Oxford
LINGUISTICS

Phonological Knowledge

Conceptual and Empirical Issues

EDITED BY

Noel Burton-Roberts, Philip Carr,
and Gerard Docherty

Phonological Knowledge

Phonological Knowledge

Conceptual and Empirical Issues

edited by

NOEL BURTON-ROBERTS,
PHILIP CARR, and
GERARD DOCHERTY

OXFORD
UNIVERSITY PRESS

7

Phonology as Cognition

MARK HALE AND CHARLES REISS

1. FORM AND SUBSTANCE IN PHONOLOGY

This chapter attempts to ground phonology within psychology. That is, we are interested in phonology as a branch of the study of mental representation, the psychology of mind. In order to develop this 'phonology of mind' we need to understand the relationship between form and substance in linguistic representation. A coherent account of this distinction has yet to be proposed for either phonology or syntax. We attempt to contribute to this necessary enquiry in the domain of phonology by first defining 'form' and 'substance', and then critiquing some recent work that implicitly or explicitly touches on the relationship between the two. We will argue that current trends in phonology fail to offer a coherent conception of form and substance and are also inconsistent with basic principles of science. Since we are not proposing a complete alternative model of phonology, we invite the reader to reflect on how our proposals could be implemented or on how our assumptions (which we believe are widely shared in principle, if not in practice) should be modified.

It has proven quite useful for linguists to conceive of a grammar as a relationship between (i) a set of symbols—entities such as features and variables, constituents such as syllables, feet, NPs, and so on, and (ii) a set of computations—operations whose operands are drawn from the set of symbols, such as concatenation, deletion, and so on. The set of symbols and relations together describe the formal properties of the system. Relevant questions in discussing formal properties include 'Is the system rule and/or constraint based?'; 'Do operations apply serially or in parallel?'; and 'Are there limits on the number of operands referred to in the course of a given phonological computation?'

The issue of substance essentially arises only with respect to the set of symbols and the extent to which their behaviour in phonological computation is driven by

This chapter expands on Hale and Reiss (2000). We are grateful to audiences at the Montreal–Ottawa–Toronto Phonology Workshop 1998 at the University of Ottawa and at the Berkeley Phonology Laboratory, as well as to Noel Burton-Roberts, Morris Halle, Bill Idsardi, Madelyn Kissock, Afton Lewis, Jean-Philippe Marcotte, and Ida Toivonen for discussion and challenging criticism that improved the chapter. The authors' names appear in alphabetical order.

what they symbolize. For the sake of simplicity we restrict ourselves in this discussion to the set of phonological primitives known as distinctive features and to the representations that can be defined as combinations of distinctive features.

We will concentrate in this chapter on this notion of substance in phonological representation. In brief, the question we are interested in is the following:

(1) Do the phonetic correlates (that is, the substance) of a particular distinctive feature or feature bundle have any non-arbitrary bearing on how that feature or feature bundle is treated by the computational system?

It is trivial to show that languages differ in that their computational systems treat specific features or feature bundles differently—for example, Standard German has coda obstruent devoicing and English does not. From this we can conclude that languages *can* treat the same symbols differently. A more challenging problem arises when we find an apparent example of cross-linguistically universal, seemingly non-arbitrary treatment of a feature or feature bundle. In such cases we must ask ourselves the following:

(2) Is the observed pattern a reflection of substantive constraints on the computational system (that is, the grammar), or is the pattern due to other causes?

Other *a priori* plausible causes include, as we shall show in what follows, the process of language change, the nature of the language acquisition device, sampling errors, and so on. From the standpoint of grammatical theory, factors such as sampling errors are obviously uninteresting. However, language change and the nature of the learning path are also, strictly speaking, not part of grammatical theory. The modular approach to linguistics, and to science in general, requires that we both model the interactions between related domains, and also sharply delineate one domain from another. Occam's Razor demands that, in doing so, we avoid redundancy and the postulation of unnecessary factors.

Even before proceeding to our argument that generalizations that bear on patterns of phonetic substance are not relevant to phonological theory as we define it, we can see that there is potentially much to gain from this modular approach in that it posits that universal phonology should be studied not just across languages, but also across modalities. What is shared by the phonologies of signed and spoken languages? We believe that phonology consists of a set of formal properties (for example, organization into syllables and feet, feature spreading processes) that are modality independent and thus not based on phonetic substance. The goal of phonological theory should be to discover these formal properties. Failure to appreciate this goal has resulted in rampant 'substance abuse' in the phonological community.

We discuss various aspects of substance abuse in Sections 2–5. In Section 6, we offer a modest contribution to a substance-free phonology. In Section 7, we return to substance with a discussion of the putative phenomenon of phonetic

enhancement in grammars. Section 8 ties together the preceding sections with arguments against functionalist 'explanation' in linguistics. We argue that *dysfunctionalist* reasoning fares as well as its better-known rival. Section 9 provides a concluding plea for a modular approach to the study of sound patterns in human languages.

2. THREE EXAMPLES OF SUBSTANCE ABUSE IN GRAMMATICAL THEORY

2.1. *Positional faithfulness in Beckman (1997)*

Beckman (1997) proposes the constraints in (3*a–b*) as members of the universal constraint set:

(3)(*a*) IDENT-σ_1(hi)
A segment in the root-initial syllable in the output and its correspondent in the input must have identical values for the feature [high].

(*b*) IDENT(hi)
Correspondent segments in output and input have identical values for the feature [high].

As Beckman explains, this set of constraints allows faithfulness to a feature, like [high], to be maintained in some contexts, but not others, since the context-sensitive constraint (3*a*) can be ranked above a markedness constraint that is violated by, say, the presence of high vowels, *HIGH, which in turn is ranked above the general constraint in (3*b*). In other words, the ranking in (4) will allow surface high vowels only in root-initial syllables.

(4) IDENT-σ_1(hi) >> *HIGH >> IDENT(hi)

This is assumed to be a welcome result: 'The high ranking of positional faithfulness constraints, relative to both the more general IDENT constraints and markedness constraints, yields the result that features and/or contrasts in *just those positions which are psycholinguistically or perceptually salient* are less susceptible to neutralization than in other locations which are not protected' (Beckman 1997: 8; emphasis in original). Beckman (1997: 5) cites more than ten psycholinguistic studies to support her claim that word-initial material is more salient than medial or final material.[1] We believe that the correct conclusion to be drawn from this psycholinguistic evidence is the *exact opposite* of that which Beckman draws.[2] Encoding the findings of psycholinguistic experimentation in the grammar is a

[1] It is unclear whether this generalization would hold, say, in a language with non-initial stress. It is also unclear whether Beckman's extension of psycholinguistic findings concerning word-initial syllable to *root*-initial syllables is justified. However, we will assume, for purposes of this discussion, that Beckman has stated the relevant generalizations correctly.

[2] We wish to stress that we are not singling Beckman out for any reason except for the fact that her paper appeared recently in a widely read journal and is well written and clear in its arguments and assumptions.

mistake because it is possible to achieve the same empirical coverage without positing new mechanisms like positional faithfulness.[3] Consider the following alternative account.

