

University of California, San Diego

From the SelectedWorks of Eric Baković

January 2012

Phonology and phonological theory

Contact
Author

Start Your Own
SelectedWorks

Notify Me
of New Work



Available at: <http://works.bepress.com/ebakovic/48>

Phonology and phonological theory

Eric Baković

UC San Diego

Phonology and phonological theory

The purpose of this chapter is to provide what I hope is a useful overview of some of the basic assumptions and results of phonology and phonological theory.¹ Phonology is the study of the sound patterns of natural human languages (henceforth just ‘languages’). The workings of a language’s sound patterns are generally assumed to be within the purview of a *phonological component* of the grammar of that language. Theories of phonology aim to adequately describe — and hope to explain — the structure of the phonological components of languages via the analysis of their sound patterns.

My focus here is on two main goals of phonological description and analysis. The first of these is the establishment of generalizations about which members of a set of posited phonological constituents are irreducibly basic and which are derived. The description and analysis of these generalizations implicates two main sets of theoretical assumptions: *representational assumptions* about what phonological constituents are and what they consist of, and *analytical assumptions* about the kinds of evidence that are brought to bear on the question of the basic vs. derived nature of a constituent.

The second main goal is the establishment of generalizations about the contexts in which phonological constituents are found and those in which they are not. In addition to the two main sets of theoretical assumptions noted above, the description and analysis of these generalizations typically implicates two other main sets of theoretical assumptions: *computational assumptions* about the mechanisms that relate representations of phonological constituents to each other, and *architectural assumptions* about how the phonological component interfaces with other grammatical components.

This chapter is largely organized by the four main sets of assumptions just mentioned.

The details of all of these sets of assumptions vary widely; I make no effort to be comprehensive in my coverage, choosing instead to aim for generality and to focus (when I do) on those assumptions I take to be more widespread among phonologists. After some very general remarks in about the goals of phonological research, I begin with discussion of some basic representational assumptions to which I will make liberal reference in the sections that follow. This is followed in by discussion of some basic analytical assumptions. Then I discuss some of the consequences of different computational assumptions, and I sketch a few alternative grammatical architectures that are generally assumed. Finally, I conclude with some summary remarks.

Goals of phonological research

Phonological research — and as a consequence, phonological theory — is driven by two explanatory goals, which I simply assume to be complementary here.² One goal is essentially cognitive: to describe and explain how the sound patterns of a language are represented and computed in the mind/brain of a speaker/hearer of that language. The other goal is essentially typological: to describe and explain how the sound patterns of all languages are fundamentally the same and how they are superficially different.

The results of the pursuit of the cognitive goal are, at least on the face of it, most in line with what is probably of most immediate interest to readers of this volume: how are the representations and computations that phonologists theorize to be in the mind/brain of a speaker/hearer involved in the processing and production of speech? But results of the pursuit of the typological goal have also formed a great deal of the foundation for cognitive theories. Phonologists have generally presupposed, at least implicitly, that observed fundamental similarities across languages directly or indirectly reflect similarities across speakers/hearers in terms of a shared genetic cognitive endowment, regardless of the specifics of what this endowment may consist of.

This presupposition in turn supports two other related assumptions that phonologists generally make: that speakers/hearers of all languages share a common vocabulary of representational constituents and computational devices, and that different speakers/hearers of the same language internalize essentially the same representations and computations. (See de Lacy 2010 for a specific probing of these and related assumptions.)

Although I am agnostic on the specifics, in what follows I generally adopt the assumptions just sketched; that typological considerations are indicative of a prior genetic endowment shared by all language users, which leads to their adoption of essentially the same (kinds of) representations and computations.

Representational assumptions

Any theory of phonology must make (and ideally, defend) assumptions about the cognitive representations of speech sounds. I discuss here a relatively conservative set of assumptions that many phonologists make about these representations.

One important assumption must be stated up front, concerning the obvious fact that there is a distinction between the cognitive representations of speech sounds and the physical properties and manifestations of the sounds themselves. I adopt here the general assumption of many phonologists that the interface between cognitive representations and physical manifestations is mediated to some significant extent by *naturalness*; that is, that cognitive representations are composed of constituents that are relatively directly motivated by and categorized according to their physical properties.

These physical properties can be divided into articulatory properties (that is, properties of speech production) and acoustic/auditory properties (that is, properties of speech perception). I focus more here on the interface between cognitive representations and articulation, but a thor-

ough investment in the naturalness assumption also requires attention to the interface between cognitive representations and acoustics/audition; cognitive representations must not only be mapped to the articulatory output produced by a speaker but must also be recovered from the acoustic/auditory input perceived by a listener.

Sustained arguments for and against naturalness can be found in Hayes et al. (2004) and Hale & Reiss (2008), respectively. Simplifying greatly here, the argument for naturalness is that the nature, frequency, and distribution of different types of phonological patterns are best explained if we assume them to be constrained by cognitive biases that directly reflect our physical limitations. The argument against naturalness, on the other hand, is essentially Occamian: that the burden of proof is on the defenders of naturalness to explain why our physical limitations are not already sufficient to explain the nature, frequency, and distribution of different types of phonological patterns.

Segmentation

Natural speech is continuous: there are typically no pauses corresponding to the spaces between written words, much less clean breaks between the speech sounds corresponding (however roughly) to written symbols. And yet a fundamental assumption of linguistic theory in general is that the cognitive representation of an utterance is categorically *segmented* at a number of levels; most significantly for phonological theory, at the level of the individual speech sound. The segmentation of speech into at least these individual speech sounds is a prerequisite to phonemic analysis, as discussed in §0.

There is some potential for disagreement about the particular segmentation of at least certain types of speech sounds; e.g., are affricates (Lombardi 1990, Clements 1999) or diphthongs (Hayes 1990, Schane 1995) one segment or two, and if one, are they simple or complex? There is

even some denial of speech sound segmentation *tout court* (Port & Leary 2005, Port 2007). I proceed here under the assumptions of the vast majority of phonologists, whereby segmentation is real and the lines are for the most part clear.

