ‘impossibility’

• Can be seen as a form of ‘inhibition’, a bias against
  – Compare unmarked forms, for which there is a preference

• What is the origin of this bias?
Markedness in grammar

- Internal encoding of external properties (e.g. frequency and phonetic ‘naturalness’ / ‘simplicity’)

- In SPE (Chomsky & Halle, 1968) markedness values were a technical device to capture the relative ‘naturalness’ of phonological structures (‘naturalness’ synonymous with cross-linguistic frequency)

- In Principles & Parameters: posits that parameter settings are not equal. Chomsky (1981: 8), 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’

- Drives acquisition: ‘the unmarked case of any parameter represents the initial hypothesis that children make about the language to be acquired’ (Kean, 1992)

- In OT, markedness constraints evaluate an output according to formal preferences, which in the main correspond to phonetic difficulty (e.g. that a nasal stop must not be followed by a voiceless obstruent)

- For Kager (1999: 11), markedness constraints in OT are ‘validated’ by cross-linguistic studies, in combination with ‘phonetic grounding’
How do we experience markedness?

- **Part of mental grammar:**
  - Language-specific. Implicit in early Prague School accounts; (no longer fashionable)
  - Universal: part of innate cognitive code for language (UG) (Chomsky & Halle onwards, including most accounts of OT)

- **Externally, as linguists! A meta-grammatical concept:**
  - Greenberg made no explicit claim about mental grammars. (see also Wurzel, 1998)
  - Markedness as a cross-linguistic observable outcome

- **General human experience:**
  - Moravcsik & Wirth (1986) claim three main markedness properties (familiarity, diversity, and simplicity) found throughout human culture, e.g. in food: everyday food is more frequent, comes in more different kinds, and is simpler than holiday food
  - Markedness in language just an instantiation of markedness in this more general sense
  - The idea that markedness is a highly general property of human culture is already found in Jakobson

- **Alternative view**
  - As part of mental ‘grammar’ (after a fashion) but one that is acquired and refined through usage and is language-specific, even while showing cross-linguistic patterns. Statistical regularities of language use intimately connected with language structure
What do we need markedness for?

• **As speakers / listeners of a language**
  – Part of architecture of grammar / our linguistic knowledge?
  – Biases the way we acquire, perceive, process and produce sounds
  – Aids acquisition?
  – Universal? / Language specific?

• **As linguists**
  – Typology: to explain common / uncommon patterns
  – Acquisition: to understand developmental patterns
  – Diachrony: to explain sound change
  – Psycholinguistics: to explain asymmetry
Some opposing views of phonological structure

- **Exogenous**
  - ‘Markedness’ = certain patterns and distributions, which arise on account of entirely (or almost entirely) external factors (phonetic facts; frequency of use etc.)
  - Phonetic determinism

- **Endogenous**
  - Phonologically determined (completely formal and abstract)

- **Exogenous masked as endogenous** (“wanting to have your cake and eat it”)
  - Encodes external factors into the grammar
  - Revised version of phonological determinism (phonetically grounded phonology)
Phonetic determinism

• Phonetic ‘facts’ (articulatory, aerodynamic, acoustic, auditory-perceptual), explain phonological structure and sound change (cf Ohala, various; Blevins, 2004; Haspelmath, 2006)

• In its extreme form, this approach finds perfect explanation in the way we hear and speak

• However, there may be a post hoc phonetic rationale for why, e.g. final obstruent devoicing is so common cross-linguistically, but there is no natural law against final voiced obstruents, and indeed they are to be found

• Within the boundaries of what is physically possible, there are many viable pathways

• Phonetic determinism may be necessary, but it is far from sufficient
An exogenous approach: Evolutionary Phonology (Blevins, 2004)

- A thesis, not against markedness per se, but against the codification of markedness in grammar

- “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)

- “Markedness and naturalness in phonology are emergent properties of synchronic systems”
Formal approach (endogenous)

- Doesn’t get very far without reference to phonetic substance

“[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)
Revised phonological determinism

Phonetically Grounded Phonology:

— "Markedness: the tendency for phonetic terms to be pronounced in a simple, natural way (as determined in part by the nature of speech articulation, acoustics, and audition, and in part perhaps by more abstract cognitive factors – all aspects of the human language faculty)." (Anderson & Lightfoot 2002: 101)

• Phonologists have sought to encode phonetic ‘ease’ in various ways: as phonetically grounded distinctive features (SPE); naturalness (Natural Phonology, e.g. Stampe, 1971; Hurch & Rhodes 1996), ‘(dis)preferred’ (e.g. Vennemann, 1988).