We know that children acquire spoken language primarily on the basis of acoustic input from speakers in their environment, with Universal Grammar (UG) providing constraints on the hypothesis space.[4] We also know that phonological contrasts are best distinguished and recalled when occurring in certain positions. Imagine a child exposed to a language $\mathcal{L}_1$ that allows high vowels in all syllables—initial, medial, and final. Imagine further that $\mathcal{L}_1$ has initial stress and that stress is realized as relatively increased duration and intensity. Given this scenario, it is easy to see that a child constructing $\mathcal{L}_2$ on the basis of ouput from $\mathcal{L}_1$ could consistently fail to acquire a contrast between mid- and high vowels in relatively short, quiet syllables (those that are non-initial and thus unstressed), but succeed in acquiring this distinction in initial syllables, which are stressed and thus longer and louder. This type of relationship between $\mathcal{L}_1$ and $\mathcal{L}_2$ is known as 'sound change' (in particular, as a 'conditioned merger'). On the other hand, it is highly implausible that an acquirer would consistently fail to correctly analyse the mid/high contrast in longer, louder (stressed) syllables, yet successfully analyse the contrast in relatively short, quiet syllables. Note that this implausibility is independent of our view of the nature of UG.

We see, therefore, that the existence of positional faithfulness phenomena can be understood as merely reflecting the nature of the learning situation and not a reflection of any grammatical principle:[5]

(5) If the acoustic cues of a given contrast in the target language are correctly analysed by the acquirer in a context where they are relatively weak, they will also be analysed correctly in a context where they are relatively strong.

Note that (5) is essentially definitional, since the strength, or acoustic salience, of a contrast is just a measure of how easy it is to perceive. What is most important to understand is that the theory proposed here is not meant to *replace* a synchronic account of the data. So, the best synchronic analysis must somehow be able to generate vowel neutralization in noninitial syllables. (5) is meant to guide us in choosing a theory of grammar in which to couch that synchronic account, but (5) is not part of the grammar. Whatever theory of phonology one adopts, it must be able synchronically to generate the type of patterns that Beckman describes, but the predictions generated by the correct theory, qua phonological theory, need not replicate the predictions derivable from (5).

[3] For other arguments against context-sensitive faithfulness, see Reiss (1996: 315).

[4] It may be a useful idealization to assume that UG does not just constrain the learning path, but completely determines it. We suspect that such a position will prove most fruitful in sketching an explicit theory of acquisition, but justification for this goes beyond the scope of this chapter.

[5] This idea is discussed more thoroughly in Hale (forthcoming *b*).

By adopting the view of sound change proposed here, we see that many supposedly phonological tendencies, or markedness patterns, are actually emergent properties—that is, epiphenomenal. 'Positional faithfulness' is due, not to the nature of *phonology*, but to the 'sifting effect' of acquisition on the incidental, arbitrary nature of the phonetic substance. Since effects such as those observed by Beckman already have a coherent extragrammatical account within acquisition theory (and it is necessary, in any event, to have an acquisition theory), building positional faithfulness into a theory of universal phonology is a misuse, or abuse, of phonetic substance in theory construction.

2.2. */r/-insertion in McCarthy* (1993)

McCarthy's (1993) discussion of intervocalic *r*-insertion in Massachusetts English is fairly well known, so an example should be sufficient for illustration. In this dialect, an underlying sequence like *Wanda arrived* is realized with a 'linking' [r]: *Wanda[r] arrived.* As McCarthy himself notes (and as discussed by LaCharité and Paradis 1993 and Halle and Idsardi 1997) 'r is demonstrably not the default consonant in English' (1993:189). That is, it is not the maximally unmarked consonant that an Optimality Theory (OT) account predicts would emerge in such a situation. In order to account for the insertion of [r] McCarthy proposes a special *rule* of *r*-insertion: 'a phonologically arbitrary stipulation, one that is outside the system of Optimality' (1993: 190). There are several problems with this proposal, many of which are insightfully discussed by Halle and Idsardi. However, we propose that one of their criticisms requires elaboration. Halle and Idsardi rightly point out that 'reliance on an arbitrary stipulation that is outside the system of Optimality is equivalent to giving up on the whole enterprise' (1997: 337), but these authors do not discuss what we consider to be the most important point: grammars do contain arbitrary processes. McCarthy's grammar has an arbitrary component (containing rules like *r*-insertion) and a non-arbitrary component (containing the substantive OT constraints). Such a theory is empirically non-distinct from the theory we propose below, which posits that *all* grammatical computations are arbitrary with respect to phonetic substance. This is because the set of phenomena predicted to exist by our theory (with only arbitrary processes) is identical to the set of phenomena predicted to exist by McCarthy's theory (with both non-arbitrary and arbitrary processes). Since McCarthy must adopt a model that allows arbitrary phenomena (like *r*-insertion), the addition to the theory of a special subcomponent to account for alleged 'non-arbitrary' phenomena violates Occam's Razor.

The diachronic source of *r*-insertion is transparent—the relevant dialects also exhibit *r*-deletion in codas, so insertion reflects rule inversion triggered by hypercorrection. Again, the diachronic facts do not make a synchronic account unnecessary, but they show us that basically idiosyncratic historical events affect specific grammars—and, in part, how they may do so.

2.3. Structural constraints on non-structures

Perhaps one of the most problematic cases of substance abuse we have come across is McCarthy's (1996) appeal to parameterized constraints to account for opacity effects in Hebrew spirantization by invoking the notion of constraint schema. McCarthy makes some reasonable simplifying assumptions in this first attempt:

> I will assume that every constraint is a prohibition or negative target defined over no more than two segments, α and β. That is, the canonical constraint is something like $*\{\alpha,\beta\}$, with appropriate conditions imposed on α and β. These conditions are as follows:
>
> (i) a specification of the featural properties of α and β as individual segments;
> (ii) a specification of the linear order relation between α and β ($\alpha < \beta$, $\beta < \alpha$, or both in the case of mirror-image rules ...
> (iii) a specification of the adjacency relation between α and β (e.g., strict adjacency, vowel-to-vowel adjacency ...)
>
> The decomposition of the conditions imposed by a phonological constraint will be crucial in accounting for the range of opacity phenomena. Even more important, though, is this: each condition—the featural composition of α, the featural composition of β, linear order and adjacency—must also name the level (underlying, surface, or either) at which it applies. Correspondence Theory allows us to make sense of conditions applying at one level or the other. As a bookkeeping device, I will state the constraints in the form of a table ... (1996: 220)

We reproduce here the schema-based constraint that McCarthy proposes to account for Tiberian Hebrew post-vocalic spirantization.

(6) Constraint for opacity in Hebrew spirantization (McCarthy 1996: 223)

*	Condition	Level
α	V	Indifferent
β	[-son, -cont]	Surface
Linear order	$\alpha > \beta$	Indifferent
Adjacency	Strict	Indifferent

As McCarthy says. 'In correspondence terms, the meaning of this constraint is this: the constraint is violated if a surface stop β or its underlying correspondent is immediately preceded by a vowel' (1996: 223).

