Uniwersytet Warszawski
Wydział Neofilologii

Joanna Zaleska
Nr albumu 194 532

Approaches to Phonological Opacity in Optimality Theory: Sympathy and DOT

Praca magisterska
na kierunku filologia
w zakresie filologia angielska

Praca wykonana pod kierunkiem
dr Beaty Łukaszewicz
Instytut Anglistyki
Wydział Neofilologii UW

Warszawa, maj 2007

Oświadczenie kierującego pracą

Oświadczam, że niniejsza praca została przygotowana pod moim kierunkiem i stwierdzam, że spełnia ona warunki do przedstawienia jej w postępowaniu o nadanie tytułu zawodowego.

Data Podpis kierującego pracą

Oświadczenie autora pracy

Świadom odpowiedzialności prawnej oświadczam, że niniejsza praca dyplomowa została napisana przez mnie samodzielnie i nie zawiera treści uzyskanych w sposób niezgodny z obowiązującymi przepisami.

Oświadczam również, że przedstawiona praca nie była wcześniej przedmiotem procedur związanych z uzyskaniem tytułu zawodowego w wyższej uczelni.

Oświadczam ponadto, że niniejsza wersja pracy jest identyczna z załączoną wersją elektroniczną.

Data Podpis autora pracy

Streszczenie

Podejścia do nieprzejrzystości fonologicznej w Teorii Optymalności: Sympatia i DOT

Celem niniejszej pracy jest przedstawienie dwóch podejść do zjawiska nieprzejrzystości fonologicznej w obrębie Teorii Optymalności (Prince i Smolensky 1993, McCarthy i Prince 1995). Rozdział pierwszy omawia podstawy Teorii Optymalności oraz definicję nieprzejrzystości według Kiparskiego (1973), ilustrując ją przykładami z języka polskiego i hebrajskiego, a następnie opisuje trudności, jakie napotyka Teoria Optymalności przy próbach modelowania nieprzejrzystości. Rozdział drugi poświęcony jest Teorii Sympatii (McCarthy 1999, 2003b). Oprócz analizy danych stanowiących typowe przykłady nieprzejrzystości, zawiera on również opis epentezy półsamogłosek oraz głosek krtaniowych w języku czeskim, która z punktu widzenia Teorii Optymalności również może być potraktowana jako nieprzejrzysta. Rozdział trzeci pokazuje, w jaki sposób ze zjawiskiem nieprzejrzystości radzi sobie Derywacyjna Teoria Optymalności (Rubach 1997, 2000a,b, 2003).

Słowa kluczowe

Nieprzejrzystość, Teoria Optymalności, Teoria Sympatii, Derywacyjna Teoria Optymalności

Dziedzina pracy (kody wg programu Socrates-Erasmus)

09.3 Lingwistyka

Table of Contents

Symbols and abbreviations

UR – underlying representation
SR – surface representation
SPE – *Sound pattern of English* by Noam Chomsky and Morris Halle (1968)
OT – Optimality Theory
DOT – Derivational Optimality Theory
// // – underlying representation
/ / – intermediate representation
[ ] – surface (phonetic) representation
C – consonant
V – vowel
– word boundary
X – skeletal slot
μ – mora
N – syllable nucleus
* – ungrammatical form
sg. – singular
pl. – plural
dimin. – diminutive
nom. – nominative
gen. – genitive

Non-IPA symbols used in transcription:

E – a yer (a fleeting vowel which shows zero-*ɛ* alternation)
š – voiceless postalveolar fricative, IPA [ʃ]
¯ – vowel length

Preface

The phenomenon of phonological opacity (Kiparsky 1973), roughly defined as the situation when the conditioning factors motivating the (non)application of a process are obscured by a different process, constituted a key object of study within derivational phonology (Chomsky and Halle 1968 and following work). Recent years saw renewed interest in opaque interactions, as they pose a major challenge to Optimality Theory (Prince and Smolensky 1993, McCarthy and Prince 1995). Standard Optimality Theory, with its two levels of representation only (input and output) fails to account for most opacity effects, as these often require additional levels of representation. Since the advent of Optimality Theory, several ways of dealing with opacity have been proposed. The aim of this thesis is to present two of these: Sympathy Theory (McCarthy 1999, 2003b) and Derivational Optimality Theory (Rubach 1997, 2000a,b, 2003).

The thesis is organised as follows. Chapter 1 offers an overview of Optimality Theory and gives a definition of opacity, subdividing it into two classes. It is then shown that neither of these can be modelled within standard Optimality Theory. Chapter 2 presents the basic principles of Sympathy Theory and shows its application to the two cases of opacity discussed in chapter 1, and then to a more complex set of data, which can also be seen as opaque from the point of view of Optimality Theory. Chapter 3 outlines the rudiments of Derivational Optimality Theory and presents Derivational Optimality Theory analyses of all the three cases of opacity. The last chapter summarises the conclusions.

Chapter 1

Theoretical background

The present chapter outlines the theoretical background necessary for the discussion of opacity within Optimality Theory. It is structured as follows. Section 1.1 introduces the basic tenets and conventions of Optimality Theory: its architecture (section 1.1.1), notation (section 1.1.2) and main constraint families (section 1.1.3). Section 1.2 is concerned with a presentation of opacity: section 1.2.1 provides a definition of opacity as it was used in the derivational framework, while section 1.2.2 demonstrates why opacity cannot be modelled within standard Optimality Theory.

1.1. Basic principles of Optimality Theory

Optimality Theory (Prince and Smolensky 1993, McCarthy and Prince 1995; henceforth OT) is a formal framework in which linguistic patterns are accounted for in terms of constraint interaction. According to the proponents of OT, Universal Grammar contains a set of well-formedness principles, or constraints, which reflect universal linguistic tendencies (Prince and Smolensky 1993: 8). These constraints are typically in conflict: they can impose incompatible demands on a specific linguistic item so that the satisfaction of one constraint entails the violation of others. As a result, it is not possible to satisfy all constraints simultaneously. Nevertheless, OT requires constraint violation to be kept to a minimum (McCarthy and Prince 1993: 1).

Conflicts between competing constraints are resolved through language-specific ranking (McCarthy and Prince 1993: 1): an individual grammar arranges the universal constraints into a strict domination hierarchy, in which each constraint takes complete priority over all lower-ranked constraints. Accordingly, in a situation where only one of two constraints can be satisfied, the higher-ranked one prevails, at the expense of the lower-ranked one. It follows from this and from the tenet of minimal constraint violation that a constraint will be violated if and only if a higher-ranked constraint is at stake.

Given the violability of constraints in OT, grammatical well-formedness cannot be defined as the satisfaction of all constraints. Instead, OT adopts a view of grammaticality in which a grammatical structure is one which best satisfies the ranked body of constraints, in that it incurs the least serious constraint violations compared to a set of competing candidate structures. Such a structure is said to be optimal. Importantly, as McCarthy and Prince (1994: 2) point out, the optimal form need not be perfect; it is merely more successful than any of its competitors, by virtue of violating constraints ranked as low as possible.

The ideas outlined above are formalised and set forth in greater detail in the following sections. Section 1.1.1 gives an overview of the basic OT machinery, section 1.1.2 presents the notation used in OT analyses and section 1.1.3 introduces two major constraint families: markedness and faithfulness.

1.1.1. OT architecture

In Optimality Theory, the lexicon is taken to store the underlying representations of morphemes of a given language. These form the input to a component called GEN (short for 'generator'), which produces a set of possible surface forms. In accordance with Freedom of Analysis, an architectural principle proposed by McCarthy and Prince (1993: 20), GEN is free to posit any amount of structure within the limits of its formal capabilities.[1] Confined only by a number of formal restrictions, GEN supplies an infinite set of linguistically possible forms, some of which are entirely dissimilar to the input.

Under Correspondence Theory, set out in McCarthy and Prince (1995) and now widely endorsed, the role of GEN is to assign a relation of correspondence between the input and the

---

[1] No broad consensus seems to exist regarding the precise nature of GEN, not least because the bulk of work in OT centres on constraints and their interaction. Nevertheless, it is generally accepted that GEN ought not to produce candidates which are impossible in all languages. More specifically, GEN is said to be restricted by the representational primitives available to it (for instance, features or prosodic categories), and by obvious structural universals (such as, "syllables must not dominate feet" or "association lines must not cross"). However, such principles are not a matter of the framework of Optimality Theory as such, but rather of the assumed theory of representation. In any case, encoding any universals in GEN is discouraged in advance (McCarthy and Prince 1993: 20, McCarthy 2002: 8f), with the burden of explanation shifted towards CON and EVAL (discussed below), even though it is in principle possible that certain restrictions form part of GEN. One attempt to find some formal criteria for distinguishing which universals should follow from GEN, and which should follow from EVAL, can be found in Rice (2005).

members of the set of possible output forms. A fairly precise, albeit somewhat simplified, definition is given in McCarthy and Prince (1995: 14[2], emphasis in the original):

(1) "Given two strings S_1 and S_2, **correspondence** is a relation $\mathcal{R}$ from the elements of S_1 to those of S_2. Elements $\alpha \in S_1$ and $\beta \in S_2$ are referred to as **correspondents** of one another when $\alpha \mathcal{R} \beta$".

Importantly, as noted by McCarthy and Prince (1994: 6), correspondence is not a function in the mathematical sense but merely a relation between the elements of S_1 and S_2. This means that elements of both S_1 and S_2 can have multiple correspondents or no correspondents at all (which will manifest itself as coalescence, deletion, epenthesis, etc.).

The possible analyses of the input compete against each other for the optimal status. The choice depends on their satisfaction or violation of a set of constraints, that is, structural requirements on output forms, which constitute the second component of the grammar, CON (for 'constraints'). Each output candidate either satisfies or violates each constraint. CON is claimed to be universal, that is, the same in every language. Particular grammars rank the constraints in a strict domination hierarchy, in which each constraint dominates all lower-ranked constraints (domination is denoted $C_1 >> C_2$, where C_1 dominates C_2). All systematic cross-linguistic grammatical differences are said to stem from differences in constraint ranking. Observe that constraint ranking entails that in every language some constraints will be ranked low to the point of being effectively inactive, while others will be undominated, and thus never violated in the language in question.

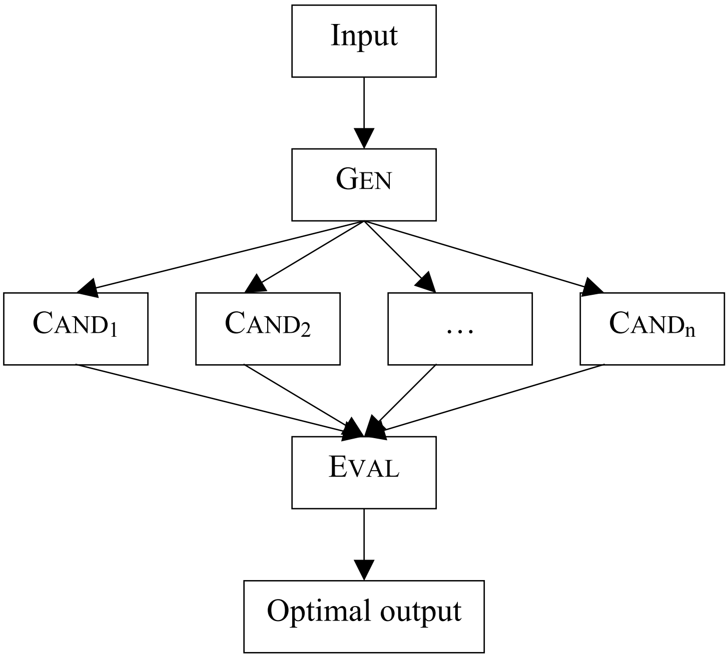
The selection of the optimal output, that is, the actual surface form, is carried out by the central component of an OT grammar, dubbed EVAL (for 'evaluator'). EVAL assesses each candidate and its associated correspondence relations to determine their relative harmony, defined as the degree of conformity with the constraint hierarchy (Prince and Smolensky 1993: 3), and imposes an order on the candidate set. If one candidate's, CAND_1, highest ranking violation (that is, the one associated with the most highly-ranked constraint that it

---

[2] Page numbers refer to the manuscript version available from the Rutgers Optimality Archive. (http://roa.rutgers.edu)

violates) is worse than the highest violation incurred by another candidate, CAND$_2$, then CAND$_1$ is less harmonic than CAND$_2$ and the violation is said to be fatal (when the uppermost violations are equivalent, the same procedure is repeated for the violations immediately below). The optimal output is at the top of the harmonic order. Such a candidate best satisfies the constraint system by virtue of incurring the least serious violations, as compared with any of its competitors. The above discussion is summarised in figure (2).

(2) Structure of optimality-theoretic grammar



One last element of Optimality Theory needs to be mentioned here, namely, the principle of strict parallelism (McCarthy and Prince 1993: 2).[3] In essence, the principle states that all candidates are evaluated in parallel, that is, all constraints apply simultaneously. This means that there are no derivational stages between the input and the output. As will be shown later, this principle has an important bearing on the issue of opacity.

This completes the presentation of the fundamentals of OT. The following section shows how the interaction of constraints in OT analyses is represented visually.

---

[3] Prince and Smolensky (1993: 5f) admit that OT is not, in principle, at odds with serialism. Nevertheless, most OT analyses rely on parallel evaluation.

1.1.2. Notation

A tableau (introduced in Prince and Smolensky 1993: 18ff) is a tool for validating ranking hypotheses and demonstrating the success of an analysis in an easy to read graphical form. Tableaux, such as the sample one shown in (3) below, are used to display all data related to the computation of the optimal surface realisation of a given input.

(3) CONSTRAINT A >> CONSTRAINT B, CONSTRAINT C

Input	CONSTRAINT A	CONSTRAINT B	CONSTRAINT C
☞ a. Candidate 1		*	**
b. Candidate 2	*!		

The input representation is given in the upper left corner and selected output candidates supplied by GEN are listed below it, along the left side of the tableau. Relevant constraints are arrayed in order of decreasing rank across the top row. A solid line between two constraint columns indicates a dominance relation; a dashed line means that there is no evidence for ranking: either one will yield the same results for the input at hand. Thus, in tableau (3), constraint A dominates constraints B and C, which are unranked with respect to each other.

The remaining rows show the success of each candidate with respect to the constraint hierarchy. Constraint violations are signalled by '*' (one for each violation) in the appropriate column. Constraint satisfaction is indicated by a blank cell. In the sample tableau, Candidate 1 incurs one violation of constraint B and two violations of constraint C while Candidate 2 only violates constraint A. However, because constraint A outranks the other constraints, this single violation proves fatal for Candidate 2, which is thus eliminated from consideration. As a result, Candidate 1, which satisfies the top-ranked constraint, emerges as the winner, even though it bears more violation marks on the whole.

In addition to the graphic conventions described above, some further notation is used redundantly for the sake of greater clarity. The '!' sign marks a fatal violation, drawing attention to the point where a candidate loses out to other possible output forms. The '☞' symbol points to the optimal candidate. Finally, shading indicates that the content of the cells is irrelevant to the evaluation process either because a given candidate has been ruled out by a higher-ranked constraint or because it is the only remaining contender.

