

Structure Preservation and Stratal Opacity in German*

Junko Ito
ito@ling.ucsc.edu

Armin Mester
mester@ling.ucsc.edu

February 26, 2000
Department of Linguistics
University of California, Santa Cruz

1. Introduction: Less is more

One of the defining characteristics of the theory of Lexical Phonology, as developed in the work of Kiparsky 1982, 1985, Mohanan 1986, Borowsky 1986, 1990, among others, was an approach that simultaneously stressed the separation and the unity of the lexical (word-level, structure-preserving) and postlexical (phrase-level, non-structure-preserving) phonology. The central idea was that even though the two constitute different modules of the grammar, the rules and their ordering are actually invariant across strata. If a rule is not seen to be active in a given stratum, this was ascribed to one of two factors: (i) the rule has been "turned off" at an earlier stratum, or (ii) the rule does not yet apply because its output is blocked by the principle of structure preservation. The overall computation of phonological form is thus partitioned into a lexical and a postlexical part, with different canons of wellformedness holding at the two levels. In this paper, we will show that within Optimality Theory (OT; Prince and Smolensky 1993) this view remains correct in essential respects, even though the formal implementation differs in important ways. The resulting model, a type of stratal OT which we will also refer to as "Weak Parallellism", is parallelist in its basic operation, but explicitly recognizes the word phonology ('lexical') and phrasal phonology ('postlexical') as separate modules that operate in sequence.¹

It will be seen that weakly parallel OT-grammars give rise to a type of stratal opacity that receives no coherent analysis in Sympathy Theory (McCarthy 1998, see also Fukazawa this volume), an extension of standard OT dedicated to the treatment of phonological opacity.² Stratal opacity is distinct from a parallelist type of opacity that flows from constraint-conjunctive sources (see Ito and Mester to appear). If this approach is on the right track, it has the prospect of reducing all phonological opacity to other factors—a welcome result allowing us to maintain a

* We would like to thank Linda Lombardi and Cheryl Zoll for detailed and constructive suggestions on an earlier version of this paper. For useful comments, we are indebted to the members of the *Special Research Project for the Typological Investigation of Languages and Cultures of the East and the West* at Tsukuba University (7/99), to the audience at the *PAIK* meeting at Kobe University (12/99), and to the participants at the *Workshop on Conflicting Rules in Phonology and Syntax* at the University of Potsdam (12/99). Special thanks are due to Luigi Burzio, Caroline Féry, Haruka Fukazawa, Shosuke Haraguchi, Bruce Hayes, Takeru Honma, René Kager, Mafuyu Kitahara, Haruo Kubozono, Gereon Müller, Akio Nasu, Sam Rosenthal, Philip Spaelti, Shin-ichi Tanaka, Markus Walter, Richard Wiese, Noriko Yamane, Teruo Yokotani, Yuko Yoshida, Hideki Zamma, and Draga Zec.

¹ This type of stratal OT was first explored in some detail by McCarthy and Prince 1993 in their analysis of the prosodic morphology of Axininca Campa and has recently also been advocated by Kiparsky 1998. Within the larger field of optimization-based approaches to phonology, this approach has a precedent in the work of Goldsmith (1990, 319-331; 1991).

² While many differences in approach and outlook remain, the arguments given here complement those of Kiparsky 1998, where a comprehensive pan-Arabic syllable typology, with analyses of opaque interactions of epenthesis, syncope, and stress, crucially involves a serial separation of lexical phonology and postlexical phonology.

simpler overall conception of the grammar without opacity-specific devices. Opacity is always the result of independently existing properties of the grammar, including both its overall architecture (lexical/postlexical serialism) and mechanisms internal to the constraint system (constraint-conjunction, see Smolensky 1995).

The background discussion in section 2 presents an overview of the sources of opacity in OT. The masking of dorsal fricative assimilation effects (responsible for the [ç]~[x] alternation) by allophonic R-vocalization in German constitutes the empirical focus of this paper, and section 3 illustrates the problem that this counterfeeding interaction poses for strictly parallel and transparent OT. Section 4 shows how a weakly parallel model of OT, distinguishing the lexical and the postlexical modules, straightforwardly accounts for the problematic interaction. Section 5 confronts Sympathy Theory with the counterfeeding problem. It turns out that, despite its initial apparent success, Sympathy faces insurmountable difficulties with the Richness-of-the-Base Hypothesis whenever, as in the case under investigation, the masking process of an opaque interaction is allophonic. The argument against Sympathy is extended in section 6, where an additional process of German phonology, g-spirantization, is brought into the picture. The resulting doubly opaque interaction cannot be resolved by Sympathy even at the price of abandoning the Richness-of-the-Base Hypothesis and stipulating constraints on inputs. Section 7 concludes the paper with a summary and a discussion of further consequences.

2. Weak parallelism and strict parallelism

A phonological system is said to contain *opacity* to the extent that it contains generalizations that make crucial contributions to the overall computation of phonological form, but are not statable as truths about outputs. This raises important issues for OT, a theory of phonological form that is based on output constraints. Since opacity involves the masking of specific generalizations that hold within a specific phonological system (i.e., not of universal constraints from the universal constraint set), it cannot simply be reduced to constraint domination. Taken by itself, constraint domination ($A \gg B \gg C \gg \dots$) can give rise to complex and convoluted surface generalizations ("C, except if B, except if A, ..."), but not to generalizations that are impossible to state as surface generalizations (see McCarthy 1998 for detailed discussion).

If OT with its output orientation is on the right track as an explanatory theory of phonology, how can opacity exist at all in phonological systems? In what sense can there be generalizations that are true even though they are not true of outputs? This problem is a real and pressing one. We will argue that the proper answer does not involve the creation of an opacity-specific mechanism grafted onto the basic model—be it Sympathy or something else—, but is already implicit in two structural properties of the grammar: (a) constraint combinatorics, which creates new composite constraints out of elementary ones, and (b) modular organization, in particular, the serial articulation into lexical and postlexical phonology as separate modules. Loosely put, our proposal is that the effects of some generalizations are masked and hidden on the surface not because there is some component of the theory responsible for the masking and hiding of generalizations, but rather as a side effect of both parallel (1a) and serial (1b) elements of the theory.

(1) a. Parallel sources of opacity:

- i. [M&_rF] conjunctions (Ito and Mester to appear): The effects of a markedness constraint M are partially blocked because the relevant constraint is in reality not the pure markedness constraint M, but rather the Local markedness-cum-faithfulness

conjunction $M\&_iF$. This restricts the markedness effects to certain derived environments (Lubowicz 1998), resulting in opacity. Informally speaking, while derived β changed into γ , underived β does not:

$/\alpha/ _ > [\beta] - > [\gamma]$
 $/\beta/ _ = > [\gamma]$

- ii. $[F\&_iF]$ conjunctions (Kirchner 1996): Local conjunctions of faithfulness constraints result in the stepwise mapping characteristic of chain shifts. Informally speaking, while underlying α changed into β and underlying β into γ , derived β does not change into γ .

$/\alpha/ > [\beta]$

$/\beta/ > [\gamma]$

$/\alpha/ _ = > [\gamma]$

- b. Serial sources of opacity

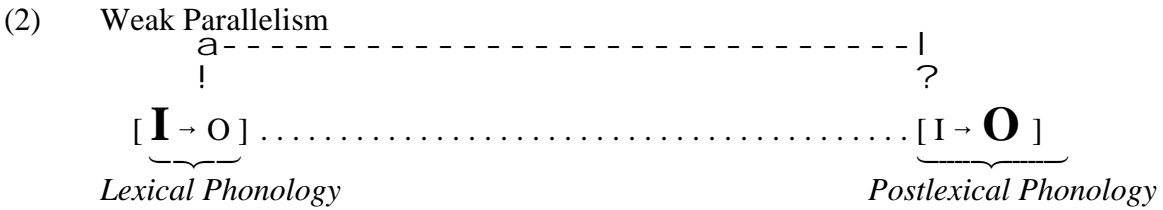
Patterns established through the constraint ranking of the lexical module are masked by the effects of the partially different constraint ranking of the postlexical module.

Viewed from this vantage point, opacity is no more than an epiphenomenon, and developing a theory with specific opacity-generating devices, in spite of ingenuity and descriptive success, is a questionable move at the level of explanatory adequacy. The sources of opacity are multiple and diverse, neither exclusively parallel nor exclusively serial; its proper treatment is therefore a strictly reductionist one. In particular, subsuming all opacity to a single factor is neither possible nor desirable, and attempts to place the explanatory burden entirely on one side, be they parallelist (McCarthy 1998) or serialist (Kiparsky 1998), cannot succeed. Opaque patterns in outputs are rather the result of the way in which the whole grammar is configured. And to the extent that the overall architecture of an OT grammar contains both parallel and serial aspects, it comes as no surprise that opaque patterns in outputs reflect this by showing a mixture of parallel and serial factors.

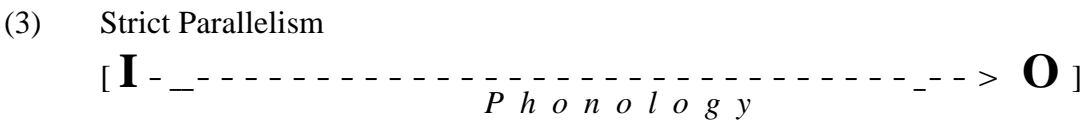
The opacity implications of the theory of $[M\&_iF]$ -conjunctions (1a) have been developed in detail in another paper (Ito and Mester to appear), and we have little to add here.³ In this paper, it is the stratal masking of generalizations (1b), in broad outline familiar from traditional phonology, that occupies our attention. Besides its opacity-related aspects, we will have occasion to explore the particular ways in which lexical and postlexical phonology show different constraint rankings.

With rule sequentialism no longer in serious consideration, some of the remaining options are outlined in (2)–(4). We adopt a basic distinction, as in (2) (dubbed "Weak Parallelism"), between lexical phonology and postlexical phonology as serially connected modules of the grammar, coupled with output-output relations (as developed in Benua 1997 and other work).

