

## The Effects of Interpretation on Form

Paul de Lacy

University of Massachusetts, Amherst

Tuesday, March 28, 2000

<delacy@linguist.umass.edu>

<http://www-unix.oit.umass.edu/~delacy>

### 0 Preliminaries

Thanks for inviting me here to talk. Please stop me at any time if you want me to repeat something, or clarify – or if you discover a damning objection to my proposals.

By the way, this handout and a draft copy of this paper will be available on the web from my webpage: it's mentioned at the end of your handout.

Also: don't follow me when I turn my page – my handout is different from yours.

### 1 Aims

I am going to start with a very brief synopsis of what I am going to propose today.

Then, I will present a problem that provides some motivation for the proposal.

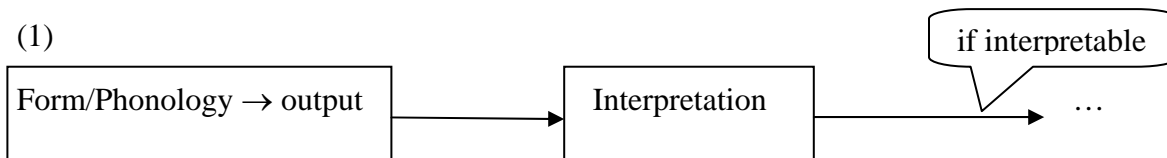
That will be followed by a more detailed discussion of the proposal and how it works.

I will then conclude this talk by talking about some implications of the proposal.

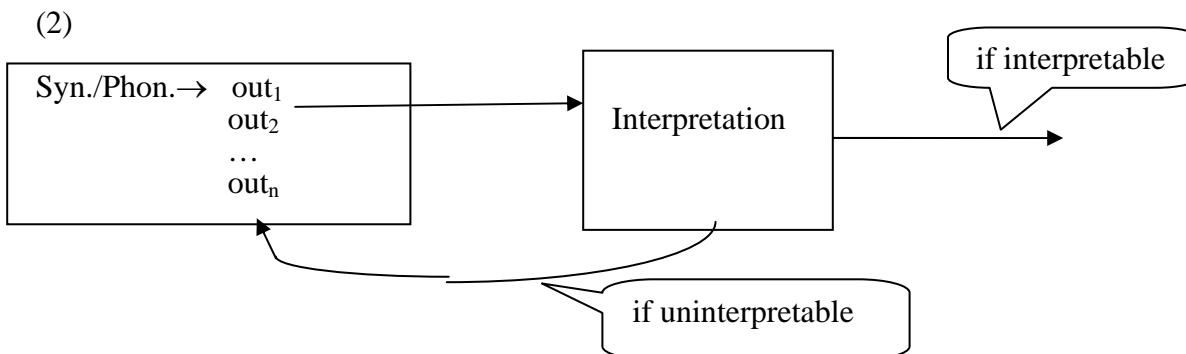
#### 1.1 Brief Synopsis

- First: a note on some terminology. I'm going to focus on the Phonological component and how it interacts with the Interpretation today, but the proposals I'm going to make can also apply to the Form. So, instead of saying "phonology & Form" constantly, I will use the term "Form" components/modules to refer to both Phonology and Form.
- Current belief: Structure-Building (Form, phonology) modules emit a single output form. This output goes into Interpretative modules, which provides some lower-level denotation, the details of which are not particularly important at the moment. The output has two potential fates: (1) it can be translated into an acceptable denotation (= 'interpretable', or (2) it cannot, for a variety of reasons = 'uninterpretable'.
- If a form is uninterpretable, that's the end of the derivation.

(1)



- The aim of this talk is to show that there is another possible – and more desirable – architecture:
- The Structure-Building modules produce an ordered set of output forms.
- The first is sent to the Interpretation.
- If the form is uninterpretable, the next output form is sent, and so on until an interpretable form is found:



- There are many issues that the graphic above does not address. We will get to them in time.
- At this point, I should point out that this proposal is *not* simply a restatement of Optimality Theory. Current OT says nothing about the interaction of Form/Phonology and Interpretation. In OT, the Form/phonology produces a single output form.

In the rest of this talk, I'm going to argue that this proposal has a variety of interesting consequences, namely:

- Casts light on why/whether the grammar is serial/parallel.
- How free GEN is.
- Casts a new light on the evolution of the language faculty.
- Deals with a certain type of conspiracy, which we will turn to now.

## 2 The Problem

Some terminology:

“Form” = Phonology and narrow-Form.

- *Ill-formedness*

Suppose a structure S can't be produced by any grammar. There are two reasons for this:

(1) S is “ungrammatical”: it fails Form constraints.

(2) S is “uninterpretable”: it cannot receive a denotation for some reason:

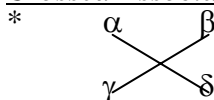
either the Form structure is ambiguous, contradictory, incomplete, or has extraneous parts.

An aside: A theory that eliminates some ill-formed structures by appealing to uninterpretability *overgenerates*. The majority of linguistic theories overgenerate, sometimes crucially (Hale et al. 197?). Montague's is an exception: every Form structure receives a denotation of some sort.

- *Uninterpretability*

Let us look at a structure that has been argued to be uninterpretable – & therefore ill-formed – in phonology:

(3) *Crossed Association Lines*



Sagey, in her 1988 *LI* squib, argues that association lines are interpreted as expressing “simultaneity of temporal realization” or something similar. In other words, the phonological structure is translated into a representation that expresses temporal precedence (unlike the phonology). Here, nodes are translated into a series of points on a continuum, etc. Most of the details are not terribly important here: the most important point is that there are conditions on the translation of the phonological structure.

So, crossed association lines present a contradiction: (1)  $\alpha$  precedes  $\beta$ , (2)  $\gamma$  precedes  $\delta$ , (3)  $\alpha$  is simultaneous with  $\delta$ , (4)  $\beta$  is simultaneous with  $\gamma$ , (5) therefore  $\delta$  precedes  $\gamma$  != *contradicts* (2).

Sagey’s proposal seems desirable because we need some notion of the ‘Interpretation’ anyway. At some point, phonological structure is going to be interpreted into a sequence of gestures. This much is inescapable for any theory of interpretation. With this idea, the sort of reasoning given above seems fairly inescapable: when ordering two simultaneous strings, the segments must not present contradictory messages.

