What Kind of Thing is a Language Faculty?!
A Critical Evaluation of OT Phonology

John Frampton
Northeastern University
Version 2, October 2001

“A psychology (rationalist, empiricist, or whatever) needs to do more than just enunciate the laws it claims that mental processes obey. It also needs to explain what kind of a thing a mind could be such that those laws are true of it; which is once again to say that it needs to specify a mechanism.”

(Jerry Fodor, 2000:17, italics in original)

For the purposes of this article, OT will refer to the theories of phonology based on an ordered list of penalty functions which grew out of Prince and Smolensky (1993, henceforth PS93, with authors P&S). Some parity of designation is useful, so RT will refer to the theories of phonology based on ordered lists of rules which grew out of Chomsky and Halle (1968, henceforth SPE). Some of the discussion is relevant to the application of optimality-based methods to syntax, but much of it is not. To the extent that syntactic applications rely on optimization over infinite sets (or sets which are effectively infinite), Section 2 is particularly relevant.

OT was introduced with shaky conceptual foundations, but the enthusiasm generated by its early successes in extending the empirical coverage of phonology into some domains which had proved resistant to RT analysis, coupled with what appeared to be a solution to the “conspiracy problem” in RT, allowed the advocates of OT to downplay fundamental conceptual questions. Now, almost a decade later, little progress has been made in providing a mechanism which supports an OT theory of phonological knowledge. At the same time, RT has successfully analyzed the phenomena which were once offered as virtual proof of the descriptive inadequacy of RT and showcased as examples of the superiority of OT. It is also now apparent to most OT practitioners that the problem of opacity is not easily circumvented in the optimality theoretic paradigm.

In this context, I will try to shed some light on what I see as the three fundamental issues separating the proponents of OT and the proponents of RT. First, and most important, the issue raised in the title question of this article. I will argue that OT has abandoned the study of the language faculty in favor of a narrowly conceived study of language. Second, the issue of opacity. The failure to come to grips with opacity is not only at the root of OT’s descriptive inadequacies, but the weakness of the OT critique of RT’s purported descriptive inadequacies. Third, the issue of conspiracy in rule systems and the role of phonotactics in RT. I will suggest that the problem of conspiracy is real, but that a solution can be found that is consistent with the RT framework. We begin with the question of opacity.
1. Opacity, descriptive adequacy

“The demand of science, in all its manifold appearances, is always the recognition of the invisible in the visible.”

Wilhelm von Humboldt, 1814.

“Viewed derivationally, computation typically involves simple steps expressible in terms of natural relations and properties, with the context that makes them natural ‘wiped out’ by later operations, hence not visible in the representations to which the derivation converges.”

Chomsky, 1995:223

The characteristic of RT derivations that Chomsky points to is called *opacity*. Opacity is expected in RT, but many phenomena whose RT accounts display opacity are impossible to account for in OT theories whose architectures resemble that proposed in *PS93*. This seems to be agreed on all sides. The first two sections of McCarthy (1998) contains an overview of the problem from an OT point of view. See Idsardi (1998) and Odden (2000) for the RT point of view. The conclusion from either point of view is that, without major revision of the *PS93* architecture, OT is descriptively inadequate.

McCarthy (1998) proposes one such revision, which introduces something like intermediate representations, but produced by auxiliary optimization processes, in order to generate opacity effects. See Idsardi (1998) for a critique. Kiparsky (1998) proposes a more natural revision by taking a half-step back towards RT derivations: a cyclic computation, but with each cycle determined by an OT grammar rather than a rule cycle. Both of these proposals need more discussion in the literature before a solid evaluation is possible. There is reason to remain skeptical that either proposal will be successful in introducing sufficient intermediate structure into the computation to produce the pervasive opacity effects which RT predicts and language appears to evidence.

In response to criticism of the descriptive inadequacy of OT, it is common for its advocates to claim that RT is descriptively inadequate as well. Part of the appeal of *PS93* and McCarthy and Prince (1995), henceforth *MP95* with authors M&P, was their success in convincing readers that certain problems which could be successfully analyzed by OT methods were resistant *in principle* to RT analysis. This is not the forum for a comprehensive and detailed analysis of all the purported counterexamples. But a discussion of some of the arguments is sufficient to cast doubt on how compelling the evidence is, particularly for their claim that these problems are problems in principle for RT.