As pointed out in Reiss (1997), this powerful constraint type has several problems. First, it compromises the OT notion of a universal, innate constraint set by allowing apparently language-specific parameterized constraints. This may not be a serious problem, since it represents an attempt to define the form of possible constraints. In other words, McCarthy could be interpreted as presenting a theory in which the intensional description of the set of constraints is universal, but

languages vary in which constraints they actually incorporate (based on evidence presented to the learner).[6]

Most relevant to our present purposes, however, is the fact that such constraints undermine implicit and explicit appeal to phonetic grounding of well-formedness constraints in McCarthy's work. For example, McCarthy and Prince (1995: 88) refer to a constraint *VgV as the 'phonologization of Boyle's Law'. It is incoherent to argue that a constraint is motivated by the facts of phonetics, when the structures that violate this constraint need not be surface structure strings. In fact, they need not exist as strings at *any level of representation.*

3. NEO-SAUSSUREANISM

The conclusion we wish to draw from the above examples and many others like them is that the best way to gain an understanding of the computational system of phonology is to assume that the phonetic substance (say, the spectral properties of sound waves) that leads to the construction of phonological entities (say, feature matrices) *never* reflects how the phonological entities are treated by the computational system. The computational system treats features and the like as arbitrary symbols. What this means is that many of the so-called *phonological universals* (often discussed under the rubric of markedness) are in fact epiphenomena deriving from the interaction of extragrammatical factors such as acoustic salience and the nature of language change. It is not surprising that, even among its proponents, markedness 'universals' are usually stated as 'tendencies'. If our goal as generative linguists is to define the set of *computationally possible* human grammars, 'universal tendencies' are irrelevant to that enterprise.

We therefore propose extending the Saussurean notion of the arbitrary nature of linguistic signs to the treatment of phonological representations by the phonological computational system. Phonology is not and should not be grounded in phonetics, since the facts that phonetic grounding is meant to explain can be derived without reference to *phonology*. Duplication of the principles of acoustics and acquisition inside the grammar constitutes a violation of Occam's Razor and thus must be avoided. Only in this way will we be able correctly to characterize the universal aspects of phonological computation.

John Ohala (e.g. 1990) has done the most to demonstrate that many so-called markedness tendencies can be explained on phonetic grounds and thus should not be explained by principles of grammar. Examples discussed by Ohala include patterns of assimilation and the contents of phonemic inventories. For an extensive bibliography on this topic, see Ohala (1998). We differ from Ohala in our

[6] McCarthy does not explicitly make this argument, but it seems to us to be a better theory than the standard OT claim that all constraints are literally present in all grammars. Of course, adopting our suggested interpretation will force OT practitioners to revise their views on acquisition and, especially, *the emergence of the unmarked.* This view of OT would also make it much closer to a theory of learned rules.

use of the term 'phonology' (which for him covers *all* aspects of the sound systems of human language), but whole-heartedly endorse his approach.

3.1. *Substance in* The Sound Pattern of English

It is obvious that our proposal runs contrary to most of the discussion in chapter 9 of *The Sound Pattern of English* (SPE) (Chomsky and Halle 1968). This chapter starts out with an 'admission' that the theory developed in the earlier chapters is seriously flawed:

> The problem is that our approach to features, to rules and to evaluation has been overly formal. Suppose, for example, that we were systematically to interchange features or to replace [αF] by [$-\alpha$F] (where α is +, and F is a feature) throughout our description of English structure. There is nothing in our account of linguistic theory to indicate that the result would be the description of a system that violates certain principles governing human languages. To the extent that this is true, we have failed to formulate the principles of linguistic theory, of universal grammar, in a satisfactory manner. In particular, we have not made use of the fact that the features have intrinsic content. (*SPE* 400)

Later in the chapter Chomsky and Halle themselves acknowledge that, with the above-quoted assertion, they are on the wrong track:

> It does not seem likely that an elaboration of the theory along the lines just reviewed will allow us to dispense with phonological processes that change features fairly freely. The second stage of the Velar Softening Rule of English (40) and of the Second Velar Palatalization of Slavic strongly suggests that the phonological component requires wide latitude in the freedom to change features, along the lines of the rules discussed in the body of this book. (*SPE* 428)

In other words, Chomsky and Halle ultimately recognize that the truly important parts of the phonology, in the sense of the ones that are unnatural, are those that cannot be derived from functional considerations of naturalness. This conclusion is echoed elsewhere: 'Where properties of language can be explained on such "functional" grounds, they provide no revealing insight into the nature of mind. Precisely because the explanations proposed here are "formal explanations", precisely because the proposed principles are not essential or even natural properties of any imaginable language, they provide a revealing mirror of the mind (if correct)' (Chomsky 1971: 44).

We propose that switching the feature coefficients as described in the first quotation might lead to the description of systems that are *diachronically* impossible human languages (ones that could never arise because of the nature of language change), but not to ones that are *computationally* impossible. The goal of phonological theory, as a branch of cognitive science, is to categorize what is a

computationally possible phonology, given the computational nature of the phonological component of UG.[7]

3.2. *Computation versus transduction: A place for substance*

It is important to note that the preceding discussion is not meant to imply that the mapping of sound to features is arbitrary. It is only the treatment of phonological representations within the computation that is arbitrary. Articulatory and acoustic substance *are* related to the representations we construct, but not within the grammar. The nature of this relationship is part of the theory of *transduction*—the mapping between the physical and the symbolic (Pylyshyn 1984). As Bregman (1990: 3) points out, 'In using the word representations, we are implying the existence of a two-part system: one part forms the representations and another uses them to do such things as calculate ...'. Bregman is concerned with the auditory system that does not have an output module—in discussing language, we also need to model output transducers that map from surface (featural) representations to articulatory gesture. For our purposes, Bregman's distinction corresponds to speech perception (construction of featural representations, ultimately from auditory signals) and grammar, which performs symbolic computation. We know from the existence of visual and auditory illusions that the transduction process is not simple. The perceptual system does not just form a direct record of physical stimuli. As Bregman points out, we know that representations are being constructed, because only then could they be constructed incorrectly, leading to illusions.

Pylyshyn (1984) provides the following discussion: 'This, then is the importance of a transducer. By mapping certain classes of physical states of the environment into computationally relevant states of a device [e.g. a human], the transducer performs a rather special conversion: converting computationally arbitrary physical events into computational events. A description of a transducer function shows how certain nonsymbolic physical events are mapped into certain symbolic systems' (p. 152). Pylyshyn points out that the '*computationally relevant* states are a tiny subset of [a system's] physically discriminable states' and that the 'former are typically a complex function of the latter' (p. 150). In (7) we paraphrase Pylyshyn's criteria for a psychological transducer (pp. 153–4)—that is a transducer from physical signals to representations.

(7) Criteria for a psychological transducer
- The function carried out by a transducer is itself *non-symbolic*; it is part of the functional architecture of the system.
- A transducer is stimulus bound, operating independently of the cognitive system.
- The behaviour of a transducer is described as a function from physical events to symbols:

[7] This argument, as well as other ideas in this chapter, was anticipated by Hellberg (1980). See also Burton-Roberts, this volume, Section 5.

(*a*) The domain of the function (the input) is couched in the language of physics.
(*b*) The range of the function (the output) must be computationally available, discrete atomic symbols (for example, feature matrices).
(*c*) The transformation from input to output must follow from the laws of physics.

This is where issues of substance arise—the physical aspects of the acoustic signal serve as the input into the transducer function. From that point on, in the manipulations of the constructed symbolic representations, substance is irrelevant to computation. Only the *formal* properties of such representations are relevant to the computational system.

It is worth contrasting Pylyshyn's well-articulated modular approach to that of Prince and Smolensky (1993), who directly reject the kind of extreme formalist position we advocate here.

