RICARDO BERMÚDEZ-OTERO & RICHARD M. HOGG

THE ACTUATION PROBLEM IN OPTIMALITY THEORY

Phonologization, Rule Inversion, and Rule Loss

Abstract. This chapter outlines Optimality Theory's contribution to research into the actuation of phonological change. We examine both phonetically driven innovation and analogical change (particularly rule inversion and rule loss).

Following Ohala, we assume that the phonologization of mechanical phonetic effects is caused by parser malfunction. It is therefore suggested that, as a theory of grammar, OT will play a secondary role in accounts of phonologization. Nonetheless, OT makes a significant contribution in this area by modeling the restrictions that universal markedness principles impose upon phonological innovation. In this connection, we argue that markedness generalizations are not mere epiphenomena of performance-driven change, and we refute the claim that inverted phonological processes are synchronically arbitrary.

In the area of analogy, the Optimality-Theoretic concept of input optimization affords new insights. We observe that most types of analogical change involve the restructuring of input representations at some level in the phonology. Restructuring usually occurs when, as a result of some independent development, learners cease to encounter positive cues to abandon their default state, in which input representations are identical with the corresponding outputs. We show that, whereas OT predicts this state of affairs, rule-based theories cannot account for the facts without imposing contradictory demands on acquisition theory.

Our discussion of analogy is illustrated with a case-study of rule inversion and rule loss in the late West Saxon dialect of Old English. The analysis is couched in the framework of interleaved OT. It is shown that, unlike strictly parallel approaches to the phonology-morphology interface, interleaved OT preserves and develops the best insights of Lexical Phonology into the life cycle of phonological processes.

Keywords: Optimality Theory, actuation, phonologization, analogy, rule inversion, rule loss, parser, input optimization, interleaving, Old English.

0. INTRODUCTION[1]

Optimality Theory (OT) is a theory of grammar in the generative sense. As such, it is designed to answer questions about the nature of linguistic competence and about the logical problem of language acquisition. In other words, OT offers a view of how linguistic knowledge is represented in the mind, and purports to show how it is possible in principle for such knowledge to be acquired (see e.g. Tesar & Smolensky 2000). Although proposals advanced in the pursuit of these goals must ultimately prove consistent with what is reliably known about language use, it falls beyond the province of a theory of grammar to account for the circumstances of performance. It is nonetheless a widespread and uncontroversial assumption that factors pertaining

D. Eric Holt (ed.), Optimality Theory and Language Change, 91-119.

to the arena of use (Hurford 1990:§2.1) often play a large causal role in the development of grammatical innovations. It follows, therefore, that the contribution of OT to the study of language change in general, and to a solution of the actuation problem (Weinreich, Herzog & Labov 1968:102) in particular, must of necessity be partial and limited. This assertion is not meant as a criticism of OT, or of generative linguistics in general; it is simply a reminder that progress in our understanding of language change is unlikely without a careful and principled division of labor between grammatical theory, on the one hand, and disciplines that focus on the arena of use, on the other.

The relevance of this division of labor comes to the fore when one addresses the actuation of innovations in the phonological component of the grammar. It is generally believed that a substantial proportion of phonological innovations, specially among those yielding neogrammarian sound changes, consists of the phonologization of nongrammatical phonetic effects. In this connection, Ohala (e.g. 1989, 1992, 1993) has persuasively argued that phonologization is the result of erroneous parsing: in a typical hypocorrection scenario, the learner misinterprets some distortion of the incoming speech signal (e.g. a mechanical coarticulation effect) as realizing some property of the output of the target grammar. This approach to phonologization places the onus of explanation firmly on a model of the parser, rather than on the theory of grammar. If this view is correct, then OT should not be expected to play a leading role in explaining why any particular instance of phonologization should occur in a given place at a given time.

In §1.1, however, we echo Jakobson's (1929) and Kiparsky's (1988, 1995) argument that grammatical conditions, both universal and language-specific, affect the incidence of phonologization. During language acquisition, in particular, the output of the parser must be filtered in accordance with markedness principles; if phonologization were blindly actuated by purely phonetic factors, then sound change could result in the violation of phonological universals. In this connection, OT eliminates the need to run global checks to ensure that derivations comply with universal markedness generalizations, as the latter are directly built into CON. Furthermore, the Jakobson-Kiparsky argument directly contradicts Hale & Reiss's (2000) assertion that markedness generalizations are epiphenomena of cumulative performance-driven change.

These facts notwithstanding, it appears that OT's contribution to the actuation problem should be assessed mainly with respect to phonological innovations that are independently known not to be caused by phonetic factors. In this set belong many of the instances of phonological change traditionally labeled as analogical. In this chapter we shall focus our attention on two representative types: those characterized in rule-based taxonomies as involving rule inversion and rule loss. Examples are given in §1.2 and §1.3, where we provide evidence to confirm that phonetics plays no significant role in the actuation of such innovations. We shall observe that, like most types of analogical change, rule inversion and rule loss involve the restructuring of input representations at some level in the phonology. Hence, OT can contribute to our understanding of such developments by establishing when and how input restructuring takes place (§2.1).

In this respect, the concept of input optimization[2] (Prince & Smolensky 1993:§9.3; Inkelas 1995; Itô, Mester & Padgett 1995) proves illuminating. If the learner selects those input representations that lead to the minimal violation of faithfulness constraints, then input and output representations will be maximally similar. Therefore, in the absence of robust evidence to the contrary (such as may be provided by alternations), input optimization compels the learner to adhere to the default assumption that, at each level of the phonology, input representations are identical with the given output. Thus, in OT the formal requirements of optimality (i.e. minimal constraint violation) closely match the functional demands of learnability (i.e. minimal abstractness).

In contrast, rule-based phonological frameworks have historically struggled to strike a satisfactory balance between learnability and abstractness. Provisions designed to limit abstractness include the Alternation Condition, Strict Cyclicity, and the Strong Domain Hypothesis. Notably, McMahon (2000a) has made a serious attempt at reasserting the Alternation Condition as a force driving input restructuring, whilst acknowledging that the Condition must be understood as a characterization of the learner's strategy rather than as a formal constraint on grammars. We shall demonstrate, however, that all such stipulations suffer from serious empirical or theoretical difficulties (§2.2). A strong reading of the Alternation Condition such as McMahon proposes would for example forbid underspecification wherever it is not warranted by alternations. This would prevent the use of structure-building lexical rules to capture the predictable phonological properties of nonalternating lexical items and would thereby deprive rule-based theory of its answer to the duplication problem. In this sense, McMahon's proposal to curb the learner's power in line with a strengthened Alternation Condition addresses the causes of input restructuring, but compromises the synchronic adequacy of rule-based theory.

In sum, this chapter outlines OT's contribution to research into the actuation of phonological change. We suggest that accounts of phonetically driven innovation will primarily draw on models of the parser, rather than on theories of grammar. Nonetheless, OT casts new light on the way in which universal markedness principles constrain phonologization; relatedly, we shall demonstrate that OT is correct in predicting that the format of inverted phonological processes is constrained by markedness. Furthermore, the Optimality-Theoretic concept of input optimization takes center stage in the analysis of phonological innovations involving input restructuring, such as rule inversion and rule loss. In this area, OT's advantage over previous rule-based frameworks lies in its success at encoding grammatical generalizations – whether static or dynamic – without appealing to underspecification and lexical minimality, which presuppose an overpowerful learner.

The chapter closes with an illustrative case-study drawn from the West Saxon dialect of Old English (§3). During the tenth century, the phonology of this dialect underwent a series of developments that culminated in the loss of a process deleting certain inflectional vowels, whilst a number of epenthesis rules suffered inversion and were replaced by syncope. Our analysis will be couched in terms of an interleaved or cyclic OT model. As our discussion will make apparent, interleaved implementations of OT enjoy a significant advantage over strictly parallel alternatives (notably OO-correspondence and Sympathy Theory) in that input restructuring at one phonological stratum can be seen to actuate concomitant changes at higher levels in the grammar. More generally, interleaved OT continues to build on the insights into the life cycle of phonological processes gained by research in the Lexical Phonology (LP) tradition.

