Derivations and Constraints in Phonology

✦

Edited by
IGGY ROCA

CLARENDON PRESS · OXFORD
1997

3

The Contents of Phonological Signs: A Comparison Between Their Use in Derivational Theories and in Optimality Theories

SYLVAIN BROMBERGER and MORRIS HALLE

1. Introduction

In all the exchanges that we are aware of between proponents of Rule-based Derivational phonology (DT, from here on) and proponents of Optimality phonology (OT from here on), the participants seem to take for granted that the phonological symbols on which both sides rely mean the same thing in both kinds of theories. In other words, they seem to take for granted that switching between OT and DT analyses has no effect on the semantic values that attach to such symbols as 'æ', 'œ', '[+ voice]', '[labial]', and so on: that these and other phonemic and feature symbols carry the same information, have the same semantic content, regardless of whether they occur in the context of a derivation or in the context of a tableau. In this chapter we want to question this presumption of semantic invariance between OT and DT. We will, in fact, argue that many—perhaps all—such symbols mean different things in the two kinds of contexts.

Questioning the presumption of semantic invariance between DT and OT is not an innocuous move. It has deep consequences for how we should go about gauging the relative merits of contributions to the two approaches. To see why, we need but imagine a world much like this one, but in which proponents of DT and proponents of OT use symbols that not only have different meanings but also look different. In such a world we wouldn't begin to compare their contributions without first finding out whether their respective predicates were designed to cover the same sorts of entity and, if so, whether on the basis of similar or different aspects. We would need that knowledge to determine whether the two approaches are in conflict, and if so on what points, or whether they are compatible—and if so whether because they supplement each other, or because their assumptions are mutually irrelevant. In other words, in such a world, we would have to become explicit about how the meanings of their symbols are related before embarking on a critical comparison or even deciding on appropriate grounds for comparison: truth, plausibility, simplicity, generality, convenience, explanatory depth, all of these, some of these? But questioning the presumption of semantic invariance is tantamount to looking at the real world as differing from that imaginary one in only a superficial respect, namely in that the proponents of the two approaches happen—perhaps for historical reasons—to use symbols of the same shape. They might as well have used symbols of different shapes. In fact that might have been less misleading. And so, in this the real world as in our imaginary one, we have to determine how the meanings that attach to their symbols are related before engaging in critical comparison or even deciding on what grounds they should be compared. Of course, even if the symbols have different meanings in the two approaches, it does not follow that these meanings must be totally unrelated or that the two approaches cannot stand in any kind of conflict. In the last part of the chapter we will, in fact, bring out points of contact between these semantic values and give a reason why a DT approach might be preferred to an OT one. It does follow that any critical comparisons should take into consideration the difference between the semantic values that each assigns to the same signs as well as the range of phenomena they appear to be covering at any particular time.

This chapter differs thus from others in this collection in that it focuses on the symbols used in the analysis of linguistic facts rather than on specific analyses. Phonological theories, as usually presented are—from a philosophical perspective at least—frustratingly unspecific about the semantics of their notation and hence about what exactly they describe. That is—up to a point—a good thing: issues about the meanings and references of the terms one uses can be tedious, elusive, suspect, divisive, inhibiting, and empirically unproductive, to say the least. And since one can achieve much in phonology, as in other sciences, while ignoring such issues, they can also seem pointless. There is, furthermore, the danger of addressing them prematurely, while the discipline and its connection to neighboring ones is still in too much flux. But ignoring such issues also has a downside: it leaves the *truth conditions* of phonological claims unspecified, that is, what the world must contain and be like if these claims are true, and what it might not contain or not be like, if they are false. That is a loss. At best it leaves us with only a partial understanding of these claims, at worst it invites confusion about their relationship to each other and to the evidence that allegedly supports them. Thus to see clearly what, if anything, about our understanding of reality is at stake in a debate such as the debate between the proponents of DT and OT, we need to know as explicitly as we can how their symbols relate to that reality, and not just how their adherents analyse representations whose links with it are left in the dark. And what could be of more interest in such a debate than its bearing on our understanding of what is real? So this chapter should be read—like

other papers on which the two of us have collaborated—with foundational rather than empirical questions in mind. The two families of questions must obviously be brought together eventually. But much preliminary work needs to be done first. This chapter belongs to the too often neglected foundational side.

In what follows we first present certain assumptions which need to be made explicit and justified; we next examine consequences of these assumptions for the semantics of DT; we then examine corresponding consequences for OT and bring out that the semantic values of similar phonological symbols are not the same in both contexts. We end with a discussion of how the two approaches might nevertheless be critically compared. Throughout the discussion we rely on extremely simplified versions of DT and OT, on the ground that these already contain all the elements needed to make our point. We limit ourselves to phonemic and feature symbols for the same reason.

2. Assumptions

We begin by putting forth three assumptions on which our discussion depends. We think that these assumptions are unproblematic and, in principle at least, widely accepted. But since they are seldom openly stated, and are often disregarded in practice, they may be more controversial than we think.

I. Our *first assumption* is that phonological symbols, that is, phoneme symbols, feature symbols, prosodic segmentation brackets, stress diacritics, and so on, can be replaced without loss of meaning by *predicates*, can be thought of as abbreviations for, as convenient stand-ins for, predicates. We will limit our attention here to phonemic and feature symbols, but we think that this assumption can eventually cover other symbols as well.

The term 'predicate' is used with different connotations in different theoretical settings. We use the term as it is used in predicate logic. In predicate logic, expressions called predicates are characterized principally by two traits that matter to us. (*a*) Predicates do not denote individual objects or events, but are *true of* them. So, for instance, the predicate 'hot' does not denote anything but is true of each hot thing, and fails to be true of cold things. (*b*) Predicates are associated with *satisfaction conditions*, conditions that define what they are true of. So, for instance, the predicate 'hot' is associated with the condition or property of being hot and is true of all and only objects that satisfy that condition. In short, the predicate 'hot' is true of objects that are hot.

Predicates, as we use the term, are thus to be distinguished from so-called *individual constants*, that is, from names or singular terms like 'Hillary Rodham Clinton', or '5', from expressions that, unlike predicates, purport to denote, to name individual objects, rather than to be true of them. The difference becomes obvious when we think of what is involved in specifying the semantics of the two classes of expressions. Giving the semantics of an individual constant requires one to specify its reference, if any, whereas giving the semantics of a predicate requires one to specify its satisfaction conditions.

In canonical notation—when their syntax is made explicit—predicates combine with two other kinds of expressions to form statements: individual constants and quantifiers. So if, in accordance with a well-known notational system, we let 'H' stand for the predicate 'hot' and 'a' stand for the name of some object a, then 'Ha' states that a is hot, and '(∀x)Hx' states that everything is hot.

In what follows we will not use the notational system used in the example above, but will use *lambda notation* instead. Though this notation may seem rebarbative at first, it is actually quite easy to decode and will be much easier to relate to standard phonological symbols than other notations. So, instead of using 'hot' or 'H' as above, we use 'λx[hot x]' to write the predicate, and 'λx[hot x](a)' to state that a is hot, and '(∀y){λx[hot x](y)}' to state that everything is hot.[1]

As can be seen from this example, the lambda notation lets one incorporate familiar terms (e.g. 'hot') in more formal symbols. This turns out to be very helpful when we want to display the predicate nature of phonological terms. So instead of the equivocal symbol 'æ' in isolation—which underspecifies whether it stands for an individual constant or a predicate—we will use 'λx[æ x]'; and instead of '[+round]'—whose logical form is totally obscure—we will use 'λx[+round x]'.[2] Thus to say of something called 'α' (we postpone for the time being saying what α might be) that it is [+round], we will write 'λx[+round x](α)', and to state that something is [+round] we will write '(∃y){λx[+round x](y)}'.

The reasons for holding that phonological symbols are stand-ins for predicates in the above senses are straightforward. The alternative is to hold that they are singular terms, that is, names. But what could they be names of? Not of particular fragments of individual utterances produced at some specific time and specific place by some specific person. Even if one could make sense of that idea—for instance, of the idea that '[+round]' or even 'æ' is the name of a part of an utterance, which utterance among the zillion that have been produced or that will some day be produced would it be the name of? Nor can they plausibly be names of abstract objects, of objects located neither in space nor in time. That would not be compatible with the fact that phonology is an empirical science and thus dependent on causal interactions between observers and what is under investigation. Abstract objects cannot stand in any such

[1] Predicates in lambda notation combine with quantifiers and other logical operators in the same way as predicates in the standard predicate notation.

[2] In this system of notation, variables can be freely substituted for each other. Thus 'λx[+round x]' and 'λy[+round y]' are synonymous. It is also best, to avoid confusion, not to use the same variables after quantifiers and after lambdas.

relation to observers. They are, by definition, causally inert. Other objects may come to mind, but none of those that we are aware of seems to survive close scrutiny. Furthermore, the symbols we need must express something that distinct actual utterances can have in common, that can be aspects shared by distinct utterances. Unless one is willing to indulge in weird metaphysical ontologies, predicates seem best suited for that function.

Phonologists, obviously, do not display their symbols in ways that commit them to a predicate interpretation. Nor do they manipulate their symbols in ways that look like any of the known predicate calculi. But that by itself does not entail that the predicate interpretation is unwarranted. There are historical reasons for their reliance on a more neutral system of notation. And there is also a good rationale for this practice: using and manipulating predicate notation would greatly complicate their presentations and their computations. Using predicate notation—and paying the cost in complexity—is useful when we need to be forthcoming about the truth conditions of phonological claims, that is, when we need to be *explicit* about their semantic value. On most other occasions one need not be so finicky, and other notational devices are more effective.