I will consider two examples: the interaction of footing and syllabification in Tongan (*PS93*:28) and the interaction of nasalization and reduplication in Malay (*MP95*:41). The form of the argument against RT is the same in each case. It is argued that since RT is forced to employ mechanisms which apply in some order, it is forced to postulate two separate rules in each case. It is then shown that neither order of application of the two rules can produce the desired outcome. There is a common thread running through both arguments; a surfacist orientation implicitly excludes the possibility that the workings of a computation are more complex than is directly revealed at the surface. [I take the term “surfacist” from Prince’s dissertation (1975:81), in which he refers to certain “die-hard surfacist” medieval grammarians.]
We begin with Tongan syllabification. The fact is that in Tongan CVV syllabifies as a single heavy syllable, except when preceding a final light syllable. At the right edge of a word, CVVCV syllabifies as CV.V.CV, with CVV split into a pair of light syllables. Mester (1991) realized that the exceptional syllabification in this position was due to the demand for a final binary foot in Tongan, interacting with Syllable Integrity (SI), which prevents feet from splitting syllables. Unless one looks past the surface, this is inexplicable in RT. Syllabification must occur before footing, since it determines the units that are footed. At the point that syllabification occurs, there is no foot to trigger anomalous syllabification. Then, when footing takes place, SI blocks splitting the resulting CVV syllable. PS93 builds a passionate argument against “bottom-up constructionism” around the Tongan example.

The error in the criticism is its failure to recognize derivational opacity. Because of this failure, P&S tacitly assume that syllable structure, once built, cannot change and that a surface condition (SI in this case) cannot be violated in pre-surface representations. If we remove the surfacist blinders, we are led instead to conclude that the derivation must look something like (1), where µ indicates a mora and feet are parenthesized.

\[
\begin{align*}
\mu & \mu & \mu & \mu & \mu \\
\mu & \mu & \mu & \mu & \mu \\
\mu & \mu & \mu & \mu & \mu \\
\mu & \mu & \mu & \mu & \mu \\
CVVCV & \rightarrow & CVV & . & CV & \rightarrow & CVV & . & CV & \rightarrow & CV & . & V & . & CV \\
\end{align*}
\]

SI is violated, but the resulting defect is easily repaired by straightforward CVV \rightarrow CV.V syllable reorganization. Frampton (2001) develops an autosegmental theory of syllable structure in which the repair is simply delinking.

One can imagine a variety of other repair operations which might apply in some language in configurations parallel to the Tongan configuration, with conflicting syllabification and footing demands. Three possibilities are given below, with t-epenthesis and a-epenthesis used illustratively to represent generic consonant and vowel epenthesis.

\[
\begin{align*}
& \mu & \mu & \mu & \mu & \mu & \mu \\
& \mu & \mu & \mu & \mu & \mu & \mu & \mu & \mu \\
& \mu & \mu & \mu & \mu & \mu & \mu & \mu & \mu & \mu \\
& \mu & \mu & \mu & \mu & \mu & \mu & \mu & \mu & \mu & \mu \\
& CVV & . & CV & \rightarrow & CV & . & CV & \rightarrow & CV & & V & . & CV \\
& CVV & . & CV & \rightarrow & CV & . & t & V & . & CV \\
& CVV & . & CV & \rightarrow & CV & . & C & a & . & CV \\
\end{align*}
\]

SI is violated, but the resulting defect is easily repaired by straightforward CVV \rightarrow CV.V syllable reorganization. Frampton (2001) develops an autosegmental theory of syllable structure in which the repair is simply delinking.

\[
\begin{align*}
& CVV & . & CV & \rightarrow & CV & . & CV & \rightarrow & CV & & V & . & CV \\
& CVV & . & CV & \rightarrow & CV & . & t & V & . & CV \\
& CVV & . & CV & \rightarrow & CV & . & C & a & . & CV \\
\end{align*}
\]

A prohibition against vowel initial syllables in a language would block (1), but would allow (2b). As far as I know, (1) and (2a) are attested in various languages (Tongan, Fijian, Southern Paiute); but (2b) and (2c), with epenthesis, are rare or perhaps even absent altogether. From the perspective of RT, this is not surprising. The more complex operations, which involve epenthesis, are not attested, but the much simpler repair operations are attested. OT has no account of the absence of the epenthetic repair possibilities. Suitable rankings of penalty functions would certainly allow either of these possibilities. Rather than a demonstration of the inadequacy of RT, what is revealed is a weakness in OT. Computational complexity has no place in
a theory which disdains mechanism. OT has no substitute for RT arguments which are based on the relative complexity of repair rules.

Now consider the interaction of nasalization and reduplication in Malay. The facts are straightforward. Nasalization spreads rightwards in Malay, with details of locality not relevant here. The verb root anem predictably surfaces as anêm. The puzzle is that when reduplicated, the surface form is ânêmânêm. The initial vowel nasalizes in spite of the fact that it follows a nasal in neither surface nor underlying representation. MP95 contends that RT is incapable in principle of analyzing the phenomenon because the environment for nasalization of the initial vowel is present neither in the form anem nor the form anemanem. Again, surfacist blinders lead one astray. Reduplication appears to be copying operation on the surface, so it is natural (but unscientific) to suppose that it is simply a copying operation. If reduplication were a unitary copying phenomenon, RT would indeed be incapable of explaining the Malay nasalization fact. Nasalization could accomplish the required nasalization of the initial vowel neither before nor after copying.