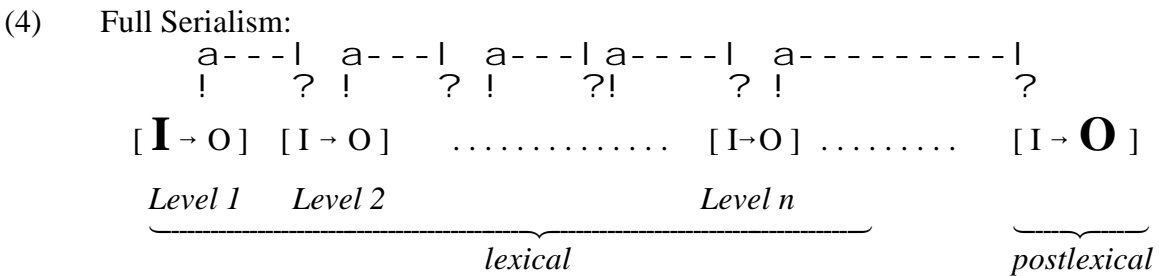
³ Except for taking note of the general pattern that emerges when the two cases in (1a) are placed side-by-side: $[M\&_iF]$ -conjunctions lead to opacity-inducing restrictions of M to derived elements, whereas $[F\&_iF]$ -conjunctions lead to opacity-inducing restrictions of some markedness constraint to underived elements.



We thus do not subscribe to Strict Parallelism (3), where no distinction is made between the lexical and postlexical modules of phonology. We are not familiar with an explicit proposal (let alone a worked-out analysis of a significant portion of the phonology of some language) along such strictly parallelist lines, but this is the approach literally implied by parallelist credos encountered in the recent literature.



On the other hand, we also do not subscribe to the multi-layered cascade of word-internal levels found in Full Serialism (4).⁴



Multiple and serially connected word-internal levels made sense in the theoretical environment of Full Serialism's intellectual predecessor, namely classical Lexical Phonology and Morphology, where level and rule serialism were involved in virtually every explanation the theory was offering: cyclic vs. non-cyclic rule application at a level, the strict cycle condition, structure-building vs. structure-changing operations, underspecification, cyclic default rules, to name a few. All of these devices and principles were intrinsically connected with serialism and constituted its theoretical lifeline. Most of them, however, are long gone from the theoretical scene, having being replaced in OT by other— and superior—means of analysis and explanation, including local constraint conjunction, Output-to-Output faithfulness, and others. It would thus be surprising to find the multiple word-internal levels characteristic of classical lexical-phonological theory and practice still holding on to their former positions.

⁴ The proliferation of word-internal levels was a problem persisting throughout the development of Lexical Phonology, where it was all too easy for analysts to circumnavigate problems by postulating level distinctions without independent support. Part of the reason was that phonology was held responsible for all kinds of sound alternations, including cases nowadays mostly analyzed as allomorph selection.

3. The problem: opaque [ç] in [+back] contexts

The case of phonological opacity to be considered in this paper involves the process of R-vocalization in German, which partially masks the context of a second alternation, namely, the dorsal fricative assimilation process responsible for the distribution of palatal [ç] and velar [x].

3.1 R-vocalization

R-vocalization replaces /R/ in syllable codas by [ɐ̯]—in Moulton's (1962) words, a "lower mid unrounded vowel between central and back". Its syllable peak counterpart [ɐ̯] (which arises, derivationally speaking, out of contraction of [ɐ̯] with schwa) is found, e.g., in the second syllable of ['mʊtɐ̯] 'mother' (*Mutter*).

(5) **R-Vocalization:** R → ɐ̯ / __ C₀] σ

Examples of R-Vocalization appear in (6), where onset [R], as in the plural form [ty:RƏn], corresponds to postnuclear [ɐ̯], as in the singular form [ty:ɐ̯].⁵

(6)	[R]		[ɐ̯]	
Türen	[.tyː.RƏn.]	Tür	[.tyːɐ̯.]	<i>door, pl/sg</i>
Ohren	[.oː.RƏn.]	Ohr	[.oːɐ̯.]	<i>ear, pl/sg</i>
Flure	[.fluː.RƏ.]	Flur	[.fluːɐ̯.]	<i>corridor, pl/sg</i>
Meere	[.meː.RƏ.]	Meer	[.meːɐ̯.]	<i>sea, pl/sg</i>
Tore	[.toː.RƏ.]	Tor	[.toːɐ̯.]	<i>gate, pl/sg</i>
hören	[.høː.RƏn.]	hört	[.høːɐ̯t.]	<i>hear, inf/3sg</i>
dürre	[.dʏ.RƏ.]	dürrste	[.dʏɐ̯s.tə.]	<i>dry, driest</i>
zerren	[.tsɛ.RƏn.]	zerrt	[.tsɛɐ̯t.]	<i>pull, inf/3sg</i>
irren	[.ɪ.RƏn.]	irrt	[.ɪɐ̯t.]	<i>err, inf/3sg</i>

Vocalized R contracts in various ways, sometimes optionally, sometimes obligatorily, with preceding low and lower-mid back vowels (i.e., [ɑ:, ɒ, ɔ]), often resulting in a long version of the preceding vowel. This is especially clear for [ɔ], where we find numerous examples such as *Horn* [hɔ:n] 'horn', *Dorf* [dɔ:f] 'village', or *Wort* [vɔ:t] 'word' that testify to the distributional generalization. Relevant alternations appear in cases where [ɔ:] (from /ɔR/) corresponds to [œɐ̯] (from /œR/, with /œ/ induced by umlaut).

⁵ The examples in this paper are taken from the speech of the second author. Apart from some regional specifics, the forms are representative of the speech of most parts of Northern Germany. A full investigation of the distribution of [R] and [ɐ̯]—which we are not undertaking in the present context—would need to attend to various further details. In the style of speech under consideration here, all coda- R's are replaced by [ɐ̯], irrespective of the weight of the preceding nucleus. In more formal speech, there is a distinction between obligatory vocalization after long V and diphthongs, as in the pronoun *wir* [vi:ɐ̯] 'we' (*[vi:R]) and optional vocalization after a short V, as in *wirr* 'confused', where both vocalized [vi:ɐ̯] and nonvocalized [vi:R] are encountered (see Krämer 1981, Hall 1993, Wiese 1996 and references cited there for details.)

(7)		[ɔ:]		[æɐ̯]	
	Dorf	[.dɔ:f.]	Dörfer	[.dœɐ̯.fœ.]	<i>village, sg/pl</i>
	Wort	[.vɔ:t.]	Wörter	[.vœɐ̯.tœ.]	<i>word, sg/pl</i>

For the low vowel [ɑ, ɑ:], contraction appears optional, as the examples in (8) show.

(8)	scharren	[.ʃɑ.RƏn.]	scharr	[.ʃɑɐ̯.]	~	[.ʃɑ:]	<i>dig, inf/imp</i>
	Scharen	[.ʃɑ:RƏn.]	Schar	[.ʃɑ(ɨ)ɐ̯.]	~	[.ʃɑ:]	<i>crowd, pl/sg</i>
			cf. Schah			[.ʃɑ:]	<i>shah</i>
			Anarchie	[,ɑnɑɐ̯'çi:]	~	[,ɑnɑ:'çi:]	<i>anarchy</i>

The contractions in (7) and (8) are caused by constraints against sequences of all-too-similar vocoids (*[ɔɐ̯] and *[ɑɐ̯]); our formal treatment of R-vocalization abstracts away from these additional effects.

The analysis of the allophonic alternation between consonantal [R] as the basic variant and vocalic [ɐ̯] in (9) follows standard lines: The vocalic variant violates a context-free markedness constraint against low glides (*ɐ̯) and appears only through the force of the top-ranked contextual constraint *CODA/R, a coda condition against uvular [R] (more exactly, an M&M conjunction *CODA&_l*R).⁶ Faithfulness ranks lowest. [ɐ̯] is the closest vocalic counterpart of [R] (see Krämer 1981), and we use IDENT(CONSONANTAL) as a stand-in for all faithfulness constraints violated in the R→ɐ̯ change.

(9)	*Coda/R	(markedness constraint against consonantal R in the coda)
	*ɐ̯	(markedness constraint against low glide)
	Ident(cons)	(faithfulness to input specifications for [consonantal])

Using [.ty:ɐ̯.]~[.ty:RƏn]. 'door, pl/sg' as an example, the correct forms are derived as in (10): The candidate with vocalic [ɐ̯] wins out in the singular, where the R-candidate violates top-ranking *CODA/R (10a). But in the plural (10b), the segment in question comes to occupy onset position, making *CODA/R irrelevant. The decision comes down to the level of context-free markedness, where the ɐ̯-candidate, with its low glide, receives a fatal violation mark.

⁶ For discussion and analysis of similar types of coda conditions, see Smolensky 1995 and Ito and Mester to appear.

(10) a.

/ty:R/ Tür 'door'	*Coda/R	*ɤ̘	Ident(cons)
☞ ty:ɤ̘		*	*
ty:R	*!		

b.

/ty:R+ən/ Türen 'doors'	*Coda/R	*ɤ̘	Ident(cons)
.ty:ɤ̘ən.		*!	*
☞ .ty:Rən.			

The fact that faithfulness (here, IDENT(CONS)) ranks lowest makes the process allophonic: Input specifications for [consonantal] are noncontrastive, as far as the distribution of [R] and [ɤ̘] is concerned. Richness-of-the-Base (Prince and Smolensky 1993) demands, therefore, that it should be irrelevant for the selection of the correct output candidate whether /ɤ̘/ or /R/ is posited in the input. Output constraints are sufficient to make the correct choice irrespective of input specifications. (11) shows that input specification as /ɤ̘/ or /R/ indeed does not matter, the same results are obtained as before.

(11) a.

/ty:ɤ̘/ Tür 'door'	*Coda/R	*ɤ̘	Ident(cons)
☞ ty:ɤ̘		*	
ty:R	*!		*

b.

/ty:ɤ̘+ən/ Türen 'doors'	*Coda/R	*ɤ̘	Ident(cons)
.ty:ɤ̘ən.		*!	
☞ .ty:Rən.			*

3.2 [ç~x] allophony

The second alternation involved, the well-known dorsal fricative allophony in (12), concerns palatal [ç] (as in *möchte* [mœçtə] 'wants to') alternating with velar [x] after back vowels (as in *mochte* [mɔxtə] 'wanted to').⁷

⁷ For [x~ç], the 'derived contrasts' of the *Kuchen~Kuh-chen* 'cake/little cow' variety made famous by Bloomfield 1930 are grammatically and prosodically conditioned, see Borowsky 1990, Merchant 1996 and references cited there. Concerning the precise place of articulation of what we transcribe as [x], Wiese 1996 (referring to Kohler 1977) is undoubtedly right in pointing out that the segment in question in many contexts contains a uvular component—in particular, after the low vowel [ɑ] and optionally also after lax [ɔ] and [ʊ]. Since our interest here lies not in coarticulatory detail, but rather in the overall distinction between front and back articulation of the dorsal fricative, we continue to use [x] as a broad transcription.