II. Our *second assumption* may seem self-evident, but will turn out to have rather surprising consequences. It is that *within any theoretical approach to phonology*, be it DT or OT—or any other approach for that matter—any particular phonological symbol stands for the same predicate in *all* contexts. By 'same predicate' we mean predicate having the *same defining satisfaction conditions*. Phonological symbols such as 'æ' and '[+round]' occur standardly in the description of underlying representations, in the descriptions of surface representations, in the descriptions of items in the lexicon, in the statement of rules, in the formulation of constraints, and so on. Our second assumption is that the semantic content—the satisfaction conditions—of any given symbol (now thought of as predicates) must be the same in all these contexts *within any single theoretical approach*, and a fortiori within any given analysis adopting that approach. In other words, we assume that within DT and within OT phonological symbols are used unambiguously.[3]

On the other hand, we do *not* assume that similar phonological symbols stand for the same predicates *across theoretical approaches*, across DT analyses and OT analyses. That, obviously, is the very assumption that we intend to deny. And there is no a priori reason to assume that it is true, even if it may have gone unquestioned in the literature. The assumption about unambiguity that we are making, on the other hand, follows from the presumption that DT and OT analyses, whatever their shortcomings, are both at least semantically

[3] This assumption should actually be slightly modified for descriptions of items in the lexicon. But this modification would not crucially affect our present discussion, and will be ignored here. We will come back to it on some other occasion.

coherent. Derivational analyses are committed to semantic unambiguity, by the nature of rules, which are blind to the derivational history of their inputs and outputs and must apply whatever that history. The symbols in the statements of rules and the symbols in the formulation of the various levels of a derivation must thus be taken to have the same meaning. Optimality analyses are committed to semantic unambiguity if by nothing else, at least by the prominence they give to faithfulness constraints. Symbols in the statements of these constraints and in the formulations of the inputs and outputs of GEN must be taken to have the same meaning.

III. Our *third assumption, this-worldly realism,* is trickier to state succinctly without getting lost in labyrinthine philosophical questions and qualifications. Its basic tenet is, however, clear. We assume that the statements[4] of a plausible phonological theory—and we deem both OT and DT to be plausible—*purport to be true*, and furthermore purport to be about *things in this our actual spatio-temporal world.* This double assumption, note, in turn entails the further assumption that phonological predicates, insofar as they are ever true of anything, are true of datable, placeable things such as, for instance, individual actions, or events, or mind/brain phases, or specific people at specific times—in short, true of things in this the actual world.

One way of making clear the point of that assumption, and of bringing out, perhaps, its controversial nature, is by contrasting it with two alternatives which some may deem plausible.

One alternative—which is sometimes labelled *instrumentalism*—would flatly deny that all the statements of DT and OT purport to be true. It would hold instead that DT and OT provide essentially symbol-manipulating recipes that put *linguists* in a position to compute certain outputs from certain inputs in more or less economical ways but without purporting to contain information beyond that contained in these inputs and outputs.[5]

Instrumentalists would look, for instance, at a DT derivation as a computation performed by *a linguist* (qua theoretician, not qua speaker) who, by manipulating certain symbols in accordance with certain rules, was able to calculate in a more or less efficient way the description of a surface representation from the description of an underlying representation. Such instrumentalists would deny that anything could be claimed for the derivation itself beyond the fact that it enabled a practitioner to pair the right descriptions. They would deny that any truth claim need be warranted on behalf of intervening representations, or on behalf of their ordering, or on behalf of the rules, and so on. For such instrumentalists it would make sense to ask whether the description of the underlying representation in a derivation was true or warranted; it would

[4] Except perhaps certain conditionals. But those should be irrelevant for the rest of the discussion.
[5] There are some who would even hold that not even the inputs but only the outputs purport to be true.

also make sense to ask whether the description of the surface representation was true or warranted; but it would make no sense to ask whether the steps in the derivation represented anything real. And such instrumentalists would, in turn, look at a derivational *theory*—that is, at any specific version of DT—as simply a recipe, or set of recipes, for constructing derivations. They would deny that the statements of the theory had more than instrumental value, were either true or false. They would interpret the DT agenda as that of providing computational recipes. Instrumentalists expect efficiency of phonological theories, not truth or insight.

Similarly, such instrumentalists would look at an OT tableau as no more than a device that a linguist (again, qua theoretician, not qua speaker) might use to calculate the description of a surface representation from the description of an underlying representation; they would deny that any truth claim about the tableau would be warranted beyond the fact that it enables *linguists* to pair the right descriptions. And such instrumentalists would look at any specific version of OT as a recipe, or set of recipes, for constructing tableaux with no more than instrumental value, and they would interpret the OT agenda as that of providing such recipes.

Instrumentalism has many attractions. Like many related forms of empiricism, it has the attraction of seeming to minimize what one needs to claim on behalf of a theory to deem it acceptable. That is the point of exempting statements of a theory from the demand that they be true. An instrumentalist is someone who asks only that a theory 'work', never mind whether it is true, false, probable, or meets other controversial demands.

Although we think that instrumentalism is ultimately incoherent and untenable, we cannot make the case here on the basis of the rough sketch we have just given. Suffice it then to say that instrumentalism would turn phonology into something that we—and many others—would find of little interest. The fact that many proponents both of DT and of OT aspire to rules or constraints that have psychological reality—that represent something real and not just computationally useful—indicates that our attitude is widely shared.

A second alternative to this-worldly realism is the view that phonology is not about things in the actual spatio-temporal world, but is about abstract, non-spatio-temporal objects, a version of a more general view often called *Platonism*, according to which there are abstract particulars, abstract individuals. One version of phonological Platonism would hold that the symbols of phonology are names of abstract objects, the sort of things often called *types*.[6] We have

[6] In contrast to *tokens*. Tokens are actual utterances produced at specific times at specific places by specific speakers. They are datable and placeable. Types, if there are types, are not. Tokens are said to *realize* or *instantiate* types by those who hold that there are types, though they also allow for types that are not instantiated by any tokens. It is important, in this connection, not to confuse the notion of an *occurrence* of a type within another type and the notion of a token *instantiating* a type. Suppose that there are such things as types, then the sentence type (1), *Joe's cat hates Mary's*

already mentioned and dismissed that version when arguing that phonological symbols are stand-ins for predicates and not for names. But another form of Platonism might hold that these symbols, though stand-ins for predicates, stand in for predicates that are *true of* abstract types, not of concrete particulars. So, for instance, Platonists of the sort we have in mind would hold that '[+round]' can indeed be rendered as 'λx[+round x]' without semantic loss, but would hold further that the sort of things of which 'λx[+round x]' is true are abstract types, not particular utterances produced at a specific time, at a specific place, by a specific speaker, or some other spatio-temporal thing, but by which they would mean some non-spatio-temporal things residing perhaps where infinite widthless Euclidean lines and pure numbers reside. We do not wish to deny here that there are abstract objects. There may well be such things. Maybe some branches of mathematics and of metaphysics study some of them. We just don't think that phonology can be about any of them. As we mentioned before, that would be incompatible with the fact that phonology is an empirical science.

Admittedly, phonologists seldom mention specific concrete entities. But, as we shall see, that does not entail that their theory must countenance abstract types.

3. Derivational Theory

We now turn to some of the consequences of these assumptions for Derivational Theory. As we mentioned earlier, we will limit ourselves to a very simplified version of DT that already contains the elements we need for our discussion. Our focus will be on the consequences of our assumptions for the meaning, the semantic content, of the phonological symbols in the context of that simplified version of DT. These consequences should carry over to more sophisticated versions. Since we assume that these symbols are stand-ins for predicates, our focus will be on the *satisfaction conditions* of these predicates. What are these satisfaction conditions like? Our worldly realism requires that

cat, contains two occurrences of the type (2), *Cat*, but contains no tokens of (2), no instantiations of the type (2), since such instantiations must be in space and time but types cannot be. On the other hand, if you now produce a token of (i.e. an instantiation of) (1), then that token will contain two tokens (two instantiations of) (2), but no occurrences of it.

The type/token distinction was first named by C. S. Peirce (1958:iv., 423): 'there is but one word "the" in the English language; and it is impossible that this word should lie visibly on a page or be heard in any voice, for the reason that it is not a Single thing or Single event. It does not exist; it only determines things that do exist. Such a definitely significant Form, I propose to term a *Type*. A single event which happens once and whose identity is limited to that one happening or a Single object or thing which is in some single place at any one instant of time, such event or thing being significant only as occurring just when and where it does, such as this or that word on a single line of a single page of a single copy of a book, I will venture to call *Token*.'

For further discussions, see Bromberger (1989), Bromberger and Halle (1992), Hutton (1990), Katz (1990), Peirce (1958), and references therein.

they be met by spatio-temporal things. Since actual utterances produced by actual people at actual places and actual times are crucially pertinent aspects of that reality, let us look at them first. And, to fix ideas, let us concentrate on one such utterance, one produced by SB in Colchester on 1 September 1995 around 3 p.m. local time, which we transcribe here in standard English orthography:

(1) Canadians live in houses.

Because this chapter is being written in 1996, that utterance is now history, is now gone and beyond our perception. But DT can nevertheless still associate a derivation with it. In rough outlines, that derivation would look like (2):

(2) {[kænæd-i-æn], Noun . . . } + {[z], Pl . . . }+{[liv], Verb . . . } + {[in] Prep
. . . } + {[hɨːs] Noun . . . } + {[z] Pl . . . }
. . . .
. . . .
kənéydiyənzlɪvɪnháwzəz

A fuller exposition would, of course, contain more steps and additional details, and might contain references to rules, and so on. We set all that aside for a longer discussion at another time, and as not essential to our current topic.