Raimy (2000) takes a more scientific view and breaks with the assumption that reduplication is a unitary copying phenomenon. He proposes that underlying representations should be viewed not as simple sequences, but as elements bound together by precedence links, as in:

(3) \[ \# \rightarrow a \rightarrow n \rightarrow e \rightarrow m \rightarrow \% \]

A vocabulary item that specifies total reduplication does not immediately trigger copying, but only alters the precedence relations, as shown below:

(4) Morphology: \[ \# \rightarrow a \rightarrow n \rightarrow e \rightarrow m \rightarrow \% \]

Crucially, the operation which linearizes (4) is a separate operation. Crucially as well, there is no requirement that this operation applies immediately after the morphology creates (4). Phonological operations can apply before linearization.

By avoiding the seduction of surface appearance and decomposing reduplication into component parts, Raimy immediately creates the context for an account of the Malay facts. In (4), the initial vowel has dual predecessors. The question of rule application in ambiguous environments is familiar from the study of Geminate Inalterability (Hayes, 1986; Schein and Steriade, 1986). Analogous to considerations discussed in those studies, if nasalization is an operation for which it is sufficient that one of its multiple environments triggers nasalization, then the derivation will precede as in (5), assuming that nasalization precedes linearization.

(5) a. Nasalization: \[ \# \rightarrow ã \rightarrow n \rightarrow ë \rightarrow m \rightarrow \% \]

b. Linearization: \[ \# \rightarrow ã \rightarrow n \rightarrow ë \rightarrow m \rightarrow ã \rightarrow n \rightarrow ë \rightarrow m \rightarrow \% \]

Nasalization of the initial vowel is opaque, carried out in an environment which is not present in the surface representation.

It goes without saying that many details of Raimy’s analysis, which extends to a broad array of reduplicative phenomena in many languages, are not dealt with here.
But the idea should be sufficiently clear to see that even examples which MP95 offers as virtual proof of the inadequacy of RT are no such thing — once surfacist blinders are removed. See Raimy (2000) and Frampton (2001) for further RT-analyses of reduplication.

2. Computation in Theory and Mind

I assume without discussion that theories of mental processes are theories of mental computation. Among the claims which a theory of phonology can make is that certain theory-internal computations should be taken to be real, meaning that the computations should be taken to correspond to mental computations. OT and RT both assume that there is a function $F$ which maps inputs to outputs, but offer different theories of $F$, and make different claims about how the theory internal computations correspond to mental computation. RT proposes a theory of how knowledge of $F$ is stored in the brain (an ordered list of rules) and a theory of the mechanism the brain uses this knowledge to compute $F(\alpha)$. OT also proposes a theory of how knowledge of $F$ is stored in the brain (an ordered list of penalty functions); but, except for one strand of OT that will be discussed later (Primitive Optimality Theory), does not propose a theory of the mechanism which the brain employs to put this knowledge to work in computing $F(\alpha)$. In order to understand the reason that OT fails to propose a theory of mechanism, we need to look at OT’s theory of $F$ in some detail.

In addition to input structures $R_{in}$, which are given by the lexicon and morphology, OT defines a class of structures $E$ which involve both input and output. The structures $E$ are easiest to describe from a derivational point of view. They contain the input structure and output structure and a partial record of how the output structure was formed from the input structure. PS93 proposed that $E$ consisted of annotated structures of the form $(m)bit[a]$, which is interpreted as $mbit \rightarrow bita$, with $m$-deletion and $a$-epenthesis. This $E$-structure is inadequate for many phonological phenomena (vowel harmony, metathesis, reduplication, etc.), but was quickly generalized in MP95 to two-level representations of the form $\langle mb_1i_2t_3|b_1i_2t_3a \rangle$ in place of $(m)bit[a]$. Both input and output structures are included, with coindexing recording a partial history of the individual terms in the structures. Aside from extending the empirical coverage of OT to a much wider range of phonological phenomena, this allowed much more derivational history to be expressed (and therefore subject to control by optimality-methods). $katbu \rightarrow katubu$ with $u$-epenthesis, for example, could be distinguished from $katbu \rightarrow katubu$ with $x$-epenthesis and subsequent leftward spreading of the final vowel: $\langle k_1a_2t_3b_4u_5|k_1a_2t_3u_6b_4u_5 \rangle$ (epenthesis) versus $\langle k_1a_2t_3b_4u_5|k_1a_2t_3u_5b_4u_5 \rangle$ (spreading, with the final and penultimate vowels coindexed). OT based on the MP95 proposal for the structure of $E$ will be called OT/CT (Correspondence Theory) in what follows. Primitive Optimality Theory (OTP) (Eisner, 2000) advances a third proposal for the structure of $E$, based on temporal alignment of multiple tiers.