(12) Dorsal Fricative Allophony			ç	→	x	/[V, +back] ___	
	[ç]						[x]
Licht	[lɪçt]	'light'	Bucht	[bʊxt]	'bay'		
Nächte	[nɛçtə]	'nights'	Nacht	[nɑxt]	'night'		
möchte	[mœçtə]	'wants to'	mochte	[mɔxtə]	'wanted to'		
Bücher	[byːçɐ]	'books'	Buch	[buːx]	'book'		
nächste	[nɛːçstə]	'next'	nach	[nɑːx]	'after'		
Eiche	[aɪçə]	'oak'	Bauch	[baʊx]	'stomach'		
euch	[ɔʏç]	'you'	auch	[aʊx]	'also'		

Palatal [ç] is the basic variant and occurs, to the exclusion of [x], in initial position (13a) and post-consonantly (13b), i.e., everywhere except after back vowels.⁸

- (13) a. [ç]e'mie 'chemistry' * [x]emie
 [ç]i'rurg 'surgeon' * [x]irurg
 '[ç]arisma 'charisma' * [x]arisma
 '[ç]thonisch 'belonging to the earth' * [x]tonisch
- b. sol[ç] 'such' *sol[x]
 Kol'[ç]ose 'kolchoz' *Kol[x]ose (but cf. Russian [kal'xos])
 man[ç]e 'some' *man[x]e
 Bron'[ç]itis 'bronchitis' *Bron[x]itis
 Kir[ç]e 'church' *Kir[x]e
 dur[ç] 'through' *dur[x]

Our basic analysis, which follows Merchant 1996 (see also literature cited there), establishes [ç] as the basic variant in relation to [x] by means of the two ranking relations in (14).⁹

- (14) *x > { *ç, Ident(back) } /zɔlx/ → [zɔlç], *[zɔlx] 'such'
 /zɔlç/ → [zɔlç], *[zɔlx]

As has often been pointed out (see (11) above), it is irrelevant for allophonic pairs such as [ç]~[x]

⁸ As Hall 1992 observes, initial [x] is found only in a handful of unassimilated loans such as *Junta* or *José*. Word-initial orthographic <ch> is often pronounced [ʃ] or [k] (the latter case includes *Charisma* for many speakers).

⁹ The precise factors underlying the status of [ç] as the basic variant in relation to [x] in Standard German warrant further study: [x] is arguably the cross-linguistically more frequent segment, and dialects of German that do not participate in the dorsal fricative alternation invariably have [x] and not [ç] (cf. Swiss German [dɪx], [mɪx] for Standard German [dɪç], [mɪç] 'you', 'me').

which one of the two is posited in the input form: higher-ranking constraints obliterate the force of IO-Faithfulness, as far as the palatal/velar distinction is concerned. In conformity with Richness-of-the-Base, inputs are free to contain either /ç/ (as in (15a)) or /x/ (as in (15b)), the correct result is obtained with the ranking in (14) (this also holds for inputs with a segment underspecified for backness).

(15) a. With /ç/ as input:

/zɔlç/	*x	*ç	Ident(back)
zɔlx	*!		
☞ zɔlç		*	*

b. With /x/ as input:

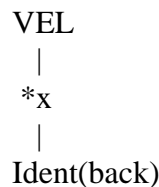
/zɔlx/	*x	*ç	Ident(back)
zɔlx	*!		*
☞ zɔlç		*	

Completing the analysis, the contextual variant [x] emerges under the pressure of the top-ranking syntagmatic velarization constraint VEL (16), a member of the family of feature agreement conditions resulting in assimilation. (16) disallows the sequence "back vocoid followed by voiceless palatal fricative".

(16) VEL: *ç / [-cons, +back] ____ "[-cons, +back] must not be followed by [ç]."

The ranking in (17)¹⁰ derives the overall distribution of the two dorsal fricatives, as illustrated in (18).

(17) [ç~x] allophony:



(18) a. /x/ as input:

/buːx/	VEL	*x	Ident(back)
☞ buːx		*	
buːç	*!		*

b. /ç/ as input:

/buːç/	VEL	*x	Ident(back)
☞ buːx		*	*
buːç	*!		

¹⁰ IDENT(BACK) here refers to consonantal backness. The vowel feature is protected by a higher-ranking constraint, possibly IDENT(VPLACE) (or different features are involved for consonants and vowels, see Ní Chiosáin and Padgett 1997, this volume for discussion).

3.3 Counterfeeding interaction

Placing the two alternations under discussion side-by-side, we note a potential feeding relationship: R-vocalization results in a [+back] vocoid, and dorsal fricative allophony produces a special allophone [x] after [+back] vocoids. The dorsal fricative exhibits opaque (counterfeeding) behavior in such interactions, as shown in (19), with the output sequence [... ɤ̥ ...]. Here we find the palatal variant [ç] in the context "[–cons, +back] ___", i.e., precisely where the variant [x] would be called for.

(19)	fürchten	[fʏ ɤ̥ tə̃n]	'fear'	*[fʏ ɤ̥ tə̃n]
	Lerche	[lɛ ɤ̥ çə̃]	'lark'	*[lɛ ɤ̥ xə̃]
	Pferch	[pfɛ ɤ̥ ç]	'corral'	etc.
	Zwerchfell	[tsvɛ ɤ̥ çfɛl]	'diaphragm'	
	Kirche	[kɪ ɤ̥ çə̃]	'church'	
	Storch	[stɔːç]	'storck'	
	horchen	[hɔːçə̃n]	'listen'	
	Lurch	[lʊ ɤ̥ ç]	'lizard'	
	durch	[dʊ ɤ̥ ç]	'through'	

The distributional opacity is especially tangible in cases such as *Storch* [stɔːç] 'storck' and *horchen* [hɔːçə̃n] 'listen', with long [ɔː] through contraction (see (7) and (8) above). All the examples in (19) show a dorsal fricative remaining palatal (front) in spite of a preceding back vocalic environment, i.e., behaving as if it stood in post-consonantal position.¹¹

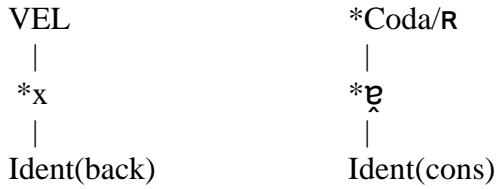
The basic interaction is shown in rule form in (20), where R-vocalization counterfeeds dorsal fricative assimilation (Wiese 1996: 257-258).

(20)		/dʊ ɤ̥ ç/	durch 'through'
	dorsal fricative assimilation:	ç → x / [–cons, +bk] ___	–
	R-vocalization:	R → ɤ̥ / ___ C ₀] σ	ɤ̥
			[dʊ ɤ̥ ç]

The two constraint subhierarchies are repeated in (21).

¹¹ There are in fact dialects showing [x] in these contexts, with [kɪ ɤ̥ xə̃] for 'church', [hɔːxə̃n] for 'listen' etc. Such 'transparent' dialects do not belong to Low German, whose consonantism is altogether different because it was not affected by the High German sound shifts (with forms such as [kɛ ɤ̥ kə̃n] for 'church').

(21) [ç~x] allophony: [R~ʁ] allophony:



The two subhierarchies interact as in (22): Here *[dʊʁx] emerges as the wrong winner (☹!!) which is not defeatable by elementary phonological means. It has no violations of the two high-ranking constraints (VEL and *CODA/R), whereas all other candidates, including the actual winner [dʊʁç] (☺), violate either one of these undominated constraints.

(22) Transparent analysis, with incorrect winner:

/dʊʁç/ 'through'	*Coda/R	VEL	*ʁ	*x	Id(cns)	Id(bk)
dʊʁç	*!					
☺ dʊʁç		*!	*		*	
dʊʁx	*!			*		*
☹!! dʊʁx			*	*	*	*

4. The lexical/postlexical divide

How, then, do we deal with this type of opacity problem in OT? As pointed out in the introduction, the answer is not likely to be found in some particular theoretical device whose exclusive function it would be to produce opaque patterns. Quite apart from problems of descriptive adequacy and completeness that existing proposals are beset with (see the discussion below in sections 5 and 6 and in Ito and Mester to appear), the rationale for the very existence of such a device would remain obscure: Why would Universal Grammar contain a component that is exclusively concerned with the hiding of generalizations?

Given the architecture of an OT grammar assumed here, opacity effects can arise in (at least) two separate ways. First, the interaction of constraints in optimality-theoretic phonology is not restricted to a ranking imposed on a given set of basic constraints, but the basic constraints can also interact by combining with each other, resulting in more complex constraints (as argued by Smolensky 1995 and others). This gives rise to parallel opacity effects. Second, traditional phonology is correct in making a basic distinction between lexical and postlexical phonology as different parts of grammar that are governed by partially different, and serially connected, constraint systems. This gives rise to stratal opacity.

Once we take the separation between the lexical and the postlexical modules of phonology seriously, the problematic interaction turns out to be a straightforward case of stratal opacity whose occurrence is arguably predictable on independent grounds (similar to some of the cases argued for in Kiparsky 1998). On the one hand, the [x~ç] alternation is a classic example of a lexically established allophonic distribution. Ever since Bloomfield 1930 its sensitivity to

lexicon-internal morphological structure is well-known, as shown by examples such as *Kuchen/Kuh-chen* ('cake/'little cow': [ku:xən]/[ku:çən], see also note 7), and by the general fact that the less restricted variant [ç] is the only one that appears morpheme-initially, irrespective of the backness of preceding vowels (*Indo-[ç]ina*, etc.). The two variants, palatal and velar, are acoustically and articulatorily far separated, we are not dealing with a gradient coarticulatory effect. The alternation is also never optional, even in fast or casual speech. Rather, the two variants [ç] and [x] themselves undergo further coarticulatory adjustment in the postlexical phonology. On the other hand, [R~ ʁ] allophony is a partially optional (see note 5) and strongly phonetically conditioned process. It shows no morphological sensitivities, and the allophone transcribed as [ʁ] is subject to an extensive degree of variation and malleability (Krämer 1981, Wedel in prep.), as is typical of processes that are under detailed phonetic control. All these properties are hallmarks of the postlexical module.