1.1.3. Major constraint families

Strictly speaking, particular constraints do not pertain to Optimality Theory understood as a formalism for the expression of conflict resolution among linguistic patterns.[4] Nevertheless, since, on the whole, most analyses assume the same substantive theories as regards the constraints, it seems apposite to present here, as part of the theoretical background, the taxonomy of constraints which will be used in this thesis.

Constraints fall into two broad categories: markedness and faithfulness. Constraints of the first type prohibit or require certain structures, thus formalising the notion of markedness (Jakobson 1962, Trubetzkoy 1939/1969, SPE 1968: chapter 9). Generally speaking, a phonological structure is more marked when it is more complex than an alternative structure. Marked members of oppositions tend to be avoided cross-linguistically (with their presence in a language presupposing the presence of the unmarked structure) and to appear later in language acquisition. In OT, markedness constraints, which take the form of well-formedness conditions on output forms, include feature constraints (which ban single features, e.g. *[-high], 'do not be [-high]'), segment inventory constraints (which ban combinations of features, e.g. *[ʔ]) and others (e.g. ONSET, 'syllables must have onsets').

Constraints belonging to the second class militate against disparities between the input and the output. Under Correspondence Theory (discussed above in section 1.1.1), faithfulness constraints are held to evaluate the correspondence relation $\mathcal{R}$ and penalise any unfaithful mapping. To be more specific, feature identity constraints require that correspondent segments are identical in terms of a given feature or node, in effect penalising feature changes. Other faithfulness constraints penalise non-one-to-one correspondence (thus prohibiting deletion, epenthesis, coalescence and fission) and compare the precedence structure of S_1 and S_2 (militating against metathesis).

Precise formulations of the constraints mentioned above will be introduced in the course of the thesis as they become relevant to the discussion of the data.

---

[4] For a discussion of the status of OT as a linguistic theory, see Mohanan (1997).

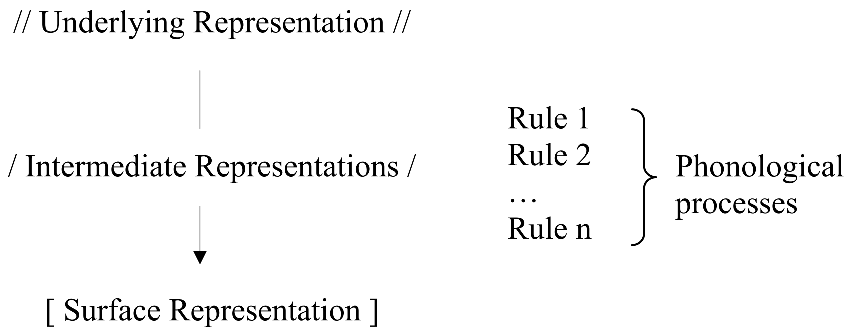
1.2. Opacity

Simply defined, an opaque generalisation is one which is partially obscured at the surface level. Opacity constituted a central object of study in early generative phonology (Chomsky and Halle 1968, hereafter SPE, and their followers), where it was related to the mechanism of rule ordering. Section 1.2.1 briefly outlines the organisation of a rule-based framework of phonology and sketches out the definition of opacity as used within this framework. In section 1.2.2, opacity is redefined within the OT framework. It is also shown that opacity poses a problem for standard OT.

1.2.1. Opacity in a rule-based framework

In derivational models of phonology, linguistic patterns are accounted for by the application of a sequence of extrinsically ordered rewrite rules, which convert underlying representations into surface representations (SPE 1968: 65f). Each rule takes as its input the representation as modified by all preceding rules, makes one structural change (if the input matches the rule's structural description) and passes the result onto the next rule (SPE 1968: 341). The organisation of phonology in the derivational theory is illustrated in (4).[5]

(4)



Opacity effects occur when a rule either creates a structure that would have triggered an earlier rule (so that at the surface level it looks as if a rule did not apply where it should) or when it removes the context of a rule (so that it looks as if a process applied where it should not). The most widely cited definition of opacity was formulated by Kiparsky (1973: 79):

---

[5] Throughout the thesis, double slashes will be used for underlying representations, single slashes for intermediate stages, and square brackets for phonetic representations.

(5) “A phonological rule R of the form *A* → *B* / *C__D* is opaque if there are surface structures with any of the following characteristics:

a. instances of *A* in the environment *C__D*

b. instances of *B* derived by R that occur in environments other than *C__D*”.[6]

The situation described in (5a) may arise in two ways, which can be schematically represented as follows:

(6) a.	// CXD //	b.	// CAX //
A → B / C__D	----	A → B / C__D	----
X → A / __D	CAD	X → D / __#	CAD

In both cases, the structural description of the opaque rule is not met at the stage when the rule applies. In (6a), which represents a so-called chain shift effect, the input to the first rule appears only after the rule has applied. In (6b), it is the context for the application of the first rule that is created *post factum*. In consequence, the surface forms remain unaffected by the first rule even though they match its structural description. This type of interaction will be referred to as **underapplication opacity**.

A typical example of such opacity is taken from Polish[7] (Rubach 1986: 251f, 275ff), where the rule of yer deletion creates the context for nasal assimilation only after the latter rule has applied. In Polish, a nasal consonant acquires the place of articulation of the following stop or affricate. However, forms in which a nasal and a noncontinuant obstruent are separated by a yer[8] in the underlying representation fail to undergo this process, even

---

[6] Kiparsky (1973: 79) adds another clause, “instances of *B* not derived by R that occur in the environment *C__D*”, thus defining neutralisation as opaque. Since this type of interaction does not pose a problem for OT, it will be ignored in subsequent discussion.

[7] To be precise, only the so-called Warsaw dialect, spoken in north-eastern parts of Poland, exhibits opacity with respect to nasal assimilation and yer deletion. In the so-called Cracow dialect (south-western Poland), the processes apply transparently, with yer deletion feeding nasal assimilation. (Rubach 1986: 275f)

[8] In Polish, the “fleeting vowels”, known as yers, surface as [ɛ] when followed by a word boundary or another yer later in the word (with the intervening consonants being transparent) and delete elsewhere. Underlying yers will be transcribed as capital letters and the difference between palatalising and non-palatalising yers will be ignored.

though at the surface level they match its structural description. Consequently, there is a [nk] vs. [ŋk] opposition, as illustrated by the data in (7).

(7) a. Irenka [irɛ**n** + **k** + a] ‘Irene’ (dimin.), compare Irenek [irɛn + ɛk] (gen.pl) vs.
rę̨ka [rɛ**ŋk** + a] ‘hand’
b. ganku [ga**nk** + u] ‘porch’ (gen.sg.) – ganek [ganɛk] (nom.sg.) vs.
bank [ba**ŋk**] ‘bank’
c. trunku [tru**nk** + u] ‘drink’ (gen. sg.) – trunek [trunɛk] (nom.sg.) vs.
punkt [pu**ŋk**t] ‘point’

A derivational explanation for this opposition is rule ordering.[9] Nasal assimilation applies at a stage when the nasal and the obstruent are not adjacent in *Irenka*, as there is a yer between them (as evidenced by the gen.pl form *Irenek*, in which the yer appears phonetically as [ɛ]). The subsequent deletion of the yer obscures why the earlier rule failed to apply. The effect is the underapplication of nasal assimilation. Derivations for *Irenka* ‘Irene’ (dimin.) and *ręka* ‘hand’, (7a), are shown in (8).

(8)

UR	// irɛn + Ek + a //	// rɛnk + a //
Nasal Assimilation	----	/ rɛŋk + a /
Yer Deletion	/ irɛn + k + a /	----
SR	[irɛnka]	[rɛŋka]

As for clause (b) of Kiparsky’s definition of opacity, here, again, there are two possibilities, presented in (9) below.

(9)

a.	// CAD //	b.	// CAD //
A → B / C__D	CBD	A → B / C__D	CBD
D → E / __#	CBE	D → E / B__	CBE

---

[9] Note, however, that Rubach’s (1986: 276f) autosegmental analysis does away with the extrinsic ordering illustrated here.

The common denominator in both cases is that a rule eliminates the context that caused an earlier rule to apply, thus obscuring the reason why it applied. This type of interaction will be referred to as **overapplication opacity**.[10]

An example of overapplication opacity is found in the interaction of epenthesis and glottal stop deletion in segolate nouns in Tiberian Hebrew.[11] Consider the data in (10).

(10) a. Epenthesis: Ø → e / C__C#
// melk // → [melex] ‘king’
b. [ʔ]-deletion: ʔ → Ø / $\overset{\text{coda}}{\_|\_}$
// qaraʔ // → [qara] ‘he called’
c. Interaction: epenthesis → [ʔ]-deletion
// dešʔ // → [deše] ‘tender grass’ (cf. [dāšʔū] ‘they sprouted’)

In Tiberian Hebrew, a rule of epenthesis (10a) inserts a vowel into final clusters, and another process deletes *ʔ* in coda position (10b). In forms where both rules can apply (10c), epenthesis seems to overapply because the epenthetic vowel is not followed by a consonant. A derivational explanation for this sort of opacity is ordering epenthesis before deletion. The deletion rule counterbleeds epenthesis, as shown in (11a) below. The order of the rules is crucial here. If [ʔ]-deletion applied first, it would bleed the insertion rule, and the surface form would be *[deš]. Both rule orders are given in (11). The following section will present an OT analysis of the same set of data to show why opacity cannot be modelled within standard OT.

---

[10] The terms *overapplication* and *underapplication* were introduced by Wilbur (1973) in her work on reduplication, and subsequently extended to other opaque phenomena by Laura Benua (1997). Other terms used in the literature to refer to these two types of opacity are *non-surface-true* and *non-surface-true apparent opacity* (e.g. McCarthy 1999) and *counterfeeding* and *counterbleeding* opacity (where the two terms refer to the potential effect of a certain ordering of rules; see Kiparsky [1968: 196] for a precise definition). Note, however, that while underapplication opacity does indeed correspond to counterfeeding rule orders (i.e. ones where rule A fails to feed [create new input for] rule B; if the order of the rules were reversed, rule A would feed rule B), overapplication opacity does not necessarily follow from a counter-bleeding rule order. The rules in (9a) do stand in a counterbleeding relation: if the second rule applied first, it would bleed (that is, remove the input for) the other rule. The rules in (9b), however, do not stand in a counterbleeding relation, as the opposite order of these rules would not result in bleeding: rule D → E / B__ would simply not apply. As it stands, the rules are in a feeding relation: the first rule provides input for the second one. This observation runs counter to the frequent assertion that feeding and bleeding orders are transparent and counterfeeding and counterbleeding orders are opaque (e.g. Kenstowicz 1994: 98f, Kager 1999: 375; see also Baković 2007).

[11] All Tiberian Hebrew data and generalisations come from McCarthy (1999: 333), after Malone (1993).

(11)	a. counter-bleeding rule order		b. bleeding rule order	
	UR	// dešʔ //	UR	// dešʔ //
	Epenthesis	/ dešeʔ /	[ʔ]-deletion	/ deš /
	[ʔ]-deletion	/ deše /	Epenthesis	----
	SR	[deše]	SR	* [deš]

1.2.2. Opacity in OT

Opaque interactions, such as the ones presented above, in which the conditioning factors that motivate a process are obscured later on in the derivation, turn out to be problematic for Optimality Theory. The difficulty arises from the fact that OT grammars are surface-oriented and, in consequence, favour transparent outcomes. This is not to say that the theory has no means for handling generalisations that do not hold true for surface forms. In a way, output forms are always opaque, since they always violate some constraints. This can be readily expressed in terms of constraint domination. However, when a crucial generalisation takes effect at an intermediate level (in derivational terms), there is no way to model that: markedness constraints can refer to surface structure only and faithfulness constraints can only compare the input with the output.

From an OT perspective, underapplication opacity can be seen as the existence of surface forms which appear to have gratuitous markedness violations: illicit structures appear in the output even though they are prohibited in similar forms. This can be shown using the data from Warsaw Polish ([7] above).[12]

Nasal assimilation is driven by a high-ranked markedness constraint requiring homorganicity between adjacent consonants. It can be assumed here that the relevant constraint is NASASSIM (‘a nasal stop followed by an obstruent must have the same place of articulation’). To allow the change of the place of articulation when the nasal-obstruent cluster is not homorganic in the underlying representation, NASASSIM must outrank IDENT$_{IO}$(Place) (‘correspondent segments are identical in terms of place of articulation’; after

---

[12] Note that it is beyond the purview of this thesis to present a full-fledged OT analysis of the Polish data, especially given that both nasal assimilation and yer deletion require a rather complex OT account. Consequently, only the candidates directly relevant to the question of opacity will be discussed. Furthermore, a number of different constraints motivating the two processes have been proposed. Since the choice between the competing analyses is tangential to the present discussion, provisional constraints will be used instead.

McCarthy and Prince 1995: 16). Tableau (12) illustrates the interaction of these two constraints for the word *ręka* ‘hand’ (7a).

(12) NasAssim >> Ident$_{\text{IO}}$(Place)

// rɛnka //	NasAssim	Ident$_{\text{IO}}$(Place)
a. rɛnka	*!	
☞ b. rɛŋka		*

The faithful rendering of the nasal-obstruent cluster in (12a) runs afoul of NasAssim. As a result, candidate (12a) is eliminated and candidate (12b) emerges as the winner, even though it breaches Ident$_{\text{IO}}$(Place).

The constraint ranking established above predicts incorrect results when applied to *Irenka* (7a). Tableau (13) illustrates the problem.[13]

(13) NasAssim >> Ident$_{\text{IO}}$(Place)

// irɛnEka //	NasAssim	Ident$_{\text{IO}}$(Place)
☜ a. irɛnka	*!	
👎 b. irɛŋka		*

The actual output, candidate (13a), flouts NasAssim and is thus incorrectly eliminated from the competition, as indicated by the ‘☜’ symbol. The symbol ‘👎’ indicates the selection of an ungrammatical candidate, (13b), which fares better on the high-ranked markedness constraint.

Tableaux (12) and (13) show a ranking paradox, by which nasal-obstruent clusters separated by yers at the underlying level seem to require a different constraint ranking (Ident$_{\text{IO}}$[Place] >> NasAssim) than nasal stops and obstruents which are adjacent in the underlying representation (NasAssim >> Ident$_{\text{IO}}$[Place]). While it is clear that the different behaviour of the nasal-obstruent clusters stems from different underlying representations, no normal calculation can render both (12b) and (13a) optimal as there are no constraints that could be added to the computation to derive the correct results. Any faithfulness constraint referring to the yer could not be linked in a straightforward manner to the quality of the

---

[13] The fully faithful candidate [irɛnEka], not included in (13), is discussed in chapter 2.

preceding nasal stop. Markedness constraints, on the other hand, can only refer to the output, so any such constraint would fail to distinguish between *ręka* and *Irenka*.

As shown above, the effect of underapplication opacity is that the attested output is more marked than it should be. Overapplication opacity, on the other hand, results in gratuitous faithfulness violations: surface forms seem to have changed without motivation. The problem can be illustrated using the Tiberian Hebrew data ([10] above).

To begin with the process of epenthesis (10a), a form with an epenthetic vowel runs against the DEP$_{IO}$(V) ('do not insert vowels') constraint from the DEP family of constraints (McCarthy and Prince 1995). The violation is compelled by a higher-ranked *COMPLEX constraint, prohibiting more than one segment in any syllable position node (Prince and Smolensky 1993: 96). The interaction of these two constraints is shown in tableau (14).