'We urge a reassessment of this essentially formalist position. If phonology is separated from the principles of well-formedness (the "laws") that drive it, the resulting loss of constraint and theoretical depth will mark a major defeat for the enterprise' (Prince and Smolensky 1993: 198, see also p. 3). This view of the goals of phonology stems from a failure to observe the critical transducer versus grammar distinction—that is, from extensive 'substance' abuse. It is also at odds with the well-established goals of cognitive science in general: 'if we confine ourselves to the scientific and intellectual goals of understanding psychological phenomena [as opposed to predicting observed behaviour] one could certainly make a good case for the claim that there is a need to direct our attention away from superficial "data fitting" models toward deeper structural theories' (Pylyshyn 1973: 48). As our discussion of markedness below will indicate, we do not believe that any 'principles of well-formedness' exist, aside from those that constrain the set of possible representations. That is, we find the evidence for markedness-based constraints to be unconvincing.

The 'principles of well-formedness' that Prince and Smolensky refer to and adopt as the basis of OT constraints are merely derived from the heuristic devices that constitute the intuitions of an experienced linguist. For example, we may intuitively believe that a sequence like [akra] will more likely be syllabified as [a.kra] rather than as [ak.ra] in a random sample of grammars, although both syllabifications are found, for example, in the Ancient Greek dialects. Lacking information to the contrary, it may be useful to assume that the more common syllabification is present in a new, unfamiliar language. This will allow the formulation of hypotheses that may then be tested, and the guess will turn out to be correct more often than not, if our intuitions have any basis. However, it is a mistake to assume that our intuitions reflect the nature of the system we are studying in any direct manner. The intuition that heavy things fall faster than light things is very useful when someone drops something from a window, but the intuition needs to be transcended to understand the workings of gravity. Heuristics are used by the analyst to make useful guesses about data, and guesses

can be wrong. This is why OT constraints need to be violable, unlike all other scientific laws.

The pervasiveness of such 'data-over-principles' approaches to phonology can be appreciated by the following quote from an influential pre-OT paper: 'The goal of phonology is the construction of a theory in which cross-linguistically common and well-established processes emerge from very simple combinations of the descriptive parameters of the model' (McCarthy 1988: 84). By concentrating on what is 'common', rather than what is possible, phonology will provide (or rather has provided) plentiful material for descriptive work at some level of sophistication, but it is clear that no science should be concerned with making it particularly simple to express that which happens often. The goal of any science is to define a coherent domain of enquiry and to establish a common vocabulary for *all* events in that domain. This involves reducing the common *and* the rare events (for example, planetary motion and the Big Bang) to special cases of an abstract set of primitive notions. All of this suggests that, while a change of course for phonological theory was definitely needed in the early 1990s, OT has been a change in exactly the wrong direction.

3.3. *Acoustophilia: a warning*

Sapir (1925: 37) points out that 'it is a great fallacy to think of the articulation of a speech sound as [merely] a motor habit'. A corresponding error is committed in many of the studies (e.g. Flemming, forthcoming) that argue for the increased use of acoustic information to model human phonological computation. This work tends to establish units of analysis in terms of measurements taken over the acoustic signal itself. We believe that this technique shows the negative effects of 'acoustophilia'—the mental state arising from the deep and abiding satisfaction that comes from having *something* concrete to measure, in this case the acoustic signal. There is, we believe, a fairly serious difficulty with such an approach: we know with a great deal of confidence that human perception does not show the kind of direct dependency on the signal that the methodology of the acoustophiliacs requires.[8] This attitude towards the study of language echoes the overly positivist brand of empiricism adopted by the behaviourists, an attitude that was already discredited in the 1950s.

An example may make this clearer. Flemming (forthcoming) argues from an examination of F_2 interactions in an experimental setting that it is necessary to have the grammar generate a statistical pattern that forms a reasonable match to his experimental results. A parallel from the field of the cognition of vision would examine the properties of an image as measured with, for example, a photometer,

[8] Since phonetic substance provides the raw material for phonological theory construction, selective use of fine-grained acoustic data can give rise to insights into the nature of phonological computation. We recognize the significant body of work done on the phonetics/phonology interface with reference to acoustic studies. Keating (1988), which uncovers interesting phonetic regularities, but maintains a theory of phonology that makes no direct reference to this phonetic substance, is a brilliant example.

and require of us that our 'grammar of vision' generate a representation like that measured on the page. So, in Fig. 7.1, it would require—since the triangle we see is of precisely the same colour and brightness as the background (as can be verified by the use of a photometer)—that we construct a human visual system that does *not* see the triangle projecting from the page. This is of course the wrong result—the human visual system, given the input in Fig. 7.1, constructs a 'percept' that is very different from the patterns we might infer from photometric readings (see Hoffman 1998). The difficulty that this presents to more acoustically oriented approaches to *phonology* is fairly obvious: it is often claimed, on the basis of some physical measurement of the signal, that something is 'difficult' or 'easy' to perceive (auditorily), 'salient' or not so salient. Again, note that the edges and inside of the perceived triangle have absolutely no physical properties to distinguish them from the background. What the visual example in Fig. 7.1 shows us is that measurements taken over the raw data presented to the human auditory system should not be taken as direct evidence for what kind of data actually arrive at the *linguistic* processing system.

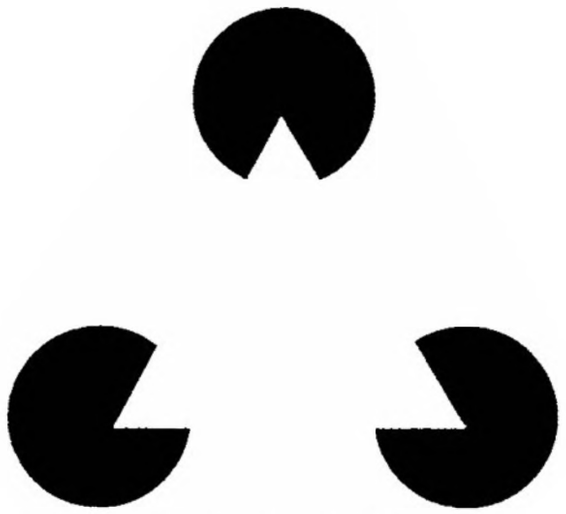

Fig. 7.1. Triangle constructed by visual system

Turning to the domain of auditory perception, it is a well-known result of psychoacoustics that the relationship between, say, intensity of a signal and *perceived* loudness is non-linear: doubling the physical intensity of a signal does not create a signal that is judged to be twice as loud. As we move further from the physical signal, to auditory perception and on to the construction of linguistic representations, things become even less clear. In particular, when several distinct and independent cues interact in the signal (as in the cases discussed by Steriade 2000), we cannot conclude without detailed and extremely difficult studies of the nature of auditory perception that we understand the way these cues interact to form an auditory percept. It is yet more difficult to determine how these auditory percepts get organized into *linguistic* (that is, featural, symbolic) representations. These topics will provide psychologically oriented phoneticians and their colleagues with challenging research projects for years to come. However, the

questions and the answers we hope to get are only distally related to the subject matter of phonology.

Part of the confusion in this area stems from the fact that discussion of 'output' forms often fails to distinguish between the output of the grammar (a feature-based representation) and, say, the output of the speaker (an acoustic or articulatory event). As demonstrated most clearly by our ability to construct 3D representations based on a black-and-white pattern on a printed page, there is a vast gap between physical stimuli/outputs and the internal (cognitive) representations that relate to them. Therefore, even if phonologists had a metric of the complexity or difficulty inherent in interpreting or creating certain physical stimuli or outputs (which they do not), it is apparent that there is no reason to believe that such a scale would translate straightforwardly to a markedness scale for representations. There is no reason to believe that the *representation* of the act of pushing a bar of gold is more difficult or complex or marked than the representation of the act of pushing a feather (cf. Burton-Roberts, this volume).

