Clarifying opacity

WILLIAM J. IDSARDI

Abstract

Various issues regarding phonological opacity are clarified. Data is drawn from Tiberian Hebrew to illustrate the nature and extent of opacities in phonological grammars. Various approaches to opacity within Optimality Theory are discussed and criticized using the Tiberian Hebrew data. It is concluded that no current approach to opacity within Optimality Theory covers the range of facts that rule-based theories explain. Implications for future research and models of phonological acquisition are also considered.

1. Introduction

In this article, I seek to clarify some of the issues surrounding opacity, derivations and Optimality Theory (OT), and to argue that rather than being a "residual issue" (Kager 1999: chapter 9), this is the single most important issue in current phonological theory. In this short article, I will provide:
- a brief overview of how opacity effects arise in OT and Generative Phonology (GP);
- data from Tiberian Hebrew which show how extensive and pervasive opacity is;
- a brief survey of the responses to opacity effects in OT and how they fail on the Tiberian Hebrew data;
- an argument that extensions to OT invalidate previous work on learnability in OT.

2. What is opacity?

There have been a fair number of definitions of opacity, some formulations linked very closely with the prevailing theory at the time, others less so. One common approach is to provide a multi-faceted definition, so as to elucidate (or at least name) various sub-types. I think that the multiplicity of names has instead had the
effect of further obfuscating the issue, and so I would like to provide the simplest possible definition:

(1) An opaque generalization is a generalization that does crucial work in the analysis, but which does not hold of the output form

Kager's choice of opacity sub-types is among the better ones proposed, limiting himself to just two opposing types: non-surface-apparent and non-surface-true. Kager's explanation of these two types are given in (2).

(2) a. "We call a generalization non-surface-apparent if it takes effect at a level concealed at the surface. A set of forms undergo a process, although they fail to match its structural description at the surface... The effect is an overapplication of [the process]." (p 373)

b. "A generalization is non-surface-true if it has cases of non-application at the surface that are controlled by a non-surface level. This is the logically opposite situation of the previous cases A set of forms fail to undergo a process even though their surface forms match its structural description... The effect is an underapplication of [the process]." (p 374)

I will use the terms overapplication opacity and underapplication opacity to refer to these two types. Overapplication opacity presents a problem for OT as a surface form seems to have changed without motivation, that is, the form appears to have gratuitous faithfulness violations. Underapplication opacity results in gratuitious markedness violations; illicit structures appear in the output even though they are banned in similar forms. As Kager explains, overapplication opacity corresponds to counter-bleeding rule orders and underapplication opacity to counter-feeding rule orders.

The rule of Epenthesis in Tiberian Hebrew, which breaks up final consonant clusters, exhibits both kinds of opacity, so it is an excellent choice to illustrate the problems (and it is the same problem discussed by McCarthy 1999, 2000). In this article, I consider only a few processes in Tiberian Hebrew and ignore many auxiliary changes. The formulations in (3) are also not precise, much more extensive discussion can be found in Prince (1975) and Malone (1993). For instance, Apocope is in fact part of a more extensive metrically-based process of vowel deletion. For the present purposes, we are interested in the following processes:

(3) Stress: stress the final syllable if closed, otherwise the penultimate
Lengthening: stressed vowels are lengthened in open syllables (and in
CVC#)
Gemination: nC → CC for non-guttural C
Degemination: CC → C / #
Clarifying opacity

Epenthesis: insert /ɛ/ before an unsyllabifiable consonant
Umlaut: a → e / _ C ɛ within a foot
ʔ-deletion: delete ? in syllable coda position
Sprantization: singleton non-emphatic stops → fricatives / V _
Apocope: delete word-final short unstressed vowels
ɔ → a: ɔ → a / _ CC
Lowering: ε → a / _ [RIR] ɔ

Derivations for two words are shown in (4)

(4) UR tæn? ‘basket’ gəmɔl-ti ‘you (sg. fem.) did, accomplished’
Stress, etc: tán? gəmɔlti
Epenthesis: tánɛʔ?
Umlaut: tɛnʔ?
ʔ-deletion: tɛnɛ
Apocope: gəmɔlt
ɔ → a: gəmɔlt
SR [tɛnɛ] [gəmɔlt]

In the form [tɛnɛ] the epenthesis process is in counter-bleeding order with ʔ-deletion, and so epenthesis appears to overapply from the perspective of the surface form. We can no longer see what motivates epenthesis once /ʔ/ is lost. From an OT point of view, a form such as *[tán] would be preferable, as it would be more faithful to the underlying form. The actually occurring form has two faithfulness violations (Max-C and Dep-V), whereas the alternative *[tán] has only one, Max-C. The apparently gratuitous Dep-V violation must be explained in some way other than the surface phonotactic patterns in the language, as there are words of this shape, for example [gán] ‘garden’.