1. THE ACTUATION OF PHONOLOGICAL INNOVATION: PARSERS AND GRAMMARS

1.1 Phonologization

For well over a century, research in historical phonology has been informed by the observation that, in numerous instances of sound change, the innovative grammar replicates a previously unintended phonetic effect. In recent decades, Ohala has conducted a vigorous program of research into the causes of phonologization (see e.g. Ohala 1989, 1992, 1993). We interpret his results as indicating that phonologization generally arises through a malfunction of the phonological parser. By 'phonological parsing' we understand the mapping of an acoustic speech signal onto the overt part of a surface phonological representation. Admittedly, it is not easy to specify the precise nature of the device or set of devices that effect this mapping. Minimally, parsing presupposes the transduction of the acoustic stimulus into some sort of pattern of neural activation such as may provide the basis for computation; in this connection, see Pylyshyn (1984) and Hale & Reiss (2000). However, there is much more to phonological parsing: notably, we shall see below that parsing is sensitive to language-specific conditions and, in this respect, is not strictly stimulus-bound. The parser could accordingly be regarded as belonging to the class of cognitive systems that Fodor calls 'input analyzers' and that operate on the output of transducers (Fodor 1983:41, 48-49). Nonetheless, we would wish to see the output of the phonological parser as consisting of overt forms in the sense of Tesar & Smolensky (2000); we regard the computation of covert structural relationships (i.e. those lacking phonetic exponence, such as syllabic and metrical constituency) as a task for the grammar; cf. Fodor (1983:note 28). In this sense, the term 'feature extractor' approximately describes the notion of phonological parser that we will we have in mind in the following discussion; see e.g. Fitzpatrick & Wheeldon (2000:140-142) for a sketch of a model of speech perception that similarly decouples feature extraction from the computation of covert structure in the grammar.[3]

Adapting Ohala's arguments to this conceptual framework, one must assume that parsing involves an element of correction, whereby during feature extraction the parser automatically filters out unintended distortions of the speech signal caused by predictable performance factors (physiological, aerodynamic, acoustic, etc.). The relevance of correction to language acquisition and language change is self-evident. If the parser fails to rectify the aberrations introduced into the speech signal during performance, then the learner's trigger experience will shift in relation to the grammatical output of the adult generation. More specifically, the learner will interpret an unintended distortion of the speech signal as manifesting a property of surface representations in the target grammar. In such cases, an unintended phonetic effect is said to become phonologized through hypocorrection.

As an example, let us consider velar softening, i.e. the process whereby velar stops become affricates before high front vowels (Ohala 1989:185-186, 1992:319-321). In the articulation of a sequence such as [ki], the front of the tongue rises in preparation for the high front vowel as the velar closure is released. As a result, the stop burst is accompanied by relatively loud noise, for it takes place in a narrowed channel with high air velocity. When correction is successfully applied, the parser automatically compensates for such effects. However, when the noisy release of the stop is misinterpreted as the physical exponent of an affricated segment in the surface phonological representation, we have hypocorrection.[4]

Ohala's work suggests that research into the properties of the parser can yield deductive-probabilistic explanations of aspects of sound change, such as the relative incidence of different types of phonologization (see e.g. Ohala 1993:§6.6). Notably, particular forms of hypocorrection can be assigned relative probabilities either by induction from experimental evidence (e.g. Winitz, Scheib & Reeds 1972) or by deduction from first principles (e.g. Hayes 1999:§10). On this basis, sound changes actuated by highly probable hypocorrection processes can be predicted to occur more frequently than those driven by comparatively improbable types of misparsing. Such predictions can be made precise through dynamical system modeling (e.g. Pulleyblank & Turkel 1996; see further Nigoyi & Berwick 1997). Consider the following examples:

• Confusion matrices for stop+vowel syllables indicate that the probability of [ku] being misperceived as [pu] is low compared with the likelihood that [ki] will be misheard as [ti]; Winitz, Sheib & Reeds (1972) quantify the difference as 0.24 against 0.47. This asymmetry would explain the fact that sound changes involving velar coronalization (1a) appear to be recorded far more frequently than developments involving velar labialization (1b); see Ohala (1989:182-185).

(1) (a) *Coronalization*:
West Germanic */kiːðɑn/ > Old English /tʃiːdɑn/ 'chide'

(b) *Labialization*:
Proto-Bantu */-kumu/ > West Teke /pfumu/ ‘chief’

• Archangeli & Pulleyblank (1994) report a case of convergent drift in the E̘doid language family (Niger-Kordofanian, Benue-Congo; SW Nigeria; see Elugbe 1989). The protolanguage had a symmetrical 5+5 ATR-harmony system, as shown in Figure 1. Among its present-day daughters, however, approximately 95% lack the low advanced vowel [a̘], whilst a smaller proportion (around 60%) have lost the high retracted vowels [i̙] and [u̙] (Elugbe 1982). The comparative instability of [a̘] could plausibly be imputed to a greater liability to misparsing (Pulleyblank & Turkel 1996:678). Bermúdez-Otero (1999:§4.1.2) suggests that this hypothesis can be submitted to experimental testing: if correct, then a confusion matrix for a 5+5 ATR-harmony system should show that the rate of misperception of [a̘] exceeds that of [i̙] and [u̙] by an appropriate margin.

i̘ i̙ u̘ u̙

e̘ e̙ o̘ o̙

a̘ a̙

Figure 1. 5+5 ATR-harmony system in Proto-E̘doid.

Admittedly, language-particular developments remain unpredictable, partly owing to the influence of social factors on propagation. In crosslinguistic studies, however, it should be possible to control for the effects of sociolinguistic evaluation by using large enough language samples, given that social selection operates on grammatical variants regardless of their content (Croft 1995:524).

If our interpretation of Ohala’s proposals is on the right track, then one must look to models of the parser – rather than to theories of grammar – for insights into the actuation of changes involving phonologization; the rise of OT as a model of phonological competence is unlikely to impinge upon this research program. However, this conclusion needs to be qualified in the light of evidence showing that the incidence of phonologization processes is constrained by grammatical factors, both universal and language-specific. In other words, the grammar would appear to exert a top-down selective pressure on the output of the parser (Kiparsky 1988:372, 391; 1995:§2). If so, grammatical theory must after all play a role in accounting for innovations in which phonetic effects become phonologized.

One line of argument leading in this direction originates with Jakobson (1929). Jakobson pointed out that phonological universals constrain the range of innovations that are possible at any given point in the history of a language; but, if phonologization were actuated by purely local phonetic factors, then certain global universal restrictions could not endure through time. Kiparsky (1995:641) provides a striking example: if lenition occurred without regard to grammatical principles, then it would be possible for a sequence of context-free lenitions to give rise to a language whose segment inventory lacked oral stops; yet languages of this nature are obviously not attested. It cannot be denied that, in numerous cases, compliance with phonological universals is precisely what one would expect in the light of the phonetic factors driving phonologization. From an acoustic viewpoint, for example, segmental contrasts are more robustly cued (and are therefore more resistant to loss by hypocorrection) in prevocalic than in preconsonantal position; accordingly, one does not expect a series of conditioned mergers to give rise to a situation where the coda supports more phonemic oppositions than the onset, in violation of phonotactic universals (Steriade 1997). Nonetheless, Kiparsky's context-free lenition scenario highlights the fact that the phonetic factors triggering hypocorrection do not always conspire to enforce phonological universals; in other circumstances, phonologization could, if given a free rein, create impossible phonological systems. Indeed, even in cases where a phonological universal appears firmly rooted in phonetic substance, the match is rarely perfect. Take again the example of the syllable coda as the preferred site of neutralization. In Eastern Andalusian Spanish, a process of aspiration cancels phonemic oppositions between obstruents in the coda. As Gerfen (2001) demonstrates, however, the phonetic grounding of this phenomenon is relatively indirect: notably, in this dialect of Spanish there are coda segments subject to aspiration that would *not* be predicted to undergo neutralization purely on the grounds of phonetic cue impoverishment, *pace* Steriade (1997).

Jakobson's reasoning is compelling,[5] but it does not require that grammatical principles should be imported into the parser; on the contrary, his argument is compatible with the presumption that parsing is primarily stimulus-bound and independent from grammatical conditions (but see below for a qualification of this view). It suffices that, whenever the data supplied by the parser prove incompatible with principles of Universal Grammar (UG), the Language Acquisition Device (LAD) should fail to respond; if so, then misparsing will never trigger grammatically impermissible innovations. Interestingly, this proposal is far easier to implement in OT than in rule-based phonology. In rule-based frameworks, phonologization is typically taken to involve the addition of a new rule at the end of the grammar (see e.g. King 1969, 1973). To ensure that the output of the new rule complies with UG restrictions, particularly implicational universals, the learner has to run a global check over derivations from existing input forms; rule-based theories offer no alternative to such global checks because they treat surface generalizations as purely epiphenomenal. In OT, in contrast, universal constraints on inventories are directly incorporated into the formulation of CON (Prince & Smolensky 1993:ch. 9; Pulleyblank 1997; Kager 1999:§1.7, §1.8): using Kiparsky's example above, it should be impossible to rank the universal constraint set in such a way as to forbid the presence of oral stops on the surface.