4. EXPLANATORY INADEQUACY

What are the implications of our view that phonology should be all form and no substance? In particular, does this conclusion about the nature of phonological operands have any positive implications for phonological theory? We think that there is one clear conclusion to be drawn. Since, as we have argued, languages appear to vary in some arbitrary ways (for example, inserting [r] and not, say, [t]), it is necessary to develop a theory that allows for such variation. In other words, the child should be equipped with a universal computational system and a set of primitives that can be modified upon exposure to positive evidence. For this reason, we believe that current versions of OT, which assume a universal set of (phonetically) substantive constraints (for example, *VoicedCoda, Lazy, and so on) do not shed light on the nature of grammar. A set of constraint templates, with principles of modification from which the learner can construct the necessary constraint inventory for the target language, may prove to be more useful. Similarly, a rule-based theory equipped with a set of principles for defining possible rules would also allow for the type of stipulative, cross-linguistic variation we have argued is necessary. Note that, given an explicit theory of acquisition, such a 'nativism *cum* constructivism' view of phonology is well constrained: UG delimits the set of possible rules or constraints; the data determine which rules or constraints are actually constructed.

In order to appreciate the fact that positing the type of substantive constraint found in the OT literature adds nothing to the explanatory power of phonological theory, consider the situation in which a learner finds himself or herself. Equipped with an OT-type UG, a child born into a Standard German-speaking environment 'knows' that voiced coda obstruents are 'marked'. However, this child never needs to call upon this knowledge to evaluate voiced coda obstruents, since there are

none in the ambient target language. In any case, by making use of positive evidence, the child successfully acquires a language like German. Born into an English-speaking environment, the child again knows that voiced coda obstruents are marked. However, the ambient language provides ample positive evidence that such sounds are present, and the child must override the supposed innate bias against voiced coda obstruents in order to learn English. So, this purported UG-given gift of knowledge is either irrelevant or misleading for what needs to be learned. Our substance-free theory of phonology shares with OT-type theories a reliance on positive evidence. The two theories have the same empirical coverage, since we also assume that both English and German are acquired. The difference is that we leave out of the genetic inheritance 'hints' that are irrelevant or misleading. We find our solution to be more elegant. Once again, note that this argument is equally applicable to markedness theories of all types, not just those couched within OT. Since markedness cannot have any bearing on learnability, it is probably irrelevant to any explanatorily adequate theory of grammar. We thus propose banishing markedness from consideration in future linguistic theorizing.[9]

5. DISCUSSION

The substance-abuse approach has been criticized for cognitive science in general by Pylyshyn (1984: 205 ff.). Pylyshyn describes a box emitting certain recurrent patterns of signals. He then asks what we can conclude about the nature of the computational mechanism inside the box, based on the observed pattern of output. The answer is that we can conclude nothing, since the observed patterns may reflect the nature of what is being computed (in his example, the output is a Morse Code rendering of English text, and the observed regularity is the 'i before e, except after c' rule), not the nature of the computer. In Pylyshyn's words, 'the observed constraint on [the system's] behavior is due not to its intrinsic capability but to what its states represent.' If we are interested in studying the phonology 'computer' then we need to distinguish a possible phonological computation from an impossible one. The set of attested phonological patterns and their distribution may be somewhat skewed by the sifting effect of language change. Real explanation of the nature of phonological computation requires us to see beyond such epiphenomena as 'markedness tendencies'.

[9] In fact, there are two distinct types of markedness in the phonological literature. This chapter is concerned with substantive markedness. Simplicity or evaluation metrics of the *SPE* symbol-counting type can be seen as measuring 'formal' markedness. We believe that the best approach to such formal requirements is to build them into the language acquisition device (LAD). Under this view, learners never compare extensionally equivalent grammars for simplicity or economy, they just construct the one that is determined by the LAD. There is, then, no reason to introduce the terms 'simplicity' and 'economy' into the theory, since they are contentless labels for arbitrary (i.e. not derivable) aspects of the LAD. For a concrete example of how we think the characterization of the LAD should be approached, see Hale and Reiss (forthcoming).

We believe that the current impregnation of the architecture of the phonological 'virtual machine' with phonetic substance represents a step backward for phonological theory. Phonologists should now call upon their impressive success in amassing descriptions of individual phonological 'programs' and aim for a more abstract, but deeper understanding of phonological computation.

Pylyshyn's example raises the question of whether constraints are appropriate elements for the construction of grammars at all. By defining grammars via constraints—that is, in negative terms—we are drawn into the problem of *inductive uncertainty*. In general, science works in terms of positive statements. A physical or formal system is defined in positive terms by a list of primitive elements, operations, relationships, and so on. The set of impossible chemical or physical processes, for example, is infinite, and so is the set of impossible linguistic structures.

Consider the question of hierarchical structure in syntax. Let us imagine that we want to express the claim that all structure is hierachically organized as a trait of UG. How should this proposal be formulated? If one seeks to characterize UG by listing constraints on the set of possible languages, then one might say something like 'Flat structure is not possible'. Since UG is instantiated in real brains, it must consist of a finite set of characteristic features. Note, however, that, using such negative constraints, we would actually need an infinite set of statements to characterize UG. This is because it is also the case that 'No language marks past tense by having the speaker eat a banana after uttering the verb', and 'No language requires that listeners look at a square to interpret iterativity', and so on are also true statements about human language. In other words, there is an infinite set of constraints on the set of possible languages.

These examples are, of course, preposterous, because in practice the constraints are stated in terms of a (usually implicit) universe of discourse. For example, the universe of discourse of linguistic theory does not include bananas, eating, seeing or squares. Therefore, a constraint is interpretable only in the context of a list of positive statements (such as a list of primitive elements like phonological distinctive features, and primitive operations like Move) that define the universe of discourse of any formal system.

We see, then, that a theory that formulates linguistic universals in terms of constraints must *also* contain a vocabulary of elements and operations in which those constraints are expressed, or to which they refer. This vocabulary of items and processes is presumably based on empirical observations and inferences. Consider a simpler alternative.

If our current hypothesis concerning UG is stated only in *positive* terms, without *negative* constraints, we can achieve a more economical model. The positive terms are just those entities and operations (features, deletions, insertions, Merge, Move, and so on) that have been observed empirically or inferred in the course of model construction. When faced with a phenomenon that is not immediately amenable to modelling using existing elements of the vocabulary, scientific methodology

(basically Occam's Razor) guides us. We must first try to reduce the new phenomenon to a description in terms of the vocabulary we already have. If this can be shown to be impossible, only then can we justify expanding the vocabulary.

Thus, a 'constraining approach' to UG, stated in terms of what is disallowed, requires a set of constraints and a vocabulary that defines the universe of discourse in which the constraints are valid. The alternative proposed here requires only the vocabulary of possible entities and operations, along with the metatheoretic principle of Occam's Razor. The alternative is thus more elegant and should be preferred.

In more concrete terms this means that our theory of UG should consist of the minimum number of primitives that we need to describe the grammars we have seen. Note that we should not be influenced in our search by preconceived notions of simplicity. For example, if we know that we need hierarchical structure for some phenomena, but there exist other phenomena that are ambiguous as to whether they require flat or hierarchical structure, then we should assume that the ambiguous cases also have hierarchical structure. If our current theory of UG contains an operation that only generates hierarchical structure from the primitive elements, constraints against flat structure will be superfluous. In fact, positive statements like 'structures are organized hierarchically' and 'all branching is binary' are also superfluous to grammar modelling (assuming they are correct), since they are just a reflection of how structure-building operations work.