In contrast, in the form [gəmɔlt] the problem is that epenthesis has underapplied, failing to break up the final [lt] cluster, and leaving a gratuitous markedness violation (See Idsardi (1998: 68–69) for arguments supporting /-ti/ as the second person feminine singular ending. The underapplication of epenthesis cannot be explained by phonotactic patterns or by faithfulness based on the segments present in the surface form, for there are minimally different words, such as /dal-t/ → [dʃɛtθ] ‘door (fem.)’ in which epenthesis does apply (along with sprantization of the final /θ/).

The other important fact illustrated by this small data set is that a single process (here apocope) can interact in a variety of different ways with many other processes. The form [tɛnɛ] itself illustrates this sort of complex interaction as this form is also opaque with regard to apocope, having a final short unstressed vowel. Thus, the three processes of epenthesis, apocope and ʔ-deletion yield complex interactions with multiple opacities even within a single form. We have discovered that epenthesis is in both counter-feeding and counter-bleeding relationships with
other rules, and so at the surface shows both underapplication and overapplication opacities. In addition, ephenthesis is in counter-feeding order with the stress placement processes, as final closed syllables are normally stressed. Therefore words like [dɛlɛθ] also have opaque stress patterns. In fact, the stress patterns show both underapplication and overapplication opacities simultaneously. Having stress on the first vowel is an overapplication opacity, as we cannot explain from the surface form alone why the vowel has received stress, while not having stress on the final vowel is an underapplication opacity. Tiberian Hebrew also has a process which lengthens stressed vowels in open syllables. Ephenthesis is also in counter-feeding order with this lengthening rule, so that the short [ɛ]’s in [dɛlɛθ] and [tɛne] are also opaque, violating the markedness constraint Stress-to-Weight.

These opacities simply compound as we consider more forms and processes of the language. It will suffice at this point to consider one additional form of Tiberian Hebrew, [zɛrɛθ] ‘seed’, whose derivation is shown in (5). (The importance of this form for phonological derivations is also discussed by McCarthy 1994: 209–212.) This word shows the operation of a process of vowel lowering whereby vowels are lowered when they are followed by a pharyngeal consonant in the same syllable. Vowel lowering serves to make the umlaut process opaque, as we can no longer see the motivation for fronting the vowel in the initial syllable (compare [zarɛθ] ‘our seed’ where the conditions for umlaut are not met).

(5)  

| UR    | zarɛθ | ‘seed’ |
| Stress: | zárɛθ |
| Ephenthesis: | záɛrɛθ |
| Umlaut: | zéřɛθ |
| Lowering: | zɛrɛθ |
| SR    | [zɛrɛθ] |

The correct form is [zɛrɛθ], even though the alternative *[záɛrɛθ]* would be more faithful to the input, by preserving the quality of the initial vowel. Modern Hebrew (that is, Contemporary Standard Israeli Hebrew, Boazky 1978) increases the opacity even further, deleting /ɛ/, so that the modern form is [zɛrɛθ], making vowel lowering opaque also. So the correct form can be opaque in several ways at once: in stress pattern, vowel length, vowel quality, umlaut and ephenthesis.

We can now quite vividly see that there is an intricate web of relations between the processes, so that opacity is not a simple phenomenon, but a highly interactive one. Ephenthesis has relationships with several other processes, and all of these interactions must be examined to arrive at the correct grammar. This is, in fact, an entirely typical situation. Languages exhibit multiple opacities, not just simple ones, and any theory of phonology must be able to explain why opacity is ubiquitous throughout human languages.
3. More sources of opacity in Lexical Phonology

Rule-based theories of phonology, such as Lexical Phonology, have additional resources besides rule-ordering, and these mechanisms can also serve as separate sources of opacity. For example, rules which apply only in derived environments will have underapplication opacities at the surface in non-derived environments. Lubowicz (1998) employs constraint conjunction in order to analyze derived environment effects in OT.