Incidentally, the Jakobson-Kiparsky argument bears directly on the autonomy of phonology and, in particular, on the status of markedness constraints in OT. Haspelmath (1999) and Hale & Reiss (2000) argue that implicational universals are epiphenomenal properties of language, arising from the cumulative effect of performance-driven change. On these grounds, Hale & Reiss exclude markedness constraints from UG, declaring them to be redundant and in breach of Ockham's razor. The Jakobson-Kiparsky argument, however, shows that UG must supply the learner with markedness principles capable of filtering the innovations triggered by misparsing; otherwise, blind phonologization could lead to the violation of global phonological universals. In OT, of course, universals emerge from CON by factorial typology. For further arguments against the claim that markedness constraints are mere epiphenomena of performance-driven change, see Bermúdez-Otero & Börjars (in preparation). Note, however, that the Jakobson-Kiparsky argument does not require CON to be innate and biologically transmitted (on this subject, see e.g. McMahon 2000b:ch. 5). The argument simply demands that CON should be universal, finite, and available to the learner at the onset of grammar construction. These criteria are compatible with noninnatist approaches such as Hayes's (1999) theory of inductive grounding, where the learner is equipped with an algorithm for constraint discovery that guarantees universality.

Research has further shown that language-specific grammatical conditions can also exert a top-down selective influence upon processes of phonologization. Notably, performance distortions affecting a given phonetic variable appear far more likely to undergo phonologization when that variable is already involved in the realization of a phonological contrast in the grammar. In other words, phonological oppositions can have a priming effect upon phonologization (Kiparsky 1995:§2.1). Tonogenesis provides a good example. For apparently physiological reasons, voiceless consonants raise the fundamental frequency (F_0) of a following vowel; see Löfqvist *et al.* (1989) and Ohala (1993:240, 269). This F_0 disturbance may be reinterpreted as the realization of phonological tone. An example is given in (2), where the Northern Kammu forms are innovative in relation to their cognates in Southern Kammu; the data are drawn from Ohala (1989:181; 1993:240), who cites Svantesson (1983:69).

(2)	*Southern Kammu*		*Northern Kammu*	
	klaaŋ	>	kláaŋ	'eagle'
	glaaŋ	>	klàaŋ	'stone'

It turns out, however, that F_0 differences associated with voicing distinctions tend to be phonologized only in languages that already possess lexical tone contrasts or that are in contact with tonal languages (Matisoff 1973, Svantesson 1989). This observation demonstrates that language-specific grammatical conditions have a significant effect upon the relative probability of phonologization processes;

predictions derived solely from the impact of phonetic factors upon the parser are bound to prove inadequate.[6]

In this sense, the evidence of phenomena such as tonogenesis indicates that, if phonologization is indeed caused by misparsing, particularly through hypocorrection, then the parsing process cannot after all be strictly stimulus-bound. This conclusion agrees with experimental phonetic evidence suggesting that speakers of different languages deploy partially different strategies in parsing: according to psycholinguistic research summarized in Cutler & Mehler (1993), for example, French speakers rely on a syllable-based speech segmentation strategy for lexical access, whilst English speakers use foot-based segmentation. Thus it appears that in the process of phonological acquisition the learner not only constructs a grammar but also attunes the parser's response to different phonetic variables according to their involvement in phonological patterns (for references, see Hume & Johnson 2001:§2.1, §4.2). This assumption would account for the existence of crosslinguistic variation in the relative probability of hypocorrection events.[7] Unfortunately, the coevolution of grammars and parsers during phonological acquisition has scarcely been explored in the generative tradition; it is therefore unclear at present whether OT could make a significant contribution in this area (although see Boersma 1998).

To conclude, mechanical phonetic effects become phonologized through parsing errors. The actuation of such innovations involves a host of performance factors lying beyond the purview of grammatical theory. Nonetheless, the Jakobson-Kiparsky argument indicates that universal markedness principles supplied by UG (innately or via inductive grounding) exert a filtering effect upon the output of the parser. In this area, OT eliminates the need to run costly global checks over derivations to ensure compliance with implicational universals; the latter emerge directly from the formulation of Con. Considerable uncertainty remains, however, over the mechanisms whereby language-specific grammatical properties affect the behavior of the parser.

1.2 Rule inversion

If our proposals are correct, phonologization is actuated in the mapping from the acoustic speech signal to the set of overt forms that provide the input to grammatical acquisition. It has long been recognized, however, that phonological change may be caused by nonphonetic factors. Innovations of this nature are traditionally discussed under the heading of analogy. In rule-based taxonomies of change (e.g. Kiparsky 1968, King 1969), most of the relevant instances fall in categories such as rule reordering, rule inversion, and rule loss – as opposed to rule addition at the end of the grammar. Here we shall concentrate on rule inversion and rule loss. We will argue that the theory of grammar must play a key role in accounting for such innovations, because the actuating factors are located in the process whereby a grammar is constructed on the basis of overt surface representations in the primary linguistic data (PLD).

Rule inversion (Vennemann 1972) affects the relationship between input and output representations within the grammar. Let the grammar generate a surface alternation between two elements α and β respectively occurring in the environments $[\_]_A$ and $[\_]_B$. Assume that, in the initial state of affairs, α is a faithful realization of the input representation common to both elements; if so, β will be generated by a mapping $\alpha \rightarrow \beta$ in environment $[\_\_]_B$. Rule inversion takes place if the learner reanalyzes β as being identical with the input representation and concomitantly posits an innovative mapping $\beta \rightarrow \alpha$ in environment $[\_]_A$; see Figure 2.

	Stage I	*Stage II*
input	α ↓ ↘	β ↙ ↓
output	$[\alpha]_A$ $[\beta]_B$	$[\alpha]_A$ $[\beta]_B$
mapping	$\alpha \rightarrow \beta$ / $[\_\_]_B$	$\beta \rightarrow \alpha$ / $[\_\_]_A$

Figure 2. Rule inversion.

The classic example of rule inversion is the rise of 'hiatus rules' (Vennemann 1972:§2). At stage I, the grammar contains a process deleting an underlying consonant *C* in coda position; in certain lexical items, the process creates an alternation between ∅ in the coda and *C* in (postvocalic) onset position. At stage II, the input representation of those items is restructured by eliminating *C*, and an innovative rule is posited that inserts *C* into postvocalic empty onsets (i.e. in hiatus environments). The development of intrusive /r/ in nonrhotic dialects of English is commonly taken to instantiate this scenario (Vennemann 1972:216), although the claim has lately stirred a great deal of controversy; see McMahon (2000a:ch. 6) for a restatement of the case for inversion and a survey of alternative analyses.

What causes rule inversion? In the light of Figure 2, the actuating factor must be some circumstance prompting the learner to model the input representation on β rather than α. A number of more or less speculative hypotheses have been advanced in this area. Vennemann (1972:§8) surmises that the most frequent alternant is likely to be preferred as the input representation. This is empirically plausible in the case of hiatus rules, particularly English /r/-intrusion: in English, a word-final segment is three times less likely to be followed by a vowel than by a consonant or pause (Bybee 1998:73). Moreover, relatively frequent alternants should have a better chance of becoming cognitively entrenched early in the process of lexical acquisition. Vennemann further suggests that input representations tend to be modeled on unmarked, morphologically basic categories. Asymmetries between paradigm members are well documented in work on analogical change (e.g. Kuryłowicz 1949, Mańczak 1958, Lahiri 1982, Lahiri & Dresher 1983-84, Dresher

2000). In this connection, the choice of input may also plausibly be affected by the prevalent direction of alternations in the language (cf. the notion of 'system congruity' in Natural Morphology; Wurzel 1989).