Speech sounds are themselves assumed to be complex constituents, made up of *features* that cross-classify them, as discussed in §0. Speech sounds are also assumed to be grouped into higher-order constituents such as syllables, as discussed in §0.

Distinctive features and (natural) classes

The production of an individual speech sound typically involves the coordination of various articulatory events, each contributing in some way to the overall acoustic signature of the sound. Speech sounds can be cross-classified according to these physical properties; some subset of the labels resulting from these cross-classifications are generally assumed to be part of our cognitive representations of speech sounds, typically going by the name of *distinctive features* (or often simply ‘features’). A rarely defended but practically always implicit assumption among feature theorists is that these features are innately available to language learners, but some recent work suggests that they are instead best understood to be emergent properties of learned phonological patterns (Mielke 2008; see also several of the contributions to Ridouane & Clements 2011).

One important function of incorporating distinctive features into a theory of phonological representation is to provide a (physically grounded) vocabulary for distinguishing speech sounds from one another, and in fact the ‘distinctive’ part of ‘distinctive features’ is meant to highlight this function. For example, a [p] is distinct from a [b] in that the former is a *voiceless* bilabial stop while the latter is a *voiced* bilabial stop. Both speech sounds are classified as having a bilabial place of articulation (meaning they are articulated with the lips) and as stops (meaning that

they involve a complete constriction at the specified place of articulation), but [b] involves vibration of the vocal folds (= voiced) while [p] does not (= voiceless). Especially if this physical difference between [p] and [b] serves some contrastive function in a given language — that is, if there are words the meanings of which are distinguished solely or primarily by whether one of the speech sounds is a [p] or a [b] — then this is taken to be evidence that ‘voicing’ is a distinctive feature of that language. The more contrasts in voicing, the greater the evidence: if [p] and [b] are distinct, then so might [t] and [d] or [k] and [g] be, since the members of these pairs also differ from each other solely in terms of voicing.

Another important function of distinctive features is to provide a vocabulary for describing *natural classes* of speech sounds; that is, sets of speech sounds that pattern together in the phonologies of some languages.³ Making reference to the features that define natural classes makes it possible to state generalizations about patterns that would only be less adequately described by listing the members of the sets of speech sounds involved in those patterns. For example, a language may distinguish voiced stops [b,d,g] from voiceless stops [p,t,k] before vowels but not elsewhere; reference to the natural class of ‘stops’ leads to a more adequate description of this distribution.

Theories of distinctive features may differ in terms of the precise set of natural classes that are able to be described. Some of these differences are *ad hoc*, but some are due to more principled differences in the set(s) of physical properties to which features are allowed to make reference. For example, some feature theories give more (or even exclusive) weight to the articulatory events of which speech sounds are composed (Chomsky & Halle 1968, Halle 1983, 1995, Clements & Hume 1995, Halle et al. 2000), while others give some more weight to the acoustic contributions that these (clusters of) articulatory events make (Jakobson et al. 1952, Flemming

1995, Stevens 2002, 2003).

Still other differences of this type may be due to a principled difference between the number of ‘values’ that distinctive features may have, and how computations may make reference to those values. Features may be unary (or ‘privative’), having only one value — the presence of the feature — to which reference may be made. Or, features may be binary (or ‘equipollent’), having two values (typically ‘+’ or ‘-’); or ternary, having three values (‘+’ / ‘-’ / absent); or multivalued in some further form. Phonologists typically assume some combination of these distinctive feature types, such that some features are unary (e.g. those denoting a consonant’s place of articulation; Sagey 1986, Clements & Hume 1995), others (e.g. the feature denoting a speech sound’s voicing) are binary (Wetzels & Mascaró 2001) or ternary (Inkelas 1994), and still others are multivalued (e.g. the feature denoting the relative height of the tongue body during the production of a vowel; Clements 1991, Parkinson 1996).

Theories of distinctive features may also differ in terms of the relative autonomy enjoyed by those features. Earlier, ‘segmental’ theories (Jakobson et al. 1952, Chomsky & Halle 1968) adopted an essentially classificatory approach, where each speech sound is composed of an unstructured list of its features. Later, ‘autosegmental’ theories (Goldsmith 1976, Clements 1985, Sagey 1986, McCarthy 1988, Clements & Hume 1995) granted speech sounds more structure and features more independence, for instance by allowing a single instance of a feature to span more than one speech sound.

This distinction affects the character of the computations involved in describing phonological patterns involving features. For example, a language may distinguish among several places of articulation in nasal consonants [m,n,ŋ] (= the class of consonantal speech sounds that involve a lowering of the soft palate so that air escapes through the nasal cavity) when these appear

before vowels, but these place of articulation distinctions may be neutralized before other consonants such that a nasal must have the same place of articulation as the following consonant (e.g., bilabial [m] before [p,b]). This pattern would be characterized as a *copying* of the place of articulation feature of the consonant to the nasal in segmental theories and as *spreading* of the place of articulation feature from the consonant to the nasal in autosegmental theories.⁴

One argument for the general autosegmental approach is computational simplicity: possible types of computations are limited to insertion, deletion, and spreading of features. Another argument is representational naturalness: for example, a shared place of articulation feature between a nasal and a following consonant reflects the phonetic reality that a single articulatory gesture spans the nasal+consonant cluster. This latter argument extends to a similar representation of long vowels and long (= geminate) consonants: a single speech sound ‘melody’ shared between two ‘timing units’ (McCarthy 1979, Clements & Keyser 1983, Hayes 1989, Perlmutter 1995).

Groupings of speech sounds

As some of the examples sketched above have shown, the distributions of (features of) speech sounds are influenced by the position of those speech sounds relative to other, neighboring speech sounds. Sometimes this influence is relatively direct, for instance when the place of articulation of a nasal is directly determined by the place of articulation of a following consonant. But other times this influence appears to be less direct, and appears to require reference to something other than the neighboring context. In many such cases, crucial reference is made to higher-order constituents; these constituents serve as domains of or boundaries for phonological patterns. The level of general acceptance of each of these posited higher-order constituents has depended on the degree to which they have been motivated by independent strands of converging

evidence, both within a single language and across multiple languages.