Rules subject to the Structure Preservation Condition may block a process from creating a segment, even though that segment may be allowed at the surface. A case like this from Modern Hebrew (Bolozy 1978) is discussed in Idsardi (1997), and briefly recapitulated here. Modern Hebrew has a low-level process of voicing assimilation in fast speech. Modern Hebrew has also restructured the phonemic system somewhat from that of the Biblical period, so that it includes the fricative phonemes /ʃ v x/ but not */θ ð y/. Concomitantly with the phoneme restructuring, the post-vocalic spirantization of stops has been restricted to apply only to /p b k/ but not to /t d g/. The only reasonable explanation is that spirantization has changed to become structure preserving in Modern Hebrew. Thus, we have derivations such as those in (6) (the UR's are simplified somewhat here).

(6)  

<table>
<thead>
<tr>
<th>UR</th>
<th>ji-kbōf</th>
<th>ji-gsós</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spirantization</td>
<td>jixbōf</td>
<td>blocked by Structure Preservation Condition</td>
</tr>
<tr>
<td>Voicing assimilation</td>
<td>ji-ybōf</td>
<td>jiksós</td>
</tr>
</tbody>
</table>

The Structure Preservation Condition on spirantization results in two kinds of opacity. The ban against */y/ is opaque in [jiybōl] because y is created through voicing assimilation, creating an overapplication opacity at the surface. In [jiksōs], we have an underapplication opacity of spirantization, as we have an instance of a stop immediately following a vowel, a stop whose fricative counterpart [x] is a member of the underlying inventory, and thus from the surface point of view, having [x] would satisfy structure preservation.

Rule-based Lexical Phonology predicts such cases, for in that theory structure preservation “turns off” at some point, and later rules are much less likely to preserve inventory structure. Thus, rule-based Lexical Phonology predicts opacities based on Structure Preservation. In fact, Lexical Phonology even prefers such languages, for lexical rules are normally structure preserving and postlexical rules are normally not structure preserving (see Kiparsky 1993 for additional examples). This example shows that structure preservation effects cannot be evaluated on the surface forms alone, and therefore structure preservation creates additional cases of opacity. In a similar way, elsewhere condition relations between rules can also interact to produce additional opacities. Unfortunately, such cases are too involved to be reviewed in this article.
4. OT has opacity capacity

One common misconception in the on-going debate about opacity is the belief that standard OT is inherently incapable of handling any kind of opacity. This is not true. Even in the original formulation of OT (Prince and Smolensky 1993) the theory has a mechanism for handling opaque generalizations – generalizations that do not hold of surface forms – namely, constraint domination. For example, if we have an input /kæt/ and if Max-C >> NoCoda then [kæt] > *[kæ], and thus NoCoda will be an opaque generalization for this form. But, depending on the rest of the constraints and their rankings, NoCoda may actually do crucial work in the analysis of the language, perhaps serving to eliminate other candidates such as *[kæt]. But, in any event, NoCoda is not surface-true. OT does, however, claim that opacity will be minimized, that is, that constraint violations will be minimal. To his credit, Kager does not fall into this common trap.

(7) "Of course, the fact that a generalization is not surface-true is, by itself, not immediately problematic for OT. Output forms are always in violation of some constraints" (p. 372)

So the basic way to handle opaque generalizations in OT is to find a dominating constraint which forces the violation of the constraint in question. Ideally, the dominating constraint should be a simple faithfulness or markedness constraint, but as we shall see, other, novel types of constraints are required.

5. OT responses to opacity

There have been several proposals for different ways to respond to the "crisis of opacity" within OT.

5.1. Denial and reanalysis

One possibility suggested to me by several people is that opacity is really an overrated phenomenon in generative phonology. Instead of resorting to opaque analyses, perhaps we can find non-opaque reasons for the changes that occur. I know of two concrete proposals for the [tɛnɛ] case, one by Balcaen and Hall (1999) and the other by Bruening (1999).