• In OT, phonetic optimality emerges from satisfying highest ranked constraints

• More or less unmarked ‘really boils down to more or less easy for the human brain’ (Mayerthaler 1987: 27)

• Recently, the encoding of system-external pressures on phonological structure has been referred to as ‘grounding’ (e.g. Archangeli & Pulleyblank 1994, Bermudez-Otero & Borjars 2006)

• Even while markedness ‘transparently reflects’, ‘is based on’, or ‘is rooted in’ (speakers’ knowledge of) phonetic difficulty, it is still part of the phonological system (e.g. Hayes & Steriade, 2004)
Phone, c
difficulty
>
grammar

Case-study: encoding of voicing aerodynamics

- Voiced geminate obstruents are ‘marked’ (rare)

- No (known) language bans just voiceless geminates

- Presence of a voiced obstruent geminate in a given language implies that of the corresponding voiceless geminate

- Phonetic factors:
  - Duration of oral closure
  - Size of cavity behind oral constriction

- Both lead to a scale of difficulty in maintaining voicing \( [g: < d: < b: < g < d < b] \)

- Become enshrined as a property of grammar
Language-specific phonetic factors

• “There is some evidence that languages indeed deploy phonological constraints based on the conditions set up by language-specific phonetic factors” (Hayes and Steriade, 2004: 20)

• Standard Thai, CVR syllables have richer tone-bearing possibilities than CV:O. The latter cannot host LH or M tones. Navajo is almost the opposite: CV:O can host any phonemic tone, but CVR cannot host HL or LH.

• All other things being equal, V is a better host for tone that R. But, at equal sonority levels, the longer sonorous rhyme is the better carrier. So it all depends on phonetic duration. In Navajo, CVR and the V: portion of CV:O are very close in duration, and since V: is more sonorous than VR, CV:O is the better host. In Thai, long vowels are dramatically shorter in closed syllables, resulting in CV:O having a considerably less sonorous rhyme than CVR.

• Language-specific difference of allophonic detail – degree of shortening in closed syllables – is apparently the source of a major phonological difference.
Extreme phonetic determinism masking as phonological

• West Coast Optimality Theory (Kirchner 1997; Flemming, 1995):
  – Phonemes not viewed as underlying entities, but instead emerge epiphenomenally from the interaction of phonetically grounded marking constraints

  – Kirchner (1997) proposes a particularly radical use of scales, with continuously valued phonetic functions (e.g. degree of articulatory laziness) figuring directly in the OT grammar
Some problems

• You can make endogenous accounts sufficient, if you go far enough... But is it **necessary**?

• Does an individual need to have knowledge of cross-linguistic patterns and distributions, even when they run counter to her own language? Isn’t it more likely that she knows what is ‘normal’ for her language, and what isn’t?

• For example: the existence of final voiced obstruents is a problem for cognitive determinism because they violate abstract markedness.

• OT gets round this by ranking constraints and then allowing low-ranking constraints to be violated

A neat fix?

• Maybe, but 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’

• For innatist versions of the 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)”
Why look for phonetic explanation?

• Discover what linguistic systems are not responsible for (Anderson, 1981): “to isolate the core of features whose arbitrariness from other points of view makes them a secure basis for assessing properties of the language faculty itself” (1981: 497). In other words, we get a better understanding of phonology by knowing what is not in phonology.

• But do we have to encode this grammatically?

• Hayes and Steriade argue that our very views of phonology change. Cue-based theory is non-arbitrary (but also teleological).
Against reification of explanation

• Haspelmath (2006) argues that we do not need markedness for:

1. Description of particular grammars
   • markedness does not lead to greater elegance of description

2. Description of UG
   • McCarthy (2002: 2): “One of the most compelling features of OT, is the way it unites description of individual languages with explanation in language typology... the grammar of one language inevitably incorporates claims about the grammars of all languages.”
   • ‘markedness’ phenomena are ultimately due to substantive factors

3. As a meta-grammatical label
   • He claims this leads to confusion, since it is not a transparent label

4. As an explanatory concept
   • Ultimate explanation is in terms of substantive factors outside the language system. If one posits an abstract intermediate ‘explanatory’ level between the phenomena and the real explanatory factors, one bears the burden of proof that such a level is needed
Why should we want to reify?