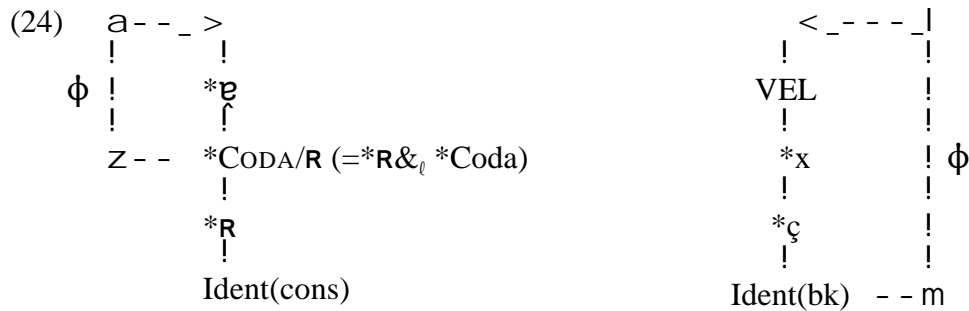
Let us, then, follow traditional wisdom for a moment and consider the consequences of assigning [ç~x] allophony and [R~ ʁ] allophony to different modules of phonology, treating the first one as lexical, and the second one as postlexical. The lexical constraint ranking is given in (23), where (23a) shows the constraints responsible for the distribution of [ʁ] and [R], (23b) the constraints responsible for [x] and [ç].

- | | | | |
|---------|--|----|---|
| (23) a. | $ \begin{array}{c} *ʁ \\ \\ *Coda/R (= *R \&_l *Coda) \\ \\ *R \\ \\ Ident(cons) \end{array} $ | b. | $ \begin{array}{c} VEL (= *ç / [+back, -cons] ___) \\ \\ *x \\ \\ *ç \\ \\ Ident(back) \end{array} $ |
|---------|--|----|---|

Turning to (23a) first, *ʁ is ranked highest in the lexical phonology, above *CODA/R—therefore no *ʁ appears in lexical outputs. This establishes the lexical segment inventory as containing consonantal [R] but no vocalic [ʁ]. The overall effect recaptures the lexical-phonological idea that the lexical phonology is "structure-preserving" (see Kiparsky 1985 and numerous later works). The lexical segment inventory is not separately stipulated, but emerges directly from the lexical constraint ranking itself, which does double duty in accounting for alternations as well as defining the lexical inventory. The remainder of the subhierarchy (23a) is straightforward: The ranking *CODA/R » *R is intrinsic since the conjoined constraint includes the simplex constraint *R,¹² and *R » IDENT(CONS) holds by M » F as a default ranking principle (see Hayes 1999, Ito and Mester 1999, Prince and Tesar 1999). Most aspects of the constraint hierarchy in (23b) have been justified above (see section 3). *ç » IDENT(BACK) again holds by M » F default.

The ranking in the postlexical phonology is shown in (24), where "ϕ" is used to highlight how postlexical ranking differs from lexical ranking.

¹² One advantage of such conjoined constraints is that their ranking in the constraint hierarchy is immediately restricted in certain respects. On the other hand, a non-conjoined coda condition could potentially be ranked anywhere in the hierarchy.



The postlexical phonology differs from the lexical phonology in two mutually independent respects. On the one hand, the contextual markedness constraint *CODA/R has moved above the context-free markedness constraint *ϕ. As a consequence, [R~ϕ] allophony is postlexically operative. In classical lexical-phonological parlance, "structure preservation has been turned off", and [ϕ] is now found in codas in place of [R]. On the other hand, the faithfulness constraint IDENT(BACK) has come to dominate the velarization constraint VEL. This is the OT-correspondent of a rule being "turned off": It freezes the effects of [x~ç] allophony in their lexical form, the alternation is an exclusively lexical phenomenon and is not at work at the postlexical level. A side effect of these two general properties of the postlexical phonology (heightened influence of contextual markedness, and preservation of broad allophone distributions determined lexically) is that the results of lexical velarization are preserved irrespective of postlexical vocalization—in other words, opacity. Similar cases of opaque [ç] are found through phrasal vowel elision, such as *tu ich nicht* 'I don't do that' pronounced as [tu:ɪçnɪç] (without elision) or, in fast speech, as [tu:çnɪç] (with elision), but never as *[tu:xnɪç]. This stratal scenario is the clue to the opaque appearance of [ç] in [+back] contexts.

(23) and (24) show that the differences between lexical and postlexical phonology are not restricted to the ranking of faithfulness constraints, but also concern the ranking of markedness constraints with respect to each other (in particular, contextual and contextfree markedness). The relation between the serially connected modules is thus a very different one from morpheme class distinctions within the phonological lexicon (native vs. foreign, etc.): The latter is arguably reducible to differential faithfulness ranking (Ito and Mester 1995), which can in turn be given a strictly parallel correspondence-theoretic analysis (Pater 1995, Fukazawa et al. 1998, Ito and Mester 1999). It makes sense that the lexical/postlexical phonology divide is of a very different nature from such vocabulary strata distinctions: It is based not on the series of contacts of a language with other languages (which is from a linguistic point of view accidental), but on substantive and universal facts relating both to domains (words vs. phrases) and to the role of gradient phonetic factors.¹³

We will now give a concrete illustration of this analysis by returning to the example [dʊϕç] 'through' familiar from above (see (20)-(22)). The structure-preserving effect of (23a) is that [ϕ] is banned from lexical outputs: only [R] is found, in any position. This means, in particular, that the lexical output is [dʊRç], not *[dʊϕx] or *[dʊϕç]. Completing the overall picture, (25) and (26) show how the two successive constraint systems of lexical and postlexical phonology result in the overall opaque mapping /dʊRç, dʊϕç, dʊRx, dʊϕx/ → [dʊRç] → [dʊϕç]. Laying the

¹³ See Bemudez-Otero 1999 for further discussion and motivation.

groundwork for later argumentation, (25) highlights the fact that this result holds regardless of the way in which the input is specified for allophonically varying features, i.e., it stands in full agreement with the Richness-of-the-Base Hypothesis.¹⁴

(25) Lexical:

a. /dʊRç/	*ɸ	*Coda/R	*R	Id(cons)	VEL	*x	*ç	Id(back)
b. /dʊRx/								
c. /dʊɸç/								
d. /dʊɸx/								
dʊRç		*	*	* _{c,d}			*	* _{b,d}
dʊRx		*	*	* _{c,d}		*!		* _{a,c}
dʊɸç	*!			* _{a,b}	*!		*	* _{b,d}
dʊɸx	*!			* _{a,b}		*!		* _{a,c}

(26) Postlexical:

/dʊRç/	*Coda/R	*ɸ	*R	Id(cons)	Id(back)	VEL	*x	*ç
dʊRç	*!		*					*
dʊRx	*!		*		*!		*	
dʊɸç		*		*		*		*
dʊɸx		*		*	*!		*	

While (25) shows the typical rich input structure, with all conceivable variants present for non-contrastive properties, lexical outputs are "phonemic", as far as [R~ɸ] is concerned: [dʊRç] is the winner, and no [ɸ] appears in any output, even if it is posited in the input. On the other hand, lexical outputs are by no means strictly phonemic, as shown by the [x~ç] relation, where both variants are already found in their appropriate places.

The postlexical phonology derives the result in a straightforward way, as shown in (26), where [dʊɸç] emerges as the winner: R-Vocalization is postlexically in force because of *CODA/R » *ɸ » IDENT(CONS), whereas postlexical outputs are conservative with regard to the distribution of [x~ç] because of IDENT(BACK) » VEL.¹⁵

In terms of the overall theory, we have arrived at a picture in which some crucial insights of Lexical Phonology have found a place in an OT model that makes a systematic and serial distinction between a lexical and a postlexical module of phonology. The two levels show different constraint rankings not only in terms of M-F relations, but also in terms of M-M relations. Concerning the latter, postlexical 'processes' are 'turned on' by promoting contextual

¹⁴ Faithfulness violation marks restricted to specific input variants are indexed in the appropriate way.

¹⁵ It should be clear that the stratal separation between lexical and postlexical phonology seen in (25) and (26) cannot be mimicked by means of OO-constraints, as in the work of Benua 1997 and others, since there is no *phonetic* output with coda-R that could serve as a reference point. This just echoes the point, made by McCarthy 1998 and others, that opacity cannot be reduced to OO-effects.

markedness over context-free markedness, undoubtedly reflecting the strongly phonetic orientation of the phrasal system. In other cases, postlexical processes might be 'turned on' by demoting relevant faithfulness constraints. Lexical outputs are characterized by 'structure preservation' (limitation to a restricted inventory of elements and structures), which is however not an all-or-nothing principle, since in each individual case everything depends on the ranking relations of the constraints involved.

5. The failure of Sympathy

Having demonstrated that weakly parallel OT provides a general and rather straightforward way to deal with opaque interactions with allophonic masking processes, it is time to turn to Sympathy Theory (McCarthy 1998), one of the most interesting proposals put forth in recent OT-work in the area of opacity. Can strictly parallelist OT, if enriched by Sympathy, come to terms with the problem posed by allophonic masking processes? Our ultimate conclusion will be a negative one, and we will build our case in a step-by-step fashion. After discussing background issues concerning the general coverage of Sympathy Theory in section 5.1, we return to the German opacity problem involving allophonic alternations in section 5.2.

5.1 Questions of coverage

Modifying the relevant portions of Kiparsky's (1973, 79) rule-based definition and building on McCarthy's recent work (see McCarthy 1998), opaque patterns in outputs can be characterized in OT as in (27).

(27) Let P be a phonological process induced by

- M: * A/C_D a markedness constraint against A in the context C_D.
- F: *(A_i,B_o) a faithfulness constraint against in-element A_i corresponding to out-element B_o.
- M » F *A/C_D dominates *(A_i,B_o).

P is *opaque* (opposite: *transparent*) to the extent that there are outputs of the forms (i) CAD ("counterfeeding", see (28a)) or (ii) CBG, with B derived from A and A→B due to P ("counterbleeding", see (28b)).