Returning to one of the examples sketched above: a distinction between voiced stops [b,d,g] and voiceless stops [p,t,k] that holds before vowels may be neutralized when the stops appear elsewhere; that is, before any other consonant (regardless of that consonant's voicing) or at the end of the word. Because the voicing of the following consonant has nothing to do with this neutralization, and because being at the end of the word also conditions it, how to characterize the neutralization in terms of the neighboring context is not obvious.⁵ Many phonologists assume that the higher-order grouping of speech sounds into syllables is responsible: to be 'before a consonant or at the end of the word' is to be at the end of a syllable (Kahn 1976, Blevins 1995).

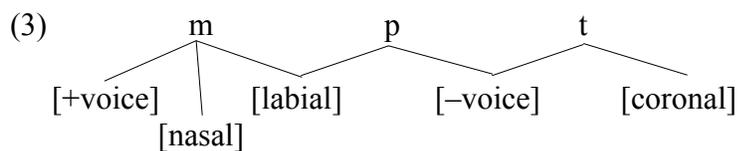
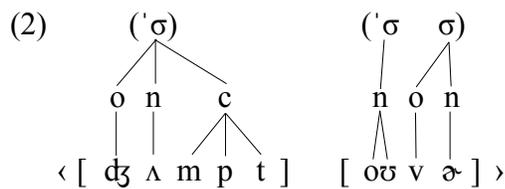
The syllable is more than a notational convenience; a syllable-based description can capture a generalization that may be more difficult to capture otherwise. For example, it is reasonable to assume that possible word-initial consonant clusters are also possible syllable-initial clusters, meaning that word-internal clusters that are also found word-initially are grouped together in the same syllable while those that are not also found word-initially are divided between syllables. This has an obvious impact on the 'before a consonant → at the end of a syllable' entailment in languages with word-initial consonant clusters: some consonants before other consonants are syllable-final while others are not. And, indeed, languages with word-initial consonant clusters that neutralize voicing distinctions as described above typically do so only syllable-finally; that is, voicing distinctions are neutralized only in clusters that are not found word-initially.

Higher-order constituents other than syllables are also generally assumed to exist. Some of these are subsyllabic constituents such as onsets (syllable-initial consonants and clusters), nuclei (vowels and vowel-like elements), and codas (syllable-final consonants and clusters) (Clem-

ents & Keyser 1983, Davis 1985, Blevins 1995), as well as the timing units that are held to be responsible for vowel and consonant length. Others are suprasyllabic; for example, pairs of syllables may be grouped into metrical feet (Lieberman & Prince 1971, Hayes 1995), primarily motivated by (and held to be responsible for) the rhythmic properties of word stress. There are also even larger constituents, such as the phonological (or ‘prosodic’) word (Selkirk 1980), the phonological phrase (Nespor & Vogel 1986), and the intonational phrase (Pierrehumbert & Beckman 1988).

The following diagrams provide a sense of how all of these representational pieces fit together. In (1) is a high-level overview of an utterance consisting of two intonational phrases (delimited by curly brackets), the first consisting of three phonological phrases (delimited by angled quotation marks), and the first of these consisting of three prosodic words (delimited by square brackets). In (2) is a close-up view of the second phonological phrase, showing the constituents of speech sounds into subsyllabic constituents (o = onset, n = nucleus, c = coda), themselves grouped into syllables (= σ), which are themselves grouped into metrical feet (delimited by parentheses). Finally, in (3) is an even closer look at some of the features that make-up of the coda cluster [mpt].

(1) *The quick brown fox jumped over the lazy dog, didn't it?*
 {<[ðə'kwɪk] ['bɹɪəʊn] ['fɑ:ks]> <['dʒʌmpt] ['oʊvə]> <[ðə'leɪzi:] ['dɑ:g]>} {<['dɪŋt ɪt]>}



Analytical assumptions

As noted in §0, I focus here on two main goals of phonological description and analysis: making generalizations about which members of a set of posited phonological constituents are basic and which are derived, and making generalizations about the contexts in which phonological constituents are found and those in which they are not. These goals are closely linked in ways that will be made clear in this section.

The focus of the discussion in this section will be on the most basic of the phonological constituents identified in §0, speech sounds and distinctive features. Indeed, higher-order constituents are typically defined primarily in terms of their utility in describing and analyzing the distributions of speech sounds and features and only secondarily in terms of their own distributions and status as basic vs. derived.

Inventories

Even a cursory description of the phonological patterns of a language will typically include a description of an *inventory* of the speech sounds used in the language. This inventory will typically be defined at two levels: the set of speech sounds used *systematically*, and the proper subset of speech sounds used *distinctively*. (There are a number of theoretical assumptions that can be tied up in the difference between the distinctive and systematic levels of phonological description and analysis; I side-step many here.)

For example, the inventory of speech sounds in English includes the bilabial stops [p^h], [p], and [b], where [p^h] indicates aspiration.⁶ All three sounds are used systematically in English; compare [p^hæn] ‘pan’, [spæn] ‘span’, and [bæn] ‘ban’. But not all three are used distinctively: the meanings of two English words may be distinguished solely or primarily by the difference between [p^h] and [b] — to wit, ‘pan’ vs. ‘ban’ above — or by the difference between [p] and [b]

— [k^hæp] ‘cap’ vs. [k^hæb] ‘cab’ — but not by the difference between [p^h] and [p]. The difference between [p^h] and [p] (that is, the aspiration feature alone) is thus not distinctive in English; it is derived rather than basic.

Of course, every actual production of a speech sound differs in some way from any other production of a speech sound, even of one that is systematically ‘the same’ speech sound. A systematic speech sound is thus an idealization; all else being ideally equal, the [p^h] of [p^hæn] ‘pan’ is systematically the same as the [p^h] of [p^hit] ‘pit’, both differ systematically from the systematically equal [p] of [spæn] ‘span’ and [spit] ‘spit’, and all of these differ systematically from the systematically equal [b] of [bæn] ‘ban’ and [bit] ‘bit’.