#### (4) *Status of NCC*

I hate to break the flow of the discussion, but I should take a moment to talk a little bit about the current status of the NCC, and whether it really applies. Some have suggested that the only tier that has temporal order specified is the melodic one. Order on other tiers is derived by inspecting association lines. If so, then crossed lines are irrelevant, and so is the NCC. However, this view has come under fire from a few different places:

- (1) Floating tones: tones can float even in the output (triggering downstep). Without explicit order between tones, the position of the floating tone would be indeterminate.
- (2) Internally complex segments: If a single segment contains two contradictory specifications – e.g. affricates – they cannot be ordered [might not be a bad thing].
- (3) Floating prosodies: Some have suggested that prosodic elements may exist independent of segments. See Prince & Smolensky (1993) for  $\mu$ ’s, for example.

We could probably argue about whether crossed lines and the NCC is really relevant until the cows come home. For this talk, though, I will assume that crossed lines are a Form problem. I will talk about the NCC here today because it has been discussed so much in the past. In any case, there are many other constraints that will also illustrate the general problem that I want to describe here. So, without further ado, here is the problem:

#### (5) *Redundancy*

I will now state – or rather restate – an *intuition* that has been said many times over the past several years:

- “If a structure  $\Sigma$  is uninterpretable, then it would be redundant to have a Form constraint  $*\Sigma$ ”. “Why ban a structure in the Form that is uninterpretable anyway?”

#### In present terms:

“Since crossed association lines are uninterpretable, it seems redundant to posit an NCC constraint.”

This sort of statement lay behind Sagey’s (1988) reasoning. And it is very popular, though, usually it is inverted to something like: “If a structure  $\Sigma$  is uninterpretable, then we can simplify the Form by eliminating  $*\Sigma$  constraints.” This sort of reasoning is a type of rule/constraint ‘look-ahead’ that has been invoked in a lot of recent work too: where the reason for avoiding some structure is because it fatally violates some constraint in a subsequent module.

There is certainly a *feeling* that if we can show/claim that a structure is uninterpretable, we can simplify the Form. I will return to why that's a good idea much later on.

At this point, I will simply assume that such a statement is reasonable: redundancy is a bad thing. Getting rid of it should be a priority, and a way to evaluate a good theory. Certainly, analogous cases have shown up in the history of linguistic theorizing. For example, in some serial theories of phonology two types of rules were invoked: *lexical redundancy* rules limited the form of lexical entries, while *phonological rules* acted on lexical items. In SPE, Chomsky & Halle noted that the lexical redundancy rules and phonological rules looked suspiciously similar. Later work showed that in some/many cases, the phonological rules were duplicating the effects of the lexical redundancy rules: the 'Duplication Problem'. Here, we have here a potential duplication problem of another kind: this time it is with the Form and Interpretation: we don't want to duplicate a restriction that already applies in the Interpretation.

On the other hand, the sort of duplication problem presented here has a somewhat different character to it, but one that is discussed at the end of this talk – when we talk about the origins of the language faculty.

For the moment, I will assume that it is undesirable to have a Form constraint that bans an uninterpretable structure.

#### (6) *Duplication*

The problem that now confronts us is that in some situations it seems absolutely necessary to have a Form constraint against an uninterpretable configuration – in the present case, the NCC. Here's why:

#### (7) *Triggering/Blocking*

The NCC can trigger or block rules. For example, let's take a theory with highly constrained operations (e.g. Archangeli & Pulleyblank 1994). Now, let's take a fairly abstract example: a segmental metathesis operation that reverses the order of two TBUs:

(8)  $ai \rightarrow ia$

Now, taking our rather strict theory, we have the derivation as follows:

(9) (a)  $\begin{array}{cc} \text{H} & \text{L} \\ | & | \\ \text{a} & \text{i} \end{array} \rightarrow$  (b)  $\begin{array}{cc} \text{H} & \text{L} \\ \diagdown & / \\ \text{i} & \text{a} \end{array}$

The structure in (b) is the minimal response necessary to satisfy the rule  $ai \rightarrow ia$ . But now we are left with (b), what can we do with it? Without an NCC, the answer is quite simply nothing. Without the NCC, the grammar does not 'know' that (b) is an undesirable structure. The derivation should proceed blithely on, ultimately crashing at the Interpretation. If (b) is later transformed into something interpretable in the Form, this would be quite coincidental and fortuitous (e.g. if one of the vowels or tones is deleted).

The problem is that – in real life – the grammar avoids crossed lines as in (b) like the plague. It does so by either *blocking* the process from applying, leaving the form in (a), or by *triggering* some repair:

- (10) (a) Repair 1: Stability
- |   |   |
|---|---|
| H | L |
|   |   |
| i | a |
- (b) Repair 2: Full Metathesis
- |   |   |
|---|---|
| L | H |
|   |   |
| i | a |

So, we need an NCC, it seems.

We don't have any principled way to escape this problem. Any attempt to get around it without invoking the NCC has to rely on a suspiciously fortuitous confluence of several different rules and operations. This plainly seems wrong.

(11) *Economy*

Theories that appeal to some notion of economy or harmony in determining the most optimal output run up against the same problem. In fact, it's even more acute. In OT, for example, 'blocking' and 'triggering' are effected by constraint ranking. In fact, the terms 'blocking' and 'triggering' are a bit hard to translate directly into OT terms since they are so process-based. Basically, if a process P (which is roughly equivalent to a markedness constraint in OT) is *blocked*, this means that all candidates that do not undergo P and are more faithful to the UR are more harmonic than candidates that do undergo P. For triggering, the opposite holds.

Significant problems arise when S is better than every other possible form. For example, suppose that some constraint \*ai resulted in a the following candidates:

- (12) (a) Crossed-Lines
- |   |   |
|---|---|
| H | L |
| \ | / |
| / | \ |
| i | a |
- (b) Stability
- |   |   |
|---|---|
| H | L |
|   |   |
| i | a |
- (c) Metathesis
- |   |   |
|---|---|
| L | H |
|   |   |
| i | a |

Without the NCC, which candidate is best?