Thus, rule inversion is actuated by factors bearing on the selection of input representations within the grammar; phonetics does not play a role. This assertion can be clearly established in the subset of rule inversion cases where the phonological and lexical distribution of alternants remains unchanged. In the case of /r/-intrusion, inversion became manifest when /r/ started to occur as a hiatus breaker in words where etymological linking /r/ had previously been absent. Nonetheless, the symptoms of inversion can be far subtler. In §3, for example, we discuss a case of rule inversion in Old English whereby the vowels inserted in obstruent+sonorant clusters under certain conditions were reanalyzed as underlying; thereafter, vowelless clusters were derived by an innovative process of syncope. Crucially, the surface distribution of vowels in the environment [–son]_____[+son] remained unaltered. Inversion can be detected through the effects of input restructuring upon other rules in the grammar: in certain inflectional forms, the restructured inputs should have fed an independent process of apocope; however, their surface correspondents showed that they did not, and apocope was consequently lost. Therefore, since inversion affected neither the realization nor the distribution of the formerly anaptyctic vowels, we can safely conclude that phonetic factors were not at play in the actuation of the change.

We have argued that, diachronically, rule inversion is actuated by nonphonetic factors; but does the historical origin of inverted rules have a bearing on their synchronic properties? This question turns out to have profound implications for the theory of grammar, for by definition inverted rules must, like all others, comply with UG principles. In some cases, inverted rules turn out to have exactly the same format as phonological processes created by phonetically driven change: e.g. in our Old English case-study, the innovative syncope rule is indistinguishable from a process originating in hypocorrection. It has been argued, however, that rule inversion may produce synchronically arbitrary effects. In particular, a number of authors claim that the choice of epenthetic consonant in hiatus rules arising via inversion cannot be synchronically motivated. In the case of English /r/-intrusion, this view is adopted by Vennemann (1972), McCarthy (1993), Blevins (1997), Halle & Idsardi (1997), Hale & Reiss (2000), and McMahon (2000a, b). The consequences for OT are momentous: if grammars may contain parochial stipulations that cannot be derived from universal markedness constraints, then either the theory is fundamentally misconceived or else its formal resources must be expanded.

However, several authors have countered this attack by suggesting that English /r/ has a special qualitative affinity with the preceding vowels in the intrusion environment, particularly /ə/ (e.g. Broadbent 1991; Donegan 1993; Harris 1994; McMahon, Foulkes & Tollfree 1994; Baković 1999; Gick 1999; Giegerich 1999; Gick 2002; Gick, Kang & Whalen 2002).[8] Notably, Gick (2002) adduces evidence from X-ray cinematography to argue that American English schwa does not lack place features, as is generally suggested, but, like /r/, involves a gesture of tongue root retraction causing mid-pharyngeal constriction. If this is correct, then one can argue that, synchronically, either faithfulness or contextual markedness constraints (Kager 1999:125) are instrumental in the selection of epenthetic /r/. Certain English dialects spoken in the Northeast of the United States (Gick 1999, personal communication) provide empirical support for this approach to inverted hiatus rules. In these dialects, /l/ undergoes vocalization in the coda, but is retained in onset position (linking /l/). This alternation between /l/-vocalization and linking /l/ creates the conditions for the development of intrusive /l/ via rule inversion. Significantly, whilst the crucial alternation is found after the entire vowel set, in many of these dialects intrusive /l/ only occurs following /ɔː/: e.g. *law*[l]-*abiding*, but *the bra is*. Magnetic resonance images obtained by Gick, Kang & Whalen (forthcoming) show that, in this dialect area, /l/ and /ɔː/ incorporate identical gestures of tongue dorsum backing, leading to similar vocal tract configurations in the pharyngeal and uvular regions; in phonological terms, this suggests that both segments possess identical specifications under the V-place node (see Sproat & Fujimura 1993 and Gick 1999:40 on C-gestures and V-gestures). Thus, rule inversion, leading to /l/-intrusion, has not taken place wherever /l/-vocalization alternates with linking /l/, but is restricted to those alternating environments where epenthetic /l/ is maximally similar to the preceding vowel. This proves that the distribution of the intrusive segment is synchronically controlled by constraints sensitive to the relationship between the hiatus breaker and the preceding vowel (for further discussion, see Bermúdez-Otero & Börjars in preparation).

There is a further argument disproving the arbitrariness of inverted hiatus rules. According to Vennemann (1972:216), “The particular consonant(s) introduced by a hiatus rule can only be explained historically.” If taken seriously, this claim implies that, in the appropriate circumstances, any coda deletion process can undergo inversion; the feature content of the target consonant is allegedly immaterial. This proposal makes a strong typological prediction capable of empirical disconfirmation: the class of hiatus breakers epenthesized by inverted rules must constitute a *random* subset of the class of consonants susceptible to synchronic coda deletion. To give but one example, since Oromo has a rule of preconsonantal /ɗ/ deletion (Lloret-Romanyach 1988), it should be possible in principle for the implosive [ɗ] to function as an epenthetic hiatus breaker. In practice, however, the class of consonants reported in the literature as being inserted by inverted hiatus rules appears far from random, consisting mostly of coronal sonorants, often with secondary dorsal or pharyngeal articulations: e.g. /r/ in nonrhotic English dialects and possibly in Yukulta (Blevins 1997:§5.2); /l/ in Bristol and certain American English dialects (Ash 1982; Blevins 1997:233; Gick 1999; McMahon 2000a:244, 262); /n/ in certain Dutch idiolects (Booij 1996:227). Thus, inverted hiatus rules appear to favor a small set of coronal sonorants over other types of consonant subject to coda deletion. If this impression is true to the typological facts, then we must conclude that, even in cases of rule inversion, the choice of hiatus breaker is constrained by markedness (inherent and contextual).

In sum, our conclusions in the area of rule inversion are similar to those of Jakobson (1929) and Kiparsky (1988, 1995) in respect of phonologization. Despite claims to the contrary, the format of inverted rules is subject to universal markedness principles supplied by UG (whether innately or via inductive grounding); when such principles are not complied with, the LAD blocks rule inversion. In contrast, Hale & Reiss (2000) assert that phonological markedness generalizations are epiphenomena of recurrent patterns of phonologization; accordingly, markedness constraints are banned from the grammar. Rule inversion, however, does not involve phonologization; as we have demonstrated, it is actuated by nonphonetic factors. Hale & Reiss therefore predict that inverted rules need not comply with markedness generalizations; their content will be 'arbitrary'. Typology falsifies this prediction if, as we surmise, the class of hiatus breakers epenthesized by inverted rules is a *nonrandom* proper subset of the class of coda consonants susceptible to deletion.[9] Similarly, the evidence of /l/-intrusion in American English dialects shows that the application of inverted hiatus rules is synchronically controlled by constraints sensitive to the relationship between the hiatus breaker and the preceding vowel.

1.3 Rule loss

Another type of phonological innovation commonly discussed under the heading of analogy is rule loss (see e.g. King 1969:46-51). This term may cover various forms of phonological obsolescence; here we shall focus our attention on the endogenous loss of contextual rules of full neutralization. In other words, we are interested in the process whereby a learner may acquire a grammar that permits a phonemic contrast to be realized in an environment where neutralization was enforced by the grammar of the adult generation. In this sense, rule loss involves the undoing of conditioned merger. This type of innovation must be carefully distinguished from the reversal of a near-merger, for in situations of near-merger contrasts are maintained – however tenuously – in production and are only suspended perceptually (see e.g. Labov 1994:ch. 12). More generally, it is vital to distinguish instances of categorical neutralization in the phonology from gradient phonetic effects. Additionally, we are not concerned here with exogenous developments whereby a contrast is re-established in the wake of massive lexical borrowing from a different language. Rather, our discussion will refer to rule loss in circumstances of excessive opacity.

A classic example of rule loss, in this restricted sense, is the restoration of obstruent voicing contrasts in word-final position in Yiddish (Kiparsky 1965, 1968; King 1969, 1976). As shown in Table 1, Middle High German had a process of coda devoicing that gave rise to alternations; in Yiddish, however, these alternations were leveled and voiced obstruents were reintroduced in word-final position.

Table 1. The loss of coda devoicing in Yiddish

	Alternating items		*Non-alternating items*
Old High German	*ta**g***	*ta**g**a*	*ave**g***
Middle High German	*ta**c***	*ta**g**e*	*ave**c***
Yiddish	*to**g***	*te**g***	*ave**c***
	‘day’	‘days’	‘away’

The loss of devoicing in word-final position appears to have been precipitated by a previous grammatical innovation whereby final schwas underwent apocope. The new process of final schwa deletion caused a crucial shift in the learner’s trigger experience. In Table 2, examples from King (1976:22) show how final devoicing ceased to be surface-true (McCarthy 1999:332); it appeared to underapply (Wilbur 1974). In this situation, the PLD failed to provide consistent evidence for final devoicing, and the neutralization rule was not acquired. In cases such as that of ‘day’, the leveling of alternations results from the learner’s failure to accommodate them under some phonological or morphological generalization (Dresher 1998, 2000; Lahiri & Dresher 1999:698-699).