A traditionally minded classification of opacity cases in ordered rule theory, distilled from Kiparsky 1973 and earlier work, appears in (28). In this collection, as explained in McCarthy 1998, the category "counterfeeding" (28a) splits into two subcases, "environment counterfeeding" and "focus counterfeeding". On the other hand, for counterbleeding opacity (28b) this distinction is not made since "focus counterbleeding" (28iv) is not part of the standard collection (we return to this case below).

(28)	in the environment		in the focus	
a. counterfeeding	(i) $A \rightarrow B / C_D$ $E \rightarrow D / _ \#$	/CAE/ --D [CAD]	(iii) $A \rightarrow B / C_D$ $F \rightarrow A / C_D$	/CFD/ -- ...A... [CAD]
b. counterbleeding	(ii) $A \rightarrow B / C_D$ $D \rightarrow G / _ \#$	/CAD/ ...B...G [CBG]	(iv) / C [+x,+y,+z] D / ([+x,+y] → [-z] / C_D ... [..... -z] ... [+x] → [-y] / C_D ... [..... -y...] ... [C [+x-y-z] D])	

The counterfeeding interaction in the environment (28i) is a case of underapplication. The rule $[A \rightarrow B / C_D]$ does not apply—in OT terms, the $M \gg F$ constraint interaction $[*A/C_D \gg *(A_i, B_o)]$ is not surface-true. Sympathy Theory deals with this variety by providing a means to make reference to the (losing) hyperfaithful candidate CAE (in this case, though not in general, identical to the input), in which the crucial environmental change $E \rightarrow D$ has not taken place. Sympathetic faithfulness then transfers this candidate's "A" (but not its "E") to the output. (See McCarthy 1998 for discussion and details of the technical execution).

The diagnosis for environment counterbleeding (28ii) is "overapplication"—in OT terms, the environment of $M: *A/C_D$ is not "surface-apparent". The process is opaque since there are outputs of the form CBG, with out-element B corresponding to in-element A and with the change $A \rightarrow B$ due to $M: *A/C_D$, even though the output contains no D. Candidate CBG, the actual output, has an apparently gratuitous violation of $*(A_i, B_o)$, which remains fatal in comparison with CAG wherever $*(A_i, B_o)$ is ranked. Similar to the environment counterfeeding interaction, Sympathy Theory singles out the candidate CBD as the Sympathy candidate, and sympathetic faithfulness forces this candidate's B onto the overall winner CBG.

In processual parlance, the focus of a rule can be involved in opacity in two ways. The first case is "focus counterfeeding" (28iii), leading to underapplication. This is the well-known chain-shift scenario: The individual shifts $F \rightarrow A$ and $A \rightarrow B$ occur, but they do not compose to take F all the way to B. From the perspective of Sympathy's claim to be a general model of opacity in OT, focus counterfeeding reveals a first crack in the theory since OT already has independent means to deal with it—Kirchner 1996 argues persuasively that focus counterfeeding is a constraint-conjunctive faithfulness effect preventing the two-step change $F \rightarrow A \rightarrow B$.¹⁶ The markedness constraint $*A/C_D$ is sandwiched between lower-ranking simple featural faithfulness (say, $FAITH(FEATURE X)$ and $FAITH(FEATURE Y)$) and a higher-ranking conjoined faithfulness (i.e., $FAITH(FEATURE X) \&_{IS} FAITH(FEATURE Y)$), ruling out the tautosegmental violation of both featural faithfulness constraints (see Kirchner 1996 for full exemplification). Sympathy is thus

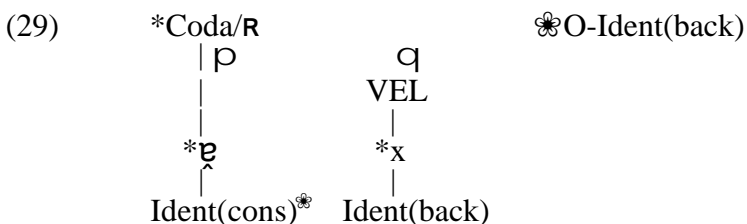
¹⁶ An alternative conception, limited to distinctive feature oppositions, has been developed by Gnanadesikan 1997. We concentrate here on the constraint-conjunctive approach because it is more general and connects to many other issues in the theory. The general issues raised by it (e.g., how to define locality for violation clustering, how to impose limits on constraint combinatorics) are not specific to focus counterfeeding, but concern important questions that have to be resolved by optimality-theoretic phonology in any case, irrespective of specific proposals how to construct complex constraints out of simple ones.

superfluous for this case of opacity.¹⁷

The three cases seen so far are presented and discussed in McCarthy 1998.¹⁸ The last category—"focus counterbleeding" (28iv)—is not part of the traditional opacity setup, neither in Kiparsky's 1973 definition nor in McCarthy's 1998 OT-adaptation of it. However, the fact that the schematic derivational scenario given in (28iv) does not qualify as opaque in the usual sense should not be the criteria for phonological opacity. Rather it suggests that mechanically transferring the notion "opacity" from a derivationalist setting to OT has its limitations.¹⁹ There are indeed cases of precisely this variety that create problems in OT in the sense that they are systematically missed by transparent constraint interactions. An actual example of focus counterbleeding will be discussed below in section 6.

5.2 Allophonic masking processes and Richness-of-the-Base

Having set the general stage for a Sympathy analysis, we now turn to the German opacity problem at hand. An appeal to Sympathy seems at first glance straightforward. Since this is a case of counterfeeding opacity, IO-IDENT(CONS) as selector singles out the faithful candidate [dʊRç] ('undoing R-vocalization', figuratively speaking), and the sympathetic faithfulness constraint *O-IDENT(BACK) transfers the [+back] property of the fricative from the *O-candidate to the overall winner. The overall ranking accomplishing this is (29), where *O-IDENT(BACK) and *CODA/R both dominate VEL.



Note the additional ranking [$*\text{CODA/R} \gg \text{VEL}$] that Sympathy needs to recruit in order to settle on [dʊɸç] (with opaque [ç]) as the winner instead of *[dʊRç] (with opaque [R]). The appearance of such additional rankings necessitated by Sympathy but without prior and

¹⁷ This is acknowledged in McCarthy (1998: 32-33), who regards the chain-shift case as irrelevant for Sympathy. At the descriptive level, this is of course true—OT already has ways of dealing with it. It is less clear when one takes Sympathy Theory serious as a general and comprehensive theory of opacity in OT. Furthermore, the availability of a constraint-conjunctive analysis does not automatically make a Sympathy approach unavailable. On the contrary, it is a straightforward matter to make the conjoined constraint the selector, and the Sympathy candidate will self-select as the overall winner under the pressure of an appropriate Sympathy-constraint. This is troubling because Sympathy here appears superimposed on a constraint-conjunctive analysis which already deals with opacity *without introducing any opacity-specific machinery*. The latter point, of central importance for the level of explanation achieved, is here highlighted by the fact that the Sympathy analysis literally contains the constraint-conjunctive analysis as a proper subpart. The case raises the suspicion that perhaps all cases of opacity are actually due not to opacity-specific machinery such as Sympathy, but to independently existing parts of OT theory.

¹⁸ Space limitations prevent us from exemplifying these cases with tableaux. The reader is referred to McCarthy 1998 for a more detailed exposition of the 'environment counterfeeding' and 'environment counterbleeding' cases.

¹⁹ This echoes a point made by Kager (1997: 497) in earlier work: "[...] one cannot, and should not, directly transfer the conceptual categories of rule-based theory into OT."

independent motivation is a general feature of Sympathy analyses (see Ito and Mester to appear). Tableau (30) illustrates this analysis.

(30)

/dʊRç/ 'through'		*Coda/R		*ɸ		Id(cns) [*]	
	⊗O-Id(bk)		VEL		*x		Id(bk)
⊗ dʊRç		*!					
☹ dʊɸç			*	*		*	
dʊRx	*!	*			*		*
dʊɸx	*!			*	*	*	*

The Sympathy analysis in (29) thus appears to work, but appearances are deceptive. The analysis is flawed by a fatal dependency on non-contrastive properties of inputs. Sympathy is reliant on the fact that the input contains /R/ and not [ɸ], which ceases to be harmless once it is recalled that both [R~ɸ] and [ç~x] are allophonic relations. This is precisely where Richness-of-the-Base demands that input specifications be free. In particular, each one of the four input options /...Rç.../, /...ɸç.../, /...Rx.../, and /...ɸx.../ should lead to the same optimal output. But this is not the case, given Sympathy's reliance on input-R: It derives phonotactically correct [...ɸç...] outputs for R-inputs /...Rç.../ and /...Rx.../ (30, 31); but /ɸ/-inputs /...ɸç.../ and /...ɸx.../ result in phonotactically impossible [...ɸx...] outputs (32a,b).

(31)

/dʊRx/	⊗O-Id (bk)	*Coda/R	VEL	*ɸ	*x	IO-Id(cns) [*]	IO-Id(bk)
⊗ dʊRç		*!					*
☹ dʊɸç			*	*		*	*
dʊRx	*!	*			*		
dʊɸx	*!			*	*	*	

(32)

a. /dʊɸç/	⊗O-Id (bk)	*Coda/R	VEL	*ɸ	*x	IO-Id(cns) [*]	IO-Id(bk)
dʊRç	*!	*				*	
☹ dʊɸç	*!		*	*			
dʊRx		*!			*	*	*
☹!!⊗dʊɸx				*	*		*

b. /dʊɸx/	⊗O-Id(bk)	*Coda/R	VEL	*ɸ	*x	IO-Id(cns) [*]	IO-Id(bk)
dʊRç	*!	*				*	*
☹ dʊɸç	*!		*	*			*
dʊRx		*!			*	*	
☹!!⊗dʊɸx				*	*		

The reason for the failure of Sympathy in (32) is easy to see: For inputs containing / \mathfrak{R} /, it is impossible to recover some underlying / \mathfrak{R} / through IO-Faithfulness in the process of \mathfrak{S} -candidate selection: No such / \mathfrak{R} / is available. Looking beyond this specific case, whose details are irrelevant for the argument, we obtain the general conclusion stated in (33).

(33) **The argument from allophony:**

Consider an alternation P counterfered by another alternation Q:

P: A → B / C__D

Q: E → D / ...