(14) *COMPLEX >> DEP$_{IO}$(V)

// melk //	*COMPLEX	DEP$_{IO}$(V)
a. melk	*!	
☞ b. melek[14]		*

Here, the fully faithful candidate (14a) is eliminated from the competition because it fatally violates *COMPLEX. As a result, candidate (14b) becomes the winner, in spite of its violation of DEP$_{IO}$(V).[15]

The process of [ʔ]-deletion (10b) is driven by a high-ranked *CODA([ʔ]) constraint ('glottal stops must not appear in coda position'). It crucially dominates the constraint violated by the attested output, MAX$_{IO}$(C) ('do not delete consonants'; after McCarthy and Prince 1995). The following tableau illustrates the computation of [qara].

---

[14] As shown in (10a), the actual output form is [melex]. Here and below, spirantisation and other processes irrelevant to the issue at hand are suppressed for the sake of simplicity.

[15] The account presented here closely follows that of McCarthy (1999). As a matter of fact, the analysis is not as straightforward as it may seem. The addition of candidates with word internal consonant deletion or vowel insertion at the word edge makes the selection of the optimal candidate far more complicated a task. Since the aim here is not so much to present a full analysis of the data as to illustrate the nature of the problem, such candidates will only be discussed at a later point.

(15) *CODA([ʔ]) >> MAX$_{IO}$(C)

// qaraʔ //	*CODA([ʔ])	MAX$_{IO}$(C)
a. qaraʔ	*!	
☞ b. qara		*

A word with an underlying complex coda containing a glottal stop (10c) surfaces without the glottal stop and with an epenthetic vowel. This cannot be accounted for straightforwardly using standard OT. A look at the tableau for [deše] in (16) reveals the problem.

(16) *CODA([ʔ]), *COMPLEX >> DEP$_{IO}$(V), MAX$_{IO}$(C)

// dešʔ //	*CODA([ʔ])	*COMPLEX	DEP$_{IO}$(V)	MAX$_{IO}$(C)
a. dešʔ	*!	*		
👎 b. deš				*
c. dešeʔ	*!		*	
☜ d. deše			*!	*

Candidates (16a) and (16c) are immediately eliminated from the competition, because they violate the high-ranked markedness constraints, *CODA([ʔ]) and *COMPLEX. The remaining two candidates both obey the markedness constraints. The transparent candidate (16b) wins the competition because it only incurs one faithfulness violation in contrast to candidate (16d), the actual output, which also breaches DEP$_{IO}$(V). In tableau (14) above, DEP$_{IO}$(V) was violated in order to satisfy *COMPLEX. In (16), on the other hand, the deletion of the glottal stop, necessitated by the high-ranked *CODA([ʔ]), has a subsidiary effect, in that it removes one consonant from a cluster, so that vowel insertion is no longer necessary and *COMPLEX is satisfied. As a result, the violation of DEP$_{IO}$(V) by candidate (16d) is gratuitous. Putting it differently, candidate (16d) is harmonically bounded by candidate (16b), which means that it cannot become optimal under any ranking of the constraints.

There have been numerous responses to the problem of opacity in phonology. Two ways to obtain the opaque forms in Polish and Tiberian Hebrew will be presented in the following chapters.

Chapter 2

Sympathy Theory

Being one of the major challenges to Optimality Theory, phonological opacity has received considerable research attention, bringing about a number of competing analytical tools to account for opacity effects in OT, notably, Output-Output Correspondence (Benua 1995, 1997), Local Constraint Conjunction (Smolensky 1995, Lubowicz 1998), Targeted Constraints (Wilson 2001) and Comparative Markedness (McCarthy 2003a). Sympathy Theory, as developed by McCarthy (1999, 2003b), is another such proposal, intended specifically as a "general model of opaque interactions within OT" (McCarthy 2003b: 25). The present chapter examines the extent to which this purpose has been accomplished. In particular, section 2.1 presents the rudiments of Sympathy Theory, which are then illustrated in section 2.2 on simple cases of phonological opacity. In section 2.3, Sympathy Theory is tested against a more complex set data. Finally, section 2.4 summarises the discussion.

2.1. Basic tenets of Sympathy Theory

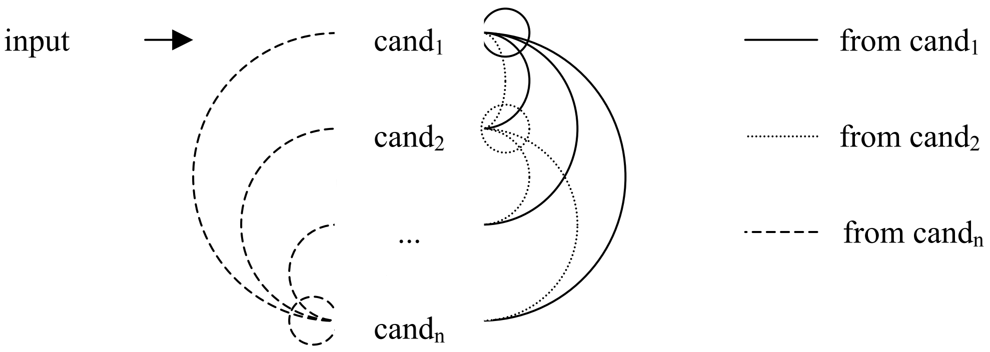
As explained in chapter 1, the problem with opaque interactions within OT is twofold. In the case of underapplication opacity, the opaque form incurs a gratuitous markedness violation. In the case of overapplication opacity, the opaque candidate is harmonically bounded by the competing transparent candidate, that is, it has a proper superset of the transparent candidate's violation marks. More precisely, it has all of the transparent candidate's violations plus a faithfulness violation. What connects these two types of interaction is the superfluous violation, avoided by the transparent candidate, which dooms the opaque candidate to failure.

Sympathy Theory suggests a way to rectify the situation by adding a new force to the system. It comes in the form of a **sympathy constraint**, which can dominate the troublesome constraints, thus providing motivation for the seemingly excessive violations. To do that, Sympathy Theory grants a special status to a selected failed co-candidate, dubbed the **sympathetic candidate**. This form is capable of affecting the shape of the outcome through a **sympathy relation**, "one of several distinct faithfulness relations provided by correspondence theory" (McCarthy 1999: 350). The following sections will take a closer look at the standard implementation of the sympathy idea, focusing, in turn, on each of the three aspects of the model highlighted in bold type above.

2.1.1. The sympathy relation

The innovation of Sympathy Theory is that it extends the correspondence relation to pairs of output candidates. McCarthy's idea is that Gen supplies correspondence relations not only from the input to all output candidates, but also from each candidate to itself and to every other form in the candidate set, as illustrated in figure (1).

(1) Inter-candidate correspondence (after McCarthy 1999: 350)



McCarthy (1999: 336) points out that in this respect Sympathy Theory is analogous to other current OT sub-theories which posit that an output form participates in a number of parallel correspondence relations: not only between the input and the output (McCarthy and Prince 1995), but also between morphologically related forms (Output-Output Faithfulness, Benua 1997) and between the base and the reduplicant (Reduplicative Identity, McCarthy and Prince 1995, 1999). The candidate-to-candidate relation is then merely a logical extension of the theory.

2.1.2. Selection of the sympathetic candidate

The sympathetic candidate, which exercises an indirect influence on the optimal output, is defined by McCarthy as "the most harmonic member of the set of candidates that obey a designated input-output (IO) faithfulness constraint, called the SELECTOR" (1999: 336, emphasis in the original). Given the arguably infinite number of output candidates, a systematic method of selecting the candidate is required. McCarthy (1999: 339) provides the following principles intended to aid the selection.

(2) a. *Harmonic evaluation*

The sympathetic candidate is the most harmonic member of the subset of candidates available under (2b).

b. *Confinement to* $C_{<+F>}$

Selection of the sympathetic candidate $\aleph_F$ is confined to $C_{<+F>}$, the subset of candidates that obey the IO faithfulness constraint F.

c. *Invisibility of sympathy constraints*

Selection of sympathetic candidates is done without reference to sympathy constraints.

In less formal terms, McCarthy's principles state that the sympathetic base should be selected through harmonic evaluation (principle 2a) from the set of candidate outputs that satisfy a certain faithfulness constraint, called the selector (principle 2b). Within the set, the sympathetic candidate is the one which best satisfies the independently motivated constraint hierarchy of a given language. McCarthy (1999: 336) remarks that in this way the selection employs the same mechanism that is used to determine the actual output, that is, Prince and Smolensky's "Harmonic Ordering of Forms" (Prince and Smolensky 1993: 73ff).

The remaining issue to be addressed is how to appoint the correct selector itself. According to McCarthy, "[t]he choice of the selector is determined on a language-particular basis" (1999: 339). However, the fact that there is no single principle to discover the selector does not mean that the selection procedure is reduced to pure guesswork. Firstly, the range of possible selectors is restricted by principle (2b) to the family of IO faithfulness constraints.[16] Secondly, McCarthy proposes some heuristics to determine the selector, arguing that it should pick out the candidate which contains the information obscured in the opaque form (McCarthy 1999: 341). Consequently, the sympathetic candidate is minimally more faithful to the input than the actual output, which, clearly enough, has to violate the selector. This, as McCarthy (1999: 337) observes, makes the sympathetic form analogous to the intermediate stage in a derivation, which also resembles the underlying representation more closely than the surface representation does.[17]

Finally, McCarthy points out that the range of possible selectors is limited even further by the fact that the selector has to be a dominated faithfulness constraint, as an undominated constraint would select the candidate that would be optimal in any case (McCarthy 1999: 341). Crucially, then, the transparent candidate competing with the opaque (and optimal) candidate must violate the selector. Otherwise, it will be selected as the sympathetic candidate and the effect of sympathy will be vacuous, which is, incidentally, the desired result when a process applies transparently.

2.1.3. Sympathy as inter-candidate faithfulness

When the sympathetic form is selected, we need a way to relate it to the actual output. The idea is that the optimal output is required to resemble the sympathetic form. Needless to say, the opaque candidate must resemble the sympathetic candidate more closely than the transparent candidate does. The resemblance is compelled by a new class of ranked, violable constraints, called **sympathy constraints**, sensitive to the aforementioned correspondence relations (by hypothesis provided by GEN), holding within the candidate set. McCarthy

---

[16] McCarthy's principle (2b) is repudiated by some researchers (Itô and Mester 1997, de Lacy 1998, Bermúdez-Otero 1999), who argue that the choice of the selector should not be confined to faithfulness constraints on account of symmetry considerations, and who present analyses in which the role of the selector is fulfilled by a markedness constraint. McCarthy (1999: 347) questions the empirical arguments, claiming that an alternative analysis is possible. He adds that, on the theoretical side, there is a negative trade-off between symmetry and restrictiveness, as expanding the set of possible selectors results in too powerful a theory. Moreover, assuming markedness constraints as potential selectors seems to be the source of the 'chaotic' behaviour described by Idsardi (1997), where adding sympathy to OT interferes with the calculation of unrelated transparent forms.

[17] McCarthy (1999: 371) notes that although the sympathetic base is typically identical to the intermediate stage in a serial derivation, this need not be the case. In cases of multiple opacity (where more than two processes interact opaquely and, consequently, more than one sympathetic base may be required), the sympathetic candidates do not always coincide with intermediate levels of serial derivation.

(1999: 351) proposes that “there are separate and therefore separately rankable faithfulness constraints on each correspondence relation”.

Under the more familiar implementation, sympathy constraints are construed as inter-candidate faithfulness constraints on the relation of the sympathetic base to the whole candidate set.[18] These can make reference to any of the properties of the sympathetic base and enforce its preservation in the output. Thus, as mentioned in section 2.1.1 above, Sympathy Theory extends faithfulness to inter-candidate relations. Phonological opacity will arise from the interaction of the standard markedness or IO faithfulness constraints and the new sympathy constraints. McCarthy (1999: 351) assumes that “the Gen-supplied correspondence relations (…) and the sympathetic faithfulness constraints on those relations are universal, though not universally active”, as sympathy effects will only be visible if a sympathy constraint is ranked high enough to outrank the relevant markedness or IO faithfulness constraints and, in consequence, to compel the seemingly gratuitous violations by the actual output. An illustration of the model presented here vis-à-vis various types of opacity will be provided in the following section.

2.2. Illustrative analyses

As noted in chapter 1, opacity involves either a gratuitous markedness or faithfulness violation (underapplication and overapplication opacity, respectively). Sympathy Theory handles both these cases by introducing a new constraint that can force the violation of the troublesome constraint.

Recall that in the case of Polish, discussed in chapter 1, which provided an illustration of underapplication opacity, it was impossible to block nasal place assimilation over underlying yers under the ranking established for nasal-obstruent clusters which were adjacent in the underlying representation. Reversing the ranking was not a viable option, since it would block nasal assimilation entirely. The incorrect evaluation is repeated below in tableau (3) for the reader’s convenience.

---

[18] McCarthy (2003b: 46ff) later departs from this model arguing that the resulting theory is too powerful, as it allows for unattested patterns of opacity to be described. Instead, he proposes a more restrictive approach, which does not look at the exact difference between the optimal output and the sympathetic candidate but places them in the relation of cumulativity. Due to space limitations, this thesis will only cover the former, more prevalent model.

(3) NasAssim >> Ident$_{IO}$(Place)

// irɛnEka //	NasAssim	Ident$_{IO}$(Place)
☜ a. irɛnka	*!	
👎 b. irɛŋka		*

As explained in chapter 1, the problem here is that standard OT has no way to link the non-application of nasal assimilation to the presence of a yer at the underlying level. As a result, the optimal output incurs a seemingly superfluous markedness violation.

Sympathy Theory provides a tool to rectify the situation by assigning special status to a certain failed candidate and enforcing faithfulness to this candidate. An appropriate sympathy constraint can override the markedness requirement, thus compelling the opaque candidate's supererogatory violation. Tableau (4) illustrates the sympathy theoretic account of underapplication opacity in Polish.

(4) **Parse(Seg)**, **❀Ident(Place)** >> NasAssim >> Ident$_{IO}$(Place), **★Max$_{IO}$(V)**

// irɛnEka //	Parse (Seg)	❀Ident (Place)	NasAssim	Ident$_{IO}$ (Place)	★Max$_{IO}$ (V)
☞ a. irɛnka			*		*
b. irɛŋka		*!		*	*
❀ c. irɛnEka	*!				☑

In this case, the sympathetic candidate (indicated with the '❀' symbol) has to be [irɛnEka] (4c), that is, a candidate in which the yer is not deleted and hence the non-application of nasal assimilation is motivated transparently. To ensure that this candidate is not selected as optimal, Parse(Seg), a well-formedness constraint which prohibits unparsed melodic segments, needs to be ranked higher than NasAssim. Candidate (4c), unlike candidates (4a,b), satisfies Max$_{IO}$(V),[19] which can therefore be assumed to be the selector (marked with the '★' symbol; obedience to the selector is indicated by '☑'). Candidate (4b) fatally violates the sympathy constraint, ❀Ident(Place), which requires that the place of articulation of output

---

[19] It is assumed here, after Yearley (1995), who offers an OT analysis of yers in Russian, that yers are represented as moraless vowels at the underlying level. Under this analysis, deleting a yer results in the violation of Max$_{IO}$(V) ('do not delete a vowel'), while yer vocalisation, used as a strategy to improve syllable structure, violates Dep$_{IO}$(μ) ('do not insert a mora'). Since the process of yer vocalisation is irrelevant for the problem of opacity, the candidate with an inserted mora, [irɛnɛka], will not be considered.

segments be identical to the place of articulation of correspondent segments in the sympathetic base. As a result, the opaque candidate (4a) is correctly chosen as optimal.