Let us first turn to the phonological symbols in the *last* line, which is presumably closest to the noises and articulatory gestures that SB actually produced that day in Colchester, and let us begin with the first of these symbols, 'k'. Most phonologists view 'k' itself as an abbreviation rendered roughly as

(3) $k =_{df}$ dorsal [−continuant, −voiced, −nasal]

Putting (3) in lambda notation—to make explicit that 'k' is a predicate—yields

(4) $\lambda x[kx] =_{df} \lambda x[\text{dorsal } x \,\&\, -\text{continuant } x \,\&\, -\text{voiced } x \,\&\, -\text{nasal } x]$

where (4), setting some technicalities aside, will be represented somewhat loosely as the conjunction of predicate (5), which has the virtue of bringing out that the feature symbols in isolation are also stand-ins for predicates

(5) $\lambda x[kx] =_{df} \lambda x[\text{dorsal } x] \,\&\, \lambda x[-\text{continuant } x] \,\&\, \lambda x[-\text{voiced } x] \,\&\, \lambda x[-\text{nasal } x]$

In other words, if we call A the thing at Colchester to which k applies, then the following statement is true:

(6) $\lambda x[\text{dorsal } x](A) \,\&\, \lambda x[-\text{continuant } x](A) \,\&\, \lambda x[-\text{voiced } x](A) \,\&\, \lambda x[-\text{nasal } x](A)$

Two questions now arise. What sort of thing does 'A' refer to in (6)? And how did A meet the satisfaction conditions of the predicate defined by (5) so as to make (6) true?

The first of these questions we will answer very simply for present purposes: 'A' refers to SB at the time of the utterance of the first segment. In other words, think of SB as a sequence of stages that make up his lifeline as an organism, and think of one of these stages as the stage he was in when he produced that initial segment in Colchester. For present purposes, these stages should be thought of as forming mereological combinations, that is, as combining with each other as parts to form larger wholes that can in turn combine with other parts or other wholes to form still more encompassing wholes. This is no doubt crude, but it will have to do for the present. Any refinement would take us far afield—and require that we greatly complicate our notation—without affecting our main point, which concerns the semantic value of the predicates. And in any case, we will eventually have to drop references to that specific event in Colchester, and generalize. So 'A' in (6) designates SB-at-a-stage-during-the-time-of-the-utterance. In what follows, instead of 'A' we will use the mnemonically more helpful 'SB', and we will use it as short for 'SB-at-a-stage-during-the-time-of-the-utterance', counting on context to foil any confusion and, in particular, any confusion due to the fact that different occurrences of 'SB' will have to refer to different stages of SB. Sometimes we will have to distinguish explicitly among different stages of SB. We will do so by using subscripts.

The second question is open to two answers.

When SB produced the first segment of that utterance, he performed a certain articulatory gymnastic: he closed his oral cavity completely with the dorsum of his tongue, he put his vocal cords in a configuration that prevented them from vibrating, he raised his velum. But that gymnastic was preceded by an 'intentional' mind/brain condition,[7] by a mental set, an impulse (it is unlikely that an exact word is available) to move these articulators through these gymnastics, a state that presupposed knowledge of English, that was expressed in the gymnastic of the articulators—was executed so to say—but needn't have been.

According to the first answer (6), or equivalently, (7),

(7) $\lambda x[kx](SB)$

is true because SB actually performed the gymnastics, and the conjuncts in (6) are predicates that describe various linguistically relevant features of the gymnastic. Thus, according to that first answer, the predicate defined by (5) was satisfied because SB performed the gymnastic, and would not have been satisfied if SB had not performed that gymnastic.

According to the second answer, (6) (and *ipso facto* (7)) is true because SB

[7] That such intentional conditions must occur was pointed out long ago by Lenneberg (1967), in his work on the synchronies of speech production. The neural paths to the various articulators being of different lengths, instructions to move them must leave the brain at different times, thus requiring that the effect be 'intended' before being accomplished.

was in the intentional mind/brain state that preceded the gymnastics, and the conjuncts in (6) describe relevant features of that state by alluding to how it would be executed, if and when executed. Thus, according to the second answer, the predicate defined by (5) was satisfied, *not* because SB performed the gymnastic, but because SB had the "intention" to perform the gymnastic.[8] The fact that he performed the gymnastic is good evidence that he had the intention, but the predicate would have been satisfied even if he had stopped short of performing the gymnastic.

The second answer is clearly more in accordance with the practice of phonology than the first. The predicates of phonology should be satisfiable even on some occasions when no gestures are produced, as in subvocal speech, mullings, silent writing, silent reading, and so on. And they should be satisfied sometimes even when the gestures that a speaker produces are not the ones that the predicate describes according to the first answer, as when that speaker makes adjustments for impediments in the mouth or noise in the environment. Thus (6) and (7) should be readable as true regardless of whether SB produced his utterance 'normally' or subvocally or read it silently, or even produced it in a slightly distorted way. Finally, as we shall see in a moment, this answer, after some modification, turns out to be compatible with our assumption that phonological symbols must be unambiguous, must have the same meaning wherever they occur in a derivation, whereas the first answer is not. So we adopt that second answer.

A similar answer is clearly forthcoming for each of the other symbols in the last line of (2). So, for instance 'ə' stands in for a predicate defined *grosso modo* as

(8) $\lambda x[\text{ə}x] =_{df} \lambda x[-\text{round } x \;\&\; -\text{high } x \;\&\; -\text{low } x + \text{back } x \;\&\; -\text{ATR } x]$

and represented for present purposes somewhat loosely as

(9) $\lambda x[\text{ə}x] =_{df} \lambda x[-\text{round } x] \;\&\; \lambda x[-\text{high } x] \;\&\; \lambda x[-\text{low } x] \;\&\; \lambda x[+\text{back } x] \;\&\; \lambda x[-\text{ATR } x]$

and

(10) $\lambda x[\text{ə}x](SB)$

that is,

(11) $\lambda x[-\text{round } x](SB) \;\&\; \lambda x[-\text{high } x](SB) \;\&\; \lambda x[-\text{low } x](SB) \;\&\; \lambda x[+\text{back } x](SB) \;\&\; \lambda x[-\text{ATR } x](SB)$

[8] In what follows we will continue to use the word 'intention', but the reader should understand that we use the term in a somewhat technical sense, without all the usual connotations. 'Impulse' might be less misleading, but not much. We assume that as the discipline progresses it will be possible to replace such infelicitous talk with talk that is about states which supervene on brain states.

is true because SB in Colchester had the articulatory intention predicated by the conjuncts in (11). And similarly, for the other symbols in that last line. So far, then, so good.

But let us now turn to the phonological symbols in the *first* line. They require that we modify these definitions. The modification will seem minor, but everything that follows hinges on it.

The first of these phonological symbols, 'k', could plausibly stand for the predicate with the satisfaction conditions assigned so far in the last line, that is, those specified in (5) and (4). On such a reading, SB met the defining satisfaction conditions of that predicate in Colchester twice: once by having the right articulatory 'intentions' (by meeting the satisfaction conditions of each of the predicates in the right-hand conjunction of (5)) *immediately* before performing the actual uttering; and once *at an earlier stage*, as part of his intention to use certain morphemes retrieved from memory in an order fixed by his syntax and his semantics. We know that he had those 'intentions' at some stage, since he carried them out eventually, so why not impute them to him from the start, so to say? But the same cannot be said about the first 'æ', or about the second one, or about the third one (nor about the 's' in the middle). The 'æ' defined on the model of (5) would be

(12) $\lambda x[\text{æ}\ x] =_{df} \lambda x[-\text{back}\ x]\ \&\ \lambda x[-\text{high}\ x]\ \&\ \lambda x[+\text{low}\ x]\ \&\ \lambda x[-\text{round}\ x]$

and thus for

(13) $\lambda x[\text{æ}\ x](\text{SB})$

to be true, there would have had to be a stage of SB at which he intended to front the dorsum of his tongue, to lower it, to unround his lips, and so on. But there is no reason at all to believe that there ever was such an SB stage in Colchester: he definitely did not execute the corresponding gymnastics there when he produced (1).

It is true that we characterized satisfaction conditions above so as to allow for unexecuted intentions in subvocal speech, and for modifications when required to overcome physical impediments. But what we are dealing with here is, from a linguistic perspective, importantly different. SB did not intend æ subvocally at all as when reading or mulling, and though he, in a sense, produced schwa sounds *instead* of æ sounds, he did not do so deliberately or because of some accidental impediment, but he did so automatically, in a way driven by the implicitly cognized rules of his acquired phonology.

So (12) won't do as a definition of the 'æ's that occur in the first line.

We now face a dilemma. 'k' appears in the last line and in the first line; 'æ' appears in the first line only. We have said that 'k' in the last line is a stand-in for a predicate pertaining to articulatory intentions—in other words, a predicate that was satisfied by an SB stage because, at that stage, SB had certain articulatory intentions. The same cannot be said about the 'æ's in the first line.

So we must either deny that 'k' in the *last* line and 'k' in the *first* line stand for the same predicate—have the same satisfaction conditions attached to them; or we must deny that 'k' and 'æ' in the *first* line stand for the same *sort* of predicates, since the predicate represented by 'æ' cannot pertain to articulatory intentions. Or we must revise our view that 'k' in the *last line* is a stand-in for a predicate pertaining to articulatory intentions.

The first option is ruled out by our second assumption: that within a given theory each phonological symbol has the same meaning wherever it occurs. The second option is not only counterintuitive, but it too is ruled out by that second assumption, since 'æ' can appear in the description of surface representations, that is, in last lines of derivations, as well as in first lines; a different example, such as 'Canada is beautiful', would have had an 'æ' in the last line as well as in the first line. That leaves us with the last option.