Languages differ in terms of one or both of these kinds of inventories. In Thai, for example, [p^h], [p], and [b] are used systematically, but unlike English, the differences between all three of these stops are also distinctive: [p^hâ:] ‘cloth’ vs. [pâ:] ‘aunt’ vs. [bâ:] ‘crazy’. In both Spanish and Tamil, [p] and [b] are used systematically, but [p^h] is not; the difference between [p] and [b] is also distinctive in Spanish ([peso] ‘weight’ vs. [beso] ‘kiss’), but not in Tamil. In Quechua, [p^h] and [p] are used systematically and distinctively ([wamp^hu] ‘opening’ vs. [wampu] ‘boat’), but [b] is not. And so on.

Distinctive speech sounds are known as *phonemes*, and the symbols representing them are typically enclosed in slashes as opposed to the square brackets reserved for systematic speech sounds (= *phones*). Thus, [b] is a phone associated with a phoneme of English that we can represent as /b/, while [p] and [p^h] are phones associated with another phoneme that we can represent as either /p/ or /p^h/ (or some voiceless bilabial stop representation unspecified for aspiration that can be unified with both).

Further analysis is typically required to determine whether two (or more) phones are as-

sociated with (= are *allophones* of) the same phoneme, and also to determine the representation of the phoneme with which those allophones are associated. The ultimate criterion is based on the assumption that the grouping of systematic phones into distinctive phonemes is psychologically real (Sapir 1949); specifically, two phones are allophones of the same phoneme in some language given behavioral evidence that native speakers of that language treat the phones as ‘the same sound’. (The most compelling kind of evidence for phonologists comes from what are called *alternations*; see §0.)

Speech sound inventories are typically not random collections; rather, they appear to be largely organized in *series* that can be defined in terms of their features (Maddieson 1984). For example, the systematic speech sound inventory of English includes a series of voiceless unaspirated stops [p,t,k], a series of voiceless aspirated stops [p^h,t^h,k^h], a series of voiced stops [b,d,g], and a series of nasals [m,n,ŋ]. Each series contains a member from a given place of articulation: bilabial [p,p^h,b,m], alveolar [t,t^h,d,n], and velar [k,k^h,g,ŋ]. Moreover, statements about the relationship between systematic and distinctive speech sound inventories typically apply in series: in English, the difference between all voiceless aspirated stops and their unaspirated counterparts is nondistinctive, while the difference between all voiced and voiceless stops is distinctive. The same considerations apply to the other languages mentioned earlier in the text: Thai distinguishes series of voiceless aspirated, voiceless unaspirated, and voiced stops; Spanish only distinguishes series of voiceless and voiced stops; Tamil uses a series of voiceless and voiced stops systematically but not distinctively; Quechua only distinguishes series of voiceless aspirated and voiceless unaspirated stops.

The generality of such statements leads many phonologists to the conclusion that the features themselves are the foci of those statements. In other words, it is not that voiced stops are

phonemes distinct from voiceless stop phonemes in English, but rather that the voicing feature *is phonemic* (= is basic) in English. Likewise, it is not that voiceless unaspirated stops are not distinct from voiceless aspirated stops in English, but rather that the aspiration feature is *not* phonemic (= is derived) in English.

Interactions

The systematic or distinctive deployment of a particular feature or speech sound can depend on a number of factors, including (but not limited to) interactions between features within the same speech sound (§0), interactions between (features of) neighboring speech sounds (§0), and interactions with higher-order constituents (§0).

Interactions between features within the same speech sound

The articulatory gesture associated with a given feature (or its acoustic/auditory consequence) may be incompatible with that of another feature to some degree. For example, complete closure of the vocal tract (as in the articulation of a stop) inhibits the maintenance of continuous vibration of the vocal folds, and this inhibition is greater for posterior places of articulation than for anterior ones (Ohala 1983). Thus it is not surprising to find languages with speech sound inventories lacking the velar (= posterior) voiced stop [g] (as in Thai) or voiced stops altogether (as in Quechua).

On the other hand, the articulatory gesture associated with a feature (or, again, its acoustic/auditory consequence) may *enhance* that of another feature (Keyser & Stevens 2006). For example, the acoustic difference between the English alveolar fricative [s] and post-alveolar fricative [ʃ] (as in [sɪn] ‘sin’ vs. [ʃɪn] ‘shin’) is enhanced by lip rounding, which is not used to distinguish English consonants from each other on its own.

Interactions between (features of) neighboring speech sounds

The realization of a featural distinction often depends on neighboring speech sounds. This was already noted in §0, with the example of a hypothetical (and in reality, quite common) language that distinguishes among several places of articulation in nasal consonants except before other consonants, where a nasal must have the same place of articulation as the following consonant. This is an example of *assimilation*, whereby one sound systematically takes on a feature of a neighboring sound; this particular type of example is known as *nasal place assimilation*. Other common types of assimilation are *voicing assimilation*, whereby the voicing of a consonant systematically matches the voicing of a neighboring consonant; *palatalization*, whereby the articulation of a consonant systematically matches the articulation of a neighboring palatal vowel (= a vowel with a raised and/or fronted tongue body); and *vowel harmony*, whereby one or more features of a vowel systematically match the same feature(s) of a neighboring vowel.⁷

Assimilations are often subject to restrictions on any or all of the following: the *target* of the assimilation (e.g., that it must be a nasal), the *trigger* of the assimilation (e.g., that it must be a stop), the *direction* of the assimilation (e.g., that the nasal must precede the stop), and the *domain* of the assimilation (e.g., within a prosodic word but not between words). Principled explanations of these restrictions (which ones are found, in what combinations, etc.) are the focus of much research in phonological theory.

A speech sound may also systematically take on a *different* feature (or a different value of a feature) than a neighboring sound; this is called *dissimilation*. Two neighboring speech sounds may systematically trade places (*metathesis*); a consonant's articulation may be systematically weakened (= shortened, or its closure made less complete) when neighboring a vowel or other weakly-articulated consonant (*lenition*); or the features of two neighboring speech sounds may

merge into a single speech sound (*fusion*).