• Is there still a case for markedness, as an abstract notion?

• Haspelmath says we do not need it for the description of individual grammars, but perhaps it tells us something about their functioning
  – Language-specific biases against and towards certain phonetic structures
  – Voiced geminate obstruents may be marked for most languages, but in those languages which have them, they are ‘natural’

• Haspelmath also argues that explanation in phonology is to be found entirely in external factors. However, phonetic determinism can only tell us about the limits and probabilities of sound systems.

• Many factors may mitigate against structures appearing in a given ‘sound’ system
  – ‘Physical’ limits
    • Articulatory / motor / mechanisms for planning execution of utterance
    • Acoustic and aerodynamic
    • Auditory / perceptual
    • Neurological capacities and demands
    • Mechanisms for accessing lexicon in production and perception
  – Pre-existing sound structure
  – Principles of form and self-organisation
Pre-existing structure: missing piece in jigsaw?

- Phonological structure also dependent on a) the material you feed in and b) the internal mechanisms of self-organisation

- In a trivial sense: if you have stops in the system and you have a process of lenition, you are likely to get fricatives, approximants etc.

- In a non-trivial sense (I): pre-existing structure can make certain pathways/strategies more likely

- In a non-trivial sense (II): pre-existing structure may present creative possibilities (cf ‘spandrels’)

Phonetic motifs (Payne, 2006; under revision)

• Low-level systematic patterns in speech behaviour: e.g.
  – lip-rounding of ESH in English;
  – greater coarticulatory vowel nasalisation in American English compared with British English;
  – velarised resonance throughout vowel paradigm in Brazilian Portuguese;
  – language-specific variation in coarticulation strategies... (overlap in /kl/ cluster much greater in Catalan than in Swedish)
  – Articulatory setting (Honkiman, 1964)

• Linguistic and non-universal: Though shaped to some degree by physics or function of speech, critically detached from ‘natural’ deterministic factors

• Source of phonological innovation – motifs can become structurally anchored in a language, facilitated by pre-existing patterns and general structural biases

• Re-incorporation (phonologisation) of phonetic substance

• Nuts and bolts of phonetic substance are universal ‘facts’ of speech production and perception, but these are filtered and harnessed in different ways cross-linguistically

• Permeability between physical, phonetic world and more abstract, grammaticalised structures
Case study: long consonants in Italian

- Multiple sources:
  - Pre-existing structure: Latin geminates
  - Cluster assimilation (lexical and post-lexical)
  - Prosodic lengthening (lexical and post-lexical)
    - Duration increase is often associated with greater prominence (there are good perceptual reasons for this)
    - Creation of post-lexical doubling, triggered by final stress, in certain varieties of Italian, not so surprising
    - Payne (2005) found compatible phonetic effects of lengthening in post-tonic position
  - Prosodically conditioned consonant lengthening exists as a phonetic motif in Pisan Italian
  - Linguistic phonetic. Not a mechanical inevitability: Payne & Eftychiou (2006) found different systematic lengthening effects in Cypriot Greek (pre-tonic lengthening)
  - Structure harmony – priming effects?
Other phonetic strategies ‘harnessed’ to cue gemination

Differences in resonance (reflecting gestural differences)
Laryngeal timing motifs?

Pre-aspiration (Sienese Italian): Stevens, 2009
Post-aspiration (Cypriot Greek voiceless stops): Armosti, 2010

Figure 2.3: Waveforms and spectrograms of four disyllabic words.
[Note: In each of the four panels, the consonants are segmented and highlighted. The words shown in each panel are (a) keříč, (b) kʰɛ̌ːɾíč, (c) teři, and (d) tʰɛ̌ːɾíčes.]
‘Phonologisation narratives’

- **Tuscan Italian**
  - Evidence of pre-aspiration (Sienese, cf Stevens 2009)
  - Post-tonic lengthening (Pisan, phonetic and phonological, Payne, 2005)
  - VC compensatory lengthening mechanism
  - Gemination only possible word-internally (ambisyllabic)

- **Cypriot Greek**
  - Post-aspiration in voiceless geminate stops (Armostis, 2010)
  - Pre-tonic lengthening (phonetic; Payne & Eftychiou, 2006)
  - CV compensatory lengthening mechanism (Armostis, 2010)
  - Gemination also possible word-initially
  - Claim that geminates are tautosyllabic and moraic (i.e. onset moraicity)
Hayes & Steriade (2004) argue that the ‘innocent misapprehensions’ (i.e. Blevins’ CCC account) cannot account for apparent directional stability in sound change. Instead would lead to random drift. But this underestimates constraining influence of pre-existing structure.