Fortunately, there is an obvious way of effecting that last option without drastically revising what we have said so far about the predicates in the last line. It is to hold that at *all* the stages at which SB undertook to produce the utterance in Colchester, he had the intention to perform certain gymnastics *unless* some rule or rules precluded them. (This sentence must be read carefully, giving 'intention' scope over 'unless'). On this revised view of things, 'k' is a stand-in not for the predicate defined by (5) but for the predicate defined by (14):

(14) $\lambda x[\underline{k}x] =_{df} \lambda x[\text{upsr dorsal}\ x]\ \&\ \lambda x[\text{upsr}\ -\text{continuant}\ x]\ \&\ \lambda x[\text{upsr}\ -\text{voiced}\ x]\ \&\ \lambda x[\text{upsr}\ -\text{nasal}\ x]$

in which 'upsr' abbreviates the clause 'unless precluded by some rule'. (Note that it occurs as part of, as inside, the predicate.) And similarly (12) is replaced by (15)

(15) $\lambda x[\underline{\text{æ}}\ x] =_{df} \lambda x[\text{upsr}\ -\text{back}\ x]\ \&\ \lambda x[\text{upsr}\ -\text{high}\ x]\ \&\ \lambda x[\text{upsr}\ +\text{low}\ x]\ \&\ \lambda x[\text{upsr}\ -\ \text{round}\ x]$

From here on, we will underline defined phonological symbols wherever their expanded definitions contain 'upsr' clauses, as we did in the case of (14) and (15).

So, on this new reading, the 'k' in the first line of (2) records that at an initial stage SB intended[9] to perform-certain-gymnastics-unless-some-rule-or-rules-precluded-this, and the 'k' in the last line records that at a final stage SB had a similar intention. The 'upsr' clauses in the more explicit versions of these occurrences happen to carry no consequences—are, so to say, barren. On the other hand, the 'æ's in the first line record that at some initial stage SB intended to perform-certain-gymnastics-unless-some-rule-or-rules-precluded-this, but the 'upsr' in the more explicit versions of these occurrences do

[9] In our quasi-technical sense of 'intend'.

have consequences, since, according to the theory, rules did intervene, and this is reflected by the absence of 'æ's in the last line.

Though we have not spelled out the lines between the first and the last line, we can safely assume that phonological symbols would occur in them as well. Some of these symbols may be identical to some appearing in the first line or in the last line, some may not. The latter possibility is, in fact, crucial to DT. But whatever the case, these symbols will all be open to the same kind of interpretation as those in the first and last line, as being stand-ins for predicates with 'upsr' clauses in the feature predicates connoting articulatory intentions.

But our account still needs to be reconciled with three other facts about (2) essential to DT: first, the fact that each phonemic symbol in (2) occurs, not in isolation, but ordered with other symbols in whole lines; second, the fact that each of these lines in turn occurs, not in isolation, but in a derivation, that is, ordered with other lines; and third, the fact that that derivation as a whole pertains not only to the utterance produced by SB in Colchester but to indefinitely many other actual and conceivable utterances.

We now turn to those three facts.

On the analysis that we are proposing, the last line of (2) as a whole—like its phonological constituent symbols—stands for a predicate, but a more complex one. The occurrence of the phonological symbols in the line signifies that the predicates they stand in for are part of the expansion of that more complex predicate—are part of its definition. And the order in which these symbols occur in the line signifies the order in which the things of which the predicates are true occurred in time.

More specifically, the last line stands in for the predicate that we abbreviate as (16),

(16) $\lambda x[\underline{\Omega}x]$

and that is defined by (17),

(17) $\lambda x[\underline{\Omega}x] =_{df} \lambda x[(\exists r)(\exists s)(\exists t)\ldots(\exists u)(\exists v)\{r < s \;\&\; s < t \;\&\ldots\&\; u < v \;\&\; \lambda y[\underline{k}y](r) \;\&\; \lambda y[\underline{\text{ə}}y](s) \;\&\; \lambda y[\underline{n}y](t) \;\&\ldots\&\; \lambda y[\underline{\text{ə}}y](u) \;\&\; \lambda y[\underline{z}y](v) \;\&\; \Sigma\{r,s,t,\ldots,u,v\} = x\}]$

Those not fully comfortable with lambda notation can get an intuitive grasp of (17)—of the satisfaction conditions it describes—by looking at a very rough paraphrase of what statement (18), built on it, asserts:

(18) $\lambda x[\underline{\Omega}x](SB)$

(18) asserts of a stage of SB (*a*) that it was made up (i.e. was the mereological sum, the Σ) of subsidiary stages (r,s,t, . . . u,v); (*b*) that these subsidiary stages occurred chronologically so that r came before s, s before t, . . . u before v; and (*c*) that the first of these stages, r, was a k̲, in other words, met the satisfaction conditions of '$\lambda x[\underline{k}x]$' given above, that the second of the stages, s, was a ə̲, in other words, met the satisfaction conditions of '$\lambda x[\underline{\text{ə}}x]$' (which we have not bothered to spell out), . . . and that the last of these stages was a z̲, in other words, met the satisfaction conditions of '$\lambda x[\underline{z}x]$'.

Turning now to the first line, it too stands for a complex predicate, and the significance of the phonological symbols in it is the same as that of the phonological symbols in the last line. In other words, each phonological symbol signifies that the predicate for which it stands occurs in the expansion (or definition) of that more complex predicate; and the order of these phonological symbols signifies the order[10] of the elements of which these phonological predicates are predicated.

However, the claim that the first line as a whole stands for a predicate requires either a long investigation into the semantics of syntactic/morphological symbols or a leap of faith to the belief that syntactic/morphological symbols—or at least those that survive to the inputs of DT phonology—are also stand-ins for predicates—predicates, moreover, that can combine with phonological predicates to form more complex predicates. We opt for faith, because at this point we have but the dimmest ideas about the satisfaction conditions of syntactic/morphological predicates. Still, without that faith, it is difficult to see how our third assumption—that the theory is a 'this-worldly' theory—can be maintained. But we leave that discussion for another occasion.

Abbreviating the predicate corresponding to the first line as (19),

(19) $\lambda x[\underline{\Gamma}x]$

we can define it in terms of syntactic and phonological predicates as (20),

(20) $\lambda x[\underline{\Gamma}x] =_{df} \lambda x[(\exists l)(\exists m)\ldots(\exists n)(\exists o)\{l \prec m \ldots\&\ldots\&\; n \prec o \;\&\; \lambda y[\{[\underline{\text{kænæd-i-æ}}], \text{Noun}\ldots\}y](l) \;\&\; \lambda y[\{\underline{z}, \text{Pl}\ldots\}y](m) \;\&\ldots\&\; \lambda y[\{[\underline{\text{hɪs}}], \text{Noun}\}y](n) \;\&\; \lambda y[\{\underline{z}, \text{Pl}\ldots\}y](o) \;\&\; \Sigma(l,m,\ldots n,o) = x\}]$

In a more explicit definition, (20) would, of course, be expanded into a still more complex definition in which the phonological symbols would be replaced by their own explicit definitions, described above, and exemplified by (14) and (15).

Here, as before, those who are not comfortable with lambda notation can get an intuitive grasp of the satisfaction conditions that the definition connotes by looking at the paraphrase of a statement about SB built on it.

(21) $\lambda x[\underline{\Gamma}x](SB)$

(21) asserts of a stage of SB (*a*) that it was made up (was the mereological sum, the Σ) of a number of subsidiary stages (l,m, . . ., n,o); (*b*) that these subsidiary stages were ordered so that l was prior to m, . . . n prior to o; (*c*) that the first of these stages was a stage of which the predicate 'being the intention

[10] We leave in abeyance here whether that ordering is temporal ordering or some other sort of ordering. That is why we use '$\prec$' instead of '$<$' in (20).

of uttering the noun pronounced [kænædiæn] unless rules or rules require modification' . . . that the last of these stages was a stage of which the predicate 'being the intention of uttering the plural morpheme pronounced z unless some rule or rules require modification'.

Let us now turn to the derivation as a whole, still read as about that utterance produced by SB in Colchester, and see what the order of the lines signifies.

Though we have not even sketched the lines that intervene between the first and last lines, if we are right so far about those two, then it is safe to assume that these other lines too are stand-ins for predicates true of SB and expandable on the model of (17) and (20). So without going into the details of such expansion, let us abbreviate these predicates as '$\lambda x[\underline{\Delta}x]$', . . . , '$\lambda x[\underline{\Lambda}x]$'. The fact that the first line precedes the second line, the second the third, . . . until we finally reach the last line, signifies two things: (*a*) that SB went through a series of stages in Colchester: first a stage that met the satisfaction conditions of '$\lambda x[\underline{\Gamma}x]$', then one that met those of '$\lambda x[\underline{\Delta}x]$', then . . ., then one that met those of '$\lambda x[\underline{\Lambda}x]$', and finally a stage that met the satisfaction conditions of '$\lambda x[\underline{\Omega}x]$'; (*b*) that each of these stages brought its successor about through a process modulated by phonological rules internalized by SB. Thus what the derivation as a whole expresses about the utterance in Colchester (we will see in a moment how to construe it more broadly)—is also expressed by the conjunction[11]

(22) $\lambda x[\underline{\Gamma}x](SB_1)\ \&\ \lambda x[\underline{\Delta}x](SB_2)\ \&\ \lambda xy[\underline{\text{motivate}}\ xy](SB_1\ SB_2)\ \&\ \ldots .\ \lambda x[\underline{\Lambda}x](SB_m)$
$\&\ \lambda x[\underline{\Omega}x](SB_n)\ \&\ \lambda xy[\underline{\text{motivate}}\ xy](SB_m\ SB_n)\ \&\ \neg\,(\exists t)\{\lambda xy[\underline{\text{motivate}}\ xy](SB_n t)\}$

In (22) '$\lambda xy[\underline{\text{motivate}}\ xy]$' is a dyadic predicate whose satisfaction conditions are met by a pair of speaker-stages if and only if the first member of that pair brings about the second member, and does this through a causal process modulated by internalized phonological rules. Note that this characterization implies that the predicate is transitive, and therefore that (22) entails (23):

(23) $\lambda x[\underline{\Gamma}x](SB_1)\ \&\ \lambda x[\underline{\Omega}x](SB_m)\ \&\ \lambda xy[\underline{\text{motivate}}\ xy](SB_1\ SB_n)$

in other words, that the stage described by the first line 'motivated' the stage described by the last line.