The problem with (b) and (c) is that they rather gratuitously deviate from the underlying form:

In (b) *i* ends up associated with *H*, whereas it was previously attached to *L*. In (c), the previous order of the tones has changed. In comparison, (a) minimally deviates from the UR: it retains the order of the tones and the associations, too. In OT terms, the crossed-lines structure *harmonically bounds* the other forms: it will always be most economical/harmonic in these situations – in OT terms, it violates a subset of the constraints that other candidates so, and in serial terms (a) applies fewer operations.

In short, we need something to ban (a) – the only possibility is an NCC constraint (either in CON or in GEN, it makes no difference).

So, without an NCC constraint, we effectively predict that metathesis is impossible.

But if we adopt the NCC constraint, we have a duplication problem: crossed lines are banned both by a Form constraint, and by the fact that they are uninterpretable.

(13) *Generalizing the Problem*

At this point, I should emphasize that this result is not a peculiarity of phonology or of crossed lines in particular. Many constraints are just like the NCC.

Many inviolable restrictions on Phonology structures have been ascribed to uninterpretability:

- Illicit feature combinations:
  - (i) [+high, +low] vowels are an interpretive (well, at least physical) impossibility. This fact has been used to block height harmony processes and to trigger featural change.
  - (ii) Nasality: In nasal harmony systems, [t] cannot be nasalized, and in some situations blocks the nasality from proceeding, while in others, it triggers non-local spreading. (Walker diss.). Again, the fact that a nasal t would be [+nasal, -nasal] and therefore uninterpretable makes a constraint against it a moot point.
  - (iii) Featural implications that were done with feature geometry: e.g. labials and dorsals don't have [anterior] specifications – only coronals. If labials had an anterior specification, this would be uninterpretable (i.e. extraneous).
- Strictly Local Feature spreading: A number of recent works have argued that feature spreading is strictly local. Gafos ascribes this restriction to an interpretive problem. However, in some cases strict feature spreading would result in an illicit feature combination – the spreading is again either blocked, or triggers a splitting of the the feature (e.g. Walker's nasal opacity problem).
- Full Specification: According to current theories (Steriade 1995, etc.), output segments must be fully featurally specified. It isn't hard to see how this could be an interpretive requirement. However, if a segment starts off underspecified, without a Form constraint requiring full specification, it should remain underspecified.
- Prosody: The presence of prosodic structure generally may well be due to interpretive requirements. However, since prosodification motivates most featural neutralization and change, the most harmonic/economical form would be one in which no prosodification took place at all.

These last two points are very important. Both full specification and prosodification are essential properties of *every* output. However, without Form constraints to require their construction, an uninterpretable output form will *always* be most economical/harmonic, making speech impossible.

### ***The Prosody Problem***

- Prosodic structure is not present in lexical entries.
    - Or it is present, but not preserved. One of these assumptions is necessary to account for the fact that prosodic structure is not contrastive (i.e. there is no language with a distinction [CVC.V] vs [CV.CV]). Therefore, either there is no underlying prosodic structure or such prosodic structure is not preserved (even further – eliminated) in the output.
    - Q: How could you ever tell whether it was [CVC.V] or [CV.CV]?
    - A: (1) Processes that refer to weight (e.g. stress), (2) phonetics: V lengthening induced by coda consonant.
  - Therefore, the most faithful and least marked/economical output form will be an unprosodified one.
  - Such a form is uninterpretable.
  - Therefore, with a single-output Form, nothing can be produced.
- Similar reasoning for binary branching.

Form is also beset by the same duplication problem. However, since this is a phonological talk, I will not dwell on Form issues here. There are a couple of very interesting cases, though, relating to binary branching and theta theory: following some

recent proposals, these should reduce to interpretive conditions, but without overt constraints on them, uninterpretable structures should always arise since they are the most economical form. [Reconstruction is also an interesting phenomenon.]

### 3 A Solution

Several entirely different solutions to the duplication problem could be entertained. The aim here is to present one that minimizes the effect of the Interpretation on the inner workings of the Form.

(14) *Multiple Outputs*

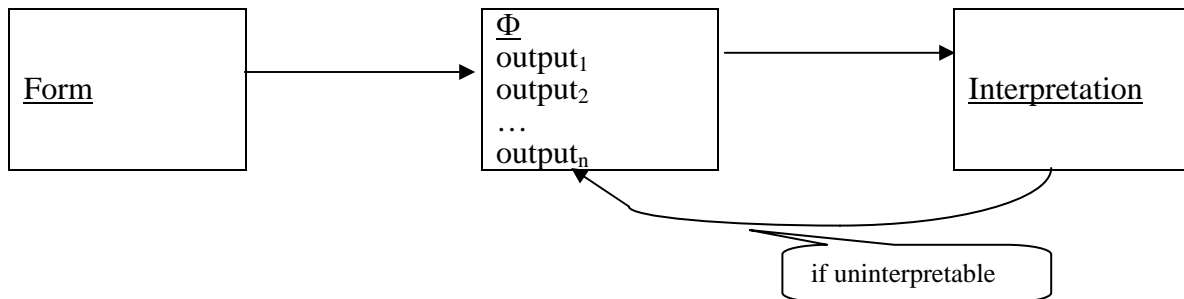
The first step is that the Form produces multiple ordered outputs. The outputs have to be ordered/ranked to make the Form deterministic.

(15)  $\Phi$

There is an algorithm  $\Phi$  which supplies the Interpretation with the first of these outputs F (then deletes it.)

(16) *Interpretability*

If F is uninterpretable,  $\Phi$  supplies the next form in the order.



(17) *An Example*

Let us take the metathesis case. The Form constraint *\*ai* has forced the most harmonic candidates to be unfaithful to the UR. As I showed before, the crossed-lines candidate will always be the highest ranked one. The ranking of other candidates will rely on the primacy of Form constraints or rules (whichever you chose):

(18) (a) Crossed-Lines > (b) Tone Stability , (c) Tone Metathesis

(a) Crossed-Lines

```

  H     L
   \   /
    \ /
   i   a
  
```

(b) Tone Stability

```

  H     L
  |     |
  i     a
  
```

(c) Tone Metathesis

```

  L     H
  |     |
  i     a
  
```

At this point,  $\Phi$  sends the most Formally harmonic form to the Interpretation – i.e. the crossed-lines candidate (a). It is uninterpretable. At this point in most theories, the derivation would come to a halt. However, with the Interpretive loop, this theory does not give up: it sends in the next most harmonic form: i.e. (b) or (c), depending on the grammar.