The various proposals for the structure of $E$ (PS93, OT/CT, and OTP) all have the property that each $x$ in $E$ is associated with a unique input representation $x_{in}$. Given $\alpha$ in $R_{in}$, the set of $x$ in $E$ such that $x_{in} = \alpha$ is called Gen($\alpha$). The name
suggests “generative,” but this is very misleading. Gen is defined extensionally, not generatively. RT, in contrast, does not directly define a class of possible output structures. Rules are constructive. The output structures are whatever structures are built by the rules.

The OT strategy for determining the input-output mapping is to identify \( F(\alpha) \) with the optimal element(s) in Gen(\( \alpha \)) under some notion of “more optimal.” A few standard definitions are needed. A penalty function is a function defined on \( E \) with nonnegative integer values. [OT analyses typically blur the distinction between constraints, which are predicates (true or false) and penalty functions, which have numeric values. The distinction is maintained here.] Given a ranked list of penalty functions \( P \), a binary relation \( \ll_P \) is defined on \( E \) by: \( x \ll_P y \) (\( x \) is more optimal than \( y \) with respect to \( P \)) if and only if there is some \( p \) in \( P \) such that \( p(x) < p(y) \) and \( p'(x) \leq p'(y) \) for all \( p' \) which are ranked more highly than \( p \). If \( S \) is some subset of \( E \), say that \( x \) is a \( P \)-optimal member of \( S \) if there is no \( x' \) in \( S \) such that \( x' \ll_P x \). Let \( \text{Opt}(S, P) \) denote the set of \( P \)-optimal members of \( S \). It is easy to show that if \( S \) is non-empty, then \( \text{Opt}(S, P) \) is non-empty. Finally, \( F(\alpha) \) is taken to be \( \text{Opt}(\text{Gen}(\alpha), P) \).

This defines an input-output mapping extensionally, but does not furnish a method for computing \( \text{Opt}(\text{Gen}(\alpha), P) \) for a given input \( \alpha \) and list \( P \) of penalty functions. The possibility of epenthesis and reduplication means that Gen(\( \alpha \)) is infinite. The fact that Gen(\( \alpha \)) is an infinite set creates significant computational problems because there are no general algorithms for finding the optimal members of an infinite set.

The definition of \( \text{Opt}(\text{Gen}(\alpha), P) \) can be unpacked somewhat using the idea of recursive filtering. If \( p \) is a penalty function on a non-empty set \( S \), let \( X/p \) be the set of all \( x \) in \( X \) such that \( p(x) \) is equal to the minimum value of \( p \) on \( X \). \( X/p \) is called \( X \) filtered by \( p \). If \( p_1, \ldots, p_n \) is a list of penalty functions on \( X \), then \( X/p_1, \ldots, p_n \) is defined recursively for \( n > 1 \) to be \((X/p_1, \ldots, p_{n-1})/p_n\). The following derivation then yields \( \text{Opt}(S, \{p_1, \ldots, p_n\}) \).

\[
(6) \quad S \rightarrow S/p_1 \rightarrow S/p_1, p_2 \rightarrow \ldots \rightarrow S/p_1, p_2, \ldots, p_n = \text{Opt}(S, \{p_1, \ldots, p_n\})
\]

This is the technique used in determining the optimal element in a “tableau.”

For a finite set \( S \), (6) is not only a definition but a method of computation. But it is instructive to think about the computation for a large tableau. Tableau computations with few candidates and few penalty functions give the illusion that the computation is nonderivational. Carrying out the computation (as a thought experiment) for tableau with 1000 candidates and 100 penalty functions, for example, will remove any illusion that the computation is nonderivational. It is necessary to go through the penalty functions one by one, starting with the most highly ranked and discard suboptimal candidates at each step. The computation stops when a single candidate remains, otherwise after all the penalty functions have been exhausted. In the later case, \( \text{Opt}(\text{Gen}(\alpha), P) \) has multiple members.