Interactions with higher-order constituents

Some of the interactions with neighboring speech sounds noted above may also involve interactions with higher-order constituents; in some cases, for example, they appear to be motivated by resulting improvements in syllable structure. There are well-documented preferences for syllables with single onset (= syllable-initial) consonants and no coda (= syllable-final) consonants (Clements & Keyser 1983), even in languages that (like English) tolerate clusters of consonants both within and across syllables; such clusters also tend to conform to similar apparent principles across languages (Selkirk 1984, Clements 1990, Morelli 1999). These preferences and principles may be systematically furthered by metathesis or fusion; vowels may be given more consonant-like articulations or consonants more vowel-like articulations; entire speech sounds may be deleted or inserted outright (as in English [weɪdʒd] ‘waged’ vs. [weɪdɒd] ‘waded’).

Positions within the syllable also play a role in the featural realization of speech sounds; in particular, syllable codas tend to license a proper subset of the set of distinctions that are licensed in syllable onsets (Harris 1983, Ito & Mester 1993, Beckman 1998). Such ‘edge effects’, whereby one side or another of a higher-order constituent licenses more or fewer contrasts than other positions, are not uncommon. Other interactions with higher-order constituents include so-called ‘minimal word effects’ (McCarthy & Prince 1995), whereby the prosodic words of a language must be of some minimum size/length (usually a metrical foot), and domain boundedness, whereby e.g. an assimilation occurs only within a particular constituent (often a prosodic word).

Computational assumptions

The various types of interactions discussed in §0 are generally assumed to be the results of computational mechanisms of the phonological component of the grammar, deriving the sys-

tematic patterns observed in *surface representations* of utterances (= the output of the phonological component) from arrays of distinctive constituents stored in *underlying representations* of lexical items (= the input to the phonological component). These representations are assumed to be the interfaces between the phonological component and other cognitive modules of linguistic knowledge; specifically, underlying representations are interfaces with the morphological and syntactic components (which encode word and phrase formation patterns) and surface representations are interfaces with the phonetic component (which encodes the gradient detail of the production and perception of speech sounds).

For some phonologists, these mechanisms take the form of *rewrite rules* that both target and change features of underlying representations in a serial procedure leading to a particular surface representation (Chomsky & Halle 1968). For others, the mechanisms take the form of resolutions of conflicts between constraints that define desirable surface configurations and constraints that penalize imperfect correspondences between underlying and surface representations; candidate surface representations are compared by the constraints in parallel, and conflicts are resolved by ranking (Prince & Smolensky 1993) or numerical weighting (Smolensky & Legendre 2006). Still others adopt other approaches, such as hybrid models with both constraints and rewrite rules (Calabrese 2005) or ones with serial candidate comparison by ranked constraints (McCarthy 2007, 2010).

Phonotactics

Evidence for these computational mechanisms is of two kinds. One kind is the straightforward distribution of (features of) speech sounds; the fact that they are found in some contexts and not in others. In these types of cases, generalized grammatical statements of these distributions — often called *phonotactic constraints* — are typically sufficient.

For example, the distributions of consonants in word-initial clusters in English are restricted in the following ways. Consider clusters of two consonants first. If the second consonant is [r], then the first consonant may be any of the stops [p,t,k,b,d,g] or fricatives [f,ʃ]. If the second consonant is [l], on the other hand, then a preceding stop may only be one of [p,k,b,g] (that is, not one of the alveolars [t,d]), and a preceding fricative may be one of [f,s] but not [ʃ]. If the second consonant is a nasal [m,n] or stop [p,t,k], then the first consonant must be [s]; in clusters of three consonants, the first must be [s], the second must be one of [p,t,k], and the third must be [r] or [l] (and the latter only if the second is not [t]). Restrictions like these are found in many different languages; they are structured, principled, and not arbitrary. Precise formal statements of phonotactic constraints expressing these restrictions differ, but phonologists generally agree that the phonological component of a language is made up in part by such statements.

Alternations

The second kind of evidence comes from *alternations*, a term that refers to examples of meaningful phonological units (= *morphemes*) the surface representations of which vary systematically according to their context within an utterance. The significance of this kind of evidence is predicated on the assumption that (the constituents of) the surface variants of a morpheme *correspond* with each other, via a single underlying representation that includes all idiosyncratic, unpredictable properties (and for some, only these).

For example, the English morpheme ‘atom’ is systematically pronounced [ˈæɾəm] when unsuffixed but [əˈtʰɑ:m] when suffixed with ‘-ic’. Three of the four speech sounds of this morpheme have imperfect correspondences between these two pronunciations: stressed [æ] with unstressed [ə], [ɾ] in the onset of an unstressed syllable with [tʰ] in the onset of a stressed syllable, and unstressed [ə] with stressed [ɑ:]. The root cause of all three is an idiosyncratic requirement

of the suffix ‘-ic’, that the syllable preceding it be stressed, trumping the conflicting requirement of the morpheme ‘atom’ that its initial syllable be stressed. Unstressed syllables in English license fewer vowel contrasts than stressed syllables, reducing many of them to a mid-central vowel [ə]; voiceless stops are aspirated at the left edges of stressed syllables but not otherwise; alveolar stops are lenited to the tap [ɾ] between vowels when the following vowel is unstressed. Reduction, aspiration, and lenition being predictable but other properties not, the underlying representation of this morpheme can be deduced to be /'ætɑ:m/. From this, the systematic surface variants are *derived* by computational mechanisms corresponding more or less to the statements of reduction, aspiration, and lenition given above.

Some alternations provide solid evidence for the conclusion that two systematic phones are allophones of the same phoneme. For example, the fact that [ɾ] in ‘atom’ corresponds with [t^h] in ‘atomic’ demonstrates that these two phones are allophones of the same phoneme in English, /t/. In Spanish, the fact that the voiced bilabial stop [b] in ['baso] ‘glass’ corresponds with the voiced bilabial fricative [β] in [mi'βaso] ‘my glass’ demonstrates that these two phones are allophones of the same phoneme, /b/ or /β/.⁸

Alternations can also provide evidence for the neutralization of particular phonemic distinctions. For example, the vowel alternations in ‘atom’ and ‘atomic’ demonstrate that the distinction between /æ/ and /ɑ:/ (['sæd] ‘sad’ vs. ['sɑ:d] ‘sod’) is neutralized to [ə] in unstressed syllables, as indeed are many of the vowel distinctions in English. Likewise, the distinction between /t/ and /d/ (['bæt] ‘bat’ vs. ['bæd] ‘bad’) is neutralized to [ɾ] between vowels when the following vowel is unstressed: ['bærə] could be either ‘batter’ or ‘badder’, modulo systematic vowel length differences before the [ɾ]. The “completeness” of these and other neutralizations has been challenged; see Flemming & Johnson (2007) on neutralization of unstressed vowels to [ə], Brav-

er (2011) on neutralization of /t,d/ to [ɾ], and Port & Leary (2005) and van Oostendorp (2008) for examination of the relevant issues more generally.