So, there is no need for a Form NCC constraint.

It is straightforward to model triggering and blocking effects in this model:

1. an uninterpretable structure  $S^U$  ‘triggers’ a repair when the output hierarchy is  $|S^U > S^I > S^{UR}|$ , where  $S^I$  is interpretable,  $S^{UR}$  is faithful to the underlying form, and  $S^I \neq S^{UR}$ .
2. an uninterpretable structure  $S^I$  blocks a process when  $|S^U > S^{UR} > S^I|$ .

In OT terms, the mechanism effectively makes a restriction on an uninterpretable structure to be ranked above every other constraint.

(19) *Modularity*

Part of the aim of this solution was to maintain modularity as much as possible. In this model, the potential effects of the Interpretation on the Form are quite limited: certainly less limited than other theories, but still quite constrained.

Most importantly, the Interpretation cannot play the role of an inter-candidate evaluation mechanism. The output order the Interpretation is provided is fixed by the Form, and the outputs are fed to it one at a time. So, the Interpretation cannot evaluate two interpretable outputs and compare their ‘ease of articulation’, etc., then decide to realize the lower-ranked one. In fact, if the Interpretation’s response to the Form is simply to “accept if interpretable and reject if not”, then its affect on the Form is very limited indeed.

(20) *The Loop*

What is essential to this model? There are two parts:

One is the loop. Without the loop, we have the same problem as before: the Form would only output one form, and if this was uninterpretable, it spell the end of the derivation. The loop is essential.

(21) *Multiple Outputs*

Apart from the loop, the requirement that the Form provides multiple outputs is important. This is obviously more akin to OT than to any other theory, so the issue deserves some mention here.

In a parallelist conception of Form, one has no choice: if the grammar produced a single output, that would most likely be uninterpretable, so the grammar would be unable to produce speech.

The first question I want to address is: “Can serialist theories employ the Dynamic Interpretation?”. It is not a simple matter:

(22) *Serialism: Persistent Interpretation*

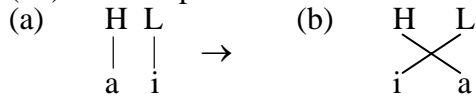
What is needed in the serialist theory is to know right away – in the course of the Form derivation – whether a form is interpretable or not. Then one can hopefully respond adequately to the problem: by triggering or blocking. However, complications arise in this model, and it is not straightforward to adapt the present proposal to fit it.

Suppose that every form produced at each step of the derivation is submitted to the Interpretation. If the form is interpretable, then the derivation can continue on in its merry way, but if uninterpretable, the derivation has to change course in some way – either reverse one step (i.e. blocking), or apply some repair (i.e. triggering).



Note: at this point, the Interpretation takes on quite a different role than in the parallelist theory: here, the Interpretation is essentially a checking mechanism, and only really ‘interprets’ when the end of the derivation has been reached.

(23) Example:



(a) is interpretable, so the next rule can apply to form (b).

(b) is uninterpretable, so some rule must repair it.

At this point, though, we have a problem: we know that (b) is uninterpretable, but not *why* it is uninterpretable. *Which part* of the structure needs to be changed? – the quality of the vowels or tones? The order of the elements? etc. The question becomes far more acute when the entire form is taken into account: i.e. the full prosodification and featural structure of *ia* – which part should we change/reorganize?

Evidently, the Interpretation not only has to tell you whether a structure is uninterpretable, but exactly which *part* is uninterpretable. If we didn’t know which part was uninterpretable, we would have to blindly & randomly apply rules until some interpretable structure was reached.

Again, the Interpretation has taken on quite a different aspect than the one used in the parallel model: instead of just rejecting or accepting, it provides detailed information on what is wrong about the form. Of course, this isn’t bad in itself, just more complex than the parallelist conception.

The question is, is this information enough? What if we reached an uninterpretable form, and were forced by the grammar to repair it, BUT such a repair led to another uninterpretable form? This arguably happens in the early stages of prosodification: each step of the prosodification structure is uninterpretable until it is all complete.

- What if [uninterpretable → uninterpretable]? How do we know we’re on the right course? All we have to guide ourselves is economy, and that might send us on an infinitude of steps to inevitable uninterpretability.

Also: question: *does every uninterpretable form trigger some repair?*

1. Interp → Uninterp → Interp → Uninterp → Interp.

- The biggest issue is the form of the repair rules. For every uninterpretable structure, it seems you need some repair rule specific to that structure. For example, for crossed lines, the repair rule would have to delete/reorder/or whatever, but for an uninterpretable feature combination, it would presumably do something else: change the feature value, delete the feature, etc. It looks like we are pairing a specific uninterpretable structure with a specific repair: in effect, *we are mentioning the uninterpretable structure in the structural description of the repair*. So, we are simply duplicating the effect of the Interpretation in the Form.

The parallelist approach avoids this because forms are not repaired, *per se*. Instead, a ‘repair’ of a form F is another candidate G that resembles F in some crucial way (usually by means of a shared input). There is no causal link between F and G.

Alternative: have a whole raft of rules. Apply randomly until an interpretable form is produced.

Problem1: when we have two repairs that *could* happen to the same structure, but don't. e.g. with floating features, the rule  $F \rightarrow \emptyset$  could be invoked, but that would also deal with crossed lines with F, whereas crossed lines with F probably employs a different repair.

**Example**

Take a lg that uses **tonal metathesis** to fix a NCC problem like we had above. Then a possible repair rule must be  $T_1T_2 \rightarrow T_2T_1$ .

However, suppose that another uninterpretable structure arises: this time with an unassociated T, due to an excess of tones *cf* TBUs (floating T= unconnected structure, therefore uninterpretable). In this case, the tone deletes. Therefore, tonal deletion must also be a possible repair.