It remains to be shown that the addition of sympathy yields correct results in the case of transparent forms. A look at the computation of *ręka* in tableau (5) reveals that this is indeed the case.

(5) PARSE(Seg), ❀IDENT(Place) >> NASASSIM >> IDENT$_{IO}$(Place), ★MAX$_{IO}$(V)

// rɛnka //	PARSE (Seg)	❀IDENT (Place)	NASASSIM	IDENT$_{IO}$ (Place)	★MAX$_{IO}$ (V)
a. rɛnka		*!	*		☑
❀☞b. rɛŋka				*	☑

Since all vowels present in the input are present in both output candidates, neither of them violates the selector. Between them, candidate (5b) is more harmonic and is, consequently, selected as the sympathetic base. Since it is trivially identical to itself, it passes on the sympathy constraint. Candidate (5a), on the other hand, differs from the sympathetic base in having an alveolar nasal instead of a velar nasal, and incurs a fatal violation of ❀IDENT(Place). All in all, Sympathy Theory proves to offer a satisfactory account of the Polish data by providing a way to differentiate between *ręka* and *Irenka*.

Turning now to overapplication, let us consider the way Sympathy Theory handles the data from Tiberian Hebrew. The approach is analogous to underapplication opacity, in that here, too, a sympathy constraint must outrank the problematic constraint. The only difference is that the constraint in question belongs to the faithfulness, and not markedness, family.

Recall that Tiberian Hebrew has a process of glottal stop deletion in the coda and vowel epenthesis into final clusters. These two processes interact opaquely giving rise to words with an epenthetic vowel which is not followed by a consonant. Tableau (6) repeats the evaluation for [deše], discussed in chapter 1.

(6) *CODA([ʔ]), *COMPLEX >> DEP$_{IO}$(V), MAX$_{IO}$(C)

// dešʔ //	*CODA([ʔ])	*COMPLEX	DEP$_{IO}$(V)	MAX$_{IO}$(C)
a. dešʔ	*!	*		
👎 b. deš				*
c. dešeʔ	*!		*	
☜ d. deše			*!	*

Sympathy Theory proposes adding a new constraint, which will outrank the problematic faithfulness constraint (here, DEP$_{IO}$[V]), thus justifying the seemingly gratuitous violation. The sympathy constraint requires resemblance to a certain failed candidate, which contains the information obscured in the actual output. Here, the sympathetic candidate has to be [dešeʔ], in which vowel epenthesis applies transparently. [dešeʔ] is the most harmonic member of the set of candidates that obey MAX$_{IO}$(C) and can therefore act as the sympathetic base. The selection of the sympathetic candidate and its influence, via a sympathy constraint, on the selection of the optimal output is shown in tableau (7).

(7) ❀**MAX(V)**, *CODA([ʔ]), *COMPLEX >> DEP$_{IO}$(V), ★**MAX$_{IO}$(C)**

// dešʔ //	❀MAX (V)	*CODA ([ʔ])	*COMPLEX	DEP$_{IO}$ (V)	★MAX$_{IO}$ (C)
a. dešʔ	*!	*	*		☑
b. deš	*!				*
❀ c. dešeʔ		*!		*	☑
☞ d. deše				*	*

Among the two candidates that satisfy the selector, candidate (7c) is the most harmonic and, consequently, becomes the sympathetic base. Now the troublesome transparent candidate (7b) is eliminated because it does not reproduce the epenthetic vowel of [dešeʔ], thus violating the sympathy constraint ❀MAX(V) ('every vowel in the sympathy candidate should have a correspondent in the output'). As a result, candidate (7d), which satisfies ❀MAX(V) emerges as the winner.

So much then for McCarthy's account of the Tiberian Hebrew data. It should be noted, however, that there are several problems with his analysis. Firstly, the inclusion of sympathy has an adverse effect on the data with glottal stop deletion only ([10b] in chapter 1). This is shown in tableau (8).

(8) ❀Max(V), *Coda([ʔ]), *Complex >> Dep$_{IO}$(V), ★Max$_{IO}$(C)

// qaraʔ //	❀Max (V)	*Coda ([ʔ])	*Complex	Dep$_{IO}$ (V)	★Max$_{IO}$ (C)
a. qaraʔ	*!	*			☑
👎❀ b. qaraʔa				*	☑
☜ c. qara	*!				*

With Max$_{IO}$(C) being the selector, candidate (8b), [qaraʔa] becomes the sympathetic base. Clearly, the attested output (8c) is less faithful to the sympathetic base than the sympathetic base itself and is, consequently, eliminated owing to its violation of the sympathy constraint.

One way to solve this problem would be to rank ❀Max(V) lower than Dep$_{IO}$(V). This, however, would mean going back to square one with the analysis of [deše], as the original purpose of introducing this constraint was to validate [deše]'s violation of Dep$_{IO}$(V).

Another possibility is adding the Align-R constraint ('the right edge of the stem coincides with the right edge of the syllable'; McCarthy and Prince 1993). It has to be ranked above *[ʔ]-Coda to eliminate the possibility of choosing [qaraʔa] as the sympathetic candidate. Such ranking allows [qaraʔ] to become the sympathetic base and leads to the selection of the attested output, [qara], as tableau (9) demonstrates. However, the ranking Align-R >> *Coda([ʔ]) would make [dešeʔ], rather than [deše], the winner in (7). Therefore, a different solution needs to be found.

(9) ❀Max(V), Align-R >> *Coda([ʔ]), *Complex >> Dep$_{IO}$(V), ★Max$_{IO}$(C)

// qaraʔ //	❀Max (V)	Align-R	*Coda ([ʔ])	*Complex	Dep$_{IO}$ (V)	★Max$_{IO}$ (C)
❀ a. qaraʔ			*!			☑
b. qaraʔa		*!			*	☑
☞ c. qara						*

The answer is to change the selector to Anchor-R (‘any element at the right edge of the input has a correspondent at the output; no deletion/insertion at the right edge’; McCarthy and Prince 1995: 123; note that its equivalent, Align-R, could not be used here, because it is not a faithfulness constraint). Using Anchor-R will not have a negative effect on the computation of [deše] (although Max$_{\text{IO}}$[C] will need to be ranked above Dep$_{\text{IO}}$[V] to ensure that [deʔ], which violates Max$_{\text{IO}}$[C], but satisfies Anchor-R, is not selected as the sympathetic base). This will also result in a correct sympathetic base in the case of [qara]. The suggested analysis is shown in tableau (10).

(10) ❀Max(V), **❀Dep(V)**, *Coda([ʔ]), *Complex >> Max$_{\text{IO}}$(C) >> ★Anchor-R, Dep$_{\text{IO}}$(V)

// qaraʔ //	❀Max (V)	❀Dep (V)	*Coda ([ʔ])	*Comp	Max$_{\text{IO}}$ (C)	★Anchor -R	Dep$_{\text{IO}}$ (V)
❀ a. qaraʔ			*!			☑	
b. qaraʔa		*!				*	*
☞ c. qara					*	*	

Anchor-R selects [qaraʔ], candidate (10a), as the sympathetic base. Since the previously introduced sympathy constraint, ❀Max(V) is now mute on the key candidates, it becomes pertinent to ensure that candidate (10b), with a paragoge vowel, does not win. Because the constraints it violates, Anchor-R and Dep$_{\text{IO}}$(V), are ranked below Max$_{\text{IO}}$(C), violated by the actual output, it is necessary to add another sympathy constraint, ❀Dep(V), that militates against vowel “insertion” with respect to the sympathetic base (‘every vowel in the output candidate should have a correspondent in the sympathetic base’). With two sympathy constraints, the evaluation is successful.

Yet another, albeit highly abstract, complication occurs in the computation of [deše] if one more candidate is brought in. Consider the following tableau.

(11) ❀MAX(V), ❀DEP(V), *CODA([ʔ]), *COMPLEX >> MAX$_{IO}$(C) >> ★ANCHOR-R, DEP$_{IO}$(V)

// dešʔ //	❀MAX (V)	❀DEP (V)	*CODA ([ʔ])	*COMP	MAX$_{IO}$ (C)	★ANCHOR -R	DEP$_{IO}$ (V)
a. dešʔ	*!		*	*		☑	
b. deš	*!				*	*	
❀ c. deše$_1$ʔ			*!			☑	*
👎 d. deše$_1$					*!	*	*
e. deʔ	*!		*		*	☑	
f. dešʔe	*!	*				*	*
☜ g. dešʔe$_1$						*	*

The problematic candidate, which fares better than the actual output on the constraints discussed so far is [dešʔe], candidate (11g). A word of explanation is in order here. The last vowel in the seemingly identical candidate (11f) does not correspond to the epenthetic vowel in the sympathetic candidate. Consequently, candidate (11f) violates both ❀DEP(V), because it contains a vowel that does not have a correspondent in the sympathetic base, and ❀MAX(V), because there exists a vowel in the sympathy candidate that does not have a correspondent in the output. However, Sympathy Theory offers the possibility of introducing candidates such as (11g), in which the final vowel does indeed correspond to the epenthetic vowel in [dešeʔ], as indicated by coindexing. Such a candidate satisfies both ❀DEP(V) and ❀MAX(V). Because of its MAX$_{IO}$(C) satisfaction, this candidate is more faithful to the input than the opaque candidate, (11d), and, as such, wins the competition. In order to eliminate it, it is necessary to introduce yet another sympathy constraint, ❀LINEARITY ('the output candidate is consistent with the precedence structure of the sympathetic base, and vice versa'; after McCarthy and Prince 1995: 123). After the introduction of the third sympathy constraint, the account of the Tiberian Hebrew data is finally complete. The following section will attempt to see whether Sympathy Theory can handle a yet more complex set of data.

2.3. Redefining opacity – glide and glottal stop insertion in Czech

The processes of glide and glottal stop insertion in Czech, analysed in this section, cannot be classified as opacity as defined by Kiparsky (1973). However, from the point of view of OT they can be seen as opaque in that in order to satisfy one high-ranked constraint two different strategies are used where only one would seemingly suffice.

2.3.1. The data and basic generalisations

Czech conforms to the universal tendency of languages to avoid onsetless syllables, exhibiting different strategies to repair such syllables depending on the position of the syllable in the word (initial or medial) and the quality of the vowels involved. Consider the following data (unless otherwise noted, the data and generalisations in this section are drawn from Rubach [2000b: 297ff]; see also the references cited therein).

(12) a. *iV* → *i.jV* d**ia**lekt [i.ja] 'dialect', patr**io**t [i.jo]'patriot'
b. *Vi* → *V.ji* kok**ai**n [a.ji] 'cocaine', hid**ui**sta [u.ji] 'Hinduist'
c. *uV* → *u.V* sil**ue**ta [u.e] 'silhouette', sit**uo**vat [u.o] 'place'
d. *Vu* → *V.u* mus**eu**m [e.u] 'museum', lyc**eu**m [e.u] 'high school'
e. *VV* → *V.V* p**oe**ta [o.e] 'poet', n**ea**ndertálec [e.a] 'Neanderthal man'
f. *#V* → *#ʔV* **A**merika [ʔa], **u**lice [ʔu] 'street'

Czech resolves a hiatus with a high front vowel *i* by inserting a front glide [j], both when the high vowel is the first (12a) and the second (12b) vowel in the sequence. Elsewhere, that is, when two non-high vowels are involved (12e), or when the high vowel is also back (12c,d), the hiatus is tolerated. Word-initially, a glottal stop is inserted before any vowel (12f).

The data provided in (13) below reveal yet another hiatus resolution strategy in Czech.[20]

(13) a. *Vi* → *Vj* fa**jn** [aj] 'fine', v**ej**.ce [ej] 'eggs'
b. *Vu* → *Vw* k**ou**.le [ow] 'ball', **au**.tor [aw] 'author'

---

[20] The fact that the words in (13a) are pronounced with a glide is indicated by their spelling. The syllabification of all the words in (13) has also been confirmed by my native speaker consultants, Kristina Krchnava and Věra Petrásová. I assume that the *Vu* sequences in (13b) are not diphthongs after Rubach (2000b: 278), who claims that in Czech, as well as Polish and Bulgarian, "Rosenthall's (1994) constraint NO-DIPTHONG (NO-DIPH) is undominated".

Note that there is a contradiction between the data in (12) and (13). More precisely, the contexts in (12b) and (13a) and in (12d) and (13b) are identical: a *Vi* and a *Vu* string, respectively. This inconsistency will be discussed later.

2.3.2. Standard OT analysis

According to Rubach (2000b: 271), the driving force behind all the processes illustrated above is the Onset (Ons) constraint, which penalises onsetless syllables (‘syllables must have onsets’; Prince and Smolensky 1993). Typical strategies that could be used to satisfy Onset include vowel deletion (so that *dialekt* [12a] would become *[da.lekt] or *[di.lekt]), diphthongisation (*[dia.lekt]), insertion (*[di.ʔa.lekt] or [di.ja.lekt]) and vowel gliding (*[dja.lekt]).

The first two processes can be discarded straightaway, as they are never used in Czech. This means that the constraints controlling these changes, namely, Max$_{IO}$(Seg) (‘do not delete segments’; McCarthy and Prince 1995) and No-Diphthong (‘diphthongs are prohibited’; Rubach 2000b: 278, after Rosenthall 1994), respectively, must be ranked above Onset, as illustrated in tableau (14).

(14) Max$_{IO}$(Seg), No-Diph >> Onset

// poeta //	Max$_{IO}$(Seg)	No-Diph	Onset
☞ a. po.e.ta			*
b. po.ta	*!		
c. poe.ta		*!	

Candidate (14a) is optimal, even though it violates Onset, because it fares better than the other candidates, which violate higher-ranking Max$_{IO}$(Seg) and No-Diph.

The insertion of [j] in *iV* and *iV* strings in (12a) can be accounted for by ranking Onset above Dep$_{IO}$(Seg) (‘do not insert segments’; McCarthy and Prince 1995).

(15) Onset >> Dep$_{IO}$(Seg)

// dialekt //	Onset	Dep$_{IO}$(Seg)
☞ a. di.ja.lekt		*
b. di.a.lekt	*!	

Although candidate (15a) violates Dep$_{IO}$(Seg), it still emerges as the winner because, unlike the fully faithful candidate (15b), it satisfies the higher-ranked Onset constraint.