For some finite subsets \( S \) of Gen(\( \alpha \)), \( \text{Opt}(S, P) = \text{Opt}(\text{Gen}(\alpha), P) \). Call this property \( P \)-sufficiency. The informal technique for computing \( \text{Opt}(\text{Gen}(\alpha), P) \) commonly used in OT papers is for the author to produce a finite subset \( S \) of
Gen(α) and to personally vouch for its P-sufficiency. Then Opt(S, P) is determined by “tableau inspection.” Sometimes there is a semi-argument that the set S is P-sufficient, sometimes not. There is never anything approaching a proof. It goes without saying that this informal method is not an algorithm in any technical sense, because it employs “choose a P-sufficient subset of Gen(α)” without giving any instructions for how this should be done.

For S = Gen(α), many intermediate stages in the derivation (6) may be infinite sets. If the theory is to explain how phonological knowledge is put to use by the speaker, or even if its ambitions are limited to enabling linguists to make firm predictions, some way must be found to calculate the optimal elements in an infinite set by finite means. Infinite sets can often be finitely described, so the problem is at least potentially solvable. If grammars, E-structures, and Gen are sufficiently restricted, the computational problem can be solved in various ways. See, for example, Ellison (1994), Bird and Ellison (1994), Tesar (1995, 1996), and Eisner (2000). Eisner’s proposal, OTP, is the most interesting because it is the only generative OT theory which makes explicit claims that the theory models the language user’s computation. The restrictions on grammars and E-structures are severe in all of these computational schemes, with OT/CT completely outside the bounds of what can be computed.

Unfortunately, space precludes a discussion of methods of computation employed. The reader who takes the mentalistic approach to the study of language seriously is encouraged to read Eisner and get some idea of the computations that OT is forced to attribute to the mind. Intermediate stages in derivations are finite automata corresponding to infinite sets of candidates. The corresponding candidate set is the set of representations which the automata “accepts” (in the technical sense of automata theory). The derivation proceeds by using penalty functions (provided they have a particular restricted character) to transform one finite-state automaton into another one. The derivations are not elegant. Eisner gives some statistics for a very simple grammar which computes footing and stress on underlying syllables, with no segmental phonology. Intermediate representations grow to automata with hundreds of states and thousands of arcs. Hand calculation is out of the question. Computer calculation, of course, is quite feasible on a grammar of this size.

Since OT appears to need the power of Correspondence Theory to achieve even a modest level of descriptive adequacy, and no computational scheme has come close to a plausible scheme for computing F in OT/CT, either for the linguist or the language user, it is not surprising that OT is silent on the question of how the phonological knowledge it proposes is, or even could be, put to use. This is certainly a curious view of knowledge.

The complexity of correspondence: The indifference of mainstream OT to addressing the question of how language makes “infinite use of finite means” is particularly clear with respect to OT/CT. The E-structure (k₁r₂o₃k₄|k₁o₃r₂o₃k₄), for example, is only a sketch of an E-structure. There are timing slots, prosodic tiers, and segmental tiers. Correspondence is possible for elements on all these tiers as well. The logic of the approach assigns indices to all the elements of the representation, with M&P even suggesting the possibility of reifying association lines and distinguishing between using an old association line or “epenthesizing” a
new one. The number of candidate structures explodes. Walther (2001) gives some 
estimates for a 7 phoneme string, with limited epenthesis and reduplication. The 
title of his paper is its conclusion: “Correspondence Theory: More Candidates Than 
Atoms in the Universe."

Of course, complexity does not correlate directly with size. But it is up to OT to 
provide a plausible mechanism for computing $F$. All indications are that OT/CP is 
a computational nightmare. The claim that this is how the language faculty works 
should not be taken lightly. OT advocates may respond that Correspondence Theory 
is not a claim about how the language faculty works. What then is it a claim about?

The big attraction of the OT/CT theory was that it allowed OT to extend 
its empirical coverage to infixing and reduplicative phenomena. But there is a 
striking downside. Under OT/CT, the simple concatenation of two strings loses any 
privileged status. Concatenation is epiphenomenal in Correspondence Theory, only a 
descriptive term. The very common occurrence of concatenation as a way to combine 
the phonological exponents of vocabulary items is, for OT/CT, the result of the fact 
that it is very common for certain penalty functions (O-Contig, I-Contig, Linearity, 
etc.) to be sufficiently highly ranked to ensure a “concatenation-like” outcome.

These penalty functions conspire to produce the appearance of concatenation.

The “virtues” of computationless phonology: The fact that OT is forced to 
rely on an informal method of making predictions which relies on linguists’ intuitions 
is an obvious defect in the theory. Nevertheless, many OT practitioners appear to 
regard it as a virtue. A typical example of this kind of thinking can be found in 
Buckley (1998), which analyzes the Manam stress system.

... I have argued for an analysis of Manam stress within OT that dispenses with 
the intermediate stages required by ... By eliminating these steps, we eliminate 
the concomitant unattested representations and capture much more effectively 
the interactions of various pressures on the output forms. ... Since there is only 
one surface representation subject to constraints, the complications and false 
starts required in a derivational analysis are avoided ...