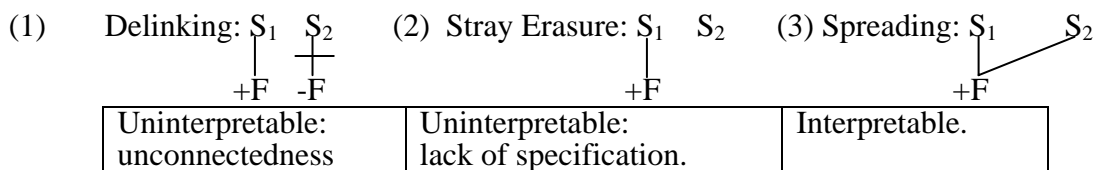
Now, if we applied these repairs randomly to an NCC violation, we should get a random application of tonal deletion or metathesis. i.e. sometimes we should get deletion of the tones and other times metathesis.

Problem 2: what if a form is at least two steps removed from an interpretable form?

**Example:**

- Early stages of prosodification?
- Two repairs: deletion or addition = both produce still uninterpretable structures.

- Assimilation:



In step (1), the second Feature is delinked, leaving it floating – an uninterpretable configuration.

To repair (1), in step (2), the delinked -F feature is deleted, leaving a non-fully specified form.

However, this forms an uninterpretable structure, that is repaired by spreading.

So, two repairs: F-spreading and F-deletion. What if we have other repairs (... highly likely), what if they applied randomly to this form? [**PVL: Which other repairs?? – e.g. epenthesis**]. Then we'd expect random application of epenthesis or spreading in this situation.

In short, this sort of model has a lot of trouble with triggering: we essentially need to pair a *particular uninterpretable structure* with a *particular repair*, which is essentially stating the repair in the syntax anyway. Nothing has really been simplified (explanation = same degree of complexity as the problem).

*Q: Is it really serialism vs parallelism that is at stake here?*

A: Yes. Inherent in serialism is a reliance of any form on the forms that preceded it in the derivation. So, we can rightly talk about a **form F** 'prompting' a change. This notion of

‘prompting a change’ – albeit due to uninterpretability – is the problem we faced with triggering. In a parallelist theory, no output form is dependent on another form for its existence, so we can’t say that the presence of some form implies the presence of another form: all forms exist for independent reasons. So, we just don’t get the triggering problems encountered above.

(24) *Serialism: The “one step back model”*

At this point, I should say that this serialist model at least could do blocking. Other models couldn’t: there are other (more straight-forward) serial theories that fare even worse: can’t do blocking or triggering.

**{TALK ABOUT THIS??}**

## 4.2 “One Step Back” Model

A failed attempt at a serialism model with the Interpretive loop is very instructive:

- (25) Suppose that every step in a serial derivation counts as an output form, with the last derived form first in the order and the first derived form last.
- (26) If the first form (surface form) is uninterpretable, take the next form, and so on.  
(i) Essentially, you take one step back in the derivation.
- (27) Problem: this gets blocking, but not triggering.
- (28) Moral: We need to know whether  $\Sigma$  is uninterpretable at the moment when it is formed.

Suppose that every step in a serial derivation counted as an output form, with the highest ranked form = the final step in the derivation. Then could we apply the Dynamic Interpretation Device to get the right result?

Unfortunately not: this model gets blocking effects, but not triggering. Suppose we got to the metathesis stage. At this point, there is nothing motivating a repair, so the derivation should retain the crossed-lines configuration right to the bitter end. In fact, from that point on, all subsequent forms would contain crossed lines and so be uninterpretable.

- *Go back until you find something interpretable, then take another path or block (grammar-specific choice).*  
No good: this would only allow blocking of a process, never triggering. So, if we went back to the crossed-lines form, you would have to go back to the preceding form, effectively blocking metathesis from ever happening. So...
- *Go back to the first interpretable-uninterpretable sequence, then either (1) go back one to the interpretable form (block) or (2) apply a repair to the uninterpretable form (trigger).*  
Still need to know *what* was uninterpretable about it. (Could do this by comparing the uninterpretable form with its preceding form and seeing what the difference is – no need for Interpretation to tell you).

- Problem: What if there never was an interpretable form? What if early prosodification led to uninterpretability? Then we'd have to go back to the start and make the same mistake all over again.

## 4 Implications

Before looking at some implications of the Dynamic Interpretation, I will briefly sum up what's happened so far:

1. The duplication problem has been solved: there's no need to state a Form constraint that bans an uninterpretable form.
2. Modularity is maintained: Interpretive information is not available in the constructing of Form forms, not cannot it determine grammaticality.
3. No new components have been added to the architecture – we needed something like  $\Phi$  anyway.
4. The major changes are (1) the loop and (2) the requirement that the Form provide multiple ranked outputs. – the latter is already a trivial emendation to OT, at least.

Now, for some implications...

### 4.1 Freedom of Analysis

The question of which inviolable restrictions hold on Form forms has been a focus of every theory. In OT, Prince & Smolensky suggest that there is 'Freedom of Analysis': operations apply very freely to form candidates. They don't go into specifics, but in practice, many assume that there are many restrictions on Form form – so-called 'inviolable constraints'. They include our friend the NCC, Headedness, Full specification, and the Prosodic Hierarchy to name but a few. In fact, there are many that aren't even discussed because they are seen as too basic or trivial: for example, the precedence relation  $R$  that holds between nodes on the same tier is assumed to be asymmetric. No-one assumes that GEN could generate a symmetric relation precedence relation.

Under the current theoretical architecture, such inviolable constraints are *needed*, otherwise we will always end up with uninterpretable outputs as the most economical/harmonic forms.

With the Dynamic Interpretation, though, that we can get rid of most, if not all of these inviolable restrictions. I've already mentioned full specification and prosodification. The need for the precedence relation to be asymmetric also follows from interpretive reasons: if it was symmetric it would be impossible to order anything:  $x < y$  and  $y < x$  are contradictory: they do not provide an unambiguous statement of which element temporally precedes the other.

Of course, there is a cost: massive overgeneration. Having a large candidate set seems to be of little concern in OT, though, so it will not concern me either.