It remains to be unravelled why [j], and not some other sound, is inserted to break *Vi* and *iV* sequences and why only sequences with [i] exhibit this kind of insertion (note that although Onset ranks higher than Dep$_{IO}$(Seg), *poeta* surfaces as [po.e.ta] (14a), and not as *[po.je.ta]). This is easily explained under the assumption, suggested by Rubach (2000b: 285), that [j] comes from spreading. Two remarks have to be made here. First, the autosegmental, X-skeletal (e.g. Steriade 1982, Levin 1985), model of representation is adopted in this analysis. In this model, the difference between [j] and [i] (and [w] and [u]) lies at the syllable tier. The glide and the vowel are held to be identical in terms of features but the former is part of the syllable margin, while the latter forms a syllable nucleus. Second, the Halle-Sagey (Sagey 1986, Halle 1992) model of feature geometry will be used. The important aspect of the model is that it views a segment as a collection of hierarchically organised features dominated by a Root node. Under these two theories, glide insertion can be seen as adding an empty X-slot to fill the onset position and then spreading the melody. Inserting any other sound, on the other hand, means that a new Root node has to be added. The difference between spreading and insertion is illustrated in (16).

(16) a. /ia/ → [i.ja]

```
x    x          x  x  x
|    |          | /   |
Rt   Rt         Rt    Rt
|    |     →    |     |
i    a          i     a
```

b. /ia/ → [i.ʔa]

```
x    x          x  x  x
|    |          |  |  |
Rt   Rt         RtRtRt
|    |     →    |  |  |
i    a          i  ʔ  a
```

Viewed in this way, [i.ja] is unfaithful to the input only by having an extra X-slot. By contrast, inserting a new Root node in [i.ʔa] entails increased markedness, as an additional violation of the *RT constraint (‘do not be a Root node’; Rubach 2000b: 285) is incurred. This is not to say, however, that such candidates are harmonically bounded by a candidate with glide insertion, as spreading incurs a violation of NO-MULTIPLE-LINK (*MULT-LINK; Rubach 2000b: 288), a constraint which militates against candidates with one melodic segment linked to more than one X-slot. The constraint must be ranked lower than *RT and ONSET for spreading to be preferred over insertion and no repair processes at all. The ranking of *MULT-LINK with respect to DEP$_{IO}$(Seg) cannot be determined on the basis of the data under consideration; any ranking of these constraints will give the same results. The above discussion is summarised in tableau (17) for the word *dialekt*. For the sake of simplicity, only the relevant part of the structure is shown.

(17) ***RT**, ONSET >> DEP$_{IO}$(Seg), ***MULT-LINK**

// dia //	*RT	ONSET	DEP$_{IO}$(Seg)	*MULT-LINK
☞ a. xxxx | | / | di a	***		*	*
b. xxxx | | | | diʔa	****!		*	
c. xxx | | | dia	***	*!		

The competing candidates are [di.ja.lekt], (17a), with the glide being an effect of spreading of the melody from the preceding vowel, [di.ʔa.lekt], (17b), with a glottal stop inserted to repair the second syllable,[21] and fully faithful [di.a.lekt], (17c). Candidate (17c) is eliminated on account of violating high-ranked ONSET. Both (17b) and (17a) violate DEP$_{IO}$(Seg) in order to comply with higher-ranked ONSET. Candidate (17b) loses out to (17a) because it incurs one

---

[21] While different consonants could also be inserted, [ʔ] is a particularly suitable candidate because it incurs a minimal violation of featural markedness constraints by virtue of having no supralaryngeal node.

more violation of *Rt, which is crucially ranked above *Mult-Link, violated by candidate (17a).

Although tableau (17) does not indicate that, *Rt must also outrank Onset to ensure the correct candidate is selected in words with no insertion. Otherwise, it would be less costly to insert a segment than to have an onsetless syllable and hence *[po.ʔe.ta] would be incorrectly predicted to be the optimal candidate, as shown in tableau (18) below (where ‘👎’ indicates the incorrectly selected winning candidate, and ‘☜’ the actual output form).

(18) Incorrect ranking: **Onset >> *Rt** >> Dep$_{IO}$(Seg), *Mult-Link

// poeta //	Onset	*Rt	Dep$_{IO}$(Seg)	*Mult-Link
☜ a. po.e.ta	*!	5[22]		
👎 b. po.ʔe.ta		6	*	

To complete the account of *VV* strings, in (12e), it should be mentioned why they are not broken by spreading, which is, as a matter of fact, what the current ranking predicts. The point is that spreading in such words would result in a non-high vowel in a syllable margin. According to Rubach (2000b: 276, after Prince and Smolensky 1993), this is not permitted in Slavic languages, which only allow high vowels to function as glides. This restriction is formalised in terms of the *M(V[-high]) constraint (‘nonhigh vowels cannot be margins’; Rubach 2000b: 276), undominated in Czech. Changing the quality of the vowel, and hence the resulting glide, is not an option, either, owing to highly-ranked feature identity constraints.

The structure of words with *uV* and *Vu* strings, such as the ones listed in (12c,d) is analogous to words with *iV* and *Vi* strings, analysed above: any vowel followed, or preceded, by a high vowel. Consequently, there is nothing in the ranking established so far to prohibit the insertion of the glide [w], as shown in tableau (19) for the word *silueta*.

---

[22] Owing to the large number of violations incurred by *Rt, integers, rather than asterisks, will be used to indicate these violations, for purposes of clarity.

(19) *Rt >> Onset >> Dep$_{\text{IO}}$(Seg), *Mult-Link

// silueta //	*Rt	Onset	Dep$_{\text{IO}}$(Seg)	*Mult-Link
☜ a. si.lu.e.ta	7	*!		
☟ b. si.lu.we.ta	7		*	*
c. si.lu.ʔe.ta	8!		*	

Candidate (19c) loses out to candidates (19a,b) on account on having an additional Root node, and thus incurring an unnecessary violation of *Rt. Candidate (19a), which is the attested output, is eliminated because it fares worse on Onset than the winning candidate (19c), in which a glide [w] is spawned by the following vowel. To prevent candidate (19b) from surfacing as the optimal output, it is necessary to add another constraint to the computation, *Onset([u]) ('[u] cannot be in the onset'; Rubach 2000b: 277), and rank it above Onset. As a result, the word *silueta* is evaluated as demonstrated in tableau (20).

(20) ***Onset([u])** >> *Rt >> Onset >> Dep$_{\text{IO}}$(Seg), *Mult-Link

// silueta //	*Onset([u])	*Rt	Onset	Dep$_{\text{IO}}$(Seg)	*Mult-Link
☞ a. si.lu.e.ta		7	*		
b. si.lu.we.ta	*!	7		*	*
c. si.lu.ʔe.ta		8!		*	

The final strategy of Onset satisfaction to be considered is gliding. Needless to say, the only vowels susceptible to gliding will be *i* and *u*, since, as noted above, non-high vowels cannot be syllable margins in Czech, nor can they be raised. Vowels preceded (*iV*, *uV*) and followed (*Vi*, *Vu*) by high vowels will be discussed separately, as they exhibit slightly different behaviour with respect to gliding.

To begin with the former, the data in (12a,c) show that prevocalic high vowels are not syllabified into the onset, so *dialekt* cannot be realised as *[dja.lekt]. Similarly, *silueta* will not be realised as *[sil.we.ta] but this might be due to the highly-ranked Onset([u]). An analogous constraint for *i* would not work, as words with initial *iV* sequences, such as the ones in (21) below, do have a [j] sound in the onset. Moreover, prohibiting [j] from appearing in onsets would also preclude glide insertion in *CiV* sequences.

(21) **ja** [ja] ‘I’, **ji**skra [ji] ‘spark’

Consequently, it would be inappropriate to try to rule out *[dja.lekt] by reference to a constraint that penalises [j] as such, for example VOWEL-NUCLEUS (‘every [-cons] segment must be linked to N’; Rubach 2000b: 274). This constraint must be ranked below ONSET to allow for gliding in *#iV* strings, as shown in tableau (22).

(22) ONSET >> **V-NUC**

// ia //	ONSET	V-NUC
☞ a. ja		*
b. i.a	*!*	
b. i.ja	*!	*

The difference between *CiV* strings (9a) and *#iV* strings (21) is that gliding the *i* in the former would result in a complex onset. The fact that this is not permitted suggests that *COMPLEX(Onset) (‘do not have complex onsets’; Prince and Smolensky 1993: 96[23]), which would be mute in tableau (22), outranks ONSET. It is not undominated, however: complex onset clusters are permitted in Czech, as illustrated by the data in (23).

(23) **tr**iumf [tr] ‘triumph’, **pr**áh [pr] ‘threshold’

The interaction of *COMPLEX(Onset), VOWEL-NUCLEUS (V-NUC), ONSET and other relevant constraints for *triumf* is shown in tableau (24).

---

[23] Prince and Smolensky propose COMPLEX, a general constraint against consonantal clusters. It will be shown below that in Czech, COMPLEX has to be split into COMPLEX(Onset) and COMPLEX(Coda).

(24) MAX$_{IO}$(Seg) >> *RT >> *COMPLEX(Onset) >> ONSET >> DEP$_{IO}$(Seg), V-NUC

// triumf //	MAX$_{IO}$ (Seg)	*RT	*COMPLEX (Ons)	ONS	DEP$_{IO}$ (Seg)	V-NUC
☞ a. tri.jumf		6	*		*	*
b. tri.umf		6	*	*!		
c. trjumf		6	**!			*
d. trumf	*!	5	*			
e. rjumf	*!	5	*			*
f. rumf	*!*	4				
g. ti.ri.jumf		7!			*	*

Candidates which use deletion (24d-f) or insertion (24g) to avoid complex onsets are immediately eliminated from the competition because they violate high-ranking MAX$_{IO}$(Seg) and *RT. The remaining candidates all have complex onsets. Candidate (24c), however, in which *i* is glided to [j], violates *COMPLEX(Ons) more severely than candidates (24a) and (24b), as it has three, and not two, consonants in the onset. Finally, between the two remaining candidates, (24a) is selected as optimal even though it violates DEP$_{IO}$(Seg) and V-NUC, because its competitor, (24b), incurs a fatal violation of higher-ranked ONSET.

Moving on now to *Vi* and *Vu* strings, it has been mentioned above that Czech uses gliding to eliminate such sequences in some words (13a,b). In the remaining words, such as the ones in (12b,d), the strategies described above are used (glide insertion in *Vi* strings and no repair in *Vu* strings). Such inconsistent behaviour suggests that either the words in (12b,d) or those in (13a,b) have to be treated as exceptions. Since it is not possible to derive diverse results from identical underlying structures, the difference between the forms must be encoded at the underlying level.

One possibility would be to try to prohibit glides in syllable codas (for example, by means of CODA[u] and CODA[i] constraints) and treat (13a,b) as exceptions. This is not a viable approach for two reasons. First, it would mean that the ubiquitous [j] sounds in Czech *Vi#* sequences, such as the ones in (25), are exceptions. Second, this solution would be difficult to express formally.

(25) r**aj** [aj] 'paradise', m**ůj** [u:j] 'my'

The reverse solution is adopted by Rubach (2000b: 278), who proposes prespecifying the vowel which resists gliding as a syllable nucleus in the underlying representation. The suggestion is made for Standard Slovak, but the same approach could be used in Czech. The difference between *kokain* (12b) and *fajn* (13a) would then be represented as shown in (26).

(26)

```
    N
    |
xxxxxx    xxxx
||||||    ||||
kokain    fain
```

The ranking established so far does not prohibit gliding in codas, so the examples in (13a,b) are accounted for. To block gliding in the exceptional cases, it is necessary to add a constraint which mandates the preservation of the underlying nucleus, Ident(Nuc) (Rubach 2000b: 278). Tableaux (27) and (28) below illustrate the difference in the evaluation of *kokain* and *fajn*. In addition, tableau (28) shows that *Complex has to be split into *Complex(Onset), ranked high to disallow gliding in *CiV* strings (compare tableau [24] above), and *Complex(Coda), ranked low to permit gliding in *Vi* strings.

(27) *Complex(Ons) >> Ons, **Ident(Nuc)** >> Dep$_{IO}$(Seg), *Mult-Link >> *Complex(Coda)

N // kokain //	*Complex (Ons)	Ons	Ident (Nuc)	Dep$_{IO}$ (Seg)	*Mult-Link	*Complex (Coda)
a. ko.ka.in		*!				
☞ b. ko.ka.jin				*	*	
c. ko.kajn			*!			*

(28) *Complex(Ons) >> Ons, Ident(Nuc) >> Dep$_{IO}$(Seg), *Mult-Link >> ***Complex(Coda)**

// fain //	*Complex (Ons)	Ons	Ident (Nuc)	Dep$_{IO}$ (Seg)	*Mult-Link	*Complex (Coda)
a. fa.in		*!				
b. fa.jin				*!	*	
☞ c. fajn						*

In both tableaux, candidate (a), with an onsetless syllable, is eliminated from consideration owing to its violation of high-ranked Onset. Candidate (27c), with gliding, loses to candidate (27b) because it does not preserve the underlying nucleus, thus fatally violating Ident(Nuc).

Candidate (28c), on the other hand, wins over candidate (28b) because there is no underlying nucleus to be kept, and consequently, it only incurs a violation of low-ranked *COMPLEX(Coda).

The data discussed thus far lend themselves to a consistent analysis. A difficulty arises when one tries to generate glottal stop insertion. The problem is best illustrated by looking at a word which has both a hiatus and a word-initial vowel, such as *idiot* ‘idiot’. Recall that in Czech, a glottal stop is inserted in word-initial position, as in (12f), and [j] insertion via spreading is used to resolve a word-medial hiatus in *iV* (12a) and some *Vi* strings (12b). The current ranking predicts no word-initial insertion before non-high vowels and *u*, and [j] insertion before *i*, so *idiot* ends up as *[ji.di.jot], as tableau (29) shows.

(29) *RT >> ONSET >> DEP$_{IO}$(Seg), *MULT-LINK

// idiot //	*RT	ONSET	DEP$_{IO}$(Seg)	*MULT-LINK
a. i.di.ot	5	*!*		
b. ji.di.ot	5	*!	*	*
c. i.di.jot	5	*!	*	*
☞ d. ji.di.jot	5		**	**
e. ʔi.di.ot	6!	*	*	
f. i.di.ʔot	6!	*	*	
g. ʔi.di.ʔot	7!		**	
h. ji.di.ʔot	6!		**	*
👎 i. ʔi.di.jot	6!		**	*

Here, all candidates that insert melody to satisfy ONSET (and, hence, also candidate [29i], which is the attested output) are eliminated from the competition. Candidate (29d), which uses spreading both word-initially and word-internally, is incorrectly selected as the winner because it is the only remaining candidate that does not violate high-ranked ONSET.

To allow glottal stop insertion, it would be necessary to rank *RT below *MULT-LINK. While such ranking would correctly account for the data in (12f), it would also have an adverse effect on other data, as hiatuses would now be resolved by inserting a glottal stop, as shown in tableau (30) for *idiot*.

(30) Onset >> Dep$_{IO}$(Seg), *Mult-Link >>*Rt

// idiot //	Onset	Dep$_{IO}$(Seg)	*Mult-Link	*Rt
a. i.di.ot	*!*			5
b. ji.di.ot	*!	*	*	5
c. i.di.jot	*!	*	*	5
d. ji.di.jot		**	*!*	5
e. ʔi.di.ot	*!	*		6
f. i.di.ʔot	*!	*		6
☜ g. ʔi.di.ʔot		**		7
h. ji.di.ʔot		**	*!	6
👎 i. ʔi.di.jot		**	*!	6

With the reverse ranking *[ʔi.di.jot], candidate (30i), is still not the winner. Together with candidates (30d), (30g) and (30h) it passes on the highest-ranked Onset, but then it incurs a fatal violation of *Mult-Link. As a result, [ʔi.di.ʔot], the only remaining candidate that satisfies *Mult-Link is selected as the optimal output.