Balcaen and Hall (1999) propose an analysis in which /ʔ/ → [ɛ] directly. With this analysis there is no Max-C or Dep-V violation, simply a fairly radical featural change in the segment. The problem here is that this creates a conspiracy between the fact that /ʔ/ changes to [ɛ] and [ɛ] is also the vowel that breaks up consonant clusters. Why should both changes select the same vowel? Furthermore, since /ʔ/ is a guttural consonant, and guttural consonants show affinities for [a] in other
circumstances (e.g., lowering), we might expect that [a] would be a more likely realization of /ʔ/ that [ε] is. But, more importantly, this analysis does not explain the other opacities in the [těne] form, namely the short stressed vowel (underapplication of lengthening), and the final short vowel (underapplication of apocope). The analysis also does not provide explanations of the other opacity effects observed above, such as the failure of epenthesis in [goymált] and the opaque umlaut in [zěraš].

Bruening (1999) proposes a different reason for [těne] and other nouns which have the underlying shape /CVCC/. These nouns are called segholates in the traditional literature due to the epenthesis of /ɛ/, which is called seghol in Hebrew. Bruening proposes that the segholates are a morphological class and have a high-ranking set of constraints which forces them to take the shape of a bisyllabic trochee. Bruening also proposes that other nouns, such as [dɔvɔ]|t] 'word', are in a different morphological class, and therefore they are not subject to the morphological trochaic constraint and thus they surface as iambic, which Bruening says is the default stress pattern. This is a more successful reanalysis, as forcing a bisyllabic trochaic shape motivates both the short vowel in the stressed syllable (contra lengthening) and the retention of the final short vowel (contra apocope). Words such as [goymált] will not be in the trochaic morphological class, and so they will prefer to end in an iamb, and this will also provide the necessary pressure against epenthesis in [goymált]. Although Bruening's analysis requires lexically-specified differences in prosodic shape preferences between words, the analysis is notable for its ability to realign several facts once a lexical difference between trochaic and iambic words is admitted. However, it is still the case that there remain opacities which this analysis does not handle, including the opaque umlaut in [zěraš]. Another case arises with /CVCC/ roots when they are augmented with the feminine /-t/ ending. We have already seen one such case, [dělčθ] 'door'. It is instructive to compare this word with a couple of minimally different ones, as shown in (8).

(8)  

<table>
<thead>
<tr>
<th>UR</th>
<th>tǎnʔ</th>
<th>dal-t</th>
<th>ban-t</th>
<th>rabb</th>
<th>Stress, etc :</th>
<th>tǎnʔ</th>
<th>dált</th>
<th>bánt</th>
<th>ráb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gemination:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degemination:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epenthesis:</td>
<td>tǎneʔ</td>
<td>dálęt</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Umlaut:</td>
<td>těneʔ</td>
<td>dělčθ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>?-deletion:</td>
<td>těne</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spirantization:</td>
<td>těne</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SR</td>
<td>[těne]</td>
<td>[dělčθ]</td>
<td>[báθ]</td>
<td>[rałv]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The underlying /n/ in /ban-t/ is motivated by other forms in the paradigm, such as the plural form [bo:nοθ] ← /ban-ot/ 'daughters' and the related form /ben/ 'son'. Notice that word-final degemination is required to explain forms with root-final geminates, such as [rałv], and that degemination must feed spirantization. What
is crucially important here is that [báθ] is not pronounced *[bɛnθ], even though this would preserve all of the consonants and it would be parallel to [dɛlθ], and it would be a bisyllabic trochee. We cannot allow [báθ] to satisfy the trochaic requirement, because then the parallel *[tán] would as well. The rule-based analysis can handle, in a unified way, the complex interaction of gemination, degemination, epenthesis and ?-deletion, whereas Bruening's explanation in terms of bisyllabic trochees necessarily falls short for at least one item.

To summarize, it is possible to realign one or even several opacities under a different generalization about the surface forms. Bruening's proposal to institute two type of feet in Tiberian Hebrew allows for flexibility and power in explaining different patterns of stress and quantity, and allows for surface-true statements for stress and quantity in many words. However, this only begins to cover the Tiberian Hebrew data, which exhibits many additional opacities. These proposals may be very promising starts, but they must be augmented to cover the rest of the data. But the proposals also offer fresh insights into the nature of these problems in Tiberian Hebrew, and thus will provide a real improvement in our understanding of these problems as we integrate the new insights into our understanding of the complete phonology.