The result, though, is worth emphasizing: with all inviolable constraints recast as Interpretive restrictions, the most harmonic form produced by the Form will be highly uninterpretable. In fact, it will be fully faithful to the input form, modulo inventory requirements.

## 4.2 Levels

To make this point more specific:

### Phonology:

Let us assume that the presence of prosodic structure is something required by the Interpretation: there are no Form constraints requiring the presence of prosodic structure. In addition, there is no underlying prosodic structure (or no faithfulness to it at all). So, the most harmonic form will have no prosodic structure because addition of such structure is marked/not economical.<sup>1</sup>

However, prosodic structure is an essential part of most featural neutralizations, including imposition of sonority-distance requirements, etc. So, the most harmonic form would not have undergone any prosody-driven neutralizations, which essentially means it hasn't undergone any context-sensitive neutralizations at all. The only sort of neutralizations we can expect are inventory restrictions:  $\|*F \gg \text{Faith}\|$  will transform every input segment into a segment that is surface-allowed in the language's inventory. In effect, then, the most harmonic/economical form is one which is fully faithful to the underlying form, modulo inventory requirements.

### S-Form

The reasoning is even more straightforward for the Form. Current assumption: there is no phrase structure underlyingly. However, let us suppose that the very presence of phrase structure – i.e. connectedness – is an interpretive requirement.

But s-Form operations crucially depend on phrase structure. Hence, the most harmonic/economical output will be one that is fully faithful to the input.

In effect, what has happened is that the major (only?) motivations for Form change have become true *output* requirements. Now, suppose we have more than one level:

1. Input → Level 1.
2. Level 1 evaluates the candidates and selects the most grammatical form.
3. Level 1's output then goes to Level 2, and so on until the final level.

We already know in advance what the output of the non-final levels is: it is the most Formally harmonic form – i.e. the fully faithful form. The output of a medial level cannot be anything else because the Interpretation does not apply to it.

Hence, **every medial level essentially restates the original input**. The only level that can do anything is the final one since the Interpretation crucially affects its output – and only its output. In short, whether the Form has  $n > 1$  levels or not is a contentless question: it doesn't matter how many levels there are, since only the last one *does* anything.

There is a way around this: require the output of every *level* to be interpretable. This means that the output of every level would be submitted to the Interpretation; if interpretable, then it is the input to the next level; if not then the next form is found.

[Of course, the Interpretation has taken on a new role: one of a checking mechanism, not of the Interpretation any more.]

However, at least we now have a testable prediction: it becomes an empirical question whether the output of every level must be interpretable. Do we have cases where the output of a level is crucially uninterpretable? (e.g. a crossed-lines structure)?

---

<sup>1</sup> Because there is no faithfulness to prosody, markedness constraints determine whether it appears. However, since there are no markedness constraints that require prosodic structure, the most economical/harmonic form will appear. We know that minimization of structure is desired, so the most economical/harmonic form would be one without any prosody at all.

### 4.3 UR and Output Disparity

Before moving onto a more obscure and heady topic, I want to point out that relegating most inviolable constraints to the Interpretation is extremely desirable for another reason: it allows significant disparity between lexical entries and output forms.

It is commonly assumed – even in OT – that lexical entries look nothing like output forms. In fact, lexical entries are also assumed to be grossly uninterpretable. For example, it is commonly assumed in phonology that lexical entries have no prosodic structure – something essential to both output forms and interpretation. Other points of difference: (1) lexical entries may lack features, but output/interpretable forms must be fully specified, (2) lexical entries may have floating features/nodes (e.g. tonal morphemes, Tongan definitive accent), but not output/interpretable forms.<sup>2</sup> This relates to the Classic Duplication Problem I mentioned earlier: although lexical redundancy rules and phonological rules showed close similarities, they also had differences – i.e. these ones.

- Tesar & Smolensky have argued that the form of a lexical entry is determined by constraint interaction: a set of potential inputs for a given output is produced, and the input that forces the output to violate fewest constraints is selected. Morphemes that exhibit allomorphy require some further emendations: effectively, a ‘compromise’ input is selected.
- If constraints on interpretability were really Form constraints stated in *GEN*, then we would expect all inputs to be interpretable. However, they are not, therefore *GEN* must be able to produce uninterpretable forms.

The fact that inviolable constraints only apply to the input set and not the output set falls out naturally from the present approach: inputs are generated freely, but outputs are constrained by interpretability.

## 5 Origins

Up to this point, I have concentrated on implications of the proposal for the form of the Form. There are other implications, too, but they relate to the origins of the language faculty as a whole. At this point, I am actually going to address why the Duplication problem I brought up earlier on really is a problem.

To refresh your memories:

I said that it was a problem if a Form constraint banned an uninterpretable form. My reason was rather weak: I appealed to a notion of ‘redundancy’ – why have two things in the grammar doing essentially the same thing? Appealing to redundancy has its place, but it is not the be-all and end-all of theorizing.

- Let us suppose that we have a single output grammar. Therefore, if  $\Sigma$  is uninterpretable, then  $*\Sigma$  is a possible constraint.
- If banning structure  $\Sigma$  is essential to interpretation, then  $*\Sigma$  is a necessary part of the Form. If the Form hadn’t evolved that way, we would have no speech. In fact,  $*\Sigma$  as

---

<sup>2</sup> These points still apply to Tesar & Smolensky’s algorithm since the morphemes I’m talking about show alternations.

an adaptation – if lg offers some sort of survival advantage, say, and  $\Sigma$  is what you need to make it a reality, the  $\Sigma$  has an evolutionary justification that far outweighs any appeal to ‘redundancy’.

- However, there are  $\Sigma$ s that are not essential to speech production. For example, the NCC is inviolable, but not every most harmonic form violates the NCC – only a subset (e.g. in metathesis) do. We can’t use the same (evolutionary) reasoning to account for the NCCs Form presence as we did for, e.g., prosodification.

- Put this way, it is something of a conundrum where the Form NCC came from and why.

We can only guess at possible origins here: perhaps FL wants to be more computationally efficient, and eliminating uninterpretable forms would somehow do this.

However, uninterpretability in single-output theories is one of the few current ways to ensure ungrammaticality/unproduceability, so the evolutionary pressure can’t be that great or pervasive.