Summing up, the two different insertion strategies create a ranking paradox, in which any constraint ranking will produce incorrect result for some of the data; ranking *Rt above Mult-Link will result in glide insertion word-intially and ranking Mult-Link above *Rt will predict word-medial glottal stop insertion. In other words, only one strategy will be used to satisfy Onset in all positions. There is nothing in the system to deal with this situation other than positing *ad hoc* constraints, such as ‘no [ʔ] word-internally’ or ‘no glides word initially’. Given that constraints should be phonetically-grounded and reflect universal tendencies, this would be a rather disappointing solution. In consequence, it is necessary to employ additional machinery. The following section is an attempt to account for the Czech insertion patterns in terms of Sympathy Theory, taking *idiot* as the main example.

2.3.3. Sympathy analysis

The first step in a sympathy theoretic analysis is the selection of a sympathetic candidate, which will influence the computation of the optimal output. Two candidates appear to be particularly appropriate to this end, one being a candidate with word-internal glide insertion only, [i.di.jot], the other one being a candidate with word-initial glottal stop insertion, [ʔi.di.ot].

As mentioned above, the sympathetic base has to satisfy a certain faithfulness constraint, the selector, which is violated by the attested output. If we first assumed that [i.di.jot] was the sympathetic base, the selector would be Anchor-L (‘any element at the left edge of the input has a correspondent at the output; no deletion/insertion at the left edge’; after McCarthy and Prince 1995: 123). The selection of the sympathetic candidate is shown in tableau (31) below.

(31) *Rt >> Onset >> Dep$_{IO}$(Seg), *Mult-Link >> **★Anchor-L**

// idiot //	*Rt	Onset	Dep$_{IO}$(Seg)	*Mult-Link	★Anchor-L
a. i.di.ot	5	**			☑
b. ji.di.ot	5	*	*	*	*
❀ c. i.di.jot	5	*	*	*	☑
d. ji.di.jot	5		**	**	*
e. ʔi.di.ot	6	*	*		*
f. i.di.ʔot	6	*	*		☑
g. ʔi.di.ʔot	7		**		*
h. ji.di.ʔot	6		**	*	*
i. ʔi.di.jot	6		**	*	*

Only three candidates, (31a), (31c) and (31f) satisfy ★Anchor-L and out of these, candidate (31c) is the most harmonic, by virtue of its minimal violation of *Rt and Onset. Consequently, it becomes the sympathetic candidate. This is where the analysis reaches a dead end, however. To obtain [ʔi.di.jot] using sympathy, it is necessary to compare all candidates to the sympathetic base using an inter-candidate faithfulness constraint. However, [ʔi.di.jot] incurs the same faithfulness violation—and more—with respect to [i.di.jot] that [ji.di.jot] does, namely Dep(Seg), which cannot therefore be used as a sympathy constraint. To eliminate [ji.di.jot] it is necessary to refer to *Mult-Link or V-Nuc, neither of which are faithfulness constraints and hence cannot serve as sympathy constraints, either. It follows from the above observations that ★ANCHOR-L cannot be the selector and [i.di.jot] the sympathetic candidate.

The second hypothesis seems more promising. Under the assumption that [ʔi.di.ot] was the sympathetic base, the selector would have to be OUTPUT CONTIGUITY, a constraint militating against medial epenthesis (O-CONTIG, ‘the portion of S_2 standing in correspondence forms a contiguous string’; McCarthy and Prince 1995: 123), and the evaluation would proceed as shown in tableau (32).

(32) ❁**DEP(Rt)** >> ONSET >> DEP$_{IO}$(Seg)[24], *MULT-LINK >>*RT, ★***O-CONTIG**

// idiot //	❁DEP (Rt)	ONSET	DEP$_{IO}$ (Seg)	*MULT-LINK	*RT	★*O-CONTIG
a. i.di.ot		*!*			5	☑
b. ji.di.ot		*!	*	*	5	☑
c. i.di.jot		*!	*	*	5	*
d. ji.di.jot			**	**!	5	*
❁ e. ʔi.di.ot		*!	*		6	☑
f. i.di.ʔot	*!	*	*		6	*
g. ʔi.di.ʔot	*!		**		7	*
h. ji.di.ʔot	*!		**	*	6	*
☞ i. ʔi.di.jot			**	*	6	*

Among the three candidates that satisfy the selector, thus violating ONSET word-medially, candidate (32e), which uses glottal stop insertion as a strategy to repair an onsetless syllable word-initially, best satisfies the constraint ranking. Candidate (32a) fares worse on ONSET, which it breaches twice, and candidate (32b) runs against *MULT-LINK. As a result, (32e) is selected as the sympathetic base.

A comparison of the two most serious contenders, (32i), the actual output, and (32g), the output which would win without sympathy, reveals that the difference between these two in relation to the sympathetic candidate is that in the latter, but not in the former, not only

---

[24] Observe that the data at hand do not reveal the exact ranking of DEP$_{IO}$(Seg), which clearly has to be dominated by ONSET but could also be ranked below *RT. With the inclusion of Sympathy, the constraint ceases to play its pivotal role in the selection of the optimal candidates, which is compelled by the following ranking: ❁DEP(Rt) >> ONSET >> *MULT-LINK >> *RT.

an X-slot, but also melody is inserted. As was mentioned above, because glide insertion in Czech is analysed as spreading from the neighbouring high front vowel onto an inserted X-slot, candidate (32i) only differs from the sympathetic candidate by having one X-slot more. Candidate (32g), on the other hand, has an extra X-slot as well as additional melody and can therefore be rejected on these grounds, by invoking the DEP(Rt) constraint (‘do not insert a root node’; after Lombardi 1998). Note that, as pointed out by Rubach (2000b: 285, ft 24), DEP(Rt) is an extension of the standard theory, because MAX and DEP are typically used to refer to full segments (which in the X-skeletal theory are identified by X-slots), and not to features or nodes.

With the aid of Sympathy Theory it is now possible to account for the two diverse insertion processes in Czech. It remains to be ascertained whether the addition of the sympathetic constraint has no untoward effects on the computation of other, transparent, data. An evaluation of one such word, *poeta* is show in tableau (33).

(33) ❀DEP(Rt) >> ONSET >> DEP$_{IO}$(Seg), *MULT-LINK >>*RT, ★*O-CONTIG

// poeta //	❀DEP (Rt)	ONSET	DEP$_{IO}$ (Seg)	*MULT-LINK	*RT	★*O-CONTIG
❀☞ a.po.e.ta		*			5	☑
b. po.ʔe.ta	*!		*		6	*

In this tableau, the actual output, candidate (33a), is the most harmonic candidate that satisfies the selector and, as such, it becomes the sympathetic base. With the candidate being fully faithful to itself, sympathy is vacuous and candidate (33a) emerges as the winner.

There is one final twist to the whole account, however. If we follow the logic of the extension of DEP and MAX to root nodes, it seems that the analysis could dispense with sympathy altogether. McCarthy and Prince’s (1995: 123) O-CONTIGUITY, which militates against word-internal epenthesis and is hence analogous to DEP, but in a different domain, could also be extended to refer to root nodes. O-CONTIG(Rt) would ban word-internal, but, crucially, not word-initial, glottal stop insertion, as shown in tableau (34).

(34) *O-Contig (Rt) >> Onset >> *Mult-Link, Dep$_{IO}$(Seg)

// idiot //	*O-Contig(Rt)	Onset	*Mult-Link	Dep$_{IO}$(Seg)
a. i.di.ot		*!*		
b. ji.di.ot		*!	*	*
c. i.di.jot		*!	*	*
d. ji.di.jot			**!	**
e. ʔi.di.ot		*!		*
f. i.di.ʔot	*!	*		*
g. ʔi.di.ʔot	*!			**
h. ji.di.ʔot	*!		*	**
☞ i. ʔi.di.jot			*	**

Under this ranking, candidates with onsetless syllables (34a-c,e) fatally violate Onset. Candidates with word-internal glottal-stop insertion (34f-h) are eliminated on account of violating *O-Contig(Rt). Word-internal glide insertion is allowed, because both Dep$_{IO}$(Seg) and *Mult-Link are ranked sufficiently low. The choice among the remaining two candidates, (34d) and (34i), is passed down to the next constraint, that is *Mult-Link. Since candidate (34d) incurs two violations thereof, candidate (34i), which only incurs one, is the winner.

2.4. Concluding remarks

This chapter has examined the success of Sympathy Theory when applied to various types of data exhibiting opacity effects. Section 2.2 has shown that Sympathy Theory provides a satisfactory account of both underapplication opacity (exemplified by Polish) and overapplication opacity (exemplified by Tiberian Hebrew). It is achieved via the introduction of a sympathy constraint, which requires resemblance to a certain failed co-candidate, thus forcing the seemingly superfluous violations incurred by the opaque candidates. Nevertheless, Sympathy Theory fails to handle glide and glottal stop insertion in Czech, which can also be seen as opaque from the point of view of OT. The problem is that no faithfulness constraints exist which could force resemblance to the sympathetic base. An extension of the theory is required to allow constraints of the Max(Feature) type. This is not an attractive solution as it would mean that two disparate mechanisms are necessary to handle opaque data in OT.

Chapter 3

Derivational Optimality Theory

As noted in chapter 1, one of the basic tenets of Optimality Theory is the principle of strict parallelism (McCarthy and Prince 1993: 2), which holds that there is no serial derivation in OT. Although in principle there is nothing in OT architecture that prevents introducing derivational levels (in fact, a non-parallel version of OT, Harmonic Serialism, was introduced and briefly discussed in one of the founding documents of OT, Prince and Smolensky 1993), most current work in OT embraces strict parallelism. Nevertheless, some researchers (Booij 1997, Kiparsky 1997, 2000, Rubach 1997, 2000a,b) argue that abandoning the principle in favour of limited parallelism would hold more advantages than disadvantages. The aim of this chapter is to present one implementation of this idea, Derivational Optimality Theory (DOT, henceforth), as developed by Rubach (1997, 2000a,b, 2003). The chapter is organised as follows. Section 3.1 introduces the core concepts of DOT. Section 3.2 presents DOT analyses of the data discussed in the previous chapters and section 3.3 summarises the discussion.

3.1. Basic principles of DOT

The basic idea of DOT is that the computation of the optimal candidate can be carried out in a stepwise manner. The phonology of a single language may consist of a number of levels, with the output of one level being the input to the following one. The evaluation at each level is fully parallel. Unlike Harmonic Serialism, DOT permits constraint reranking between levels. Different constraint hierarchies at two levels may yield different results, thus creating opacity effects. To restrict the theory, Rubach (2000b: 313) proposes the following principles:

(1) a. *Level Minimalism*
The number of derivational levels is minimal.
b. *Reranking Minimalism*
The number of rerankings in minimal.
c. *Constraint Minimalism*
The number of constraints is minimal.

Principle (1c) follows from the idea that only when the number of constraints is limited can constraint interaction be insightful. The proliferation of constraints undermines the predictive power of the theory. Rubach (2000b: 314) notes that this is the case with Sympathy Theory, where the number of faithfulness constraints is increased by two, for each IO faithfulness constraint has an equivalent sympathy constraint. In fact, if McCarthy's (1999: 351) claim that "there are separate and therefore separately rankable faithfulness constraints on each correspondence relation" is to be taken seriously, the number of faithfulness constraints actually rises by the square of the number of possible candidates, an astronomical amount.

Principles (1a) and (1b) provide an answer to the objection that DOT is not restrictive enough as, by itself, it allows completely different rankings at different levels, a situation not attested in natural languages (McCarthy 1999: 389). Rubach claims that both reranking and postulating additional levels come at a cost. As far as the former in concerned, reranking of constraints between derivational stages should only be admitted if it is justified by the analysis (Rubach 2003: 602). It seems only logical that the changes introduced at level 1 should be preserved at level 2 by means of high-ranked identity constraints. Nevertheless, reranking these and other constraints should always be motivated.

As regards principle (1a), Rubach argues that also introducing new derivational levels should be argued for. The only exception is the cut between word and sentence phonology, which have always been regarded as different. This captures the insights of Lexical Phonology (Kiparsky 1982), which made a distinction between the lexical and post-lexical levels. The introduction of any other levels needs to be motivated. Notably, in DOT, the distinction between other (word internal) levels need not coincide with morphological domains. Here DOT differs from Kiparsky's LPM-OT (Lexical Phonology and Morphology, Kiparsky 2000) where phonological strata have to coincide with morphological levels.

3.2. Analyses

This section provides DOT analyses of the data discussed in the previous chapter. Note that all the processes discussed here take place at word level. However, Rubach also provides examples of the interaction between word-level and sentence-level phonology, for example for Czech (Rubach 2000b: 299–301) and Russian (Rubach 2000a).

3.2.1. Underapplication and overapplication opacity

To begin with the processes of nasal assimilation and yer deletion in Polish, recall that in standard OT it was impossible to link the non-application of nasal assimilation to the presence of an underlying yer.

The problem can be easily resolved if two derivational levels are assumed. At level 1, yers are not deleted. This is when all processes sensitive to yers (that is, nasal assimilation, and, presumably, other processes blocked or triggered by yers) take place. This means that at this level, Max$_{IO}$(V) must outrank the constraint mandating the deletion of yers, Parse(Seg). Ident$_{IO}$(Place) must be dominated by NasAssim to allow nasal assimilation. At level 2, yers get deleted through the reranking of Max$_{IO}$(V) below Parse(Seg). At this point, nasal assimilation can no longer be an active process, so NasAssim has to be ranked lower than Ident$_{IO}$(Place), which has to be ranked high to preserve the changes that took place at level 1. The evaluation of *Irenka* is then as shown in tableaux (2a,b).

(2a) Level 1: NasAssim, Max$_{IO}$(V) >> Ident$_{IO}$(Place), Parse(Seg)

// irɛnEka //	NasAssim	Max$_{IO}$(V)	Ident$_{IO}$(Place)	Parse(Seg)
i. irɛnka	*!	*		
ii. irɛŋka		*!	*	
☞iii. irɛnEka				*

(2b) Level 2: Parse(Seg), Ident$_{IO}$(Place) >> Max$_{IO}$(V), NasAssim

// irɛnEka //	Parse(Seg)	Ident$_{IO}$(Place)	Max$_{IO}$(V)	NasAssim
☞ i. irɛnka			*	*
ii. irɛŋka		*!	*	
iii. irɛnEka	*!			

At level 1, where the ranking permits nasal assimilation, the optimal output, that is, candidate (2aiii), which keeps the underlying yer, retains the alveolar nasal because the context for nasal assimilation is not met. At level 2, candidate (2bii), with an assimilated nasal, fatally violates IDENT$_{\text{IO}}$(Place), which is now ranked higher than NASASSIM. Candidate (2biii) is eliminated owing to its violation of PARSE(Seg). As a result, candidate (2bi), in which the nasal does not assimilate to the following obstruent, emerges as the winner.

It is now necessary to show that the two-level analysis works for the transparent cases, too. The tableaux (3a,b), for *ręka* prove that it does.