This did appear in Linguistic Inquiry and apparently passed the scrutiny of their 
reviewers, so it can be taken to represent a strand of thinking in OT that is not 
simply idiosyncratic. The most complex case that Buckley considers in his paper 
has 5 candidates and 4 penalty functions, which is amenable to “optimization at a 
glance.” But surely no one can believe that this is how the brain works. There is no 
linguist inside the language faculty preselecting manageable short lists of candidates 
for the IO-computation.

Buckley, in essence, takes the point of view that it is advantageous that OT 
proposes no way to compute Opt($\text{Gen}(\alpha), P$) because this allows OT to dispense 
with intermediate representations. This is seen as advantageous; apparently because 
intermediate representations are unattested. Buckley assumes without comment that 
extremely narrow evidence is allowed for attesting to the existence of a representation. 
Only surface evidence counts. RT considers the role of intermediate representations 
in computation to be evidence for their existence. This is science. Mendel took 
the role of his proposed “factors” in explaining heritability as evidence for their 
existence. Empiricist biologists rejected the idea of such “hidden factors,” very much
as surfacist phonologists reject the idea of “hidden” intermediate representations.

The final claim, that “there is only one surface representation subject to constraints,” is so bizarre that I will leave it without comment.

Buckley did not invent this odd way of looking at grammars. P&S promote a computationless view of grammar.

I-Language: A few pages of PS93 at the beginning of Chapter 10 are devoted to “foundational issues.” Responding to criticism that OT is computational intractable, they say:

This qualm arises from a misapprehension about the kind of thing that grammars are. It is not incumbent upon a grammar to compute, as Chomsky has emphasized repeatedly over the years. A grammar is a function that assigns structural descriptions to sentences; what matters formally is that the function is well-defined.

The idea that a grammar is a recursive procedure has been a constant in Chomsky’s ideas for almost 50 years. It is the core insight of generative grammar, explaining how language makes “infinite use of finite means.” The defining characteristic of a recursive procedure is that it is a finite recipe for computation. So generative grammar certainly assumes that it is incumbent upon a grammar to compute. P&S may want to challenge the generative program, but they need to argue their case.

The quote above continues below:

The requirements of explanatory adequacy (on theories of grammar) and descriptive adequacy (on grammars) constrain and evaluate the space of hypotheses. Grammatical theorists are free to contemplate any kind of formal device in pursuit of these goals; indeed, they must allow themselves to range freely if there is any hope of discovering decent theories. Concomitantly, one is not free to impose arbitrary additional meta-constraints (e.g. ‘computational plausibility’) which could conflict with the well-defined basic goals of the enterprise.

Without doubt, theorists are free to contemplate anything they like. Ranging freely is encouraged. But after contemplation and free ranging comes a scientific proposal, which purports to be true of something in the real world and is subject to test against reality. The issue is not the spirit that scientists should bring to their speculative endeavors, but the kind of evidence that can be brought to bear in distinguishing two theories. P&S attempt to circumscribe the domain of admissible evidence so that computability has no bearing on deciding between two theories of grammar. This is a very curious stipulation. Suppose that P&S discovered a neural network that quickly and efficiently computed Opt(Gen(α), P). Certainly, P&S would quickly report on this as strong evidence in support of OT. They would be entirely correct to do so. There are very good reasons why such a result has a bearing on our evaluation of a theory of grammar. Chomsky’s well-known discussion of E-language versus I-language in Knowledge of Language (1986) is particularly enlightening on this point. Chomsky argues persuasively that the object of inquiry of linguistics must be states of the language faculty (I-languages), not the sets of expressions that these states generate (E-languages). The I-language is a real object, best thought of as a biological organ. For standard scientific reasons we abstract
away from the brain in various ways in order to gain one or another insight into I-language. But any and all evidence that bears on the resemblance of the construct arrived at in theory to the real object of interest weighs in the balance in evaluating that theory, with the obvious qualification that it is not always easy to decide what the weight of evidence is. It would be absurd to propose a theory of the heart that claimed that issues of muscle contraction have no bearing in evaluating the theory. It is absurd because muscle contraction is the mode of activity of the heart that is of interest. P&S’s contention that issues of computation have no bearing on evaluating theories of I-language are just as absurd. Computation is the mode of activity of the I-language that is of interest.

P&S argue that we should abstract away from how the brain computes. All that is left is to make a theory of the products of its computations. That is, a theory of E-language. It is not uncommon to hear the complaint that “OT is just a theory of the data.” If data is understood as E-language, the complaint is an entirely accurate criticism.