5.2. *Brute force I. Abstract structure*

One answer to opaque effects offered in the original OT formulation was an enrichment of the notion of the output structure. Prince and Smolensky (1993) proposed that deleted elements were not deleted within the phonology per se; rather, they were left unparsed, and then would be left unpronounced by general convention. Goldrick and Smolensky (1999) and Goldrick (1999) offer an updated version of this abstract structure approach, calling this approach *Turbidity*. Turbidity posits a difference between the "projection" of a segment and the "pronunciation" of a segment. Constraints can refer to a segment’s projection or pronunciation or both. Empirically, Turbidity is no different from various containment models examined in Idsardi (1998), and so the arguments given there from ?-deletion and syncope in Tiberian Hebrew apply equally to this case. Rather than rehash these arguments, we will instead consider the difference in (8) between [tɛnɛ] and [báθ]. Since the epenthesis in [tɛnɛ] is opaque, we must consider the projection of the /θ/, not just its pronunciation (which is null). But if the projections of /n/ and /l/ are considered in /ban-θ/, then *[bɛnθ]* will be chosen as optimal. In fact, the underapplication of epenthesis in this form is not explicable in terms of projection. Turbidity, like unparsed segments, can account for overapplication due to underlying segments, but it cannot deal with underapplication through bleeding when the bleeding process is itself obscured at the surface. There is no aspect of the projection of /n/ that is maintained at the surface, and therefore projection has no utility in explaining [báθ].
As a general matter, if left unchecked the postulation of additional abstract structure can provide an arbitrary amount of power. For example, if we allow each segment to carry an arbitrarily large set of subscripts we can indicate in a "single representation" the entire derivational history of each segment by encoding in the subscripts which lines in the derivation contain that segment. That is, representations can be used to mimic derivations even to the smallest detail. However, it is also true that in the history of phonology, the postulation of new (but constrained) kinds of abstract structure has been exactly what the theory needs at certain points—two excellent examples are Goldsmith (1976) (autosegmental associations) and Rainey (1990) (non-linear representations for reduplication).

\[ \text{5.3 Brute force II: Constraint proliferation} \]

Another way to handle opacity problems is to give up trying to find a unifying generalization and simply add additional constraints to handle the special cases. Frequently, these new constraints will be morphologically conditioned ones, without any independent justification. For example, consider the form [têne], with opaque epenthesis. We could postulate a constraint such as (9):

(9) No triliteral noun root can surface with the correspondent of the second radical consonant in word-final position.

This bars any [n]-final form from surfacing for /tan/. Faced with opaque umlaut in [zféra], we might propose an additional dissimilation constraint, (10):

(10) *áCa (where C is non-guttural, compare [báñal] ← /báñl/ 'lord, master')

The problem is that (9) does not connect ?-deletion with epenthesis and (10) is left unconnected to umlaut. These analyses may work, but they do not shed any light on the connections between the processes in the language. Since languages do make use of arbitrary suppletion (such as "go/went" in English) it is almost certainly impossible to construct grammatical theory so as to universally rule out analyses employing constraints similar to (9) and (10) Such analyses will manage to grind out the facts, and will be non-opaque, but will have nothing else to recommend them.

\[ \text{5.4. Two-level constraints} \]

Kager (1999) provides an excellent discussion of two-level constraints in section 9.2.2. It is sufficient to note here that umlaut in [zféra] is defined on a crucially intermediate stage of the derivation, after epenthesis (which creates the context) but before lowering (which destroys the context). The crucial information is neither present in the input, nor in the output, and such cases cannot be handled with two-level constraints.
5.5 Intermediate levels and chained OT calculations

A possibility which is gaining a significant number of adherents is to chain together several OT calculations with level-ordering as in Lexical Phonology. Kager's discussion in section 9.2.3 foreshadows many of the developments of Kiparsky (1999) and Bermúdez-Otero (1999). Kager deals very well with this idea, which marries OT with regular rule-based serialism. The conceptual problem is that this gives OT two distinct computational devices with which to do phonology: tableau calculations and procedural derivations. In addition, the levels must be empirically motivated, and not all opacity can be reduced to well-motivated level-ordering. We have already seen a simple example, the [zéraṭ] case, for which the relevant processes (epenthesis, umlaut and lowering) are all postlexical processes. Any postulation of level-ordering here would be entirely ad-hoc. Dresher (1982) contains an excellent summary of the issues regarding level-ordering and how they relate to Tiberian Hebrew.