The approach advocated here seems to be consistent with that used in science in general. If a physicist observes a constraint on the behaviour of a particle, say, then he or she posits a set of properties for that particle from which the observed behaviour emerges. The constraint thus has the status of a derivative and not primitive aspect of the theory. The arguments given here for constraints *on* grammars can be extended to apply to constraints *in* grammars as well, but this discussion is beyond the scope of the current paper (see Reiss 1999).

The issue of 'substance abuse' is closely tied to the use of constraints in phonological theory. Despite the fact that phonologists tend to characterize current debate concerning OT as a question of 'rules versus constraints', this is misleading. Many rule-based analyses make use of constraints such as the Obligatory Contour Principle (OCP). Constraints in otherwise rule-based phonologies serve two main purposes. Either they define certain structures as disfavoured or ill-formed, and thus subject to modification by rule; or they are used to block the application of a rule just in case the rule's output would be disfavoured or ill-formed. Work by Paradis (1988) and Calabrese (1988) are typical of the use of constraints as diagnostics for repair of certain structures. The rule-based account of stress systems presented by Halle and Idsardi (1995) appeals to 'Avoidance Constraints' (pp. 422 ff.) that prevent the application of rules in cases where the rules' output would be a 'disfavoured' structure. The OCP has been invoked for both of these purposes in a number of papers, most notably McCarthy (1986) and Yip (1988), who makes the following remark: 'The main

contribution of the OCP is that it allows us to separate out condition and cure. The OCP is a trigger, a pressure for change ...' (p. 74).

Given the problems with markedness theory alluded to above, note that, in the absence of a theory of disfavouredness, this approach is slightly circular: the only real evidence for the disfavoured status is that the posited rule appears to be blocked; and the reason for the blocking is that the resultant structure would be disfavoured. Halle and Idsardi point out that certain advantages derive from mixing rules with constraints in the analysis of individual languages. In general, the use of constraints allows us to formulate simpler rules. However, they note that a fully rule-based analysis is in principle always possible—Halle and Vergnaud (1987) is an example they cite. We propose that considerations of elegance for a theory of UG take precedence over elegance in the analysis of individual languages, and thus the Halle and Idsardi system, for example, should be adapted in a way that preserves its mathematical explicitness, while doing away with constraints on unattested structures. In general, a goal of future phonological research should be to take the idea of rule-based phonology seriously—by avoiding constraints altogether. Such an approach will offer a principled alternative to OT and other constraint-based models. In other words, rather than stating simple, but empirically inadequate rules, reinforced by an arsenal of language-particular or universal constraints, we should attempt to understand what kind of rules we actually need if we are to do without constraints. An example of this approach is discussed in the next section.

6. A RESULT IN THE FORMAL CHARACTERIZATION OF UNIVERSAL GRAMMAR

In order to show that there is progress to be made in the characterization of formal properties of UG consider a limited type of condition on rule application (or constraint applicability). Vowel syncope rules are found with (at least) all three of the following types of conditioning:

(8) Some conditions on vowel deletion rules (Odden 1988: 462)
  (*a*) Delete a vowel unless flanking Cs are identical.
  (*b*) Delete a vowel blindly (whatever the flanking Cs are).
  (*c*) Delete a vowel only if flanking Cs are identical.

Condition (*a*) can be restated as 'Delete a vowel if flanking Cs are *not* identical'. Thus, (*a*) demands non-identity and (*c*) demands identity of segments in Structural Descriptions (SDs). Phonological formalism must, therefore, have at least enough power to express conditions of non-identity and identity. These conditions may also be restricted to a given subset of phonological features, such as the set of Place features.

Autosegmental representation can represent (c) using linked structures—two *C*-slots may be linked to a single-feature tree or matrix. Alternatively, two slots

may be explicitly linked to separate, but identical trees/matrices. However, (*a*), the requirement of non-identity, cannot be represented using just autosegmental notation. This is because non-identity can be due to a disagreement with respect to *any* feature, and autosegmental notation does not make use of variables. In order to represent conditions of non-identity, Reiss (1999) makes use of a system of Feature Algebra (FA) incorporating the existential and universal quantifiers. FA allows the formulation of conditions that have traditionally been notated as, say, $C_1 \neq C_2$ and $C_1 = C_2$. The conditions are stated here in prose form:

(9) Attested conditions of rule application
  (i) The NON-IDENTITY CONDITION (encompasses condition (*a*))
  There exists some feature F, such that C_1 and C_2 have opposite values for F.
  (ii) The IDENTITY CONDITION (encompasses condition (*c*))
  For all features F, C_1 and C_2 have the same value.

In both conditions the set of features over which non-identity or identity is computed may be a subset of the total feature set. For example, an identity condition may be applicable only to the set of Place features in a given rule.

Reiss (1999) applies the FA formalism to data presented by McCarthy (1987), Yip (1988) and Odden (1986, 1988) in their arguments concerning the status of the OCP as a principle of grammar. The use of FA notation has several benefits. First, it provides us with counter-arguments to Yip's claim that the effects of, for example, the IDENTITY CONDITION should not be built into SDs. Secondly, it allows us to evaluate the status of constraints like the OCP in the light of data conforming to the apparently contradictory conditions (*a*) and (*c*). Thirdly, the formalism helps us to discover that two other formally similar conditions are unattested.

(10) Unattested conditions on rule application
  (i) COMPLETE NONIDENTITY CONDITION
  For all features F, C_1 and C_2 have the opposite value for F.
  (ii) VARIABLE PARTIAL IDENTITY CONDITION
  There exists some feature F, such that C_1 and C_2 have the same value for F.

The COMPLETE NON-IDENTITY CONDITION would allow a rule deleting a vowel only if flanking segments have opposite values for, say, all Place features, or even for *all* features; for example, 'Delete a vowel in the environment C_1__C_2 if C_1 is [−anterior, −labial, +dorsal] and C_2 is [+anterior, +labial, −dorsal], or C_1 is [+anterior, −labial, +dorsal] and C_2 is [−anterior, +labial, −dorsal], and so on.

The VARIABLE PARTIAL IDENTITY CONDITION would allow, say, a rule that deleted a vowel only if flanking consonants have the same value for *any* feature (perhaps in a given subset of features): 'Delete a vowel in the environment C_1__C_2 if and only if C_1 and C_2 are both [αanterior], or [αlabial], or [αdorsal], and so on'.

It turns out that, while these two conditions are apparently unattested in phonology, they are used in the interpretation of binding relations. Thus a careful consideration of the formal requirements of UG can lead to interesting results. It

should be satisfying enough to get a handle on what we know UG can do, what its formal properties are, without worrying about what it cannot. In this sense, positive characterizations of grammars are to be preferred to constraint-based ones.