Alternative approach: not tightly deterministic, but allows for a more constrained model than Evolutionary Phonology.

Variation in substance may be exogenous to language, but the mechanisms by which new forms are harnessed into the system are endogenously determined. System-conforming, but in a non-trivial and creative way.

Idea that languages not only avoid certain things (too much phonetic complexity), they also make use of things, they exploit structural possibilities, reinforce patterns (motifs).

Phonological structures are emergent. This isn’t a new idea, cf Kiparsky on diachronic conspiracies:

“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)

But in addition, there is cross-referencing over different levels of abstraction and significance.

The mechanisms by which they do so remain to be explored, but could fit into a modified, enriched exemplar framework
   – Priming effects.
Setting within an exemplar framework

• More developed for speech perception than speech production

• However, potentially very useful also for production (see Pierrehumbert, 2001)

• Weighting of exemplars, greater activation

• Think about why certain exemplars would become more activated?

• More instances (familiarity)

• But also more productive throughout the system? Would lead to contrast being used throughout consonant system

• This stretches the exemplar model beyond a simple mapping of phonetic distribution. Associations at different levels of abstraction may conspire to promote certain productions
Hale criticises Kiparsky’s account of structural preservation precisely on these grounds:

- “Changes such as “phonologisation” are not dependent upon existing representations (which the child cannot directly access), but rather represent solutions to that challenge which differ from those opted for by previous generations.”

But in the model I propose there can be a **mismatch between phonetic form and the function with which it is associated**

- For example, the phonetic form of a long consonant might be increasingly pervasive in a language, but attached to different functions (phonemic, prosodic...)

- Psycholinguistically, we might test this through priming effects – is the perception of long consonants facilitated in those languages that already have them?

- Conversely, a ‘function’ might become attached to different phonetic forms. In Sienese Italian, Stevens (2009) also reports a more general, breathy laryngeal setting. Possible interpretation is that ‘glottal abduction’, as a generalised phonetic motif, is gaining ground, but is being harnessed, structurally, to cue geminate contrast

- The use of a pre-existing ‘form’ may be conservative, but the association with ‘function’ is creative / innovative
Where does this leave markedness?

- Language-specific structural biases reflected in behavior (production and perception)
- Set within universal ‘natural’ limits (phonetic, functional, etc. – true universals that do not need to be cognitively encoded)
- Is markedness the wrong way of looking at things? a negative expression
- Biases create structural possibilities, which may be ‘exapted’ (cf Lass) for different functions
- The ‘value’ of particular (preferred) phonetic forms in a particular language goes beyond just its statistical frequency
- In a non-teleological way, contribute to mechanism of phonologisation
- Markedness is a narrative of natural history (cf Evolutionary Phonology), but one that continues to ‘resonate’ in (language-specific) grammar
Evidence of language-specific low-level phonetic / prosodic biases is proliferating, and feeding theory development on e.g. perception (cf Hawkins)

A challenge for the framework outlined here is to investigate these biases in a way which integrates perception and production. Try to match asymmetries in production with asymmetries in perception. In other words, do phonetic motifs have a psycholinguistic reality?

White, Payne & Mattys (2009) found that Southern and Northern Italian differed according to marking of prosodic boundaries, with very strong prosodic timing effects in Sicilian Italian (may contribute to percept of rhythm)

Next step is to see whether native speakers of these varieties have different perceptual capabilities with regard to judging durational differences
Conclusions

- Markedness is a description of biases in structure that are reflected in behaviour (production and perception) – *but are not simply reducible to behavior*

- Biases may be pervasive at different structural levels (leading to phonologisation)

- Avoids both phonological and phonetic determinism. There is no unitary theory

- Non-universal, natural histories of sound systems. Vignettes, pieced together post hoc

- Does not mean that sound systems are random or unconstrained

- Constrained by: physical factors, but also pre-existing structure, and laws of form

- Phonology as both internal (mental) and external (permable to phonetic substance). Compare:
  
  - For Optimality theorists, phonology is ‘big’ and substance-full, and intriniscally part of formal linguistics (and UG) INTERNAL
  
  - Others, (e.g. Carr, Burton-Roberts) consider phonology to be ‘big’ and substance-full but not formal, not part of formal linguistics EXTERNAL?