No Sympathy analysis is possible for interactions of this kind whenever

- (i) Q is allophonic and
- (ii) crucially affects P's environment.²⁰

In (33), the sympathetic way of ensuring opaque A in C_D (over transparent B) is via an IO-faithfulness-based reference to C_E—but given that the relative distribution of the allophones D and E is free in inputs (Richness-of-the-Base), no probing via IO-faithfulness secures this reference. This conclusion stands as long as two conditions hold: (i) Richness-of-the-Base remains a fundamental principle of OT, admitting rich inputs in cases of allophony,²¹ and (ii) \mathfrak{S} -selectors are restricted, as proposed by McCarthy 1998, to IO-Faithfulness constraints.²²

A significant number of opaque interactions fall into the category dealt with in this paper, where a crucial factor in the environment involves an allophonically varying property and is therefore not accessible to Sympathy's faithfulness-guided probing of inputs. Another example

²⁰ We saw earlier that when Q affects P's focus, we are dealing with the chain-shift scenario, where Sympathy is superfluous in any theory with a rich enough faithfulness component (allowing, for example, local conjunctions of markedness constraints). The allophonic argument against Sympathy also subverts any idea that rich faithfulness alone could deal with all cases of focus counterfeeding: The opacity of any alternation P counterfered by another alternation Q, where Q is allophonic and crucially affects P's focus, is in principle beyond the reach of faithfulness. As an illustration, consider the following:

a.
$$A \rightarrow \left\{ \begin{array}{l} B_1 / _ _ D_1 \\ B_2 / _ _ D_2 \\ B_3 / _ _ D_3 \end{array} \right\}$$

b.
$$\left\{ \begin{array}{l} B_3 \rightarrow B_4 \\ B_2 \rightarrow B_3 \\ B_1 \rightarrow B_2 \end{array} \right\} / C_ _$$

In such cases, /A/ is the classical input for any B_{i+1} in an output of the form $[CB_{i+1}D_i]$, with any B_n serving as an alternative input. This subverts the faithfulness-based account of chain shifts, which is dependent on the input being fixed as B_i . While this is sufficient to make the point as a *Gedankenexperiment*, empirical examples are hard to come by, given the tight specifications that any actual case would have to meet. Cases like these, if they exist, could arise out of stratal scenarios, with (a) taking place in the lexical phonology and (b) in the postlexical phonology. In that case, the faithfulness-based account of the chain shift (b) would work postlexically, since the postlexical input, identical to the lexical output, would contain the required B_i -elements.

²¹ Note that the Richness-of-the-Base is the desirable null hypothesis — all special legislation on inputs remains inferior and stipulative in comparison.

²² Ito and Mester 1997b argue that for the analysis of German hypocoristics the \mathfrak{S} -selector must be an alignment constraint. Bermúdez-Otero 1999 provides further empirical evidence that the selection of \mathfrak{S} -candidates cannot be confined to the class of IO-faithfulness constraints.

from our own work (see the Sympathy analysis given in Ito and Mester 1997a) involves the opaque interaction of two processes in Japanese: (i) the well-known Rendaku Voicing in compounds (*tome* ‘clip’, *kami+dome* ‘hair clip’), blocked in stems containing voicing obstruents (*taba* ‘bundle’, *hana+taba* ‘flower bouquet’, **hana+daba*), and (ii) Velar Nasalization, an allophonic process replacing word-medial /g/ by [ŋ] (/kagi/→*kaŋi* ‘key’, vs. word-initial *geta* ‘clogs’). In processual terms, Rendaku voicing feeds Velar Nasalization in the focus of the rule (i.e., *k*→*g*→*ŋ*: *ori+kami* → *ori+gami* → *ori+ŋami*), and is itself counterfed by Velar Nasalization when the velar nasal replacing /g/ appears in the environment (*saka+toge* → *saka+toŋe* ‘reverse thorn’). It is sufficient to note in the present context that a Sympathy analysis will stumble over the fact that only a reference to /g/ can prevent voicing of /t/ to [d] in [sakatoŋe], *[sakadoŋe], but this reference cannot be relied on, since [g~ŋ] are allophonically distributed. The case is thus entirely parallel to the German case analyzed in detail in this paper (see Ito and Mester 2000 for a full analysis of the Japanese case).

It stands to reason, and we will see below, that this result is not restricted to counterfeeding and actually holds in greater generality, affecting counterbleeding interactions as well. It is important here to look beyond the technical details in order to get to the heart of the problem, viz., Sympathy Theory’s central tenet that opacity arises through *hyper-faithfulness* to the input, which is implemented by having the Sympathy candidate selected by a faithfulness constraint F. This has non-vacuous effects only when the selector is a dominated (and hence violated) faithfulness constraint, so that the Sympathy candidate is F-wise more faithful to the input than the transparent winner (it cannot be less faithful, and if it is equally faithful, it coincides with the transparent winner). What the argument from allophony shows is that hyper-faithfulness in this sense cannot be the solution to the problem of opacity within OT. In an ironic twist, the very limitation of selectors to faithfulness that was intended to make Sympathy Theory restrictive turns out to have curtailed its ineffectiveness in a crucial way.

The flaw in Sympathy Theory that reveals itself in (33) transcends the area of opacity and affects the basic predictions that a grammar makes regarding contrasts and non-contrasts, one of its most fundamental tasks under any conception. Sympathy, via the special access provided by selector status of low-ranking IO-faithfulness constraints, has the power to ‘tunnel through’ to irrelevant and arbitrary aspects of inputs. In this way, it subverts the noncontrastive status of allophonic variants such as [ç~x] or [R~ʁ], in effect turning non-contrasts into contrasts. To illustrate, consider the two potential inputs /dʊRç/ and /dʊʁç/, non-contrastive by hypothesis. The Sympathy-induced mappings /dʊRç/ → [dʊʁç] and /dʊʁç/ → [dʊʁx] show that an OT-grammar equipped with Sympathy is able to turn these into *two different outputs*. Sympathy has thus introduced a contrast between allophones (here, [ç] and [x]) through the backdoor, against all intentions.

6. Doubly opaque interactions

An analytically still more decisive argument emerges in situations where the correct opaque winner apparently cannot be selected by Sympathy, even granting some stipulation on inputs to the effect that they must contain /R/ (and not /ʁ/) for output forms with coda-[ʁ]. If this turns out to be true, it shows that Sympathy cannot be saved even by sacrificing Richness-of-the-Base.

The relevant cases are the Northern German *g*-spirantization examples in (34), which result in mappings such as /bʊRg/ → [bʊʁç] ‘castle’. Here, we are confronted with another source of distributionally opaque [ç]: processually speaking, underlying /...Rg/ has turned into [...ʁç] in a series of steps: $Rg]_{\sigma} \rightarrow R\gamma]_{\sigma} \rightarrow Rx]_{\sigma} \rightarrow Rç]_{\sigma} \rightarrow ʁç]_{\sigma}$.

(34)	Hamburg	[hɑmbʊɤ̥ç]	(name of city); cf. hambu[ɤ̥gɪ] (adj.)
	Berg	[bɛɤ̥ç]	'mountain'; cf. Ber[gə] (pl.)
	Chirurg	[çɪɾʊɤ̥ç]	'surgeon'; cf. Chiru[ɤ̥gən] (pl.)
	birg	[bɪɤ̥ç]	'rescue', imp.
	borg	[bɔ̥ç]	'borrow', imp.
	Georg	[ge :ɔ̥ç]	(name)

The background here is a general process of *g*-spirantization in syllable codas (analyzed in detail in Ito and Mester to appear, where it is shown to present a different kind of opacity-problem for Sympathy) resulting, in conjunction with coda devoicing and the constraint configuration in (21) above, in [x] after back vocoids, and [ç] otherwise. Underlying /g/ in these cases is supported by ubiquitous alternations such as [bɛɤ̥ç]~[bɛɤ̥gə] 'mountain, sg/pl'. In rule terms, this case involves two opaque relations:²³

(35)			/bʊɾg/ 'castle'
	g-spirantization:	g → ɤ̥ / ___] _σɤ̥
	coda devoicing:	[-son] → [-voi] / ___] _σx
	dorsal fricative allophony:	[+dors,-son,+cnt,-voi] ^{↗ [-bk]} _{↘ [+bk]} / [-cns,+bk] ___ç
	R-vocalization:	R → ɤ̥ / ___] _σɤ̥...
			[dʊɤ̥ç]

(35) is doubly opaque: On the one hand, coda devoicing counterbleeds *g*-spirantization in the focus (this is an instance of the case missing from the classical collection of opacity cases discussed earlier, see (28)). On the other hand, *R*-vocalization simultaneously counterfeeds the second branch of dorsal fricative allophony and counterbleeds its first branch in the environment.

Let us consider a Sympathy approach to this deeply opaque interaction. The two relevant markedness constraints are ranked intrinsically due to their conjunctive nature as in (36).²⁴

(36)	*Coda/g	=[*Coda& _ℓ *VoiObs& _ℓ *Dorsal]:	No voiced dorsal plosives in codas
	*Coda/VoiObs	=[*Coda& _ℓ *VoiObs]:	No voiced obstruents in codas

²³ Note that over and above the fact that [ç] represents the basic variant (see (13)), there must be a rule turning (derived) [x] into [ç] everywhere except after back vowels.

²⁴ See Ito and Mester to appear for further discussion and motivation for these constraints. Lombardi, this volume, presents interesting typological arguments against positing such coda conditions on voicing. We have nothing to add here but merely note that we expect the eventual resolution of these problems to depend on further developments in the theory of feature faithfulness constraints. Consider, e.g., a theory where segment deletion/insertion always triggers Max/Dep-Feature violations, in addition to Max/Dep-Segment: In such a theory, segment-deleting/inserting candidates would in general be harmonically bounded by feature-changing candidates.

These two markedness constraints are accompanied by the usual entourage of appropriate selector-, sympathetic-, and ordinary faithfulness constraints in the tableau in (37) (from Ito and Mester to appear), which represents an attempt to capture opaque g-spirantization by means of Sympathy.

(37)

/ho:nI _g / 'honey'	*Coda/g	*Coda/ VoiObs	IO-Id(voi)*	✱O- Id(cont)	IO-Id(cont)
[ho:nI _g]	*!	*		*	
✱ [ho:nI _ɣ]		*!			*
[ho:nI _k]			*	*!	
✱ [ho:nI _ç]			*		*!

The central idea is that [ho:nI_ɣ] is recruited as the Sympathy candidate (through IO-IDENT(VOI) as selector), and even though it cannot win in the overall competition (because of *CODA/VOI OBS), it transfers the fricative nature of its closing consonant to the winning output through ✱O-IDENT(CONT), where its place of articulation is adjusted (by the constraint configuration (21) responsible for [ç~x] allophony).