Table 2. Apocope causes final devoicing to underapply in Yiddish

	Before apocope				*After apocope*			
	Sing.	Plur.	Sing.	Plur.	Sing.	Plur.	Sing.	Plur.
Nom.	*gebe*	*gebe*	*tac*	*tage*	*ge**b***	*ge**b***	*tac*	*ta**g***
Acc.	*gebe*	*gebe*	*tac*	*tage*	*ge**b***	*ge**b***	*tac*	*ta**g***
Gen.	*gebe*	*geben*	*tages*	*tage*	*ge**b***	*geben*	*tages*	*ta**g***
Dat.	*gebe*	*geben*	*tage*	*tagen*	*ge**b***	*geben*	*ta**g***	*tagen*
	‘gift’		‘day’		‘gift’		‘day’	

As in rule inversion, however, input restructuring plays a key role in the actuation of rule loss. In the Yiddish case, it is crucial that, after apocope, certain inflectional endings should have been reanalyzed as underlyingly null. This prevented learners from converging on a grammar whose derivations recapitulated history: see Figure 3, where a synchronic apocope rule counterfeeds final devoicing.

	nom.sg.	*nom.pl*
UR	/tag/	/tag-ə/
Devoicing	tak	—
Apocope	—	tag
SR	[tak]	[tag]

Figure 3. Counterfeeding derivation.

It is not unlikely that grammars of this nature were acquired whilst apocope remained variable. Observe that the implementation of this derivation poses no serious obstacles to interleaved or stratal OT (see references in §4), for devoicing clearly took place at the word level and apocope may safely be assumed to have entered the grammar as a postlexical innovation. In the event, however, schwa deletion became obligatory, and learners equated the input representation of the nom.pl. ending with its surface realization, namely /-∅/.

Like rule inversion, therefore, rule loss is ultimately driven by principles governing the relationship between input and output representations. Moreover, it is easy to demonstrate that opaque neutralization processes cannot be lost through misparsing. We assume, uncontroversially, that the implementation module (see note 4) has access to output phonological representations but cannot look back into derivations. By implication, categorical mergers leave no phonetic traces capable of affecting the parser. Furthermore, let us suppose for the sake of argument that the parser could occasionally undo the effects of a neutralization process through hypercorrection.[10] Nonetheless, the parser only hypercorrects in the presence of factors triggering a putative coarticulation effect (Ohala 1989:189-190, 1992:334, 1993:255). Hence, hypercorrection cannot undo a neutralization process whose conditioning environment fails to surface.

1.4 Recapitulation

The facts of phonological change show that universal markedness principles must be available to the learner at the onset of grammar construction, *pace* Haspelmath (1999) and Hale & Reiss (2000). Knowledge of markedness is required to filter the output of the parser so as to ensure that phonologization complies with phonological universals, particularly those imposing global requirements on phonological systems (e.g. most implicational universals). Similarly, despite contrary claims in the literature, markedness principles prevent rule inversion from creating synchronically arbitrary phonological processes. In this area, OT has achieved a significant advance in relation to rule-based grammatical frameworks. The theory directly incorporates markedness generalizations into the grammar; phonological universals emerge from CON by factorial typology. In the acquisition process, therefore, the learner is no longer compelled to evaluate grammatical hypotheses according to external markedness criteria, for compliance with universals, implicational or otherwise, is guaranteed.

Our discussion has further shown that accounts of phonologization will have to rely heavily on models of the parser. In contrast, analogical changes such as rule inversion and rule loss are actuated by factors bearing on the relationship between input and output representations within the grammar. Grammatical theory must therefore be expected to cast light on the actuation of such changes.

2. INPUT RESTRUCTURING

2.1 Input optimization

Lightfoot (1999:225) argues that, in any case of grammatical innovation, a solution to the actuation problem must consist of two elements: "(a) an account of how trigger experiences have shifted and (b) a theory of language acquisition that matches PLD with grammars in a deterministic way." Our discussion in §1.2 and §1.3 suggests that, in processes of rule inversion and rule loss, the crucial shift in the learner's trigger experience is usually caused by some independent development in the grammar of previous generations (which may itself have arisen through phonologization). In the case of English /r/-intrusion, for example, the earlier rise of /r/-deletion altered the PLD by creating a new set of alternations where the innovative unfaithful alternant was more salient than its faithful counterpart. Similarly, the loss of final devoicing in Yiddish followed the advent of apocope, which destroyed the conditioning environment of obstruent voicing alternations. According to Lightfoot's methodological guidelines, therefore, grammatical theory must explain how such shifts in the PLD triggered the ensuing changes in the grammar.

Our analysis suggests that, on both occasions, the immediate effect of the PLD shift was input restructuring: the learner posited innovative input representations identical with the corresponding outputs (or, in the case of alternating items, with the most salient output alternant). Interestingly, the state of affairs where input representations are identical with their respective outputs can be described as the initial or default position in the acquisition process: since input representations are not given in the PLD but have to be constructed, in the initial stages of lexical development learners must store observed surface forms (see Hale & Reiss 1998, Reiss this volume). It follows, therefore, that PLD shifts leading to rule inversion and rule loss remove crucial cues that the learner needs to depart from the default state. More generally, to explain phonological analogy the theory of grammar must define the range of data that can lead a learner to posit unfaithful input-output mappings; analogical innovations are predicted to occur when such data cease to be robustly represented in the trigger experience.

Defining the cues for input acquisition is to a large extent an empirical task and, as we shall see presently, one that remains surrounded by considerable uncertainty.

Nonetheless, OT has made a substantial contribution in this area through the principle of input optimization (Prince & Smolensky 1993:§9.3). Under the latter, the optimal state of an OT grammar coincides with the learner's default position in the acquisition process. This fact can be established by means of a simple deductive argument. According to the principle of output orientation, the scope of markedness constraints is limited to the output (for discussion, see Bermúdez-Otero 1999:50-52). As a result, the content of input representations does not affect the satisfaction of markedness constraints; inputs will only be relevant to the evaluation of faithfulness. Faithfulness constraints, in turn, penalize disparities between input and output representations. For any given output, therefore, the input-output mapping will be maximally harmonic if the input is identical with the output. Thus, the formal requirements of optimality in OT (i.e. minimal constraint violation) closely match the functional demands of learnability (i.e. minimally abstract inputs). This result contrasts sharply with the tensions that arise in rule-based phonology, whether in relation to the evaluation measure (see e.g. Koefoed 1974) or the use of underspecification (see §2.2).

In generative phonology it has been traditionally assumed that alternations provide learners with the best cues for non-surface-true inputs. According to Kiparsky & O'Neil (1976:550), for example, "Underlying representations that are supported by purely distributional regularities are less well entrenched [*scilicet* are more liable to change by reanalysis] than those supported by phonological alternations". Indeed, as a matter of pure logic, the underlying representation of an alternating morpheme can only be identical with one of its surface allomorphs; hence, phonological (as opposed to purely morphological or lexicalized) alternations must always involve input-output disparity.

Inkelas (1995) develops this line of thought in a highly restrictive way. In her theory of Archiphonemic Underspecification, she proposes that underlying representations are calculated separately for each individual morpheme: i.e. input optimization applies to each allomorph set and finds that input form that allows maximum harmony in the derivation of the entire set. This implies that the input representation of a morpheme *m* can differ from its surface realization only if the learner has encountered positive evidence that *m* is subject to alternation.

However, recent research suggests that Inkelas's original proposals are too restrictive. In a study of languages with root-controlled vowel harmony, for example, Harrison & Kaun (2000) used evidence from word games to probe the input representation of harmonic and disharmonic vowels in nonalternating environments (typically, in root domains). The theory of Archiphonemic Underspecification predicts that nonalternating segments will be fully specified in the input, whether they behave harmonically or disharmonically. However, Harrison & Kaun found that, in word games, harmonic root vowels usually behaved as though they were underlyingly unspecified for the harmonic feature: i.e. they patterned like alternating harmonic vowels in suffixes and unlike disharmonic segments. These results appear to indicate that learners sort segments into classes according to their phonological behavior and generalize input representations from alternating to nonalternating environments. Interestingly, this hypothesis suggests that allophonic information will not be stored in lexical representations as long as the relevant allophonic pattern produces some instances of alternation.