3. Phonotactics and Conspiracy

If several rules in an RT grammar appear to cooperate to ensure that the final representation satisfies a particular phonotactic $W$ (a well-formedness condition on surface representations), the rules are said to be part of a conspiracy. Phonotactics in the SPE formalism have no status. If outputs all satisfy some condition $W$, it is simply a fact about the E-language which the grammar generates. $W$ is not part of the I-language, and plays no causal role in the input-output relation. Conspiracies have no account, but are simply accidental facts about the I-language. [The phenomenon and the questions it raises for RT were first recognized by Kisseberth (1970).]

While some phonotactics in some languages may be accidental facts about the E-language, it seems to me that the evidence that at least some phonotactics are part of the I-language is compelling. Ito (1986) shows convincingly, in my view, that syllable well-formedness conditions are part of the I-language and that various rules are dedicated to serving these well-formedness conditions. Calabrese (1995) shows that well-formedness conditions on bundles of phonetic features are part of the I-language. Prince and McCarthy (1986) show that in many languages a minimal word condition is part of the I-language. In all of these cases, it appears that certain phonotactics play a pervasive role throughout the phonology. When we refer to surface conditions, it should be kept in mind that the surface form corresponding to some phonetic event cannot be simply read off the phonetics without some knowledge of the particular language that the expression is part of. It is worth noting that all of the phonotactics mentioned above are relatively transparent, requiring significantly less than full knowledge of the language.

P&S criticize the fact that RT grammars have no natural place for this kind of knowledge:

...one might feel compelled to view a grammar as a more-or-less arbitrary assortment of formal rules, where the principles that the rules subserve (the
‘laws’) are placed entirely outside grammar, beyond the purview of formal or theoretical analysis, inert but admired.

There are two errors in P&S’s criticism. In the first place, they imply that all rules ‘subserve’ principles external to the grammar. There is no evidence for this. On the one hand, the evidence for conspiracy is over a narrow range of phonotactics. On the other, rule systems appear to be littered with more-or-less-accidental rules. In the second place, they assume that the I-language and the grammar are identical. All theories of I-language, RT and OT included, have proposed that the I-language has mechanisms devoted to acquisition. The conversion of data to knowledge is not automatic and it is beyond question that there are recursive procedures which play a role in acquisition but do not have any role in online production. SPE, for example, proposed that there was a metric on the space of phonological grammars which was used in matching grammars to data. Such a metric is part of I-language, but is not used in online production. It plays a role in the design of the rule system, but is not part of the rule system itself.

If we correct these two errors, P&S’s claim reduces to something approximately like:

\( 7 \) I-languages contain a system of phonotactics. Grammars are more-or-less arbitrary assortments of formal rules within the envelope of possibilities imposed by the phonotactics and UG. The system of phonotactics is not part of the grammar. Certain rules can sometimes be identified as serving to ensure that the grammar produces outputs satisfying particular phonotactics.

As far as I can see, this does not contradict the RT program. It does contradict the view that all knowledge is encoded in the grammar, but that view is untenable. And it does add unanswered questions to the problems of acquisition, because it presupposes that it is not contradictory to assume both that phonotactics and rules are acquired in parallel, and that phonotactics constrain the rule system. Note in this regard that the phonotactics identified above are relatively independent of the fine-structure of the rule system, hence of the kind that one might expect to both depend on aspects of the rule system for their acquisition and to simultaneously play an important role in further rule acquisition.

Note that P&S’s criticism relates rules to phonotactics through the relation ‘subserve,’ whose meaning in this context is that the rule is the instrument or means by which the phonotactic is implemented. That is not far away from (7). No direct action of the phonotactic is involved in phonological change. The instrument of change is the rule, not the phonotactic which the rule suberves. Other authors take the relation of phonological change and phonotactics to be more direct. Kenstowicz (2001) says that “surface phonotactics constrains underlying structure as well as motivates and directs phonological processes.” This view is totally at odds with (7), which allows no direct role to phonotactics in either directing or motivating phonological processes. Phonotactics can motivate the introduction of a certain rule into the grammar, but it is the rule that does the work.

There are two further points that need clarification.
1. The design of I-language sketched above, in which phonotactics have a place outside of online grammar, stresses that phonotactics do not directly guide phonological processes. But there is room for confusion here. Consider, for example, a minimally bimoraic requirement $W$ on surface words, a phonotactic. It is entirely possible to generalize the SPE framework to allow rules of the form (8), as suggested by Sommerstein (1974).

\[(8) \quad \text{if not } W \text{ then Augment} \]

In a rule of this kind, violation of $W$ directly triggers a phonological process. But it does so not in its capacity as a phonotactic, but simply as a condition on rule application parallel to a structural description. The rule (8) has (perhaps) been incorporated into the rule system because of the phonotactic $W$, but this is a question of the design of the rule system. In (8), $W$ appears as a rule trigger, not a design condition.