Architectural assumptions

The minimal architectural assumption, shared by all phonological theories making the kinds of representational, analytical, and computational assumptions outlined thus far in this chapter, is that there are two levels of representation for the computation of an utterance. One is an underlying representation, stored separately for every morpheme, that at the very least encodes basic distinctions among phonological constituents. The other is a systematic surface representation, the end result of the changes made by the rules or constraints that apply in the course of the computation.

Elaborations of these theories differ according to whether there are significant levels of representation other than these two; these additional levels of representation are typically motivated by observed interactions between the phonological component and other components of grammar, most significantly morphology and syntax. For example, many phonologists subscribe to some variant of the theory of lexical phonology and morphology (Kiparsky 1982, to appear), whereby phonological and morphological computations are interleaved such that the product of every morphological computation (e.g., the addition of an affix to a stem) is the input to a phonological computation. Research within this theory has identified significant differences between phonological computations applying to different morphological constituents, leading to models in which there are levels of representation corresponding to these differences.

The phonological component may also interact with the morphological and syntactic components via ‘interface constraints’ regulating either the cross-component alignment of category edges (Selkirk 1986, McCarthy & Prince 1994) or the correspondence between morpho-

phonological constituents (Burzio 1994, McCarthy & Prince 1999).

There is of course also the not-insignificant matter of the interface between the systematic surface representation level of the phonological component and the ultimate phonetic output. There is mounting evidence that much of what was once relegated to a universal, noncategorical ‘phonetic interpretation module’ — and even some of what was once considered ‘low-level’ systematic phonology — is both language-specific and gradient. This has led to the development of several detailed models of the interface between phonology and phonetics, including e.g. Browman & Goldstein (1992), Zsiga (1997), Boersma (1998), Flemming (2001), Pierrehumbert (2003), and Ladd (2008).

Summary

Phonologists aim to adequately describe the structure of the phonological components of languages via the analysis of their sound patterns. The main goals of phonological description and analysis are the establishment of generalizations about the distribution and reducibility of phonological constituents in different languages. These goals have led to the development of four main sets of assumptions: representational, analytical, computational, and architectural. Phonologists generally share the assumptions laid out in this chapter, though of course the specifics differ from one theory to another.

References

- Archangeli, Diana, and Douglas Pulleyblank. 1994. *Grounded Phonology*. Cambridge, MA: MIT Press.
- Baković, Eric. 1994. Strong onsets and Spanish fortition. *MIT Working Papers in Linguistics* 23, 21–39.
- Barlow, Jessica. 2003. The stop-spirant alternation in Spanish: Converging evidence for a fortition account. *Southwest Journal of Linguistics* 22, 51–86.
- Beckman, Jill. 1998. Positional faithfulness. Doctoral dissertation, University of Massachusetts, Amherst.
- Beňuš, Štefan, and Adamantios I. Gafos. 2007. Articulatory characteristics of Hungarian ‘transparent’ vowels. *Journal of Phonetics* 35, 271–300.
- Blevins, Juliette. 1995. The syllable in phonological theory. In John A. Goldsmith (ed.), *The Handbook of Phonological Theory* (1st edition). Oxford: Blackwell, 206–244.
- Boersma, Paul. 1998. Functional phonology: Formalizing the interactions between articulatory and perceptual drives. Doctoral dissertation, University of Amsterdam.
- Braver, Aaron. 2011. Incomplete neutralization in American English flapping: A production study. *University of Pennsylvania Working Papers in Linguistics* 17.1, 31–40.
- Browman, Catherine P., and Louis M. Goldstein. 1992. Articulatory phonology: An overview. *Phonetica* 49, 155–180.
- Burzio, Luigi. 1994. *Principles of English Stress*. Cambridge: Cambridge University Press.
- Calabrese, Andrea. 2005. *Markedness and economy in a derivational model of phonology*. The Hague: Mouton de Gruyter.
- Chomsky, Noam, and Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper & Row.
- Clements, G. Nick. 1985. The geometry of phonological features. *Phonology Yearbook* 2, 225–252.
- Clements, G. Nick. 1990. The role of the sonority cycle in core syllabification. In John Kingston and Mary E. Beckman (eds.), *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*. Cambridge University Press, 282–333.
- Clements, G. Nick. 1991. Vowel height assimilation in Bantu languages. In Kathleen Hubbard (ed.), *Proceedings of the 17th Annual Meeting of the Berkeley Linguistics Society, Feb. 15–18, 1991: Special Session on African Language Structures*, 25–64.