5.6 OO-correspondence

Kager discusses the use of paradigm-uniformity constraints (or Output-Output constraints) in section 9.2.4. As McCarthy (1998) points out, such constraints offer no help in the case of [tēné] as this is the only form in the paradigm which displays the epenthesis and hence the opacity. Thus, the opacity cannot be bootstrapped from any other related form. Additional arguments from Tiberian Hebrew against paradigm-uniformity accounts are given in Idsardi (1998: 59 ff.).

5.7 Sympathy and culminativity

The most radical and imaginative new proposal for handling opacity effects within OT is McCarthy's 1998 Sympathy Theory proposal, which Kager very ably discusses in section 9.2.5. McCarthy's focus is, in fact, the Tiberian Hebrew [tēné] case. McCarthy selects a sympathy candidate with Max-C (⊗tēné?), and requires all of the vowels from the sympathy form to be present in the output form (Max-V-⊗O). We have already seen several forms which are problematic for this analysis. The forms [gɔrmált] and [bāth] cannot be handled because epenthesis underapplies in these forms, and Max-V-⊗O serves to explain only overapplication opacities, thus predicting *[gɔrméθə] and *[bénéθə]. The multiple opacities exhibited by forms such as [zéraṭ] provide a particular challenge to sympathy theory, the details of this argument can be found in Idsardi (1999).

McCarthy (1999) offers a further refinement of sympathy theory, which he calls Cumulativity. The refinement does not affect the problems raised by [gɔrmált], [bāth] and [zéraṭ]. Much of McCarthy's discussion is directed against derivational Duke-of-York gambits, in which A → B → A. In this context,
the derivation of [bâθ] is also relevant. The final /t/ geminates, and then degem-
inates in a Duke-of-York gambit so as to allow spirantization. But this Duke-of-
York analysis coheres with many other facts of Tiberian Hebrew. For instance,
we know from other forms that gemination (i.e., total anticipatory assimilation of
/n/) precedes and blocks spirantization ([yippɔːl] ← yanpol ← /ya-napol/ ‘he will
fall down’). We also know that word-final degemination precedes and feeds spi-
rantization ([rav] ← /rabb/ ‘much’). Furthermore, we know that epenthesis feeds
spirantization ([delɔθ] ← /dal-t/ ‘door’), and that degemination must precede and
therefore block epenthesis ([rav] ← /rabb/ ‘much’ not *[rɛɛvɛv]) (The alternative
analysis in which the geminate structure blocks epenthesis is not well-supported in
Hebrew. Hebrew does allow long-distance geminates morphologically, and there
is one form which exceptionally allows epenthesis in this environment, [rɛɛvɛ] ‘ter-
or’. Additionally, Hebrew in the post-Biblical period does allow epenthesis into
geminate structures more generally.) Given this array of facts, the best answer is
for gemination to feed degemination even though this results in a Duke-of-York
gambit in the derivation of [bâθ].

Given how well the derivation of [bâθ] coheres with the rest of the Tiberian
Hebrew facts, perhaps we should instead question our reaction to Duke-of-York
gambits. Turning to another science for clarification of this problem by analogy,
consider the operation of a catalyst in a chemical reaction. Given a reaction A
+ B → X + Y, the introduction of a catalyst, C, doesn’t change the fundamental
nature of the reaction, you simply get the catalyst back in the products. A + B +
C → X + Y + C. The catalyst seems to play no role. But with the proper catalyst,
the reaction requires less energy to be viable and proceeds more quickly. The
explanation is that the catalyst reacts with one of the reagents, creating a crucial
intermediate compound, but the catalyst is then “bumped out” by another reagent
and comes back in the products. The operation of catalysts in chemistry may seem
unexpected, even magical, but it is real. The lessons from chemistry are that nature
operates in ways we find surprising and that reactions must be understood both in
an overall way and in step-by-step way to fully understand them.