7. THE MIRAGE OF ENHANCEMENT

A particularly illustrative combination of what we consider to be the misuse of substantive considerations and functionalism can be found in the literature on phonetic enhancement and the maximization of contrast (e.g. Stevens *et al.* 1986). For example, the tendency of three-vowel systems to contain the maximally distinct set /i,u,a/ is taken as a reflection of a phonological principle demanding the 'best' use of the available acoustic space. Like other claims concerning markedness and UG, this pattern is no more than a tendency. However, we can show that the view of markedness as an emergent property, outlined above, can give insight into this statistical pattern. Imagine a language $\mathcal{L}_1$ that had the four vowels /i,u,e,a/. Now we know that merger of acoustically similar vowels (such as /i/ and /e/) is a common diachronic process. It would not be surprising if a learner constructing $\mathcal{L}_2$ on the basis of data from speakers of $\mathcal{L}_1$ were to fail to acquire a slight distinction and end up with a three-vowel system containing /i,u,a/. However, it is much less likely that the learner would fail to acquire an acoustically more robust distinction like /u/ versus /a/ and end up with an inventory containing, say /i,u,e/.[10] So, vowels that are close together in the acoustic space are likely to merge diachronically. Vowels which are acoustically distant are not likely to merge diachronically. The observed pattern of maximal contrast is thus not built into the phonology, but is an emergent property of the set of observed phonological systems owing to the nature of diachronic sound change.

8. FUNCTIONALISM AND DYSFUNCTIONALISM

The rise of OT has been accompanied by a revival of functionalism in phonology. In fact, there is no necessary connection between OT as a theory of computation and functionalist reasoning, and an OT proponent might invoke what we call the National Rifle Association defence ('Guns don't kill people; people kill people'): computational theories are not inherently functionalist, people are functionalist. However, the ease with which functionalist ideas can be implemented in OT has clearly invited this 'functionalist' explosion and may bear on the question of whether or not the theory is sufficiently constrained or even constrainable. Note also that the 'logic' of functionalism (namely, that *all* phenomena are explicable by

[10] Note that 'phonetic substance' may itself indicate how weak the reasoning is in this case: English [i], as well as the other front vowels, is significantly lower than Danish [i]. Why is the 'maximization of contrast' not active at the phonetic level—precisely the level that provides the alleged 'substance' (perceptual distinctness, in this case) for the functionalist claim?

reference to competition between universal, but violable, principles) is identical to the logic of OT. In this section we briefly show that the 'substance' orientation of functionalism can be turned on its head to yield a theory that we will dub 'dysfunctionalism'.

Many functionalist theories of grammar can be summarized in almost Manichean terms as consisting of a struggle between the 'competing forces' of ease of articulation (what is presumed to be 'good' for the speaker) and avoidance of ambiguity (what is presumed to be 'good' for the hearer). As an example of the former, consider Kirchner's (1997: 104) constraint 'LAZY—Minimize articulatory effort'. For the avoidance of ambiguity, consider Flemming's (forthcoming) MAINTAIN CONTRAST constraints, which are violated by surface merger of underlying contrasts.

The interplay of what is 'good for' the speaker and what is 'good for' the hearer supposedly gives rise to the patterns we see in language: sometimes mergers occur and the speaker's output is 'simplified'—potentially creating a difficulty for the hearer; sometimes the speaker maintains distinctions, perhaps producing a more 'complex' output, thus avoiding ambiguity for the hearer.[11]

The problem with this theory is that functionalist principles can be replaced by their opposites, which we will call 'dysfunctionalist' principles, with no significant change in the set of grammars predicted to exist. Consider the following principles, proposed by a linguist with a different view of human nature than the functionalists have.

(11) Principles of dysfunctionalism
OBFUSCATE: merge contrasts, use a small inventory of distinctive sounds, and so on.
NO PAIN–NO GAIN: maintain contrasts, use a large inventory, generate allomorphy, and so on.

Merger, widely attested in the languages of the world, as well as the oft-proclaimed diachronic principle that 'change is simplification', will be accounted for by the (dys)functional requirement that one should OBFUSCATE. The failure of merger, equally well attested, and the generally ignored diachronic process of 'complexification', will be attributed to the effects of the NO PAIN–NO GAIN Principle. The competition of these two 'dysfunctionalist' principles will thus lead to the exact same results as the usually cited functionalist principles. While the ultimate question of whether human beings are fundamentally lazy, but helpful, or something seemingly more perverse is intriguing, it hardly seems as though investigation into such matters should form the foundation of a theory of

[11] Further evidence for the incoherence of the functionalist position is the fact that 'careless' speech can often lead to supposedly complex outputs such as the stop cluster in [pt]*ato* for *potato*. Onset stop clusters are not found in careful speech, so it is surprising, from a functionalist perspective, that they should be found precisely when the speaker is not putting forth greater articulatory effort.

phonological computation.[12] We propose, therefore, that functionalism provides no insight into the nature of grammar. Again, we propose leaching all substance out of phonology in order better to observe the abstract computational system.

The alternative—which seems to be the focus of many current developments in phonological theory—seems clear. Given a sufficiently rich and explicit theory of the human personality (giving us principles such as 'be lazy' and 'be helpful to the listener') and the human articulatory and perceptual systems ('phonetic' substance), phonology itself will turn out to be epiphenomenal. While this seems considerably less promising to us, it has clear implications for the research strategy that phonologists should adopt. Phonologists, under such a view, should focus their energies in two domains: phonetics and the empirical explication of fundamental features of the human personality ('laziness,' 'helpfulness,' and so on).

The anti-functionalist stance taken here is, of course, not new. For example, Halle (1975: 528), points out that, 'Since language is not, in its essence, a means for transmitting [cognitive] information—though no one denies that we constantly use language for this very purpose—then it is hardly surprising to find in languages much ambiguity and redundancy, as well as other properties that are obviously undesirable in a good communication code.' Halle suggests that it is more fruitful to conceive of language as a kind of mathematical game than to concern ourselves with the 'communicative functions' approach to studying language. The latter viewpoint led to such dead ends as the application of formal information theory to natural language.

CONCLUSIONS

We are advocating that phonologists, qua phonologists, attempt to explain less, but in a deeper way. As we hope to have indicated, empirical results provided by phoneticians and psycholinguists contribute to the development of a substance-free phonology, and we look forward to important cooperation with scholars in these fields. We recognize that only they can provide explanation for many (E-language) generalizations that are striking in their statistical regularity.[13] Since

[12] The authors would be happy to provide examples—drawn from the history of linguistic theory—of the evolutionary advantages of self-interested effort (NO PAIN–NO GAIN) and OBFUSCATE. We refrain for reasons of space, fully confident that the reader will have no difficulty generating ample evidence on his or her own.

[13] But see Engstrand (1997*a*,*b*) for arguments that the statistics may be misleading. For example, the purported markedness of /p/, as evidenced by its relative rarity in voiceless stop inventories, *vis-à-vis* /t/ and /k/, is probably illusory. The overwhelming majority of the languages in a database like UPSID (Maddieson 1984; Maddieson and Precoda 1989) lacking a /p/ are found in Africa. Similarly, the languages of Africa do not 'avoid' voiced velar stops, which are also commonly assumed to be marked (see Hale, forthcoming *b*). 'Thus, it cannot be concluded that velars and bilabials constitute underrepresented members of the respective voiced and voiceless stop series. Although this pattern is to be expected from proposed production and perception constraints, it is largely overridden by areal biases' (Engstrand 1997*a*: 1).

we believe that the focus of phonological theory should be on the cognitive architecture of the computational system, we also believe that the non-substantive aspects of OT have been tremendously important for the development of the field. The best of the OT literature is far more explicit about the nature of the assumed computational system than its predecessors often were. The mere existence of such a well-developed alternative to rule-based phonology is valuable, regardless of specific formal problems (for example, synchronic 'chainshifts') or the 'substance abuse' found in any particular implementation. However, we have also raised the question of whether constraints are appropriate entities for scientific modeling, since they must always be accompanied by a somewhat redundant positive characterization of a universe of discourse.