Let us now finally turn to the third fact about (2) with which our account has to be reconciled, the fact that (2) pertains to indefinitely many actual and conceivable utterances beyond the one produced by SB in Colchester.

One way of accommodating our account so far to this third fact would make (2) true but uninteresting. It consists in interpreting (2) as expressing a statement with underspecified arguments. On this interpretation, (2) can signify what (22) signifies, but it can also signify what (22) would have signified had it

[11] We remarked above that we would use 'SB' ambiguously to refer to different stages of SB. In (22) we use subscripts to flag that different stages of SB are mentioned. Without some such device (22) would make no sense.

contained different arguments, that is, had 'SB_1' and 'SB_2' been replaced by references to other stages—either other stages of SB, or stages of some other English speaker—that brought about an utterance transcribable as (1). On this interpretation (2) is not only systematically ambiguous and vague but could not be true unless someone on some occasion had had the intention of uttering (1).

Fortunately there is another way of accommodating this third fact, a way that makes (2) unambiguous, not vague, and—importantly—more interesting because capable of being true regardless of whether anyone ever intended to utter (1). It interprets (2) as not referring explicitly or implicitly to any individual stages at all, but as signifying a *law*, a *nomological generalization*, that happens to have been instantiated by SB in Colchester, but need never have been instantiated to be true (if true). Any theory worth its salt entails indefinitely many laws over and above those that have instantiations, and this should be true of any acceptable interpretation of DT. The law expressed in our notation is

(24) $\Delta(\forall z)\{\lambda x[\underline{\Gamma}x](z) \rightarrow (\exists u)\{\lambda x[\underline{\Delta}x](u)\ \&\ \lambda xy[\underline{\text{motivate}}\ xy](zu)\}\ \&\ \ldots .$
$(\exists v)\{\lambda x[\underline{\Lambda}x](v)\ \&\ (\exists w)\lambda x[\underline{\Omega}x](w)\ \&\ \lambda xy[\underline{\text{motivate}}\ xy](vw)\ \&$
$\neg\,(\exists t)\{\lambda xy[\underline{\text{motivate}}\,xy](wt)\}\}\}$

Informally, (24) states that to *any* stage satisfying the predicate of the first line, i.e. (20) there must correspond a stage satisfying the predicate of the second line and motivated by the former . . . and a stage satisfying the predicate of the last line that is motivated by a stage satisfying the predicate of the penultimate line, but motivates no further stage. The initial symbol Δ marks the modality of what follows, namely that it is a law,[12] and not an ordinary accidental generalization, and thus that it entails counterfactual conditionals.

Since motivation, as we have defined it, is transitive, (24) entails

(25) $\Delta(\forall z)\{\lambda x[\underline{\Gamma}x](z) \rightarrow (\exists w)\{\lambda x[\underline{\Omega}x](w)\ \&\ \lambda xy[\underline{\text{motivate}}\ xy](zw)\ \&$
$\neg\,(\exists t)\{\lambda xy[\underline{\text{motivate}}\ xy](wt)\}\}\}$

which, roughly speaking, states that any stage satisfying the predicate of the first line must motivate a stage satisfying the predicate of the last line.

A proposition such as (25) can, without harm, be paraphrased as describing a relationship between an underlying representation and a surface representation. However, we must be careful not to interpret 'representation' in such a paraphrase as referring to abstract entities of some sort, or, in fact, to any entities. (25), like the law of universal gravitation, or for that matter the law of supply and demand, does not refer to any particular entity or entities, abstract or concrete. The law has a domain of spatio-temporal entities (speaker stages),

[12] It is the statement of a very idealized law, but in that respect it is very much like the statement of laws provided by other natural sciences.

but it can be true even if, by happenstance, no entity in that domain ever meets its antecedent conditions.

Before leaving DT, we should point out that generalizations such as (24) (and (25)) are not to be confused with rules. They are not rules. Rules are, in the term sometimes used by Chomsky, 'cognized' by speaker-hearers. Not so these generalizations. They are not part of a speaker-hearer's mind/brain endowment. Furthermore, unlike rules, they express truths (or falsities): if (24) is true, then our spatio-temporal world is subject to a nomological generalization about some speaker-hearers (speakers who share SB's idiolect), to which it is not subject if (24) is false. However, *the truth conditions* of (24) require that there be speakers capable of cognizing rules. But note that this requirement follows from the *meaning*, from the satisfaction conditions, of the phonological symbols in (2).

4. Optimality Theory

We now turn to Optimality Theory. OT would presumably also offer an analysis of the utterance token that SB produced in Colchester, but that analysis would be in the form of a tableau, not of a derivation. The tableau would have an input, which we will assume would be given either in a form identical to the first line of (2) or in a form that at least contains the same phonological symbols as that first line; that is, it would assign the same phonological 'underlying' structure to the utterance. The tableau would contain a set of GEN outputs, one of which—the 'winner'—would, we will assume, be identical to the last line of (2), the others of which would differ from that winner in ways that do not matter for this discussion.[13] The tableau would also contain a set of constraints the details and order of which do not matter for this discussion either.

Let us now look at the 'k' that would occur at the initial position of the 'winner'. Following reasoning similar to our reasoning about the last line of (2), we take that 'k' to be a stand-in for a predicate true of a stage of SB at Colchester, and true of that stage because it was a stage of 'intending' to move articulators in certain specific ways.

But the predicate for which 'k' is a stand-in in the tableau cannot have exactly the same satisfaction conditions as the 'k' in the derivation, the satisfaction conditions given in the definition (14). That definition presupposes the possible intervention of rules, and OT does not admit rules.

On the other hand, the satisfaction conditions of the 'k' in the tableau cannot

[13] Our assumptions that the input of GEN would be identical to the first line of the derivation (2) and that the 'winner' would be identical to the last line are not crucial to establish our claim that the symbols shared by DT and OT that share the same shape do not share the same meaning, but enable us to simplify the discussion considerably.

be those given by definition (5) either, and this for reasons similar to those that led us to reject (5) as inappropriate for the 'k' in the DT derivation. 'k' happens to occur in the 'winner' and in the input to GEN; 'æ', however, though it occurs in the input to GEN, and in some of the outputs of GEN, it does not occur in the 'winner' at all. We cannot deny that the 'k' in the 'winner' and the 'k' in the input to GEN are stand-ins for the same predicate: that is ruled out by our second assumption, the assumption that within a theory each phonological symbol has the same meaning, wherever it occurs. We cannot deny that 'k' and 'æ' are stand-ins for similar predicates, since that is not only counterintuitive but open to the same objection: in other tableaux 'æ' can appear in the description of 'winners' even if it does not happen to do so in this one.

The solution of the dilemma here is analogous to the solution of the corresponding dilemma raised by DT: the predicate for which 'k' is a stand-in in the tableau is in every respect identical to the predicate for which it is a stand-in in the derivation, except for one crucial difference: the 'unless' clauses in (14) get replaced by a different one, roughly 'unless not optimal according to the UG constraints as ranked for the language of the speaker', a clause that we will abbreviate as 'uno'. In other words, in OT the predicate for which 'k' is a stand-in is defined by (26). (From here on, we will use double underlining for OT predicates, to distinguish them from DT predicates.)

(26) $\lambda x[\underline{\underline{k}}x] =_{df} \lambda x[\text{uno dorsal } x]\ \&\ \lambda x[\text{uno } -\text{continuant } x]\ \&\ \lambda x[\text{uno } -\text{voiced } x]\ \&\ \lambda x[\text{uno } -\text{nasal } x]$

By parity of reasoning, the 'æ's in the tableau stand not for (15) but for

(27) $\lambda x[\underline{\underline{\text{æ}}}\ x] =_{df} \lambda x[\text{uno } -\text{back } x]\ \&\ \lambda x[\text{uno } -\text{high } x]\ \&\ \lambda x[\text{uno } +\text{low } x]\ \&\ \lambda x[\text{uno } -\text{round } x]$

These definitions will obviously do for the occurrences in the 'losers' as well. And similar ones, which we will not bother to spell out, are forthcoming for the other phonological symbols.

So, on the reading proposed here, the 'k' in the input to GEN records that at some initial stage SB intended to perform-a-certain-gymnastics-unless-this-would-not-be-optimal-etc. and the same letter in the 'winner' records that at a final stage SB had the identical intention. The 'uno' clauses in the definition happen to carry no consequences in these occurrences. On the other hand, the 'æ's in the input to GEN record that at some initial stage SB intended to perform-a-certain-(different)-gymnastics-unless-this-would-not-be-optimal-etc. but the 'uno' clause in the definition did carry consequences. We leave aside for the moment what the occurrence of the phonological symbols in the 'losers' signify.

The account so far—like the account at a similar stage of our discussion of the phonemic symbols in DT—still needs to be reconciled with three other important aspects of phonological symbols: first, the fact that these symbols do not occur in isolation, but occur in an ordered sequence with other symbols in whole lines; second, the fact that each of these lines occurs not in isolation but in a tableau; third, the fact that the tableau as a whole pertains not only to the utterance produced by SB in Colchester but to indefinitely many other actual and conceivable utterances. But we can be relatively brief since we can repeat—though with a few *crucial* modifications—much of the discussion on the analogous points about (2) and DT.

The 'winner' as a whole stands for a predicate similar in every respect to the one defined by (17), but with the constituents in that definition replaced by the corresponding predicates belonging to OT, that is, defined with 'uno' clauses instead of 'upsr' clauses. Using double underlining to mirror the similarities and differences, the definition becomes

(28) $\lambda x[\underline{\underline{\Omega}}x] =_{df} \lambda x[(\exists r)(\exists s)(\exists t)\ldots(\exists u)(\exists v)\{r < s \;\&\; s < t \;\&\ldots\&\; u < v \;\&\; \lambda y[\underline{\underline{k}}y](r) \;\&\; \lambda y[\underline{\underline{\text{ə}}}y](s) \;\&\; \lambda y[\underline{\underline{n}}y](t) \;\&\ldots\&\; \lambda y[\underline{\underline{\text{ə}}}y](u) \;\&\; \lambda y[\underline{\underline{z}}y](v) \;\&\; \Sigma\{r,s,t,\ldots,u,v\} = x\}]$

The informal gloss on the statement corresponding to (18), namely

(29) $\lambda x[\underline{\underline{\Omega}}x](SB)$

is similar to the informal gloss for (18), except, again, that the mentions of DT predicates must be replaced by mentions of OT predicates. We refer the readers to that previous gloss and let them make the substitutions.