5.8. Constraint conjunction

Kager discusses constraint conjunction in section 9.2.6. The idea of constraint
conjunction is that by putting together two constraints, candidates which violate
both of the constraints simultaneously can be eliminated, and a different winner
will be selected. For the [tɛnt] case we could conjoin [Max-C & NoCoda] >>
Max-C >> NoCoda. Intuitively, this would prohibit the deletion of a consonant
that resulted in the creation of a closed syllable. The form *[tɛn] violates both
NoCoda and Max-C and so would be ruled out. However, this analysis doesn’t
work for [bâθ], which eliminates /n/, but still ends up with a closed syllable. Even
if we make the constraint more specific, limiting it to /?/: [Max-? & NoCoda], it
still doesn't work as there are forms such as [qəɾtʃə] → /qəɾtʃə-/ti/ 'you fem sing named'.

On a conceptual level, constraint conjunction adds a new mode of constraint interaction to the theory. When the normal mode of constraint interaction — domination of one constraint by another — will not work, we can create an alternative form of constraint by conjoining two or more constraints, thus getting around the regular ranking interaction.

5.9 Summary

It should be clear from the discussion of the various strategies for handling opacity, which have been proposed in OT, that none of them can insightfully handle all of the opaque effects observed in Tiberian Hebrew. But, as Ito and Mester (1999) correctly point out, we shouldn't expect OT to replace the effects of rule-ordering with a single device; in fact we should expect a multiplicity of devices to replace it. But this ‘toolbox’ approach offers a non-unified alternative to what was a unified rule-based account of opaque phenomena. As we know from the OT argument based on rule-based conspiracies, unified accounts are preferable to non-unified ones (after all, why did Einstein search for a unified field theory, and why did Feynman et al. win the Nobel prize for their partial success, Q.E.D.?). As the toolbox approach also shows, rule-based phonology does not have a monopoly on conspiracies — systems of constraints are just as likely to suffer from recurring partials — overlap in several constraints which cannot be gathered together into a unified account.

There can be little doubt that the most influential responses to opacity have been sympathy theory and constraint conjunction; indeed, these are the longest subsections in Kager’s discussion of opacity. These two approaches are similar in an important way. They create a new kind of constraint interaction in addition to the normal ranking of regular constraints. Because of the additional machinery, these two modifications of OT have important implications for acquisition. The constraint demotion algorithm of Tesar and Smolensky (1998) (or its revision in Prince and Tesar 1999, i.e. Biased Constraint Demotion) cannot by itself explain the acquisition of systems which include sympathetic constraints or constraint conjunction. These systems cannot be learned by constraint demotion alone because the systems allow the learner a choice other than constraint reranking, and so when faced with recalcitrant data, the learner must choose whether to rerank constraints or instead to add a sympathetic constraint or to conjoin together two constraints. Kager also notes related learnability problems posed by constraint conjunction (p. 400).

Certainly the future response from OT proponents — in addition to inventing additional machinery not yet considered — will be to use all of the strategies outlined above. And of course, this is exactly what they should do. However, given the
lack of success, and the paucity of new generalizations or discoveries arising out of these responses, it seems clear that the "crisis of opacity" within OT points to a degenerating character of the research program.

6. Conclusions: Where do we go from here?

I am sure that the most common interpretation of my writings on OT is to conclude that I have nothing good to say about it. And more strongly than that, that my criticisms point out negative problems without offering constructive solutions. So I think many people would like to ask me the question: What aspects of OT should be retained in future theories of phonology? Certainly the notion of a choice from a (small) set of (plausible) alternatives is a powerful and attractive idea. Where OT fails is in its over-reliance on this feature of the theory, making GEN overly powerful (even infinite), and in its abandonment of the traditional derivational structure of GP, with individual forms calculated as the result of "processes". One way of salvaging part of OT is to see OT as the theory of the internal structure of a process. The way to do this is to vastly decrease the power of GEN, to provide only very simple alternatives, or the order of the single fundamental operations in Archangeli and Pulleyblank (1994). In other words, we should restrict GEN so that it can do only one thing at a time. Likewise the set of constraints should be vastly reduced so that only a few general criteria are invoked in any given process. One proposal close in spirit to these ideas is the linearization process in Raimy (1999). The problem Raimy addresses is to linearize a non-linear directed graph through the use of very general principles of maximization (use all nodes, use all arcs), minimization (find the smallest) and preferences (use more recent information). A research program which breaks phonology into a sequence of processual optimizations offers a potential synthesis in the rule-constraint dialectic.

University of Delaware

References


Goldsmith, John (1976) *Autosegmental Phonology*. PhD dissertation. MIT.


— and Computer Science Department. University of Colorado, Boulder.