It seems that the best we can do is appeal to random evolutionary mutation (*not* adaptation). Quite frankly, given the current state of evolutionary understanding, this is fine. It seems difficult to motivate every constraint from an adaptive standpoint.

- With this conclusion, there arises the question of how many inviolable non-essential constraints there are in the Form. I think there are quite a few – the NCC, strict locality of spreading, etc. As the number mounts, our appeals to random change evolution also mounts. We end up with a very fortuitous Form indeed.

The alternative is to adopt the dynamic interpretive device. This changes the whole perspective on why the Form is the way it is. For a start, there is no need to posit constraints in the Form. In effect, no constraint is *essential* to the production of language since the Interpretation weeds out the uninterpretable forms anyway. The *only* condition on the Form is that it produces *some* interpretable forms. There is even no condition on the ratio of interpretable:uninterpretable forms; the Form may wildly overgenerate, and that would present no problem to generation. There is no need to appeal to random chance acts of evolution, since there are no constraints that need this *raison d’être*.

In short, the Dynamic Interpretation device makes the evolution of language a much simpler affair: the Form does not have to be so adaptive to ‘fit’ with the Interpretation. It can be an extremely loose fit indeed.

In fact, the Form does not need to be at adaptive, nor – I surmise – is there any real *pressure* for it to be adaptive at all. As long as it produces some interpretable outputs, we’re all set.

This allows us to address a question raised by Chomsky in a recent paper (1999:1): “To what extent is the human faculty of language FL an optional solution to minimal design specifications, conditions that must be satisfied for language to be usable at all?” [Full quote in handout]

The present answer is that FL is potentially not much of an optimal solution at all. This contrasts with most current theories where FL is a very optimal solution, and the places where the Form overgenerates are crucial to the theory: a way of accounting for the illformedness of certain forms.

So, if the Form is not adaptive, nor is there any/much pressure on it to adapt to the demands of the Interpretation, we are faced with a whole raft of other questions: if the Form is not terribly adaptive, where did it come from and why is it this way? Is it an exaptation, and if so, from what? Whatever the answer is to these questions, the

Dynamic Interpretation Device seems to knock a chink in the idea that FL is very adaptive. At least *some* parts of it are not, nor should be. If so, we are essentially left without a biological source for (certain) Form constraints. Our only way of measuring whether a constraint is reasonable or not is whether it is mathematically simple and conforms to the origins of the exapted Form – of the latter we can say nothing at this time.

Not only does the present proposal not require much adaptation, it avoids the threshold problem: “How much did the Form have to evolve before it was useful to the Interpretation?” The answer is: only as much as it needed to supply *some* interpretable outputs to the Interpretation.

*Q: Why do you think that you need to justify every constraint from an adaptive/evolutionary standpoint anyway?*

A: It seems like a good idea if we’re going to say that constraints are innate. Since I’m not a functionalist/behaviourist, I have to say that any putative functional adaptation is species-level. So, constraints must have evolved, as Chomsky & Lasnik (1977) argue for the *that-trace* constraint. I’m not keen on random chance providing us with what we have. Perhaps some constraints didn’t evolve for any adaptive reason, but I think this view mounts in improbability as the constraints mount and their relation to functional concerns becomes clearer. But I’m not keen on being a functionalist/behaviourist by saying that constraints are learned, so I have to attribute functional relatedness to adaptation.

Now, adaptation comes about through a survival advantage. If a constraint doesn’t give a survival advantage, then an evolutionary argument isn’t going to hold too much water.

Therefore it’s simply better to get rid of the constraint entirely, then we don’t have to worry about where it came from. Hence the present proposal.

## FAQ

Q1: What about cases where an uninterpretable structure never triggers anything? – it always is the most harmonic form, causing the derivation to crash

A: This relates to the debate about grammaticality and OT: how do we deal with absolute unproducibility? OT took away the notion of ungrammaticality, and now I am getting rid of uninterpretability. I don’t see this as a problem. Certainly, you cannot rely on uninterpretability to ban a certain output configuration. It just makes the stakes higher.

**If a form never appears in the output, it just means that the input has neutralized to a possible form.**

Q2: What about exceptions to ‘inviolable’ principles?

A: We now have a formalist explanation for those. Here’s one of the least controversial examples: Connectedness states that everything in a structure must be connected – both ordered and dominated. But floating tones are an exception – perhaps the only exception. Some tones have been argued to be floating even in the output. How do we explain this?

The present theory actually does better than the alternatives. If there was an GEN constraint Connectedness, floating tones should be impossible. With the interpretive theory, we can say that floating tones *are* interpretable, while other floating nodes are not. There may well be – probably is – a functional reason for this. The relation between pitch change and segmental structure is fairly loose: it is often for late realization of pitch to happen, etc. Perhaps this indicates that the production of pitch is relatively independent of other gestural production. So, it is reasonable to expect that a tone that is



unconnected to a segment can still be given an interpretation. In a way, this is a *functionalist explanation*, but *the present theory is entirely formalist*. [EXPAND?]

Q3: Why not say the Interpretation *repairs* the NCC?

A: This imbues the Interpretation with a lot of power. And just too much when you think of the range of possible repairs. For example, suppose we had an NCC violation. How would the phonetics repair this? Languages deal with the violation in a large variety of different ways, so we'd have to start imputing the power of the phonology in the Interpretation.

Q4: Doesn't this the Duplication Problem only arise given a certain conception of the interpretation?

A: Any grammar that overgenerates runs the risk of duplication, and every theory of grammar (apart from Montague's) does overgenerate. Also, you may notice how commonplace it is for linguists to dismiss a structure as Formally/phonologically impossible *because* it is uninterpretable. e.g. [+high, +low] vowels.

Q5: How do you know when a Form constraint should instead be explained by Interpretive means?

A: That really comes down to your theory of Interpretation. I certainly didn't even try to present a theory of Interpretation today. Ultimately, it would be nice to have one. However, a well-worked out theory of Interpretation doesn't have to be provided before the arguments I made today could be plausible. All you need to assume is an overgenerating Form, thereby recognizing the role of uninterpretability in grammatical form. The next step is to deduce what you can about the interpretation and see if it overlaps with any Form principles.