(3a) Level 1: NASASSIM, MAX$_{\text{IO}}$(V) >> IDENT$_{\text{IO}}$(Place), PARSE(Seg)

// rɛnka //	NASASSIM	MAX$_{\text{IO}}$(V)	IDENT$_{\text{IO}}$(Place)	PARSE(Seg)
i. rɛnka	*!			
☞ ii. rɛŋka			*	

(3b) Level 2: PARSE(Seg), IDENT$_{\text{IO}}$(Place) >> MAX$_{\text{IO}}$(V), NASASSIM

// rɛŋka //	PARSE(Seg)	IDENT$_{\text{IO}}$(Place)	MAX$_{\text{IO}}$(V)	NASASSIM
i. rɛnka		*!		*
☞ ii. rɛŋka				

Here, at level 1, the optimal output, which becomes the input to level 2, is candidate (3aii), in which the nasal-obstruent cluster is homorganic. The other candidate, (3ai), is suboptimal on account of violating high-ranked NASASSIM. At level 2, candidate (3bi), which would be problematic under standard OT, is eliminated owing to its violation of IDENT$_{\text{IO}}$(Place), which requires the nasal assimilated at level 1 to keep its velar place of articulation. Consequently, candidate (3bii) is selected as optimal. Not only does it satisfy NASASSIM (which is no longer a high-ranked constraint), but also it is fully faithful to the input.

To turn now to overapplication opacity, recall that Tiberian Hebrew exhibits an opaque interaction of glottal stop deletion in the coda and vowel epenthesis into final clusters, which results in words with an epenthetic vowel which is not followed by a consonant. This is problematic for standard OT, as the opaque candidate incurs an unnecessary violation of DEP$_{\text{IO}}$(V), avoided in the transparent candidate, in which the glottal stop is deleted and there is no epenthesis.

Here, again, the solution is straightforward if a derivational step is permitted. In DOT, these two processes can be assigned to different levels: level 1 is an epenthesis level, while level 2 focuses on glottal stop deletion. If all possible candidates discussed in chapter 2 are to be considered (that is, candidates discussed by McCarthy [1999: 336] plus candidates with word-final epenthesis and methathesis), the following constraint rankings are necessary. At level 1, *COMPLEX, which is the driver for epenthesis, needs to be ranked higher than DEP$_{IO}$(V). Since glottal stops in the coda are allowed at this level, *CODA([ʔ]) must be ranked lower than MAX$_{IO}$(C), to ensure that [ʔ] is not deleted, and ALIGN-R, so that a vowel is not inserted at the end of the word (which would result in the satisfaction of both *CODA([ʔ]) and *COMPLEX). The ranking of MAX$_{IO}$(C) and *CODA([ʔ]) is reversed at level 2 to enforce the deletion of glottal stops. For the same reason, ALIGN-R needs to be ranked lower than *CODA([ʔ]). In addition, some constraints now need to become "active" to ensure that the vowel inserted at level 1 is not deleted or metathesised. These are MAX$_{IO}$(V) and LIN, respectively. The above discussion is summarised in tableaux (4a,b), which show the evaluation of [deše] 'tender grass'.

(4a) Level 1: *COMPLEX, MAX$_{IO}$(C), ALIGN-R >> DEP$_{IO}$(V), *CODA([ʔ])

// dešʔ //	*COMPLEX	MAX$_{IO}$(C)	ALIGN-R	DEP$_{IO}$(V)	*CODA([ʔ])
i. dešʔ	*!				*
ii. deš		*!			
☞iii. dešeʔ				*	*
iv. deše		*!	*	*	
v. deʔ		*!			
vi. dešʔe			*!	*	

(4b) Level 2: *COMPLEX, *CODA([ʔ]) >> DEP$_{IO}$(V), MAX$_{IO}$(V), LIN >> ALIGN-R, MAX$_{IO}$(C)

// deše$_1$ʔ$_2$ //	*COMPLEX	*CODA ([ʔ])	DEP$_{IO}$ (V)	MAX$_{IO}$ (V)	LIN	ALIGN -R	MAX$_{IO}$ (C)
i. dešʔ$_2$	*!	*		*			
ii. deš				*!			*
iii. deše$_1$ʔ$_2$		*!					
☞iv. deše$_1$							*
v. deʔ$_2$		*!		*			*
vi. dešʔ$_2$e			*!	*		*	
vii. dešʔ$_2$e$_1$					*!	*	
viii. deše$_1$ʔ$_2$e			*!			*	

At level 1, candidate (4aiii), with a final glottal stop and an epenthetic vowel breaking the final cluster, becomes the winner, as this is the only candidate that satisfies the high-ranked constraints: *COMPLEX, MAX$_{IO}$(C) and ALIGN-R. At level 2, coda-final glottal stops are no longer allowed, which eliminates candidates (4bi,iii,v). The transparent candidate (4bii) incurs a fatal violation of MAX$_{IO}$(V) because it fails to preserve the vowel inserted at level 1. Finally, candidates which keep the epenthesised vowel but metathesise it (4bvii) and/or insert another one (4bvi,viii) are eliminated owing to their violations of DEP$_{IO}$ and LIN. As a result, the correct candidate, (4biv), is selected as optimal.

Tableaux (5a,b) and (6a,b) show that the rankings established for [deše], work for the transparent cases, that is for [melex] 'king', which exhibits epenthesis only, and for [qara] 'he called', with [ʔ]-deletion only.

(5a) Level 1: *COMPLEX, MAX$_{IO}$(C), ALIGN-R >> DEP$_{IO}$(V), *CODA([ʔ])

// melk //	*COMPLEX	MAX$_{IO}$(C)	ALIGN-R	DEP$_{IO}$(V)	*CODA([ʔ])
i. melk	*!				
☞ ii. melek				*	
iii. melke			*!	*	
iv. mel		*!			
v. mek		*!			

(5b) Level 2: *COMPLEX, *CODA([ʔ]) >> DEP$_{IO}$(V), MAX$_{IO}$(V), LIN >> ALIGN-R, MAX$_{IO}$(C)

// mele$_1$k$_2$ //	*COMPLEX	*CODA ([ʔ])	DEP$_{IO}$ (V)	MAX$_{IO}$ (V)	LIN	ALIGN -R	MAX$_{IO}$ (C)
i. melk	*!			*			
☞ ii. melek							
iii. melk$_2$e			*!	*		*	
iv. melk$_2$e$_1$					*!	*	
v. mel				*!			*
vi. mek				*!			*
vii. mele$_1$k$_2$e			*!			*	

(6a) Level 1: *COMPLEX, MAX$_{IO}$(C), ALIGN-R >> DEP$_{IO}$(V), *CODA([ʔ])

// qaraʔ //	*COMPLEX	MAX$_{IO}$(C)	ALIGN-R	DEP$_{IO}$(V)	*CODA([ʔ])
☞ i. qaraʔ					*
ii. qara		*!			
iii. qaraʔa			*!	*	

(6b) Level 2: *COMPLEX, *CODA([ʔ]) >> DEP$_{IO}$(V), MAX$_{IO}$(V), LIN >> ALIGN-R, MAX$_{IO}$(C)

// qara$_1$ʔ$_2$ //	*COMPLEX	*CODA ([ʔ])	DEP$_{IO}$ (V)	MAX$_{IO}$ (V)	LIN	ALIGN -R	MAX$_{IO}$ (C)
i. qara$_1$ʔ$_2$		*!					
☞ ii. qara$_1$							*
iii. qara$_1$ʔ$_2$a			*!			*	
iv. qarʔ$_2$a			*!	*		*	
v. qarʔ$_2$a$_1$					*!	*	

In the case of [melex] ‘king’, the attested output, candidate (5aii) is selected at level 1. Then appropriate faithfulness constraints ensure that this candidate becomes optimal also at level 2. In the case of [qara] ‘he called’, the ranking at level 1 produces an output that is identical to the input (6ai). Then, at level 2, the ranking for glottal stop deletion produces the required result (6bii).

To sum up, it has been shown that DOT is able to handle both underapplication and overapplication opacity. The following section will present a DOT analysis of the non-standard opacity in Czech.

3.2.1. Glide and glottal stop insertion in Czech

Rubach (2000b: 298f) presents a DOT analysis of the Czech data. The aim of this section is to expand his analysis to include all candidates considered in chapter 2 and to present an alternative DOT account with a reversed ordering of events.

Rubach proposes that in Czech word-initial insertion and glottal stops should be banned at level 1. This can be achieved by ranking Onset lower than Align-L and *Rt[25], respectively. Glide insertion is allowed word-medially so Onset should outrank *Mult-Link and Dep$_{IO}$(Seg). At level 2, Align-L and *Rt are reranked below Onset to allow word-initial glottal stop insertion, and *Mult-Link is reranked higher up, in order to prohibit glide insertion. Crucially, *Mult-Link must be ranked higher than *Rt, so that glottal stop insertion is preferred over gliding. Rubach's proposal is illustrated by the following tableaux:

(7a) Level 1: *Rt, Align-L >> Onset >> Dep$_{IO}$(Seg), *Mult-Link

// idiot //	*Rt	Align-L	Onset	Dep$_{IO}$(Seg)	*Mult-Link
☞ i. i.di.jot	5		*	*	*
ii. i.di.ot	5		**!		
iii. ʔi.di.ʔot	7!	*		**	
iv. ji.di.jot	5	*!		**	**

(7b) Level 2: *Mult-Link, Onset >> *Rt, Dep$_{IO}$(Seg) >> Align-L

// i.di.jot //	*Mult-Link	Onset	*Rt	Dep$_{IO}$(Seg)	Align-L
☞ i. ʔi.di.jot	*		6	*	*
ii. i.di.jot	*	*!	5		
iii. i.di.ot		**!	5		
iv. ji.di.jot	**!		5	*	*

(after Rubach 2000b: 299)

If more candidates are considered, for example, ones which undo the changes introduced at the previous level and/or make new ones, faithfulness constraints militating against these changes have to be added to the computation. Tableaux (8a,b) present a DOT computation of *idiot* with a larger candidate set (which includes all candidates discussed in the previous chapters).

---

[25] Rubach uses *[cg] instead of *Rt, which means banning all glottal stops rather than glottal stop insertion (as opposed to glide insertion via spreading; see chapter 2, page 34). Since this change has no bearing upon the analysis, *Rt will be used here for the sake of consistency with the previous chapters.

(8a) Level 1: *RT, ALIGN-L >> ONSET >> DEP$_{IO}$(Seg), *MULT-LINK

// idiot //	*RT	ALIGN-L	ONSET	DEP$_{IO}$(Seg)	*MULT-LINK
i. i.di.ot	5		**!		
ii. ji.di.ot	5	*!	*	*	*
☞ iii. i.di.jot	5		*	*	*
iv. ji.di.jot	5	*!		**	**
v. ʔi.di.ot	6!	*	*	*	
vi. i.di.ʔot	6!		*	*	
vii. ʔi.di.ʔot	7!	*		**	
viii. ji.di.ʔot	6!	*		**	*
ix. ʔi.di.jot	6!	*		**	*

(8b) Level 2: *MULT-LINK, ONSET >> *RT, DEP$_{IO}$(Seg) >> ALIGN-L

// i.di.jot //	*MULT-LINK	ONSET	*RT	DEP$_{IO}$(Seg)	ALIGN-L
i. i.di.ot		*!*	5		
ii. ji.di.ot	*!	*	5	*	*
iii. i.di.jot	*!	*	5		
iv. ji.di.jot	*!*		5	*	*
v. ʔi.di.ot		*!	6	*	*
vi. i.di.ʔot		*!	6		
👎 vii. ʔi.di.ʔot			7	*	*
viii. ji.di.ʔot	*!		6	*	*
☜ ix. ʔi.di.jot	*!		6	*	*

Like in tableau (7a), at level 1, the ranking which mandates word-internal glide insertion and no ONSET satisfaction word initially yields / i.di.jot / as the optimal output. At level 2, however, the analysis goes awry, as all candidates with glides or onsetless syllables, including the actual output, (8bix), are eliminated by high-ranked *MULT-LINK and ONSET. Consequently, candidate (8bvii), with two glottal stops, is incorrectly selected as the winner.

To salvage the analysis, it is enough to add high-ranked identity constraints, which will ensure that the changes introduced at level 1 are preserved at level 2. These are MAX$_{IO}$(Seg), which will mandate the retention of the inserted glide, and IDENT$_{IO}$([-cons]), which will ensure that a glottal stop does not replace the glide word-medially. As tableau (9) shows, the computation of *idiot* at level 2 is now successful.

(9) Level 2: MAX$_{IO}$(Seg), IDENT$_{IO}$([-cons]) >> *MULT-LINK, ONSET >> *RT, DEP$_{IO}$(Seg) >> ALIGN-L

// i.di.jot //	MAX$_{IO}$ (Seg)	IDENT$_{IO}$ ([-cons])	*MULT-LINK	ONSET	*RT	DEP$_{IO}$ (Seg)	ALIGN -L
i. i.di.ot	*!			**	5		
ii. ji.di.ot	*!		*	*	5	*	*
iii. i.di.jot			*	*!	5		
iv. ji.di.jot			**!		5	*	*
v. ʔi.di.ot	*!			*	6	*	*
vi. i.di.ʔot		*!		*	6		
vii. ʔi.di.ʔot		*!			7	*	*
viii. ji.di.ʔot		*!	*		6	*	*
☞ ix. ʔi.di.jot			*		6	*	*

The current hierarchy of constraints still yields incorrect results when applied to words with a sequence of two vowels other than *i*. This is shown in tableaux (10a,b) for the word *Izrael* ‘Israel’.

(10a) Level 1: ALIGN-L, *RT >> ONSET >> DEP$_{IO}$(Seg), *MULT-LINK

// izrael //	*RT	ALIGN-L	ONSET	DEP$_{IO}$(Seg)	*MULT-LINK
☞ i. iz.ra.el	6		**		
ii. iz.ra.ʔel	7!		*	*	
iii. jiz.ra.el	6	*!	*	*	*
iv. jiz.ra.ʔel	7!	*		**	*
v. ʔiz.ra.el	7!	*	*	*	
vi. ʔiz.ra.ʔel	8!	*		**	

(10b) Level 2: MAX$_{IO}$(Seg), IDENT$_{IO}$([-cons]) >> *MULT-LINK, ONS >> *RT, DEP$_{IO}$(Seg) >> ALIGN-L

/ iz.ra.el /	MAX$_{IO}$ (Seg)	IDENT$_{IO}$ ([-cons])	*MULT-LINK	ONSET	*RT	DEP$_{IO}$ (Seg)	ALIGN -L
i. iz.ra.el				*!*	6		
ii. iz.ra.ʔel				*!	7	*	
iii. jiz.ra.el			*!	*	6	*	*
iv. jiz.ra.ʔel			*!		7	**	*
☜ v. ʔiz.ra.el				*!	7	*	*
👎 vi. ʔiz.ra.ʔel					8	**	*

At level 1, the ranking which only allows word-internal glide insertion, produces an output form with two onsetless syllables. At level 2, the ranking should be geared towards word-initial glottal stop insertion only. However, there is nothing in this ranking to prohibit glottal stop insertion word medially. This was not a problem in words where ONSET was satisfied by glide insertion but here another constraint is necessary to ensure that candidate (10bvi), with two glottal stops, does not become optimal. O-CONTIG, militating specifically against word-medial insertion, needs to be introduced to the computation. Note that the introduction of this constraint will not have an untoward effect on the computation of *idiot*, since the glide is inserted at level 1, so at level 2, O-CONTIG is satisfied. Tableau (11) demonstrates the correct level 2 evaluation of *Izrael*.