Along similar lines, Inkelas (2000) claims that learners use the observable phonological properties of newly encountered morphemes to assign them to existing lexical classes; the class to which the morpheme is assigned determines its input representation. Input-output disparities can thus arise in individual morphemes for which the learner has not yet encountered positive evidence of alternation.

In sum, the evidence adduced by Harrison & Kaun (2000) and Inkelas (2000) indicates that properties of the input acquired on the basis of alternations can be extended to nonalternating environments. Learners sort segments and morphemes into classes; optimal inputs are then computed for entire classes, rather than individual units.[11] There is, however, no compelling evidence that abstract input representations can be motivated on the basis of purely distributional evidence without the aid of alternations. Like Harrison & Kaun (2000), Yip (1996) uses noncore linguistic phenomena such as word games and poetic rhyme to probe the contents of underlying representations in Chinese dialects, where alternations are characteristically rare; her results fail to provide consistent support for non-surface-true inputs.

2.2 Restructuring in rule-based theory

In the previous section we have claimed that OT renders it relatively straightforward to account for input restructuring. In the absence of adequate cues for input-output disparity, the theory favors surface-true input representations, which arguably constitute the default option from the viewpoint of learnability. Admittedly, research is still ongoing to determine what counts as an ‘adequate cue for input-output disparity’, but our general assessment of the advantages of input optimization remains valid. A brief comparison with antiabstractness provisions in rule-based theories will serve to establish this point.

In rule-based LP, abstractness was moderated to some extent by the Strict Cycle Condition (Kean 1974, Mascaró 1976), which prevented underived lexical items from taking a free ride on cyclic rules (see e.g. Kaisse & Shaw 1985:§4.1). Strict cyclicity, however, is severely limited in scope: as Giegerich (1999) and McMahon (2000a) emphasize, a rule can avoid nonderived environment blocking simply by being assigned to a noncyclic stratum. Additionally, structure-building rules are normally exempt from strict cyclicity, which in effect allows underspecified structures to take free rides on cyclic rules applying in feature-filling fashion (see McMahon 2000a and below). More seriously, the link between cyclicity and nonderived environment blocking has failed to withstand empirical scrutiny, casting doubt on the very existence of the Strict Cycle Condition (Kiparsky 1993, Cole 1995).

Kiparsky's (1984:§1.5) Strong Domain Hypothesis also has the effect of minimizing abstractness. The Strong Domain Hypothesis stipulates that rules cannot become active at noninitial levels: if a rule applies at level *n*, then it must also apply at all levels higher than *n*. Under this provision, the operations undergone by the output of a noninitial stratum *n* are largely similar to those undergone by the input to *n* in the immediately preceding stratum *n*–1. In this sense, the Strong Domain Hypothesis indirectly expresses a general preference for input-output identity; the latter emerges as an epiphenomenon of similarities between mappings at different levels in the grammar. However, Kiparsky's proposal can only be regarded as a statement of preference, for counterexamples are known to exist (Orgun 1996a:§4.4.2). Moreover, even as a statement of preference, the Strong Domain Hypothesis can only play a limited role in promoting restructuring: the hypothesis may motivate changes in the stratal assignment of rules; but it condones opaque interactions within strata, and it cannot motivate input restructurings that leave the rule system unaffected (as is often the case in analogical extension).

McMahon (2000a) adopts a more radical stance in relation to input restructuring in rule-based LP. Notably, she advocates the reinstatement of the Alternation Condition (Kiparsky 1973) in a strengthened reading inspired by Lass & Anderson (1975:231). A recent formulation of the Alternation Condition can be found in Giegerich (1999:126):

(3) *Alternation Condition*
Structure-changing rules cannot apply to all occurrences of a given morpheme.

In this formulation, structure-changing rules are allowed to apply only if they create alternations. Note, however, that Giegerich's statement permits structure-building operations in nonalternating environments; by implication, (3) is compatible with the underspecification of predictable nonalternating structure (cf. Inkelas 1995). In contrast, McMahon (2000a) rejects underspecification altogether, claiming that it undermines other LP principles such as Structure Preservation; similarly, she suggests that the underspecification of nonalternating structure reintroduces free rides by the back door. In McMahon's framework, therefore, the force of the Alternation Condition is to penalize all rules that fail to create alternations.

To this extent, McMahon's deployment of the Alternation Condition produces results similar to those of Inkelas's (1995) version of input optimization: the underlying representation of nonalternating items is expected to be surface-true. Unlike input optimization, however, the Alternation Condition cannot be derived from first principles in rule-based theory; it is a stipulation imposed from without. This is connected to the fact that the Alternation Condition cannot be interpreted as a formal constraint on rules, since compliance can only be ascertained through a global check of derivations (Kiparsky 1982a:148, 1993:277-8; Giegerich 1999:§4.3.3). Thus, McMahon asserts that the Condition reflects a property of the learner's acquisition strategy, but concedes that it has no logical link with rule-based formalism.

Despite these drawbacks, McMahon's version of the Alternation Condition achieves the goal of providing a general motivation for input restructuring. Nonetheless, this diachronic success comes at a dear price from a synchronic viewpoint. McMahon's rejection of underspecification entails abandoning the use of structure-building rules to capture the predictable properties of nonalternating lexical items (see e.g. Kiparsky 1982a:167-8, 1982b:§3). This leaves rule-based theory on the horns of a dilemma. On the one hand, predictable nonalternating structure could be defined by morpheme structure constraints; such constraints, however, would substantially overlap with structure-changing rules, leading to a resurgence of the 'duplication problem' (Clayton 1976, Kenstowicz & Kisseberth 1977). On the other hand, predictable nonalternating structure could simply be treated as a matter of lexical stipulation. However, this alternative ultimately leads to absurd conclusions: languages with few or no alternations, such as Chinese, would be treated as having little or no phonology, in spite of possessing a rich and productive system of phonotactic regularities.

In essence, the problem lies in the fact that rule-based theories express phonological generalizations in a dynamic fashion, i.e. as operations. In nonalternating environments, however, rules can only apply if predictable structure is stripped off by underspecification. In turn, underspecification requires a powerful learner pursuing a strategy of lexicon minimization (Steriade 1995:114ff.). The diachronic evidence of input restructuring, however, calls for a weaker learner, whose default strategy is "what you see is what you get" (Anderson 1981:530). Thus, rule-based frameworks impose contradictory demands on acquisition theory. McMahon's (2000a) proposals highlight this contradiction, as she sacrifices synchronic adequacy in the interests of diachronic evidence. In OT, in contrast, this tension is resolved. Regularities are expressed in a static fashion, i.e. as constraints. Phonological generalizations can therefore be enforced in nonalternating environments without recourse to underspecification, and the principle of lexical minimality is exposed as factitious.

3. A CASE-STUDY: RULE INVERSION AND RULE LOSS IN LATE WEST SAXON

3.1 Apocope, anaptyxis, and parasiting in Ælfric's a-*stem nouns*

We shall illustrate the preceding discussion with a case-study of rule inversion and rule loss in the late West Saxon dialect of Old English. Our analysis will show that, during the tenth century, children's trigger experiences were crucially altered by the increased incidence of a process of anaptyxis that targeted coda-onset clusters with rising sonority. This development eroded previously well-established vowel~∅ alternations, prompting the reanalysis of epenthetic vowels as underlying and the replacement of anaptyxis by an inverted syncope rule. Input restructuring was simultaneously attended with the loss of an opaque process of apocope. This set of changes will be charted through its impact on the morphophonology of a subset of neuter *a*-stem nouns. Our discussion will be confined to directly relevant forms; for the wider synchronic context and historical background, see Bermúdez-Otero (2000, in preparation).

The writings of Ælfric (*circa* 950 - *circa* 1010) represent a standardized, relatively conservative variety of late West Saxon. Table 3 illustrates the morphophonological behavior of four major types of neuter *a*-stem noun in Ælfric's dialect. For each noun type we give nom.acc.sg., nom.acc.pl., and gen.pl. forms, with the gen.pl. representing all other oblique forms.