Fortified by the study of this sympathetic interaction, we are ready to turn to /bʊR_g/ → [bʊʁ_ç] with its doubly opaque [ç]: as a spirant opaquely derived from /g/, and as a palatal distributionally opaque in view of the preceding [ʁ]. In trying to account for the palatality of the opaque out-element [ç] (as opposed to a transparent velar [x]), a faithfulness-based link to the input correspondent, velar [g], holds little promise. Recall that the input is fixed as /g/ since this is a case of neutralization, with [g] and [ç~x] standing in contrast (*tauchen* [tʌʊxən] 'dive' vs. *taugen* [tʌʊgən] 'have value', *Gleichung* [glɛiçʊŋ] 'equation' vs. *Neigung* [nɛi_gʊŋ] 'inclination', etc.)

In (38), we make the Sympathy-favorable assumption that the input is fixed as the R-ful /bʊR_g/ (as against /bʊʁ_g/), i.e., for the sake of the argument Richness-of-the-Base is suspended as a principle of grammar. The basic setup in (29) for opaque [ç] after [ʁ] is of course not yet equipped to derive the correct result with [...ç], but rather produces a winner with transparent [...k] (see (38) and the reference cited above for background and details).

(38) /bʊRg/	⊗O-Id (bk)	*Coda/ VoiOb s	*Coda/R	VEL	*ɸ	*x	IO- Id(cns) [⊗]	IO- Id(bk)
bʊRg		*	*					
⊗ bʊRk			*					
bʊRɣ		*	*					
bʊRx			*			*		
bʊRç	*		*					*
bʊɸg		*			*		*	
⊗ !!bʊɸk					*		*	
bʊɸɣ		*			*		*	
bʊɸx					*	*	*	
⊗ bʊɸç	*			*	*		*	*

The sympathetic constraint configuration in (37) needs to be integrated with the one in (38), resulting in the double-Sympathy system in (39), where ⊗ and ⊖ refer to different Sympathy relations (⊗ is responsible for opaque [ç] after [ɸ], ⊖ for spirantization of /g/ in codas).

(39)	⊗O-Id(bk)	*Coda/R	*Coda/g
	ρ	!	!
	q	VEL	*Coda/VoiObs
		!	! ρ
	*ɸ	*x	⊖O-Id(cont) IO-Id(voi) [⊖]
	!	!	!
	IO-Id(cns) [⊗]	IO-Id(bk)	IO-Id(cont)

Tableau (40), besides giving a feel for the complexity of this analysis, reveals a surprising result: The analysis is still not able to select the desired winner [bʊɸç] (designated by "⊗"), but continues to produce *[bʊɸk].

*O-IdCodaVLE *ϕ *x IO-Id(O-IdCodaCodaVLE *ϕ)IO-Id(cnt)

(40) /bʊRg/

bʊRg		*						*	*	*
* bʊRk		*							*	*
bʊRɣ		*						*		*
bʊRx		*			*			*		*
bʊRç	*	*				*		*		*
bʊϕg				*		*		*	*	
⚡!!bʊϕk				*		*		*	*	
⚡ bʊϕɣ				*		*		*		*
⚡!!bʊϕx				*	*	*		*		*
☹ bʊϕç	*		*	*		*	*	*		*

Depending on whether the leftmost block of constraints (aiming to ensure [ç] after [ϕ] (from /R/) or the rightmost block of constraints (targeted towards coda [ç~x] from /g/), ranks higher, either [bʊϕk] (designated by ⚡!!) or [bʊϕx] (designated by ⚡!!) is selected as the winner.

As far as we can determine, the result in (40) reveals a fundamental inability of Sympathy to properly deal with such double-opacity cases. The R-ful *-Sympathy candidate [bʊRk] shows a final stop [k]; different from the earlier cases with underlying /ç~x/, its R is thus of no help in bringing about final palatality in the output through sympathetic faithfulness. On the other hand, the ϕ-ful ⚡-Sympathy candidate [bʊϕɣ], while showing a final fricative, knows nothing about suppressing R-vocalization and thus cannot contribute to final palatality either.

As the preceding remarks suggest, the malfunction appears to be due to the fact that the two prongs of double-Sympathy do not communicate effectively, making it unsuitable as a model of double-opacity situations. Leaving it to practitioners of Sympathy to explore further analytical options that might overcome this problem,²⁵ we draw two preliminary conclusions:

- (i) Incompleteness: Sympathy fails as a model for opacity cases where the environmental factor is involved in an allophonic alternation. The basic idea of Sympathy is here incompatible with Richness-of-the-Base.
- (ii) Double opacity: Even if Richness-of-the-Base is sacrificed in favor of Sympathy in order to

²⁵ It is unclear whether the revised "cumulativity"-based approach to Sympathy in McCarthy 1999 (intended to address "Duke-of-York"-related issues raised by Kiparsky 1998) changes matters in a decisive way. Our attempts to come up with a cumulativity analysis of multiple-opacity Yawelmani (from McCarthy 1998) using any of the versions of cumulativity presented in McCarthy 1999 have so far not been successful. As for the Richness-of-the-Base/allophony argument against Sympathy presented earlier, it appears to persist under cumulative Sympathy since the mode of Sympathy candidate selection remains unchanged.

respond to (i), Sympathy Theory still encounters serious problems in double opacity situations.

Concluding this section, it behooves us to show how the problematic double opacity case /bʊʀg/ → [bʊʁç] falls into place in stratal OT. The essential property of this example is that it combines the stratal opacity of [ç] in post-[ʁ] contexts seen earlier with a different kind of opacity, namely, the parallelist kind of opacity instantiated by German g-spirantization. Parallel opacity is created through local constraint conjunctions of the forms M&_ℓF and F&_ℓF. As shown elsewhere (see Ito and Mester to appear for a detailed study of such M&_ℓF-cases), M&_ℓF conjunctions lead to opacity in the configuration M&_ℓF » C » M (where C is antagonistic to M): Here M, in conjunction with F, appears to apply only to derived elements (F-violators), not to underived elements. F&_ℓF conjunctions lead to opacity in the configuration F1&_ℓF2 » M » F1, F2 (where M is antagonist to F1 and F2). Here M applies only to underived elements, not to derived elements, i.e., elements that have changed in violation of F1 cannot go on to also violate F2, and vice versa. This results in the familiar focus-counterfeeding chain shift scenario explored by Kirchner 1996 and others.

Given what we have seen earlier, it is clear that the allophonic type of opaque interactions exemplified by the opaque [ç] after the [+back] vocoid [ʁ] in German cannot be reduced to the parallelist variety. In both M&_ℓF and F&_ℓF conjunctions, the crucial ingredient leading to opacity is "F" (IO-faithfulness), i.e., a reference to the input. However, an allophonic masking process in an opaque interaction cannot be circumnavigated through reference to the input, here, to the underlying [ʀ] in a mapping like /dʊʀç/ → [dʊʁç]. This forestalls any constraint-conjunctive account involving faithfulness, just as it forestalls any Sympathy account.²⁶

Without further discussion, we take from Ito and Mester to appear the analysis of g-spirantization involving the M&_ℓF conjunction *CODA/DORSPLoS&_ℓIDENT(VOI) (ruling out dorsal plosives in codas that simultaneously violate voicing faithfulness, see the work cited for ranking justification and further discussion). The violation is avoided by means of spirantization (violation of lower-ranking IDENT(CONT)).

$$\begin{array}{ll}
 (41) & *Coda/VoiObs & = [*VoiObs\&_{\ell} *Cod] \\
 & ! & \\
 & [*Coda/DorsPlos\&_{\ell} Ident(voi)] & = [*Cod\&_{\ell} *DorsPlos\&_{\ell} Ident(voi)] \\
 & ! & \\
 & Ident(cont) & \\
 & ! & \\
 & *Coda/DorsPlos & = [*Cod\&_{\ell} *DorsPlos] \\
 & ! & \\
 & Ident(voi) &
 \end{array}$$

The workings of this analysis are illustrated in tableaux (42) and (43). (42) shows how input

²⁶ One could try to take the bull by its horns and suggest an M&_ℓM account (this amounts to denying that we are dealing with a case of opacity): Thus Ito and Mester 1998 pursue the brute-force strategy of positing *ʁ&_ℓx (with the locality parameter set as adjacency) as the operative constraint conjunction. Besides its ad-hoc nature, the attempt remains unconvincing because it has trouble with basic questions of distribution, which lead to the piling on of further and further restrictions and qualifications (thus the other order [xʁ] is fully wellformed, contractions as in /ʃtʊʀç/ → [ʃtʊ:ç], *[ʃtʊ:x] complicate the picture, etc.).

/g/ in the coda changes into output [ç] (because [k] violates *CODA/DORSPLoS&_lID(VOI)), whereas (43) shows how input /k/ survives unchanged in the coda.

(42)

/ho:nIg/	*Coda/ VoiObs	[*Coda/DorsPlos & _l Id(voi)]	Id(cnt)	*Coda/ DorsPlos	Id(voi)
[ho:nIg]	*!			*	
[ho:nIɣ]	*!		*		
[ho:nIk]		*!		*	*
☞ [ho:nIç]			*		*

(43)

/plastIk/	*Coda/ VoiObs	[*Coda/DorsPlos & _l Id(voi)]	Id(cnt)	*Coda/ DorsPlos	Id(voi)
[plastIg]	*!			*	*
[plastIɣ]	*!		*		*
☞ [plastIk]				*	
[plastIç]			*!		

In (44), all three parts of the analysis are combined—the two subhierarchies of (24), with lexical/postlexical ranking differences resulting in serially opaque [ç] after [ɸ], and the spirantization hierarchy in (41) with its M&F constraint resulting in parallel opacity ([ç] instead of [k] for input-/g/, but [k] for input-/k/).