Q6: What do you mean by 'The Interpretation'?

A: I have deliberately tried to keep this a little nebulous today. In fact, I want to keep it as nebulous as possible: the Interpretation is "the module to which the Form sends its output". It is quite current to say that this module isn't the end of the story: it translates Form form into another representation – for phonology, into a more gradient representation; for Form, into a predicate-argument language, and then there may be (probably are) several modules after these. The important aspects of the Interpretation are that it has no direct effect on the Form: it doesn't constrain grammaticality. In addition, it can't repair Form structures, and is in fact quite limited: it fails in the face of ambiguity, incompleteness, and perhaps extraneous elements. Apart from that, I don't think anything more needs to be said about the module for present purposes, although ultimately, a full theory of the Interpretation needs to be presented (and - of course - work in this area has been underway for quite some time).

## References

- Archangeli, D. and D. Pulleyblank (1994). Grounded Phonology. Cambridge, Massachusetts, MIT Press.
- Bach, Emmon. 1976. "An Extension of Classical Transformational Grammar." in *Problems of Linguistic Metatheory*. Proceedings of the 1976 Conference at Michigan State University, pp.183-224.
- Chomsky, N. (1965). Aspects of the Theory of Form. Cambridge, MIT Press.
- Chomsky, N. (1995). The Minimalist Program. Cambridge, MA, MIT Press.
- Chomsky, N. (1998). "Minimalist Inquiries." .

- Chomsky, N. (1999). "Derivation by Phase." *MIT Occasional Papers in Linguistics* 18.
- Chomsky, N. and H. Lasnik (1977). "Filters and control." *Linguistic Inquiry* 8: 425-504.
- Chomsky, N. and M. Halle (1968). *The Sound Pattern of English*. New York, Harper & Row.
- Clements, G.N et al. (199?) "Symposium on Non-linear phonology and phonetic interpretation."
- Dryer, Matthew S. 1988. "Universals of Negative Position." in Michael Hammond, Edith Moravcsik, and Jessica Wirth (eds.). *Studies in Form Typology*. pp. 93-124. Amsterdam: John Benjamins
- Goldrick, M. and P. Smolensky (1999). Opacity, turbid representations, and output-based explanation. Workshop on the Lexicon in Phonetics and Phonology, Edmonton, Alberta.
- Goldsmith, J. (1976). Autosegmental phonology, MIT.
- Goldsmith, J. (1979). The aims of autosegmental phonology. Current Approaches to Phonological Theory. D. Dinnsen. Bloomington, Indiana University Press: 202-22.
- Guthrie, M. (1970). Comparative Bantu. Farnborough, Hants, Gregg.
- Hale, Kenneth & Samuel Jay Keyser. 1994. "On Argument Structure and the Lexical Expression of Form Relations." in Kenneth Hale & Samuel Jay Keyser (eds). *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. pp. 53-109. Cambridge: MIT Press.
- Hammond, M. (1988). "On deriving the Well-formedness Condition." *Linguistic Inquiry* 19: 319-325.
- Hauser, Marc D. 1997. *The Evolution of Communication*. Cambridge, MA: MIT Press.
- Heim, Irene & Angelika Kratzer. 1998. *Semantics in Generative Grammar*. Malden, MA: Blackwell.
- Jackendoff, Ray. 1997. *The Architecture of the Language Faculty*. Linguistic Inquiry Monograph 28. Cambridge, MA: MIT Press.
- Kaye, J. (1974). "Opacity and recoverability in phonology." *Canadian Journal of Linguistics* 19: 134-149.
- Kayne, Richard. 1984. *Connectedness and Binary Branching*. Dordrecht: Foris.
- Kayne, Richard. 1994. *The Antisymmetry of Form*. Cambridge, MA: MIT Press.
- McCarthy, J. (1979). Formal Problems in Semitic Phonology and Morphology, MIT.
- McCarthy, J. (1986). "Lexical phonology and nonconcatenative morphology in the history of Chaha." *Revue québécoise de linguistique* 16: 209-228.
- McCarthy, J. (1986). "OCP Effects: Gemination and antigemination." *Linguistic Inquiry* 17: 207-263.
- McCarthy, J. (1998). Sympathy and phonological opacity. Amherst, MA, University of Massachusetts, Amherst.
- McCarthy, J. (1999). "Sympathy and phonological opacity." *Phonology* 16(2): xxx-xxx.
- McCarthy, J. and A. Prince (1986). Prosodic Morphology, University of Massachusetts at Amherst and Brandeis University.
- McCawley, James D. (1982). "Parentheticals and Discontinuous Constituent Structure." *LI* 13: 91-106.
- Montague, Richard. 1974. *Formal Philosophy: Selected Papers of Richard Montague*. Richard H. Thomason (ed.). New Haven: Yale University Press.

- Newmeyer, Frederick J. 1998. *Language Form and Language Function*. Cambridge, MA: MIT Press.
- Orgun, C. O. and R. Sprouse (1997). From MParse to control: Deriving ungrammaticality. Berkeley, CA, University of California.
- Pinker, S. & P. Bloom (1990) "Natural Language and Natural Selection." *Behavioral and Brain Sciences* 13:707-727.
- Prince, A. and P. Smolensky (1993). *Optimality Theory: Constraint interaction in generative grammar*. New Brunswick, NJ, Rutgers University.
- Pulleyblank, D. (1983). Extratonicity and polarity. Stanford, Stanford Linguistic Association.
- Sagey, E. (1988). "On the ill-formedness of crossing association lines." Linguistic Inquiry 19: 109-118.
- Selkirk, E. (1984). Phonology and Form: The Relation between Sound and Structure. Cambridge, MA, MIT Press.
- Steriade, D. (1995). Underspecification and markedness. Handbook of Phonological Theory. J. Goldsmith. Cambridge, MA, Blackwell: 114-174.
- Zoll, C. (1996). Parsing below the Segment in a Constraint-based Framework. Department of Linguistics. Berkeley, University of California, Berkeley.

Paul de Lacy  
University of Massachusetts, Amherst  
Tuesday, March 28, 2000  
<delacy@linguist.umass.edu>  
<http://www-unix.oit.umass.edu/~delacy>