Table 3. Neuter a-*stem noun paradigms in Ælfric*

UR	/ʃip-/	/word-/	/werod-/	/wætr-/
nom.acc.sg.	*scip*	*word*	*werod*	*wæter*
nom.acc.pl.	*scipa, -u*	*word*	*werod*	*wætera, -u*
gen.pl.	*scipa*	*worda*	*weroda*	*wætera*
gloss	'ship'	'word'	'troop'	'water'

In the *a*-stem noun class, the neut.nom.acc.pl. ending occurs in two variants: /-u/ and /-ɑ/. The former reflects the etymology of the suffix (< Germanic *-*ō*). The low variant /-ɑ/, in contrast, constitutes a morphological innovation, already attested in some early West Saxon manuscripts (Cosijn 1886:7, 15) and statistically prevalent in Ælfric's texts (Pope 1967-8:183). As shown in Table 3, both variants are subject to apocope after stems meeting certain prosodic requirements, such as *word*- and *werod*-; the prosodic conditions are discussed below. Intriguingly, however, the gen.pl. ending /-ɑ/ fails to undergo deletion in the same environments, even though it is identical with the predominant variant of the neut.nom.acc.pl. suffix.

Apocope had entered Old English in prehistoric times as one of the manifestations of a wider process of high vowel deletion.[12] At this stage, the behavior of *a*-stem noun endings was simply determined by vowel height: the original neut.nom.acc.pl. suffix /-u/ underwent apocope, whilst gen.pl. /-ɑ/ was exempt. In Ælfric's dialect, however, apocope has ceased to be solely dependent on height and can apply to low as well as high vowels. In particular, the innovative neut.nom.acc.pl. suffix [-ɑ] reflects underlying /-ɑ/; it cannot be synchronically derived from underlying /-u/ by an optional lowering process ordered after apocope. As Hogg (1997:§4, 2000:§4) points out, the *ō*-stem nom.sg. suffix /-u/ – also subject to apocope in the appropriate environments –remains unlowered: e.g. fem.nom.sg. *gifu*, ***gifa* 'gift'.[13]

The contrast between neut.nom.acc.pl. /-ɑ/~/-u/ and gen.pl. /-ɑ/ in respect of apocope reveals a stratal split in the inflectional system of West Saxon: the neut.nom.acc.pl. suffix is inserted at level I; other *a*-stem noun endings, including gen.pl. /-ɑ/, are introduced at level II. Apocope targets both high /-u/ and low /-ɑ/, but only applies at level I. There is similar evidence of phonological stratification in other dialects (Dresher 1993). In Anglian and early Kentish, for example, apocope applies to the neut.nom.acc.pl. suffix of *a*-stem nouns but fails in the 1sg.pres.ind. of strong verbs: e.g. *bīdu* 'I wait', *haldu* 'I hold' (Campbell 1959:§346, §731 note 1; Suzuki 1988). The stratal split between apocope-prone and apocope-resistant nominal inflections may in fact have taken place fairly early in the history of Old English, when vowel reduction in unstressed syllables (Campbell 1959:§369, §373) obliterated height contrasts to which apocope had initially been sensitive: e.g. the apocope-prone *i*-stem nom.acc.sg. ending /-i/ fell together with apocope-resistant dat.sg. /-æ/, both yielding [-e] (Bermúdez-Otero 2000:§§36-37, in preparation).

Let us now examine the prosodic conditioning of apocope. The evidence of Table 3 suggests that the relevant final vowels undergo deletion when they would otherwise be preceded by a heavy syllable (e.g. ***wor.d**u**) or a sequence of two lights (e.g. ***we.ro.d**u**). Thus, if stress is assigned in Old English by erecting a moraic trochee at the left edge of the domain (Hutton 1998a:§3; cf. Dresher & Lahiri 1991), then apocope can be interpreted as avoiding stray syllables. In Optimality-Theoretic terms, the constraint hierarchy of level I will rank PARSE-σ, which requires exhaustive footing (see e.g. Kager 1999:153), above the faithfulness constraint preventing vowel deletion (MAX-V; see McCarthy & Prince 1995); see Tableau 1.[14]

Tableau 1. Apocope

Level I		PARSE-σ	MAX-V
/ʃipɑ/	☞[$_{\omega}$[$_{\Sigma}$.ʃi.pɑ.]]		
	[$_{\omega}$[$_{\Sigma}$.ʃip.]]		*!
/wordɑ/	☞[$_{\omega}$[$_{\Sigma}$.wor.]d]		*
	$_{\omega}$[$_{\Sigma}$.wor.].dɑ.]	*!	
/werodɑ/	☞[$_{\omega}$[$_{\Sigma}$.we.ro.]d]		*
	[$_{\omega}$[$_{\Sigma}$.we.ro.].dɑ.]	*!	
	[$_{\omega}$[$_{\Sigma}$.wer.].dɑ.]	*!	*

At first blush, however, the absence of apocope in the nom.acc.pl. of *wæter* is unexpected, as the final syllable is not incorporated into foot structure: [$_{\omega}$[$_{\Sigma}$.wæ.te.]**.rɑ.**]. The key to this underapplication effect lies in the fact that, underlyingly, the stem ends with a sequence of obstruent+sonorant: /wæ**tr**-/. The medial vowel of *wætera* is absent at level I, being inserted by a level-II process of anaptyxis that counterfeeds apocope.

Consider first the derivation of underlying /wætr-ɑ/ at level I. The word-medial cluster /-tr-/ constitutes a potential onset in Old English. Like Common Germanic, however, Old English treats such consonant sequences as heterosyllabic when immediately preceded by a short stressed vowel.[15] Accordingly, input /wætr-ɑ/ is mapped onto output [$_{\omega}$[$_{\Sigma}$.wæt.].rɑ.]. The final syllable remains unfooted; nevertheless, apocope cannot apply because it would create a word-final consonant cluster with rising sonority. As Tableau 2 shows, such a configuration would violate either SONPK→σ, which requires sonority maxima to constitute syllable nuclei (Selkirk 1984), or NUC→V, which penalizes syllables headed by a nonvocalic segment (see Prince & Smolensky 1993:16).

Tableau 2. Apocope blocked after obstruent+sonorant clusters

Level I		NUC→V	SONPK→σ	PARSE-σ	MAX-V
/wætrɑ/	[$_{\omega}$[$_{\Sigma}$.wæt.].rɑ.] ☜			*	
	[$_{\omega}$[$_{\Sigma}$.wæt.]r]		*!		*
	[$_{\omega}$[$_{\Sigma}$.wæ.tr̩.]]	*!			*

The blocking of apocope after underlying obstruent+sonorant consonant clusters is a dialectal feature that opposes West Saxon to other Old English dialects (Hogg 1997:§3, 2000:§6): cf. nom.acc.pl. *weter* in the *Vespasian Psalter* (early Mercian), *wæter* in *Rushworth2* (late South Northumbrian). Historically, blocking is an innovation; although systematic in Ælfric, it was still variable in the early West Saxon dialect of the Alfredian translations: cf. nom.acc.pl. *wætr**u*** ‘water’ with *wolc* ‘cloud’ (underlying /wolkn-u/) in the Hatton manuscript of the *Cura Pastoralis* (Sweet 1871:373/13, 285/24). The development of blocking was a complex process involving input restructuring and the percolation of constraint rankings from level II to level I; see Bermúdez-Otero (2000:§§38-50, in preparation).

The epenthetic vowel in Ælfric’s *wæt**e**ra, -u* is inserted at level II by a process that we shall label ‘anaptyxis’; see Tableau 3. Anaptyxis is fuelled by CONTACT, the constraint that requires that sonority fall across syllable boundaries (Clements 1990, 1992; Vennemann 1988:40; Bat-El 1996:304). At level II, CONTACT outranks DEP-V, the faithfulness constraint that penalizes vowel insertion (McCarthy & Prince 1995). Interestingly, anaptyxis is moderated by PARSE-σ, for the epenthetic vowel must not give rise to an extra unfooted syllable: e.g. nom.acc.gen.pl. *wǣpna* rather than ***wǣp**e**na* ‘weapon’.