(44)

a - >					< - - - - -
!	!				!
ϕ !	*ɸ	*Coda/VoiObs		VEL	!
!	!	!		!	!
Z - -	*Coda/R	*Coda/DorsPlos& _l Ident(voi)		*x	!
	!	!	ρ	!	!
	*R	Ident(cnt)		*ç	!
	!	!		!	!
	Ident(cons)	*Coda/DorsPlos		Ident(bk)	- - m
				!	!
				Ident(voi)	

Tableaux (45) and (46) (where "....." indicates irrelevant constraints that have been omitted) demonstrate that this analysis has no trouble deriving the total mapping /bʊRg/ → [bʊɸç] (as well as the mapping /bʊɸg/ → [bʊɸç], with a different input variant). The fact that the overall winner is [bʊɸç] and not [bʊɸk] follows from the ranking *CODA/DORSPLoS&_lID(VOI) » *ç which connects two of the subhierarchies, as indicated in (44).

(45) Lexical:

/bURg/ /bUR̥g/	*Coda/ VoiObs	*Coda/ DorsPlos & Id(voi)	Id (cnt)	...	*ɸ	*Coda/ R	...	VEL	*x	*ç	Id (bk)
bURg	*!			...		*	...				
bURk		*!		...		*	...				
bURY	*!		*	...		*	...				
bURx			*	...		*	...		*!		
☞ bURç			*	...		*	...			*	*
bUR̥g	*!			...	*!		...				
bUR̥k		*!		...	*!		...				
bUR̥Y	*!		*	...	*!		...				
bUR̥x			*	...	*!		...		*!		
bUR̥ç			*	...	*!		...	*!		*	*

(46) Postlexical:

/bURç/	*Coda/ VoiObs	*Coda/ DorsPlos & Id(voi)	Id (cnt)	...	*Coda/ R	*ɸ	...	Id (bk)	VEL	*x	*ç
bURg	*!	*!	*	...	*!		...	*!			
bURk			*	...	*!		...	*!			
bURY	*!			...	*!		...	*!			
bURx				...	*!		...	*!		*	
bURç				...	*!		...				*
bUR̥g	*!	*!	*	...		*	...	*!			
bUR̥k			*	...		*	...	*!			
bUR̥Y	*!			...		*	...	*!			
bUR̥x				...		*	...	*!		*	
☞ bUR̥ç				...		*	...		*		*

7. Summary and further consequences

We have argued in this paper that Sympathy, as proposed by McCarthy (1998), in spite of its initial promise, falls short of the goal of serving as a general model of opaque interactions within OT. While Sympathy's coverage of the various cases of opacity seems at first glance to be impressively complete, closer investigation has shown that this is far from being the case. In this paper we have demonstrated that there are important classes of environment opacity cases ((28i), (28ii)) which receive no adequate analysis by Sympathy. The main argument comes from a

whole class of interactions which are bona fide cases of opacity but which cannot be understood in terms of Sympathy at all. They involve alternations whose opacity is due to "environment counterfeeding", with the environmental factor itself subject to an allophonic alternation.

What follows from the existence of such non-sympathetic opacity? In its narrowest interpretation, the result means that Sympathy Theory is not a complete model of opacity in OT. Taken more broadly, it suggests that the resolution of the whole opacity issue is to be sought elsewhere: Given that Sympathy needs to be complemented by other modes of analysis in order to deal with all cases of environment counterfeeding/-bleeding (28i,iii) and is already superseded by F&_tF constraint conjunction for focus counterfeeding (28ii), the question of Sympathy's actual contribution becomes pressing.

We have seen in this paper that a certain type of opacity encountered in natural languages, while intractable by the strictly parallelist means of sympathetically enriched OT, finds a straightforward account in a theory that makes a systematic serial distinction between a lexical and a postlexical module of phonology. In a larger perspective, this weakly parallel model of OT incorporates in a natural way some key insights of classical Lexical Phonology connected to the ideas of 'turning off' of processes and rules, and structure preservation. Lexical outputs are characterized by "structure preservation" (limitation to a restricted inventory of elements and structures), which depends on the ranking relations of the relevant constraints involved. In the postlexical module, promoting contextual markedness over context-free markedness will result in the process being 'turned on'. As we have seen, this gives rise to a stratal type of opacity, which exists independently of a parallelist type due to complex constraints created by some kinds of constraint conjunction. In special cases, the two types of opacity interact with each other, giving rise to multiply opaque patterns in outputs.

References

- Benua, Laura. 1997. Transderivational Identity: Phonological Relations between Words. Doctoral dissertation, University of Massachusetts, Amherst. [ROA-259-0498].
- Bermúdez-Otero, Ricardo. 1999. Constraint Interaction in Language Change: Quantity in English and Germanic. Doctoral dissertation, University of Manchester.
- Bloomfield, Leonard. 1930. German ç and x. *Le maître phonétique* 3, 20. 27-28.
- Borowsky, Toni. 1986. *Topics in the Lexical Phonology of English*. Doctoral dissertation, University of Massachusetts, Amherst.
- Borowsky, Toni. 1990. On the word level. In *Studies in Lexical Phonology*, edited by S. Hargus and E. M. Kaisse. San Diego, etc. Academic Press. 199-234.
- Fukazawa, Haruka. this volume. Local Conjunction and Extending Sympathy Theory: OCP Effects in Yucatec Maya.
- Fukazawa, Haruka, Mafuyu Kitahara, and Mitsuhiro Ota. 1998. Lexical stratification and ranking invariance in constraint-based grammars. Ms. University of Maryland, College Park, Indiana University, and Georgetown University. [ROA-267-0698].
- Gnanadesikan, Amalia. 1997. Phonology with Ternary Scales. Doctoral dissertation, Department of Linguistics, University of Massachusetts at Amherst, Amherst, MA.
- Hall, Tracy Alan. 1992. *Syllable structure and syllable-related processes in German*. Tuebingen. Niemeyer.
- Hall, Tracy Alan. 1993. The phonology of German /R/. *Phonology* 6 (1). 1--18.
- Hayes, Bruce. 1999. Phonological Acquisition in Optimality Theory: The Early Stages. Ms. University of California, Los Angeles. [ROA-327-0699].
- Ito, Junko, and Armin Mester. 1995. The core-periphery structure of the lexicon and constraints on reranking. In *Papers in Optimality Theory*, edited by J. Beckman, S. Urbanczyk and L. Walsh. Amherst. GLSA. 181-210.
- Ito, Junko, & Armin Mester. 1997a. Sympathy Theory and German Truncations. In Viola Miglio & Bruce Morén, eds., *Proceedings of the Hopkins Optimality Workshop/Maryland Mayfest 1997*. University of Maryland Working Papers in Linguistics. 5. 117-138. [ROA-211-0897]
- Ito, Junko, and Armin Mester. 1997b. Featural Sympathy. In *Phonology at Santa Cruz (PASC) 5*, edited by D. Karvonen, M. Katayama and R. Walker. Santa Cruz, CA. 29-36.
- Ito, Junko, and Armin Mester. 1998. [M&F] and [M&M] conjunctions. Handout of talk given at the Tübinger Silbenkonferenz, July 1998.
- Ito, Junko, and Armin Mester. to appear. On the sources of opacity in OT: coda processes in German. In *Structure and Typology of the Syllable*, edited by C. Féry and R. v. d. Vijver. Cambridge, U.K. Cambridge University Press. [ROA-347-0999]
- Ito, Junko, and Armin Mester. 1999. The phonological lexicon. In *A Handbook of Japanese Linguistics*, edited by N. Tsujimura. Oxford. Blackwell. 62-100.
- Ito, Junko, and Armin Mester. 2000. Weak Parallelism and Modularity: Evidence from Japanese. Report of the Special Research Project on the Languages of the East and West. Tsukuba University.
- Kager, René. 1997. Rhythmic vowel deletion in Optimality Theory. In *Derivations and Constraints in Phonology*, edited by I. Roca. Oxford. Oxford University Press. 463-499.
- Kiparsky, Paul. 1973. Abstractness, opacity and global rules. In *Three Dimensions of Linguistic Theory*, edited by O. Fujimura. Tokyo. TEC. 57-86.
- Kiparsky, Paul. 1982. Lexical phonology and morphology. In *Linguistics in the Morning Calm*, edited by I. S. Yang. Seoul. Hanshin. 3-91.

- Kiparsky, Paul. 1985. Some consequences of Lexical Phonology. *Phonology* 2. 85-138.
- Kiparsky, Paul. 1998. Paradigm effects and opacity. Ms. Stanford University.
- Kirchner, Robert. 1996. Synchronic chain shifts in Optimality Theory. *Linguistic Inquiry* 27. 341-350.
- Kohler, Klaus J. 1977. *Einführung in die Phonetik des Deutschen*. Berlin. E. Schmidt.
- Krämer, Wolfgang. 1981. *Akustisch-phonetische Untersuchungen zum vokalischen /R/-Allophon des Deutschen*. Hamburg. Helmut Buske Verlag.
- Lombardi, Linda. this volume. Why Place and Voice are different: Constraint-specific alternations and Optimality Theory.
- Lubowicz, Anna. 1998. Derived environment effects in OT. Ms. University of Massachusetts, Amherst. [ROA 239-0198].
- McCarthy, John J. 1998. Sympathy & phonological opacity. Ms. University of Massachusetts, Amherst. [ROA-252-0398].
- McCarthy, John J. 1999. Sympathy, Cumulativity, and the Duke-of-York Gambit. Ms. University of Massachusetts, Amherst. [ROA-315-0499].
- Merchant, Jason. 1996. Alignment and fricative assimilation in German. *Linguistic Inquiry* 27. 709-719.
- Moulton, William G. 1962. *The Sounds of English and German*. Chicago. University of Chicago Press.
- Ní Chiosáin, Máire, and Jaye Padgett. this volume. Markedness, segment realization, and locality in spreading. Ms. UC Santa Cruz. [also LRC-97-01, and ROA-188-0497].
- Pater, Joe. 1995. On the nonuniformity of weight-to-stress and stress preservation effects in English. Ms. McGill University. [ROA-107-0000].
- Prince, Alan, and Paul Smolensky. 1993. Optimality theory: Constraint interaction in generative grammar. Ms. Rutgers University and University of Colorado, Boulder.
- Prince, Alan S., and Bruce Tesar. 1999. Learning phonotactic distributions. Ms. Rutgers University, New Brunswick. [ROA-353-1099].
- Smolensky, Paul. 1995. On the structure of the constraint component Con of UG, Handout of talk at UCLA, 4/7/95. [ROA-86-0000]
- Wedel, Andrew. in prep. An acoustic study of vocalized /R/ in German. Ms. UC Santa Cruz.
- Wiese, Richard. 1996. *The Phonology of German, The Phonology of the World's Languages*. Oxford. Clarendon Press.