REFERENCES

Beckman, J. N. (1997), 'Positional Faithfulness, Positional Neutralisation and Shona Vowel Harmony', *Phonology*, 14: 1–46.

Bregman, A. (1990), *Auditory Scene Analysis* (Cambridge, Mass: MIT Press).

Burton Roberts, N. (this volume), 'Where and What is Phonology? A Representational Perspective'.

Calabrese, A. (1988), 'Towards a Theory of Phonological Alphabets', Ph.D. dissertation (MIT).

Chomsky, N. (1971), *Problems of Knowledge and Freedom* (New York: Random House).

—— (1986), *Knowledge of Language* (New York: Praeger).

—— and Halle, M. (1968), *The Sound Pattern of English* (New York: Harper & Row).

Engstrand, O. (1997*a*), *'Areal Biases in Stop Paradigms', in Papers from Fonetik 97, The Ninth Swedish Phonetics Conference*, held in Umeå, 28–30 May 1997. Reports from the Department of Phonetics, Umeå University (PHONUM), iv. 187–190.

—— (1997*b*), 'Why are Clicks so Exclusive?' in *Papers from Fonetik 97, The Ninth Swedish Phonetics Conference*, held in Umeå, 28–30 May 1997. Reports from the Department of Phonetics, Umeå University (PHONUM), iv. 191–4.

Flemming, E. (forthcoming), 'Phonetic Detail in Phonology: Toward a Unified Account of Assimilation and Coarticulation', in K. Suzuki and D. Elzinga (eds.), *Coyote Papers , Proceedings of the Arizona Phonology Conference 5, Features in Optimality Theory.* University of Arizona, Tucson.

Hale, M. (forthcoming *a*), 'What is output?' in M. Halle, C. Reiss, and B. Vaux (eds.), *Phonology 2000* (Oxford: Oxford University Press).

—— (forthcoming *b*), *Theory and Method in Historical Linguistics* (Oxford: Blackwell).

—— and Reiss, C. (2000), 'Substance Abuse and Dysfunctionalism: Current Trends in Phonology', *Linguistic Inquiry*, 31: 157–69.

—— —— (forthcoming), 'Grammar Optimization', *Language Acquisition.*

Halle, M. (1975), 'Confessio Grammatici', *Language*, 51: 525–35.

—— and Idsardi, W. (1995), 'Stress and Metrical Structure', in J. A. Goldsmith (ed.), *Handbook of Phonological Theory* (Oxford: Blackwell).

—— —— (1997), '*r*, Hypercorrection and the Elsewhere Condition', in I. Roca (ed.), *Derivations and Constraints in Phonology* (Oxford: Clarendon Press), 331–48.

—— and Vergnaud, J. R. (1987), *An Essay on Stress* (Cambridge, Mass.: MIT Press).

Hellberg, S. (1980), 'Apparent Naturalness in Faroese Phonology', *Nordic Journal of Linguistics*, 3: 1–24.

Hoffman, D. D. (1998), *Visual Intelligence* (New York: Norton).

Keating, P. A. (1988), 'Underspecification in Phonetics', *Phonology*, 5: 275–92.

Kirchner, R. M. (1997), 'Contrastiveness and Faithfulness', *Phonology*, 14: 83–111.

LaCharité, D. and Paradis, C. (1993), Introduction: The Emergence of Constraints in Generative Phonology and a Comparison of Three Current Constraint-Based Models', *Canadian Journal of Linguistics*, 38: 127–51.

McCarthy, J. (1986), 'OCP Effects: Gemination and Antigemination', *Linguistic Inquiry*, 17: 207–630.

—— (1988), 'Feature Geometry and Dependency: A Review', *Phonetica*, 45: 84–108.

—— (1993), 'A Case of Surface Rule Inversion', *Canadian Journal of Linguistics*, 38: 169–95.

—— and Prince, A. (1995), 'Faithfulness and Reduplicative Identity', in J. Beckman, L. Walsh-Dickey, and S. Urbanczyk (eds.), *University of Massachusetts Occasional Papers (UMOP)* 18 (Amherst), 249–384. Rutgers Optimality Archive, http://ruccs.rutgers.edu/roa.html.

—— (1996), 'Remarks on Phonological Opacity in Optimality Theory', in *Proceedings of the Second Colloquium on Afro-Asiatic Linguistics.* J. Lecarme, J. Lowenstamm, and U. Schlonsky (eds.), *Studies in Afro-Asiatic Grammar, Papers from the Second Conference on Afro-Asiatic Languages, Sophia, Antipolis, 1994* (The Hague: Holland Academic Graphics), 215–43.

Maddieson, I. (1984), *Patterns of Sounds* (Cambridge: Cambridge University Press).

—— and Precoda, K. (1989), 'Updating UPSID', *Journal of the Acoustical Society of America*, suppl. 1, Vol. 86, S19.

Odden, D. (1986), 'On the Obligatory Contour Principle', *Language*, 62: 353–83.

—— (1988), 'Antiantigemination and the OCP', *Linguistic Inquiry*, 19: 451–75.

Ohala, J. J. (1990), 'The Phonetics and Phonology of Aspects of Assimilation', in J. Kingston and M. Beckman (eds.), *Papers in Laboratory Phonology I* (Cambridge: Cambridge University Press).

—— (1998), 'A Bibliography of the Phonetics of Sound Change', Department of Linguistics, University of California, Berkeley.

Paradis, C. (1988), 'On Constraints and Repair Strategies', *Linguistic Review*, 6: 71–97.

Prince, A. and Smolensky, P. (1993), *Optimality Theory: Constraint Interaction in Generative Grammar* (Rutgers Center for Cognitive Science, Technical Report, 2; Rutgers Univesity, New Brunswick, NJ).

Pylyshyn, Z. (1973), 'On the Role of Competence Theories in Cognitive Psychology', *Journal of Psycholinguistic Research*, 2: 21–50.

—— (1984), *Computation and Cognition: Toward a Foundation for Cognitive Science* (Cambridge, Mass: MIT Press).

Reiss, C. (1996), 'Deriving an Implicational Universal in a Constrained OT Grammar', *NELS* 26: 303–17.

—— (1997), 'Unifying the Interpretation of Structural Descriptions', *West Coast Conference on Formal Linguistics*, 15: 413–27.

Reiss, C. (1999), 'Philosophical and Empirical Reasons to Ban Constraints from Linguistic Theory', MS Concordia University.

Sapir, E. (1925), 'Sound Patterns in Language', *Language*, 1: 37–51. Repr. in Martin Joos (ed.), *Readings in Linguistics* (Washington: American Council of Learned Societies, 1957), 19–25.

Steriade, D. (2000), 'Paradigm Uniformity and the Phonetics–Phonology Boundary', in M. B. Broe and J. B. Pietrehumbert (eds.), *Papers in Laboratory Phonology V* (Cambridge: Cambridge University Press), 313–34.

Stevens, K., Keyser, S. J., and Kawasaki, H. (1986), 'Toward a Phonetic and Phonological Theory of Redundant Features', in J. Perkell and D. Klatt (eds.) *Symposium on Invariance and Variability of Speech Processes* (Hillsdale, NJ: Lawrence Erlbaum).

Yip, M. (1988), 'The Obligatory Contour Principle and Phonological Rules: A Loss of Identity', *Linguistic Inquiry*, 19: 65–100.