  - Blevins considers phonology to be minimal, and substance-free; sound patterns are phonetically and diachronically determined EXTERNAL
Thank you!
Parallels with Evolutionary Biology

- Mapping: lack of isomorphism between phonetics and phonology / between genotype and phenotype
- Phonological determinism / genetic determinism. Phonetic substance in phonology / sociobiologists
- Phonetic determinism / extreme ‘environmentalism’. Evolutionary Phonology / neo-Darwinian ultra-selectionists
- Generative ‘entrenchment’
- Wimsatt (1986, 1991)
- percolation of structural properties
- no absolutely or purely innate or absolutely or purely acquired traits
- Analogue to innateness is generative entrenchment
- Differentiation proceeds from the general to the particular."
- The greater conservatism of features at earlier developmental stages implies that, on the average, features which are expressed earlier in development are, probabilistically speaking older and most likely to be more widely taxonomically distributed than features which are expressed later in development.
- RS depends on their also being phonemic geminates, but not vice versa. Deletion of geminate contrast would be more ‘catastrophic’ than deletion of RS
- Different degrees of entrenchment – but also of ‘productivity’
- Robustness: Permits hidden genetic variation to accumulate, and possibly to serve as a source of new adaptations and evolutionary innovations (Kitano, 2004)
- Co-extension
acquisition

- early input through sensory channels has an ambiguous status
- Its deprivation often appears not just to deprive the organism of some information which it has to learn later, but of a **capacity** for acquiring experience through that sensory channel
- Loss of capacity rather than loss of information
- Early experience may be required for the development of this capacity, but as such, it performs a function more like **food** than like **information**
- it is quite clear that information acquired from the environment **can** have a profound effect if it is deeply generatively entrenched relative to subsequent behavior, and on this analysis, if it is so generatively entrenched, it is "innate". This explains the ambiguous role of early experience discussed in item (6) of section 2 above. **If the early experience which is withheld in a deprivation experiment has a generative role with respect to a wide range of subsequent experience in that sensory modality, its loss will produce such far-reaching consequences that it would readily be described as a loss of capacity.**
Minimal requirements of ‘innateness’

• (1) The acquisition of that kind of information at that stage of development is deeply **generatively entrenched** with respect to subsequent behavior.
• (2) The developmental program is **designed** to receive information of that sort at that stage of development.
• (3) The information must be of a relatively **specific** sort.
• (4) The environment of the developing organism is a **reliable source** of the required information at that stage of development.
levels of internal constraints

• Multiple
• Evolved phenotypes ‘carry information about’ the ecologies in which they evolved (Dennett), but also carry information about the *internal* organisation of the creatures that have them (e.g. genotypic and ontogenetic structures)
• Endogenous effects
• Internal constraints and dynamics filter what selection can act upon and to what extent it can do so
Cascade effects

• Selection of one trait may entail the selection of several traits
• Initial ‘selection’ obscured
• Side effects / spandrels
• Same phenotype may be the result of quite different genes or gene complexes (convergence)
• Different phenotypes may be the result of the same genes or gene complexes (differential gene regulation)
• In language, same phonology can be expressed with different phonetic means, and different phonology can lead to same phonetic means
• Different sources of gemination (prosodic lengthening; assimilation)
• Different manifestations of gemination (pre-aspiration; post-aspiration; duration; different resonance)
Structural “conflict”

• What does it mean to say that a phonological system has inherent conflict, and that such conflict triggers change?
• Or that sound change has an inherent direction?
• At any one time, the users of a language are blissfully unaware of these conflicts and of any direction. They do not direct their productions towards a specific phonological goal.
• And even if there were a goal towards which sound change were directed, that would mean that, once arrived at that goal, there would be no more change. There is no evidence that this will ever happen.