(11) Level 2: O-CONTIG, MAX$_{IO}$(Seg), IDENT$_{IO}$([-cons]) >> *MULT-LINK, ONS >> *RT, DEP$_{IO}$(Seg) >> ALIGN-L

/ iz.ra.el /	O-CONTIG	MAX$_{IO}$ (Seg)	IDENT$_{IO}$ ([-cons])	*MULT-LINK	ONS	*RT	DEP$_{IO}$ (Seg)	ALIGN-L
i. iz.ra.el					**!	6		
ii. iz.ra.ʔel	*!				*	7	*	
iii. jiz.ra.el				*!	*	6	*	*
iv. jiz.ra.ʔel	*!			*		7	**	*
☞ v. ʔiz.ra.el					*	7	*	*
vi. ʔiz.ra.ʔel	*!					8	**	*

DOT proves to provide a satisfactory account of the Czech data. The solution adopted by Rubach mimics the rule ordering of a derivational account shown in (12a), where glide insertion applies first (the equivalent of level 1) and glottal stop insertion applies second (the equivalent of level 2). However, the reverse order, shown in (12b), is also possible.

(12)	a.			b.	
	UR	// idiot //		UR	// idiot //
	[j]-insertion	/ idijot /		[ʔ]-insertion	/ ʔidiot /
	[ʔ]-insertion	/ ʔidijot /		[j]-insertion	/ ʔidijot /
	SR	[ʔi.di.jot]		SR	[ʔi.di.jot]

Translated into DOT, the above would mean that at level 1, only word-initial insertion should be allowed (which could be achieved via the ranking of ONSET below CONTIGUITY but above Dep$_{\text{IO}}$[Seg]) and glottal stop insertion should be preferred over glide insertion (which would mean ranking *Mult-Link above *Rt). At level 2, O-Contig, *Mult-Link and Dep$_{\text{IO}}$(Seg) need to be ranked below Onset to permit word-internal glide insertion as an onset-providing strategy. *Rt, on the other hand, should be ranked above Onset to ensure that glottal stops are not inserted word medially. Additionally, appropriate faithfulness constraints should be ranked high to prevent the reversal of the changes introduced at level 1. These would be Max(Seg), which would protect the X-slot inserted at level 1, and Ident$_{\text{IO}}$([cg]), which would preserve the word-initial glottal stop. Tableaux (13a,b) and (14a,b) show the computation of *idiot* and *Izrael* under the alternative approach.

(13a) Level 1: O-Contig >> Onset >> Dep$_{\text{IO}}$(Seg), *Mult-Link >> *Rt

// idiot //	O-Contig	Onset	Dep$_{\text{IO}}$(Seg)	*Mult-Link	*Rt
i. i.di.ot		**!			5
ii. ji.di.ot		*	*	*!	5
iii. i.di.jot	*!	*	*	*	5
iv. ji.di.jot	*!		**	**	5
☞ v. ʔi.di.ot		*	*		6
vi. i.di.ʔot	*!	*	*		6
vii. ʔi.di.ʔot	*!		**		7
viii. ji.di.ʔot	*!		**	*!	6
ix. ʔi.di.jot	*!		**	*!	6

(13b) Level 2: Ident$_{\text{IO}}$([cg]), Max$_{\text{IO}}$(Seg) >> *Rt >> Ons >> Dep$_{\text{IO}}$(Seg), *Mult-Link, O-Contig

// ʔi.di.ot //	Ident$_{\text{IO}}$ ([cg])	Max (Seg)	*Rt	Onset	Dep$_{\text{IO}}$ (Seg)	*Mult -Link	O-Contig
i. i.di.ot		*!	5	**			
ii. ji.di.ot	*!		5	*		*	
iii. i.di.jot		*!	5	*	*	*	*
iv. ji.di.jot	*!		5		*	**	*
v. ʔi.di.ot			6	*!			
vi. i.di.ʔot		*!	6	*	*		*
vii. ʔi.di.ʔot			7!		*		*
viii. ji.di.ʔot	*!		6		*	*	*
☞ix. ʔi.di.jot			6		*	*	*

(14a) Level 1: O-Contig >> Onset >> Dep$_{IO}$(Seg), *Mult-Link >> *Rt

// izrael //	O-Contig	Onset	Dep$_{IO}$(Seg)	*Mult-Link	*Rt
i. iz.ra.el		**!			6
ii. iz.ra.ʔel	*!	*	*		7
iii. jiz.ra.el		*	*	*!	6
iv. jiz.ra.ʔel	*!		**		7
☞ v. ʔiz.ra.el		*	*		7
vi. ʔiz.ra.ʔel	*!		**		8

(14b) Level 2: Ident$_{IO}$([cg]), Max$_{IO}$(Seg) >> *Rt >> Ons >> Dep$_{IO}$(Seg), *Mult-Link, O-Contig

// ʔiz.ra.el //	Ident$_{IO}$ ([cg])	Max (Seg)	*Rt	Onset	Dep$_{IO}$ (Seg)	*Mult-Link	O-Contig
i. iz.ra.el		*!	6	**			
ii. iz.ra.ʔel		*!	7	*	*		*
iii. jiz.ra.el	*!		6	*		*	
iv. jiz.ra.ʔel	*!		7		*	*	*
☞ v. ʔiz.ra.el			7	*			
vi. ʔiz.ra.ʔel			8!		*		*

In both cases, at level 1, a candidate with word-initial glottal stop insertion is selected as optimal, since all other candidates either violate O-Contig (if a segment is inserted word-medially) and/or *Mult-Link (if the candidate features a glide) or incur too many Onset violations (if no change at all is made). At level 2, high-ranked Ident$_{IO}$([cg]) and Max(Seg) cause the elimination of all candidates in which the word-initial glottal stop is removed or replaced with a glide. Then high-ranked *Rt eliminates the candidates with word-internal glottal stop insertion. In the case of *Izrael*, this is enough for the correct candidate (14bv) to be selected as optimal. In the case of *idiot*, the final selection is performed by Onset. It is violated by candidate (13bv) but not by candidate (13bix), which becomes optimal.

Since both DOT accounts presented here are consistent with the data, it seems impossible at this stage to give preference to one of them without reference to additional data or theoretical guidelines.

3.3. Concluding remarks

As shown in the present chapter, DOT is capable of accounting not only for opaque interactions as defined by Kiparsky (section 3.2.1) but also for opacity under the extended definition (section 3.2.2), as it allows to assign the processes that interact opaquely to different levels. Although DOT's hybrid nature (combining parallel evaluation and derivational levels) has been criticised, the theory has its advantages as it obviates the need to resort to other supplementary theories (such as MAX(Feature)) and it allows to keep the number of constraints to a minimum.

Conclusions

This thesis has sought to present two alternative approaches to the phenomenon of phonological opacity in Optimality Theory: Sympathy and DOT. It has been shown that both theories provide adequate accounts of underapplication (illustrated by yer deletion and nasal assimilation in Polish) and overapplication opacity (illustrated by epenthesis and glottal stop deletion in Tiberian Hebrew). However, when the notion of opacity is expanded to include glide and glottal stop insertion in Czech, a sympathy theoretic account becomes impossible and recourse to other auxiliary theories is necessary, which calls into question the status of Sympathy as *the* mechanism for handling opacity. DOT, on the other hand, which sacrifices the principle of strict parallelism, is able to account for the Czech data. An additional advantageous effect of DOT is that the number of constraints is reduced.

References

Baković, Eric. 2007. A revised typology of opaque generalizations. Ms., University of California, San Diego.

Benua, Laura. 1995. Identity effects in morphological truncation. In *Papers in Optimality Theory*, ed. by Jill N. Beckman, Laura Walsh Dickey, and Suzanne Urbanczyk. Amherst: University of Massachusetts, GLSA. 77–136.

Benua, Laura. 1997. Transderivational identity: phonological relations between words. Doctoral dissertation, University of Massachusetts, Amherst.

Bermúdez-Otero, Ricardo. 1999. Constraint interaction in language change: quantity in English and Germanic. Doctoral dissertation, University of Manchester and Universidad de Santiago de Compostela.

Booij, Geert E. 1997. Non-derivational phonology meets Lexical Phonology. In *Derivations and constraints in phonology*, ed. by Iggy Roca. Oxford: Oxford University Press. 261–288.

Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English*. New York: Harper & Row.

de Lacy, Paul. 1998. Sympathetic stress. Ms., University of Massachusetts, Amherst.

Halle, Morris. 1992. Phonological features. In *International encyclopedia of linguistics*, ed. by William Bright. Oxford: Oxford University Press. 207–212.

Idsardi, William, J. 1997. Sympathy creates chaos. Ms., University of Delaware, Newark.

Itô, Junko and Armin Mester. 1997. Sympathy theory and German truncations. In *University of Maryland working papers in Linguistics 5. Selected phonology papers from the Hopkins Optimality Theory Workshop 1997/University of Maryland Mayfest 1997*, ed. by Viola Miglio and Bruce Morén. College Park: University of Maryland. 117–139.

Jakobson, Roman. 1962. *Selected writings 1: phonological studies*. The Hague: Mouton.

Kager, René. 1999. *Optimality Theory*. Cambridge: Cambridge University Press.

Kenstowicz, Michael. 1994. *Phonology in generative grammar*. Cambridge, Mass. and Oxford: Blackwell.

Kiparsky, Paul. 1968. Linguistic universals and linguistic change. In *Universals in linguistic theory*, ed. by Emmon Bach and Robert T. Harms. New York: Holt, Rinehart and Winston. 170–202.

Kiparsky, Paul. 1973. Abstractness, opacity, and global rules. In *Three dimensions of linguistic theory*, ed. by Osamu Fujimura. Tokyo: TEC. 56–86.

Kiparsky, Paul. 1982. From cyclic phonology to lexical phonology. In *The structure of phonological representations, volume 1*, ed. by Harry van der Hulst and Neil Smith. Dordrecht: Foris. 131–175.

Kiparsky, Paul. 1997. LP and OT. Handout, LSA Summer Linguistic Institute, Cornell University, Ithaca, N.Y.

Kiparsky, Paul. 2000. Opacity and cyclicity. *The Linguistic Review* 17. 351–365.

Levin, Juliette. 1985. A metrical theory of syllabicity. Doctoral dissertation, MIT, Cambridge, Mass.

Lombardi, Linda. 1998. Evidence for Max Feature constraints from Japanese. In *University of Maryland working papers in linguistics* 7. 41–62.

Lubowicz, Ania. 1998. Derived environment effects in OT. In *Proceedings of the 17th West Coast Conference on Formal Linguistics*, ed. by Kimary Shahin, Susan Blake and Eun Sook Kim. Stanford, CA: CSLI. 451–465.

Malone, Joseph L. 1993. Tiberian Hebrew phonology. Winona Lake, IN: Eisenbrauns.

McCarthy, John J. 1999. Sympathy and phonological opacity. *Phonology* 16. 331–399.

McCarthy, John J. 2002. *A thematic guide to Optimality Theory*. Cambridge: Cambridge University Press.

McCarthy, John J. 2003a. Comparative markedness. *Theoretical Linguistics* 29. 1–51.

McCarthy, John J. 2003b. Sympathy, cumulativity, and the Duke-of-York gambit. In *The Optimal Syllable*, ed. by Caroline Féry and Ruben van de Vijver. Cambridge: Cambridge University Press. 23–76.

McCarthy, John J., and Alan Prince. 1993. Prosodic Morphology: Constraint interaction and satisfaction. Ms., University of Massachusetts, Amherst, and Rutgers University, New Brunswick, N.J.

McCarthy, John J., and Alan Prince. 1994. The emergence of the unmarked: Optimality in Prosodic Morphology. Ms., University of Massachusetts, Amherst, and Rutgers University, New Brunswick, NJ.

McCarthy, John J., and Alan Prince. 1995. Faithfulness and reduplicative identity. In *University of Massachusetts occasional papers in linguistics 18*, ed. by Jill N. Beckman, Laura Walsh Dickey, and Suzanne Urbanczyk. Amherst: University of Massachusetts, GLSA. 249–384. Also available as ROA 60-0000, Rutgers Optimality Archive, http://roa.rutgers.edu.

McCarthy, John J. and Alan Prince. 1999. Faithfulness and identity in Prosodic Morphology. In *The prosody-morphology interface*, ed. by René Kager, Harry van der Hulst, and Wim Zonneveld. Cambridge: Cambridge University Press. 218–309.

Mohanan, Karuvannur P. 1997. Is OT a theory? Ms., National University of Singapore, Singapore.

Prince, Alan, and Paul Smolensky. 1993. Optimality Theory: Constraint interaction in generative grammar. Ms., Rutgers University, New Brunswick, N.J., and University of Colorado, Boulder.

Rice, Curt. 2005. Freedom does not imply anarchy. Paper presented at the Workshop on Freedom of Analysis, Tromsø, September 1-2, 2005.

Rosenthall, Sam. 1994. Vowel/glide alternation in a theory of constraint interaction. Doctoral dissertation, University of Massachusetts, Amherst.

Rubach, Jerzy. 1986. Abstract vowels in three-dimensional phonology: the yers. *The Linguistic Review* 5. 247–280.

Rubach, Jerzy. 1997. Extrasyllabic consonants in Polish: Derivational Optimality Theory. In *Derivations and constraints in Phonology*, ed. by Iggy Roca. Oxford: Oxford University Press. 551–581.

Rubach, Jerzy. 2000a. Backness switch in Russian. *Phonology* 17. 39–64.

Rubach, Jerzy. 2000b. Glide and glottal stop insertion in Slavic languages: A DOT analysis. *Linguistic Inquiry* 31. 271–317.

Rubach, Jerzy. 2003. Duke-of-York derivations in Polish. *Linguistic Inquiry* 34. 601–629.

Sagey, Elizabeth. 1986. The representation of features and relations in non-linear phonology. Doctoral dissertation, MIT, Cambridge, Mass.

Smolensky, Paul. 1995. On the internal structure of the constraint component Con of UG. Handout for a talk presented at University of California, Los Angeles, April 7, 1995.

Steriade, Donca. 1982. Greek prosodies and the nature of syllabification. Doctoral dissertation, MIT, Cambridge, Mass.

Trubetzkoy, Nikolai S. 1939/1969. *Principles of phonology* (translation of *Grundzüge der phonologie*). Berkeley and Los Angeles: University of California Press.

Wilbur, Ronnie B. 1973. The phonology of reduplication. Doctoral dissertation, University of Illinois, Urbana.

Wilson, Colin. 2001. Consonant cluster neutralisation and targeted constraints. *Phonology* 18. 147–197.

Yearley, Jennifer. 1995. Jer vowels in Russian. In *Papers in Optimality Theory*, ed. by Jill N. Beckman, Laura Walsh Dickey, and Suzanne Urbanczyk. Amherst: University of Massachusetts, GLSA. 533–571.