The input to GEN as a whole stands for a predicate very similar to the one defined by (20), but again with the phonological constituents replaced by the corresponding phonological predicates of OT, defined with 'uno' clauses instead of 'upsr' ones. Using double underlining to mark the difference, the definition of that predicate is (30):

(30) $\lambda x[\underline{\underline{\Gamma}}x] =_{df} \lambda x[(\exists l)(\exists m)\ldots(\exists n)(\exists o)\{l \prec m \ldots\&\ldots\&\; n \prec o \;\&\; \lambda y[\{[\underline{\underline{\text{kænæd-i-æ}}}], \text{Noun}\ldots\}y](l) \;\&\; \lambda y[\{\underline{\underline{z}}, \text{Pl}\ldots\}y](m) \;\&\ldots\&\; \lambda y[\{[\underline{\underline{\text{hɪs}}}], \text{Noun}\}y](n) \;\&\; \lambda y[\{\underline{\underline{z}}, \text{Pl}\ldots\}y](o) \;\&\; \Sigma(l,m,\ldots n,o) = x\}]$

And the informal gloss for

(31) $\lambda x[\underline{\underline{\Gamma}}x](SB)$

is similar to the informal gloss for (21), except that the mention of DT predicates must be replaced by mentions of OT predicates. The structural similarities between these OT and DT definitions and their glosses, however, should not blind us to the differences. These differences raise an important question that we will take up later, namely whether (21) and (31) could be true together. Standard notation implies that they must be true together. We will argue that they can't be.

We turn now to the relationship between (31) and (29) expressed by the fact that the former corresponds to the input and the latter to the winning output of GEN. That relationship is reminiscent of the relationship between the first and the last line of the derivation, a relationship between something describable as an underlying representation and something describable as a surface representation. But it cannot be the relationship expressed by the dyadic predicate '$\underline{\text{motivate}}$' in (23). That relationship is a transitive relation whose occurrence is modulated by phonological rules. The relationship expressed by the positioning of inputs and winning outputs of a tableau is neither transitive nor modulated by phonological rules. Furthermore, the relation expressed by the dyadic predicate '$\underline{\text{motivate}}$' holds only between stages that satisfy DT predicates. Not so the relationship that concerns us now. So let us introduce a new dyadic predicate. To flag its similarity with '$\lambda xy[\underline{\text{motivate}}\ xy]$' we will use a similar notation, and to flag its roots in OT, we will use double underlining:

(32) $\lambda xy[\underline{\underline{\text{motivate}}}\ xy](SB_1\ SB_2)$

In short, then, the tableau asserts the analogue of (23), namely (33):

(33) $\lambda x[\underline{\underline{\Gamma}}x](SB_1) \;\&\; \lambda x[\underline{\underline{\Omega}}x](SB_2) \;\&\; \lambda xy[\underline{\underline{\text{motivate}}}\ xy](SB_1\ SB_2)$

We can now account for the fact that the tableau is relevant beyond the event in Colchester in a very straightforward way. It—like the derivation of DT—can be construed as expressing a law, a nomological generalization, that happens to have been instantiated in Colchester, that may have been instantiated on many other occasions, but that might never have been instantiated while still being true. That nomological generalization in lambda notation is

(34) $\Delta(\forall z)\{\lambda x[\underline{\underline{\Gamma}}x](z) \rightarrow (\exists w)\{\lambda x[\underline{\underline{\Omega}}x](w) \;\&\; \lambda xy[\underline{\underline{\text{motivate}}}\ xy](z\ w)\}\}$

We have not said anything so far about the 'losing' outputs of the tableau, but it is easy to see what the tableau—interpreted as significant beyond Colchester—asserts through them. If we let '$\lambda x[\underline{\underline{\Delta}}x]$' abbreviate one of the predicates expressed by one of these outputs, then the tableau also expresses not only (34), but also (35):

(35) $\Delta(\forall z)\{\lambda x[\underline{\underline{\Gamma}}x](z) \rightarrow \neg(\exists w)\{\lambda x[\underline{\underline{\Delta}}x](w) \;\&\; \lambda xy[\underline{\underline{\text{motivate}}}\ xy](z\ w)\}\}$

which essentially states that stages which satisfy the conditions of the input predicate never bring about stages satisfying the conditions of the 'losing' predicates.

The DT derivation (2)—interpreted as expressing a family of generalizations—expresses no generalization analogous to (35). On the other hand, the OT tableau—interpreted as expressing a family of generalizations which includes the single positive (34) and many negative ones like (35)—expresses nothing analogous to (24). That difference, in many eyes, is at the heart of what differentiates DT and OT. But it should not make us overlook the less openly displayed difference on which we have dwelt so far: that the predicates even in generalizations that look similar, such as (34) and (25), do not have the same satisfaction conditions, and thus that the truth conditions of even similar-looking generalizations are deeply different.

5. Comparing DT and OT

If what we have said so far is right, one might conclude that contributions to DT phonology and contributions to OT phonology are what has come to be called in some circles 'incommensurable', and that the switch from DT to OT by many linguists constitutes what has been called a 'paradigm shift', a kind of switch that is said to mark many so-called 'revolutions' in the history of the natural sciences.

The notion of incommensurability was introduced in contemporary philosophy of science by Thomas Kuhn and Paul Feyerabend. Roughly speaking, two theories are said to be incommensurable when no logical contradictions or entailments can exist between their respective statements, and thus when it is in principle impossible to establish that if one is right the other must *ipso facto* be wrong. This will happen, according to Kuhn and Feyerabend and others—as we read them—when the terms used by the two theories differ in meaning so that the satisfaction conditions of the predicates of one theory can be met independently of the satisfaction conditions of the predicates of the other theory. In other words, two theories T_1 and T_2 are incommensurable, on the view we are describing, if it is possible for something to meet the satisfaction conditions of any predicate belonging to T_1 without it following that something (the same thing, or something else) meets, or fails to meet, the satisfaction conditions of any of the predicates belonging to T_2. In such a situation the generalizations couched in the vocabulary of T_1 and those couched in the vocabulary of T_2 would clearly be logically independent, would neither logically entail nor logically exclude each other.

It is of course easy to come up with plausible examples of theories that stand in such a relationship of incommensurability: Keynesian economics and quantum mechanics, for instance. The interest in the notion is generated, however, not by such boring examples of theories that don't even seem to share terminology, but rather by Kuhn's, Feyerabend's, and other people's claim that certain historical theories which *seem* to share terminology, and which *seem prima facie* commensurable and even mutually exclusive, actually do not share terminology and are incommensurable, are, in principle, actually compatible. A frequently alleged instance is that of Aristotelian and Copernican astronomy. Both rely on the term 'planet' and thus seem to say mutually contradictory things about a same set of objects. Closer scrutiny allegedly shows that these theories use the term with different meanings to denote a different range of objects, and thus link the term 'planet' with distinct and independent satisfaction conditions. Other examples involve theories using such terms as 'atom' and 'energy' and 'gravitation'. Those are the sort of examples implicated in what Kuhn and others after him have called 'paradigm shifts' in the history of science.

We have argued in the two previous sections that in spite of appearances to the contrary, DT and OT do not share terminology, except at a very superficial notational level. Admittedly, we have limited ourselves to a few terms in that terminology, but the considerations we have put forth can obviously be expanded to reach beyond these examples. And nothing we have said implies that relations of mutual implications or exclusions exist across the terminologies. So there is no reason to hold that if, for instance, '$\lambda x[\underline{\Gamma}x]$' is true of anything, then '$\lambda x[\underline{\underline{\Gamma}}x]$' must *ipso facto* also be true of something, perhaps the same thing. For all we have said so far, there is a possible world in which the first is true of some SB stage but the second one is not, and there is another possible world in which the second is true of some stage of SB but the first one is not. And there is a third possible world in which both are true of some SB stages, possibly, though not necessarily, the same stage.

There is nevertheless a strongly felt conviction abroad that the two theories could not be true together, or, more narrowly, that the laws implicit in a derivation such as (2), namely law (24) or even (25), and laws implicit in the corresponding tableaux, such as (34), cannot be true together. There could conceivably be a plausible theory that combines aspects of OT and DT, that is, that uses both sets of predicates, but there will never be a plausible theory that entails both (24) and (34). And this not because these two laws are *logically* incompatible, could not 'in principle' both be true. Nor even because such an ecumenical theory would be uneconomical. The felt conviction—which we share—goes deeper. We now want briefly to describe the ground of that conviction.