- Clements, G. Nick. 1999. Affricates as noncontoured stops. In Osamu Fujimura, Brian D. Joseph, and Bohumil Palek (eds.), *Proceedings of LP '98: Item Order in Language and Speech*. Prague: The Karolinum Press, 271–299.
- Clements, G. Nick, and Samuel J. Keyser. 1983. *CV Phonology: A Generative Theory of the Syllable*. Cambridge, MA: MIT Press.
- Clements, G. Nick, and Elizabeth Hume. 1995. The internal organization of speech sounds. In John A. Goldsmith (ed.), *The Handbook of Phonological Theory*. Oxford: Blackwell, 245–306.
- Davis, Stuart. 1985. Topics in syllable geometry. Doctoral dissertation, University of Arizona.
- de Lacy, Paul, ed. 2007. *The Cambridge Handbook of Phonology*. Cambridge: Cambridge University Press.
- de Lacy, Paul. 2010. Phonological evidence. In Steve Parker (ed.), *Phonological Argumentation: Essays on Evidence and Motivation*. London: Equinox.
- Flemming, Edward. 1995. Auditory representations in phonology. Doctoral dissertation, UCLA. [Published 2002, New York: Routledge.]
- Flemming, Edward. 2001. Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology* 18, 7–44.
- Flemming, Edward. 2005. Deriving natural classes in phonology. *Lingua* 115, 287–309.
- Flemming, Edward, and Stephanie Johnson. 2007. Rosa's roses: Reduced vowels in American English. *Journal of the International Phonetic Association* 37, 83–96.
- Gafos, Adamantios I. 1996. The articulatory basis of locality in phonology. Doctoral dissertation, Johns Hopkins University. [Published 1999, New York: Routledge.]
- Goldsmith, John A. 1976. Autosegmental phonology. Doctoral dissertation, MIT.
- Goldsmith, John A., Jason Riggle, and Alan C. L. Yu, eds. *The Handbook of Phonological Theory* (2nd edition). Oxford: Blackwell.
- Hale, Mark, and Charles Reiss. 2008. *The Phonological Enterprise*. Oxford: Oxford University Press.
- Halle, Morris. 1983. On distinctive features and their articulatory implementation. *Natural Language and Linguistic Theory* 1, 91–105.
- Halle, Morris. 1995. Feature geometry and feature spreading. *Linguistic Inquiry* 26, 1–46.
- Halle, Morris, Bert Vaux, and Andrew Wolfe. 2000. On feature spreading and the representation of place of articulation. *Linguistic Inquiry* 31, 387–444.

- Harris, James W. 1983. *Syllable Structure and Stress in Spanish: A Nonlinear Analysis*. Cambridge, MA: MIT Press.
- Hayes, Bruce. 1989. Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20, 253–306.
- Hayes, Bruce. 1990. Diphthongisation and coindexing. *Phonology* 7, 31–71.
- Hayes, Bruce. 1995. *Metrical Stress Theory: Principles and Case Studies*. Chicago: University of Chicago Press.
- Hayes, Bruce, Robert Kirchner, and Donca Steriade, eds. 2004. *Phonetically-Based Phonology*. Cambridge: Cambridge University Press.
- Inkelas, Sharon. 1994. The consequences of optimization for underspecification. *Proceedings of the Northeast Linguistics Society* 25, 287–302.
- Ito, Junko, and R. Armin Mester. 1993. Licensed segments and safe paths. *Canadian Journal of Linguistics* 38, 197–213.
- Jakobson, Roman, Gunnar Fant, and Morris Halle. 1952. *Preliminaries to Speech Analysis*. MIT Acoustics Laboratory Technical Report #13.
- Kahn, Daniel. 1976. Syllable-based generalizations in English phonology. Doctoral dissertation, MIT.
- Keyser, Samuel J., and Kenneth N. Stevens. 2006. Enhancement and overlap in the speech chain. *Language* 82, 33–63.
- Kiparsky, Paul. 1982. From cyclic phonology to lexical phonology. In Harry van der Hulst and Norval Smith (eds.), *The Structure of Phonological Representations, Part I*. Dordrecht: Foris, 131–175.
- Kiparsky, Paul. *Paradigms and Opacity*. Stanford: CSLI Publications (to appear).
- Ladd, D. Robert. 2008. *Intonational Phonology*. Cambridge: Cambridge University Press.
- Liberman, Mark, and Alan Prince. 1977. On stress and linguistic rhythm. *Linguistic Inquiry* 8, 249–336.
- Lisker, Leigh, and Arthur S. Abramson. 1971. Distinctive features and laryngeal control. *Language* 47, 767–785.
- Lombardi, Linda. 1990. The nonlinear organization of the affricate. *Natural Language and Linguistic Theory* 8, 375–426.
- Lozano, María del Carmen. 1979. Stop and spirant alternations: Fortition and spirantization processes in Spanish phonology. Doctoral dissertation, The Ohio State University. [Distributed by Indiana University Linguistics Club, Bloomington, IN.]

- Maddieson, Ian. 1984. *Patterns of Sounds*. Cambridge: Cambridge University Press.
- McCarthy, John J. 1979. On stress and syllabification. *Linguistic Inquiry* 10, 443–465.
- McCarthy, John J. 2007. *Hidden Generalizations: Phonological Opacity in Optimality Theory*. London: Equinox.
- McCarthy, John J. 2010. An introduction to Harmonic Serialism. *Language and Linguistics Compass* 4, 1001–1018.
- McCarthy, John J., and Alan S. Prince. 1994. Generalized alignment. In Geert Booij and Jap van Marle (eds.), *Yearbook of Morphology 1993*. Dordrecht: Kluwer, 79–153.
- McCarthy, John J., and Alan S. Prince. 1995. Prosodic morphology. In John A. Goldsmith (ed.), *The Handbook of Phonological Theory*. Oxford: Blackwell, 318–366.
- Mielke, Jeff. 2008. *The Emergence of Distinctive Features*. Oxford University Press.
- Morelli, Frida. 1999. The phonotactics and phonology of obstruent clusters in Optimality Theory. Doctoral dissertation, University of Maryland.
- Nespor, Marina, and Irene Vogel. 1986. *Prosodic Phonology*. Dordrecht: Foris.
- Ohala, John J. 1983. The origin of sound patterns in vocal tract constraints. In Peter F. MacNeilage (ed.), *The Production of Speech*. New York: Springer-Verlag, 189–216.
- van Oostendorp, Marc. 2008. Incomplete devoicing in formal phonology. *Lingua* 118, 1362–1374.
- van Oostendorp, Marc, Colin J. Ewen, Elizabeth Hume, and Keren Rice, eds. 2011. *The Blackwell Companion to Phonology*. Oxford: Blackwell.
- Padgett, Jaye. 2002. Feature classes in phonology. *Language* 78, 81–110.
- Parkinson, Frederick. 1996. The representation of vowel height in phonology. Doctoral dissertation, The Ohio State University.
- Perlmutter, David. 1995. Phonological quantity and multiple association. In John A. Goldsmith (ed.), *The Handbook of Phonological Theory*. Oxford: Blackwell, 307–313.
- Pierrehumbert, Janet B., and Mary E. Beckman. 1988. *Japanese Tone Structure*. Cambridge, MA: MIT Press.
- Pierrehumbert, Janet B. 2003. Probabilistic phonology: Discrimination and robustness. In Rens Bod, Jennifer Hay, and Stefanie Jannedy (eds.), *Probabilistic Linguistics*. Cambridge, MA: MIT Press, 177–228.
- Port, Robert F. 2007. How are words stored in memory? Beyond phones and phonemes. *New Ideas in Psychology* 25, 143–170.