2. The proposal above for incorporating phonotactics into the I-language, but outside the online grammar, make phonotactics part of the acquisition module. There have been suggestions over the years that language acquisition and transmission play a role in the appearance of conspiracies in stable human languages. For recent comments, see Reiss (2000), which elaborates a suggestion of Kiparsky (1973). Reiss argues that phonotactics are indeed accidents and that the appearance of what look like conspiracies is a result of the process of language transmission. I am not convinced that all conspiracies arise in this way, but the line of argument is certainly worth some thought and discussion.

**Some remarks on teleology**: There is something very human and very comforting about teleological explanation. OT offers such comfort, because it gives the illusion that phonological processes are happening for an immediate reason. The metaphor is an input that is striving to conform as best it can to certain conflicting desiderata. Change/difference happens for a reason.

The metaphor for derivational grammar offers none of the comfort of teleology. A sequence of mindless operations applies — what happens happens. This is a mechanical world. But some teleological comfort can be found in the acquisition process sketched above. It is not overly misleading to view the I-language as striving toward the goal of matching (in some sense) the grammar to the linguistic environment, constrained to the rule systems that UG permits and mediated in part by phonotactic generalizations that have been inferred (in some way) from the linguistic environment.
4. Conclusion

Generative linguistics grew out of Chomsky’s insight that the theory of computation developed by Gödel, Turing, Church, Post, and others in the 1930s and 40s provided the basis for understanding how a finite mechanism can underlie the infinite expression of language. It answered the question of what kind of thing the language faculty is. Looked at a suitable level of abstraction, it is a recursive procedure.

OT retreats from the question of accounting for “the infinite use of finite means.” Surprisingly, considering the central role this question had played in the development of generative grammar and to the contribution of generative grammar to cognitive science generally, the shift in position is not addressed in PS93. Chapter 10, which addresses foundational issues, has a subsection called “The Connectionism Connection, and other Computation-based Comparisons.” This section is all that PS93 has to say about the mechanism that might underlie OT. What is most remarkable about the section is the way that mechanism is downplayed, being referred to as an “implementational issue.” The central task of our field is to discover how a particular component of the brain works. It is bizarre to call this an implementational issue. [See Bromberger and Halle (1997) for an early criticism of OTs failure to address the question of mechanism.]

OT offers no direct answer to the question of what kind of a thing a mind might be such that its products (E-language) are describable in OT terms, but does suggest that it is a connectionist network. All brands of connectionism that are discussed by P&S are explicitly rejected by them — for good reason. Although the idea of interacting penalty functions has a ready analog in connectionist theories of computation, the idea of strict domination (one violation of $p$ is worse than any number of violations of $p'$) is alien to connectionist computation. Nevertheless the section ends with the suggestion that OT can be implemented with connectionist computation. P&S say that OT can be seen “as opening the way to a deeper understanding of the relation between the cognitive realm of symbol systems and the subcognitive realm of activation and inhibition modeled in connectionist networks.” Earlier in the section, P&S give a reassurance that “computationalists have always proved resourceful,” suggesting that the mechanism problem will soon be solved. On the one hand, P&S reject known connectionist models (since they are demonstrably inadequate). On the other, they attach OT to unknown connectionist models (which, of course, cannot be shown to be inadequate).

Phonologists, particularly those who are uncomfortable with the idea that their field may be retreating from the study of mind to the study of “linguistics” narrowly conceived, should ask themselves if the promise of a connectionist substrate for OT is justified. A decade has gone by since the reassurance, and little has been forthcoming. OT remains without a plausible mechanism. It remains only a theory of E-language, not an explanation of what kind of thing a mind might be such that it produces E-language. At best, it has traded the conspiracy problem not only for the opacity problem, but even more significantly, for the “mechanism problem.”
Notes

1. Thanks to Sylvain Bromberger, Sam Gutmann, and Morris Halle for our many discussions of OT. Thanks to Noam Chomsky for discussion of the relation between scientific theory and physical reality. Thanks to Charles Yang for many discussions of issues of computation and language and for reminding me of the quotation from Fodor which inspired the title for the paper. Thanks to Sam Epstein for supplying the quotation from Humboldt. Of course, none of these people should be held responsible for my conclusions.

References


Reiss, Charles. 2000. The OCP or NoBanana: Philosophical and empirical reasons to ban constraints from linguistic theory. Ms., Concordia University, Montreal.
Tesar, Bruce. 1996. Computing optimal descriptions for Optimality Theory: Grammars with context-free position structures. Proceedings of the 34th Annual Meeting of the ACL.