To do this we need a third set of predicates, predicates that belong to neither DT nor OT as we have described these so far. Fortunately, we can describe their satisfaction conditions very quickly: their satisfaction conditions are exactly the same as the satisfaction conditions of predicates belonging to DT and exactly the same as the satisfaction conditions of predicates belonging to OT, but without any 'unless' clauses. In fact, we have already defined some of them. So, for instance, this set of predicates includes (5), repeated here as (36),

(36) $\lambda x[kx] =_{df} \lambda x[\text{dorsal } x]\ \&\ \lambda x[-\text{continuant } x]\ \&\ \lambda x[-\text{voiced } x]\ \&\ \lambda x[-\text{nasal } x]$

and it includes (9), repeated here as (37),

(37) $\lambda x[ə x] =_{df} \lambda x[-\text{round } x]\ \&\ \lambda x[-\text{high } x]\ \&\ \lambda x[-\text{low } x]\ \&\ \lambda x[+\text{back } x]\ \&\ \lambda x[-\text{ATR } x]$

and so on. To mark that a predicate belongs to this third set, we will write it with *no* underlining, as above. And though this may be somewhat misleading, we will refer to such predicates as 'phonetic' predicates. They include defined ones as well, and in particular (38):

(38) $\lambda x[\Omega x] =_{df} \lambda x[(\exists r)(\exists s)(\exists t) \ldots (\exists u)(\exists v)\{r < s \;\&\; s < t \;\&\; \ldots \;\&\; u < v \;\&\; \lambda y[ky](r) \;\&\; \lambda y[\text{ə}y](s) \;\&\; \lambda y[ny](t) \;\&\; \ldots \;\&\; \lambda y[\text{ə}y](u) \;\&\; \lambda y[zy](v) \;\&\; \Sigma\{r,s,t, \ldots, u,v\}=x\}]$

Note that no underlining occurs in either the definiendum or the definiens. To get an intuitive grasp of the satisfaction conditions sketched in (38), simply go back to the gloss for (17) and make the obvious adjustment by mentally deleting all underlinings.

With the help of these predicates we can state a set of laws entailed by DT laws and by OT laws. To fix ideas, let us look at two specific ones.

The first entailed by (25) of DT, namely (39):

(39) $\Delta(\forall z)\{\lambda x[\underline{\Gamma}x](z) \rightarrow (\exists w)\{\lambda x[\underline{\Omega}x](w) \;\&\; \lambda xy[\underline{\text{motivate}}\; xy](zw) \;\&\; \neg(\exists u)(\lambda xy[\underline{\text{motivate}}\; xy](wu)) \;\&\; \lambda[\Omega x](w)\}\}$

states that any stage that satisfies the predicate of the first line of the derivation (2) motivates (in the DT sense) a stage that satisfies the predicate of the last line (and thus is a stage that does not motivate further stages) and that *also* satisfies the 'corresponding' *phonetic* predicate '$\lambda x[\Omega x]$'.

The second one, entailed by (34) of OT,

(40) $\Delta(\forall z)\{\lambda x[\underline{\underline{\Gamma}}x](z) \rightarrow (\exists w)\{\lambda x[\underline{\underline{\Omega}}x](w) \;\&\; \lambda xy[\underline{\underline{\text{motivate}}}\; xy](z\, w) \;\&\; \lambda x[\Omega x](w)\}\}$

states that any stage that satisfies the predicate of the input of the tableau motivates a stage that satisfies the predicate of the winner and that *also* satisfies the 'corresponding' *phonetic* predicate '$\lambda x[\Omega x]$'.

But if it is possible for both (39) and (40) to be true—if the derivation and the tableau on which they are grounded can both be valid—that would mean that stages satisfying '$\lambda x[\Omega x]$' can be explained in three different ways. They can be explained as brought about through a process that is an instantiation of law (39), or as brought about through a process that is an instantiation of law (40), and through a process (or processes) that is the instantiation of both laws more or less simultaneously. More concretely, it would mean that the utterance by SB in Colchester could have been brought about by one of two independent processes, either of which would have been sufficient to bring the utterance about, or that it was brought about through the joint operation of both, though either would have been sufficient. But that seems highly implausible. Not logically impossible. But implausible. We would have a theory that allows for 'overdetermination'. Overdetermination is rare and implausible in the absence of strong evidence, but is, admittedly, not impossible.[14]

[14] Scepticism about overdetermination has deep roots and a long history. Thus we find that Sir Isaac Newton gave the following rule in his 'Rules of Reasoning in Philosophy': 'Rule I: We are to admit no more causes of natural things that such are both true and sufficient to explain their appearance' (Newton 1934: 398). Contemporary philosophers discuss the notion mostly in connection with issues in the philosophy of mind (the rationality of allowing mental causation as well as physical causation to account for human behavior and attitudes). For references to the relevant literature, see e.g. Yablo (1992), who summarizes the relevant principle succinctly: 'If an event x is causally sufficient for an event y, then no event x* distinct from x is causally relevant . . . ', or Kim (1979).

What has been said about this pair of derivation and tableau can, obviously, be repeated and generalized for any pair of derivations and tableaux with superficially (i.e. when put in standard notation) identical last line and 'winner'. So, if there are such pairs, anyone sceptical about overdetermination (and most scientists are) must conclude that DT and OT cannot both be right.

We can, of course, envisage things unfolding in such a way that DT and OT somehow end up entailing only distinct phonetic predictions, and thus avoid the issue of overdetermination. But this is very improbable and at most of abstract speculative interest—and, in any case, would make it even less likely that they could both be right. Such a situation could theoretically take two forms.

In the first form, DT and OT would entail *only* laws whose formats are respectively like those of (39) and (40), with phonological predicates in the antecedent that look identical when translated in standard phonological symbols, but with *phonetic* predicates in the consequent that exclude each other—that cannot be simultaneously satisfied by anything—that look very different when translated in standard phonological symbols. In other words, DT would yield only derivations with first lines identical (in standard notation) to the input of tableaux that OT generates, but with last lines that are different (in standard notation or in phonetic predicate notation) from the winners in the tableaux. In such an unlikely eventuality, DT and OT would essentially be making conflicting predictions about what speakers can intend to pronounce. Very strictly speaking, they could still be *logically* compatible, but any sensible linguist would judge that one of them, *at most*, could be right.

In the second form, DT and OT would entail only laws whose formats are again respectively like (39) and (40), but this time not only with phonetic predicates that look different when translated in standard symbols but also with phonological predicates in the antecedent that look different when translated in standard phonological symbols. In other words, DT would yield only derivations whose first lines are unlike the input of any tableau provided by OT, and whose last line is unlike the winner of any tableau provided by OT. In this imaginary situation DT and OT would neither be redundant nor make conflicting predictions. But—assuming that each made at least some verified predictions, and neither made disconfirmed ones—they would then at best each be demonstrably incomplete, and neither could be right.

In short, then, unless one is willing to countenance overdetermination, it would be a mistake to argue that DT and OT could both be simultaneously right, both give us a correct view of reality—from the fact that they rely on predicates with entirely different satisfaction conditions.[15]

But if DT and OT cannot both be right, it may now look as if the debate

[15] We had to rely on a third set of predicates, *predicates belonging to neither DT nor OT*, to make this point about incommensurability. Whether analogous devices are available in the case of other scientific theories that have been alleged to be incommensurable is an interesting question that we can obviously not pursue here.

between their respective proponents should—in principle at least—be decided in a fairly obvious way, eventually, though not immediately. Proponents of DT will continue for a while to come up with derivations that entail laws like (39), proponents of OT will continue for a while to come up with tableaux that entail laws like (40), and the group that eventually comes up with the largest number of confirmed laws (while being charged with no disconfirmed ones)—and seems likely to continue to do so—will carry the day. Along the way, quibbles may arise about what counts as confirmation and disconfirmation of this or that law (in its guise of derivation in the case of laws like (39) or in its guise of tableaux in the case of laws like (40)) and how much weight should be attached to this or that case, but these can be settled in the way scientists usually settle such questions. Note, by the way, that the phonetic predicates (in their guise of last line of derivations and of 'winners' in tableaux) in these laws would play an indispensable role in this way of conducting the debate.

However, as scientists, we not only seek theories that beget large numbers of confirmed generalizations, we also want them to beget explanations.

These two objectives can—up to a point—be met in one fell swoop, since by providing generalizations one often also provides grounds for answers to why-questions. When and how is a complicated matter, but without going into it we can see that, in the instances that interest us, this will often be the case. So, for example, in the situation that we have envisaged above, DT and OT would each provide an answer to

(41) Why do English speakers say 'kənéydiyənzlɪvɪnháwzəz', that is, get to a stage satisfying 'λx[Ωx]' and not 'kænæydiænzlivinhɨ:səz', a stage satisfying some other predicate?

DT would do so by subsuming the fact in the question under (39) and OT by subsuming it under (40). The two answers would be different, and what we said above about overdetermination also means that they could not both be true. Nevertheless, OT and DT would each be able to come up with an answer simply because each is able to come up with an appropriate generalization. And so the theory that eventually yields the largest number of confirmed generalizations will also be the one that eventually yields answers to the largest number of why-questions like (41).

However, why-questions like (41) are not the only explanatory questions we expect phonology to answer. We also expect it to answer explanatory *what*-questions, and in particular questions like (42):

(42) What is the sequence of the stages traversed (the set of predicates satisfied) by a speaker in the course of producing utterances satisfying e.g. 'λx[Ωx]' from mnemonic elements merged and structured by the syntax of that speaker?

DT provides such answers, and is designed to provide further ones since it is designed to yield laws like (24) which predict sequences of stages. OT, on the

other hand, does not provide such answers, and is designed not to do so, since it is explicitly designed not to yield laws about sequences of stages, but instead to yield only laws like (34) which, like (25), concern 'motivation' relations between underlying stages and surface ones, and are silent about intervening ones. OT therefore, no matter how successful in providing generalizations like (34), will forever leave us in the dark about a family of questions to which DT can, at least in principle, yield answers. That is a good reason for pursuing DT rather than OT. And since a good reason for pursuing DT rather than OT is also a good reason for accepting the presumptions of DT over the presumptions of OT, it is a good reason for deciding the debate between them immediately in favor of DT without waiting to see which one is likely to accumulate the largest number of generalizations.

Two rejoinders will no doubt come to mind. The first rejoinder allows that DT is designed to answer questions such as (42) that OT cannot, even in principle, answer; but insists that the situation is symmetrical, since no theory can answer every question and OT, for its part, is designed to answer a type of question that DT cannot answer, of which an example, somewhat roughly put, would be

(43) In what respect is an utterance satisfying 'λx[Ωx]' optimal?