- Port, Robert F., and Adam P. Leary. 2005. Against formal phonology. *Language* 85, 927–964.
- Prince, Alan, and Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. [Published 2004, Medford, MA: Blackwell.]
- Ridouane, Rachid, and G. Nick Clements, eds. 2011. *Where Do Phonological Features Come From? Cognitive, Physical, and Developmental Bases of Distinctive Speech Categories*. Amsterdam: John Benjamins.
- Sagey, Elizabeth. 1986. The representation of features and relations in nonlinear phonology. Doctoral dissertation, MIT. [Published 1990, New York: Garland Press.]
- Sapir, Edward. 1949. The psychological reality of phonemes. In David G. Mandelbaum (ed.), *Selected Writings of Edward Sapir in Language, Culture and Personality*. Berkeley and Los Angeles: University of California Press, 46–60.
- Schane, Sanford A. 1995. Diphthongization in Particle Phonology. In John A. Goldsmith (ed.), *The Handbook of Phonological Theory*. Oxford: Blackwell, 586–608.
- Selkirk, Elisabeth O. 1980. The role of prosodic categories in English word stress. *Linguistic Inquiry* 11, 563–605.
- Selkirk, Elisabeth O. 1984. On the major class features and syllable theory. In Mark Aronoff and Richard Oehrle (eds.), *Language Sound Structure*. Cambridge, MA: MIT Press, 107–136.
- Selkirk, Elisabeth O. 1986. On derived domains in sentence phonology. *Phonology Yearbook* 3, 371–405.
- Smolensky, Paul, and Géraldine Legendre. 2006. *The Harmonic Mind: From Neural Computation to Optimality-Theoretic Grammar*. Cambridge, MA: MIT Press.
- Steriade, Donca. 1999. Phonetics in phonology: the case of laryngeal neutralization. In Matthew K. Gordon (ed.), *Papers in Phonology 3*. UCLA Working Papers in Linguistics 2, Department of Linguistics, UCLA, 25–146.
- Steriade, Donca. 2000. Paradigm uniformity and the phonetics-phonology boundary. In Michael Broe and Janet Pierrehumbert (eds.), *Papers in Laboratory Phonology V: Acquisition and the Lexicon*. Cambridge University Press, 313–334.
- Stevens, Kenneth N. 2002. Toward a model for lexical access based on acoustic landmarks and distinctive features. *Journal of the Acoustical Society of America* 111, 1872–1891.
- Stevens, Kenneth N. 2003. Acoustic and perceptual evidence for universal phonological features. In *Proceedings of the XVth International Congress of Phonetic Sciences, Barcelona*, 33–38.
- Wetzels, W. Leo, and Joan Mascaró. 2001. The typology of voicing and devoicing. *Language* 77, 207–244.

Zsiga, Elizabeth C. 1997. Features, gestures, and Igbo vowels: An approach to the phonology-phonetics interface. *Language* 73, 227–274.

¹ It is of course not possible to delve as deeply as one would like into any of the areas covered here, and I attempt to counterbalance the resulting superficiality by citing some of the relevant literature. (It is also not possible to reach out as broadly as one might like; I'm afraid that I make far less of an effort to remedy this deficiency.) My citation bias is in many cases at least modestly geared toward more recent work, in large part because previous work can be tracked down more easily that way than vice-versa. I somewhat purposely do not cite the individual contributions to three recently published compendia of phonological research — de Lacy (2007), van Oostendorp et al. (2011), and Goldsmith et al. (2011) — and instead encourage the interested reader to consult any or all of these volumes for further detailed discussion. The second is especially comprehensive and accessible, both by design and in fact.

² These goals are sometimes assumed to be in conflict, but usually only when a particular set of theoretical assumptions made with one goal in mind is questioned/argued against with the other in mind.

³ But see Flemming (2005), where it is argued that natural classes derive not from the content of the set of available features but from the content of the set of constraints making reference to those features.

⁴ Sets of features (like individual places of articulation) that behave as a class in this way are assumed to be grouped together in some theories of autosegmental representation. See Padgett (2002) for a review, and for arguments that these featural groupings are not best captured representationally.

⁵ Not obvious, but not impossible: Steriade (1999), for example, notes that a distinction between voiced and voiceless consonants before vowels is easier to perceive — and thus more

likely to be maintained — than the same distinction before consonants and utterance-finally.

Failure to maintain a distinction utterance-finally is then carried over to word-final position via paradigmatic analogy (Steriade 2000).

⁶ Standard feature theories typically categorize these three (types of) speech sound multi-dimensionally, with [p^h] and [p] together as voiceless and [p] and [b] together as unaspirated. But it is also possible to categorize them on a unidimensional scale based on what is known as *voice onset time* (VOT; Lisker & Abramson 1971), the time between the release of a stop and the onset of voicing on a following vowel (or other sonorant). The production of [b] involves a systematically shorter (typically negative) VOT than the production of [p], which involves a systematically shorter VOT than the production of [p^h].

⁷ ‘Neighboring’ is to be interpreted loosely here, as vowels typically harmonize across intervening consonants. This means either that these consonants must be ignored in some way (see Sagey 1986, Archangeli & Pulleyblank 1994, and Clements & Hume 1995 for various proposals) or that they must participate in the harmony in some articulatorily measurable way (Gafos 1996; see also Benus & Gafos 2007).

⁸ This phoneme is traditionally assumed to be a basic stop (Harris 1969), but there are good reasons to think that it is instead a basic fricative or approximant (Lozano 1979, Baković 1994, Barlow 2003).