However, there is an important difference between (43) and (42). All phonologists must admit that (42) is a legitimate question, whatever their views about specific derivational accounts of this or that surface representation. That is, all phonologists who are realists and who agree that so-called underlying representations and surface representations are actually implicated in real time in the production of utterances—whether proponents of DT or OT—implicitly accept the presuppositions of questions like (42): that there are stages of the sort assumed by that question. But the same is not true of (43). From the point of view of DT, the notion of optimality is at best suspect, and the fact that creative and gifted linguists have come up with tableaux should eventually turn out to rest on aesthetically intriguing, but ultimately accidental, epiphenomena that need to be explained but that have little if any explanatory depth themselves. One can think of analogies in other disciplines. So, for instance, the predictive power of the laws of geometric optics is explainable ultimately in terms of the mechanisms and laws of the wave theory of light, but not vice versa.

The second rejoinder denies that OT cannot offer answers to questions like (42). It can, though these answers have to supplement the information contained in tableaux and be built with the help of algorithms such as those proposed by Tesar—or certain variations thereof. But if one allows that such algorithms may be—and for all we know actually are—implemented in the brain of speakers and could actually be invoked in speech performance, then, the second rejoinder goes, OT offers answers to questions like (42) that are at least as cogent as the answers offered by DT.

This is not the place for a detailed discussion of Tesar's and Tesar-like algorithms. The pertinent facts about them, for what concerns us here, are certain overall characteristics that—as far as we understand them—these algorithms all share. They all describe computational processes that build in a finite—though normally very large—number of steps the 'winner' output for any given input to a tableau; in other words, these computational processes halt after a finite interval with the description of the 'winner' when started with an appropriate set of ordered constraints and an appropriate 'underlying' string. Such a process typically begins operating at the left edge of the input 'underlying' string, and produces a finite number of new strings, say, *n* new strings, after scanning that left edge and performing on it whatever operations the phonology allows. The process then 'selects' from among these *n* new strings one in particular as a temporary 'winner', the one that is optimal so far according to the ranked constraints. It can do so in a finite number of steps, since *n* is finite and the number of constraints is finite. It then takes that temporary 'winner' as input, turns to the next segment, and performs on that segment whatever operations are allowed by the phonology, thereby producing a set of, say, *m* new strings, all alike at the left edge, but different at the second segment. It then selects from among these *m* strings as new temporary 'winner' the one that is optimal, so far again, according to the constraints. It then compares that tentative 'winner' with the previous 'losers' to determine whether one of the latter might not be 'better' because, for instance, the constraint it violates has a lower weight than the combination of constraints violated by the new tentative winner. If so, it takes the latter as new input for operations on the second segment and discards that second tentative 'winner'. If not, it proceeds to the next segment of the second tentative 'winner', and repeats the process of selecting, going back, perhaps discarding and starting all over again, or going on, until it reaches the right edge, where it will eventually come to a halt. In short, any system, whether a brain or an artefact, that implements such an algorithm must go through a very large number of stages, each satisfying OT predicates, but most of which will be of no consequence, before settling on what it will articulate.

Whether or not Tesar-type algorithms can equip OT with search procedures that do the work of tableaux depends on the nature of the constraints that OT eventually adopts. But let us assume for the sake of this discussion that there are such algorithms. In other words, let us assume that OT can, in principle, and could eventually, in fact, come up with an answer to questions like (42) built by applying such algorithms. That should not end the matter. We still have to compare the tenor of such putative answers with that of the putative answers provided by DT. And for this we must compare DT derivations, for

plausibility and simplicity, not with tableaux, but with descriptions of Tesar-type productions that generate the 'winner' of tableaux from inputs to GEN.

Plausibility is a matter about which people can, of course, disagree. But Tesar-type accounts are prima facie much less plausible as psychological accounts than derivations. Off-hand, we can think of no other non-volitional psychological system that proceeds along such searches, that massively assumes, stores, discards, re-assumes, re-discards states along the way to action, all this without new inputs ever intervening.

When it comes to simplicity, it seems obvious that DT accounts will always be simpler than accounts based on Tesar-type procedures. Tesar-type procedures require a much larger number of intervening stages and operations than do the derivations of DT, and involve greater amounts of redundancy and detours. This is somewhat ironical, since it was misgivings about intermediate representations that motivated much of the interest in OT to begin with.

Conclusion

We want to stress that our comparison of DT and OT, and our reason for viewing DT as more promising than OT, were based on a number of essentially non-empirical considerations. We started from three assumptions. The first was that phonological symbols stand for predicates. The second was that phonology is about things in this our spatio-temporal world, and more specifically about speaker–hearer stages. The third was that phonological symbols are used unambiguously within any theory, within DT and within OT, though not necessarily across them. We argued—not from specific empirical considerations but from considerations about the character of the theories—that the predicates on which DT and OT rely come with distinct satisfaction conditions that converge at some points but not at all points. It is *only* because of this convergence that the two theories can be compared along empirical lines at all. In the course of our discussion we added a fourth assumption: that the phonetic characteristics of utterances are not 'overdetermined', are not brought about by two or more independent processes any one of which would have been sufficient. That was an assumption based on admittedly somewhat inchoate considerations which we—and presumably others—find compelling and which justify the view that the competition between DT and OT is real and is about what we believe about the actual world. The debate between DT and OT will have to be settled—if it is ever settled—by appeal to how well they predict and answer legitimate questions about specific empirical evidence. But it will never be settled cleanly until and unless we become more explicit about the validity of these non-empirical considerations and, in particular, about their consequences for the meaning each theory implicitly assigns to the phonological symbols they both share.

Acknowledgements

We are grateful to Thomas Green, Ned Hall, Jim Harris, Bill Idsardi, Michael Kenstowicz, Russell Norton, Iggy Roca, Philippe Schlenker, and the participants in the Essex Workshop on Derivations and Constraints in Phonology in September 1995 for comments on previous versions of this chapter and other forms of help. Morris Halle gratefully acknowledges the assistance of the British Academy for support towards his travel expenses.

References

Bromberger, S. (1989). 'Types and tokens in linguistics', in A. George (ed.), *Reflections on Chomsky* (Oxford: Blackwell), 58–89. Also in Bromberger (1992).

—— and Halle, M. (1986). 'Why phonology is different', *Linguistic Inquiry* 20(1): 51–70.

—— —— (1992). 'The ontology of phonology', in Bromberger, *On What We Know We Don't Know*. Chicago: University of Chicago Press/Stanford: Center for Study of Language and Information, 209–28.

Feyerabend, P. K. (1962). 'Explanation, reduction, and empiricism', in H. Feigl and G. Maxwell (eds.), *Minnesota Studies in the Philosophy of Science 3*. Minneapolis: University of Minnesota Press, 28–97.

—— (1975). *Against Method*. London: New Left Books.

Hutton, C. M. (1990). *Abstraction and Instance: The Type–Token Relation in Linguistic Theory*, Oxford: Pergamon.

Katz, J. J. (1990). *The Metaphysics of Meaning*. Cambridge, Mass.: MIT Press.

Kim, J. (1979). 'Causality, Identity, and Supervenience in the Mind–Body Problem', *Midwest Studies in Philosophy 4*. Minneapolis: University of Minnesota Press 31–50.

Kuhn, T. S. (1970). *The Structure of Scientific Revolutions*, 2nd edn. Chicago: University of Chicago Press.

—— (1983*a*) 'Commensurability, comparability, communicability', in P. D. Asquith and T. Nickles (eds.), *Proceedings of the 1982 Biennial Meeting of the Philosophy of Science Association* vol. ii. East Lansing, Mich.: PSA, 669–88.

—— (1983*b*). 'Rationality and theory choice, *Journal of Philosophy* 80: 563–70.

—— (1990). 'Dubbing and redubbing: the vulnerability of rigid designation, in C. W. Savage (ed.), *Scientific Theories*. Minnesota Studies in the Philosophy of Science 14. Minneapolis: University of Minnesota Press, 298–318.

Lenneberg, E. H. (1967). *Biological Foundations of Language*. New York: Wiley.

McCarthy, J., and Prince, A. (1995). 'Faithfulness and Reduplicative Identity', in J. N. Beckman, L. W. Dickey, and S Urbanczyk (eds.), *Papers in Optimality Theory* (University of Massachusetts Occasional Papers 18). Amherst: GLSA, 249–384. Also to appear in R. Kager, H. van der Hulst, and W. Zonnereld (eds.), *The Prosodic Morphology Interface*. Cambridge: Cambridge University Press. Also in Rutgers Optimality Archive, http: //ruccs.rutgers.edn/roa.html.

Newton, I. (1934). *Mathematical Principles of Natural Philosophy*, trans. A. Motte and F. Cajori. Berkeley, Calif.: University of California Press.

Peirce, C. S. (1958). *Collected Papers of Charles Sanders Peirce*, iv, ed. C. Hartshorne and P. Weiss. Cambridge, Mass.: Harvard University Press.

Prince, A., and Smolensky P. (1993). 'Optimality Theory: Constraint Interaction in Generative Grammar'. MS, Rutgers University and the University of Colorado. Rutgers Center for Cognitive Sciences RUCCS-TR-2. To appear, Cambridge, MA: MIT Press.

Smolensky, P. (1995). 'On the explanatory adequacy of Optimality Theory', talk at the conference on Current Trends in Phonology, Abbaye de Royaumont, France, June 1995.

Tesar, B. (1995*a*). *Computing Optimal Forms in Optimality Theory: Basic Syllabification*. Technical Report CU-CS-763-95, Feb. 1995. Dept. of Computer Science, University of Colorado, Boulder. Also in Rutgers Optimality Archives ROA-52, http: // ruccs.rutgers.edu/roa.html.

—— (1995*b*). 'Computational Optimality Theory'. Dissertation, Dept. of Computer Science, University of Colorado at Boulder. Also in Rutgers Optlmality Archive ROA-90, http: //rucs.rutgers.edu/roa.html.

Yablo, S. (1992). 'Mental causation', *Philosophical Review* 101 (2): 245–80.