Tableau 3. Anaptyxis

Level II		PARSE-σ	CONTACT	DEP-V
[$_\omega$[$_\Sigma$.wæt.].rɑ.]	[$_\omega$[$_\Sigma$.wæ.te.].rɑ.] ☜	(*)		*
	[$_\omega$[$_\Sigma$.wæt.].rɑ.]	(*)	*!	
[$_\omega$[$_\Sigma$.wæːp.].nɑ.]	[$_\omega$[$_\Sigma$.wæːp.].nɑ.] ☜	(*)	*	
	[$_\omega$[$_\Sigma$.wæː.].pe.nɑ.]	(*)*!		*

Historically, anaptyxis entered the language considerably later than apocope. Thus, the stratal ascription of the two processes in Ælfric's grammar is consistent with their relative age: the older process occupies a higher level (see Table 4 below). Indeed, anaptyxis remained variable throughout the Old English period: in Ælfric, for example, one finds both dat.pl. *næglum* and *næg**e**lum* 'nail' (Pope 1967-8:11/36, 11a/138). Additionally, anaptyxis competed with a rival repair strategy of gemination (Campbell 1959:§408): e.g. in the Hatton manuscript of the early West Saxon *Cura Pastoralis*, dat.sg. *wæt**e**re* occurs alongside *wætre* and *wæ**tt**re* (Sweet 1871:293/8, 261/8, 309/7). In late West Saxon, though, the incidence of anaptyxis rose steadily during the tenth century (see §3.3). Throughout, the probability of application was directly governed by sonority and followed the cline -C_r- > -C_l- > -C_n- (see Cosijn 1883:149-151; Brunner 1965:§160).

As shown in Table 3, nom.acc.sg. *wæter* also surfaces with an epenthetic vowel. This epenthetic vowel is inserted by a process traditionally known as 'parasiting', which repairs word-final consonant clusters with rising sonority. Like anaptyxis, parasiting must be assigned to level II; it cannot apply to bare stems at level I, as this would pre-empt optional anaptyxis in forms with level-II endings, incorrectly ruling out unepenthesized variants such as *næglum* (see above). This implies that DEP-V dominates NUC→V and SONPK→σ at level I, but is demoted in the level-II hierarchy. Unlike anaptyxis, however, parasiting is allowed to create unfooted syllables, and so NUC→V and SONPK→σ must outrank PARSE-σ: e.g. nom.sg. [$_\omega$[$_\Sigma$.wæː.].pen.] 'weapon' from underlying /wæːpn/. The operation of parasiting is illustrated in Tableau 4.

Tableau 4. Parasiting

Level II		NUC→V	SONPK→σ	PARSE-σ	DEP-V
[$_\omega$[$_\Sigma$.wæt.]r]	[$_\omega$[$_\Sigma$.wæ.te.]r] ☜				*
	[$_\omega$[$_\Sigma$.wæt.]r]		*!		
	[$_\omega$[$_\Sigma$.wæ.tr̩.]]	*!			
[$_\omega$[$_\Sigma$.wæ:p.]n]	[$_\omega$[$_\Sigma$.wæ:.].pen.]☜			*	*
	[$_\omega$[$_\Sigma$.wæ:p.]n]		*!		
	[$_\omega$[$_\Sigma$.wæ:.].pn̩.]	*!		*	

From a diachronic viewpoint, parasiting is older than anaptyxis but younger than apocope. Indeed, apocope feeds parasiting in those dialects in which the former is not blocked after obstruent+sonorant clusters (Brunner 1965:§148, Anm. 2; Campbell 1959:§574.3; Dresher & Lahiri 1991:279-281; Keyser & O'Neil 1985:141-142; Luick 1964:§304). Again, these facts accord well with the stratal location of apocope and parasiting in Ælfric's grammar (see Table 4 below).

In sum, in Ælfric's dialect the morphophonology of the *a*-stem nouns under consideration is governed by the grammatical system represented in Table 4.

Table 4. The morphophonology of neuter a-*stem nouns in Ælfric*

Level	*Morphology*	*Phonology*
I	neut.nom.acc.pl. /-u/~/-ɑ/	apocope (blocked after obs+son clusters)
II	other *a*-stem noun endings	parasiting, anaptyxis

In Figure 4 we give a representative sample of derivations.

		nom.sg.		*nom.acc.pl.*		*gen.pl.*	
UR		/wætr-/	/werod-/	/wætr-/	/werod-/	/wætr-/	/werod-/
L I	*morphology*	—	—	wætr-ɑ	werod-ɑ	—	—
	phonology	—	—	—	werod	—	—
L II	*morphology*	—	—	—	—	wætr-ɑ	werod-ɑ
	phonology	wæter	—	wæterɑ	—	wæterɑ	—
		wæter	*werod*	*wætera*	*werod*	*wætera*	*weroda*

Figure 4. Derivation of neuter a-*stem nouns in Ælfric.*

3.2 Apocope lost and anaptyxis inverted

As we noted above, Ælfric's works are linguistically conservative. In other late West Saxon texts, the system of alternations depicted in Table 3 shows signs of incipient obsolescence. The gloss to the *Regius Psalter*, for example, contains forms such as *word**u*** 'words' and *weorc**u*** 'works', where the neut.nom.acc.pl. suffix /-u/ surfaces in the apocope environment after monosyllabic heavy stems (Hogg 1997:121). Disyllabic stems are subject to the same development, and so plural forms like *werod**u*** begin to appear; indeed, this item already occurs once in the Ælfrician manuscripts alongside prevalent *werod* (Pope 1967-8:18/405). It is therefore clear that, as late West Saxon evolved, apocope became subject to loss.

There are also good reasons to believe that, at the same time as apocope was lost, stems ending in an obstruent+sonorant cluster underwent restructuring, and anaptyxis was replaced by an inverted rule of syncope. Observe that, with the demise of apocope, the inflectional paradigm of *wæter* becomes identical with that of *scip*, *word*, and *werod*; compare Tables 3 and 5.

Table 5. Neuter a-*stem noun paradigms after the loss of apocope*

nom.acc.sg.	*scip*	*word*	*werod*	*wæter*
nom.acc.pl.	*scipa, -u*	*worda, -u*	*weroda, -u*	*wætera, -u*
gen.pl.	*scipa*	*worda*	*weroda*	*wætera*

The new set of surface forms provides no grounds to posit /wætr-/ as the lexical representation of the stem. One can therefore safely assume that the medial vowels formerly inserted by parasiting and anaptyxis have now been incorporated into the underlying representation: i.e. /wæter-/. In line with Richness of the Base (Prince & Smolensky 1993:§3), a systematic restructuring of underlying representations involves a concomitant change in the grammar:[16] the ranking {NUC→V, SONPK→σ} » DEP-V percolates upwards from level II to level I, thereby prohibiting final obstruent+sonorant clusters in level-I domains. Given this ranking at level I, both input /wætr/ and input /wæter/ yield output [wæter], satisfying Richness of the Base; but input optimization favors /wæter/, this avoids faithfulness violations.

The *Ormulum* (South Lincolnshire, *circa* 1180) provides solid independent evidence that stems ending in an obstruent+sonorant cluster were indeed forbidden in early Middle English – although, admittedly, this evidence casts light on late West Saxon developments only indirectly, as Orm's East Midlands dialect derives from an Anglian ancestor. In the *Ormulum*, the composition of underlying representations can be diagnosed by reference to the application of a level-I process of closed syllable shortening (CSS), which is counterfed by level-II syncope (Bermúdez-Otero 1999:§4.2.3). A sample of derivations is given in Figure 5. (Note that final consonants are extrasyllabic.)

	nom.acc.sg.		*infinitive*	*nom.acc.pl.*	*gen.sg.*	*3sg.pret.*
UR	/tɑːkən-/	/hefən-/	/deːm-/	/tɑːkən-/	/hefən-/	/deːm-/
L I *morphology*	—	—	—	—	—	deːm-də
CSS	—	—	—	—	—	dem.də
L II *morphology*	—	—	deːm-ən	tɑː.kən-əs	he.fən-əs	—
syncope	—	—	—	tɑːk.nəs	hef.nəs	—
SR	[tɑː.kən]	[he.vən]	[deː.mən]	[tɑːk.nəs]	[hev.nəs]	[dem.də]
spelling	*tákenn*	*hefenn*	*demenn*	*tacness*	*heffness*	*demmde*
gloss	‘token’	‘heaven’	‘deem’	‘token’	‘heaven’	‘deem’

Figure 5. Syncope counterfeeds closed syllable shortening in the Ormulum.

In nom.acc.sg. [tɑː.kən], the second vowel must already be present at level I, for the root-vowel would otherwise be subject to closed syllable shortening. Similarly, the plural /tɑːkən-əs/ only undergoes syncope at level II, whence the underapplication of shortening. The ranking {Nuc→V, SonPk→σ} » Dep-V is therefore in force at level I in Orm’s early Middle English grammar.

Returning now to late West Saxon, after the restructuring of stems ending in obstruent+sonorant clusters a form such as *næg**e**lum* ‘nail’ dat.pl. becomes a faithful correspondent of the new underlying representation, i.e. /nægel-/. By implication, variants such as *næglum* are now derived by means of an optional level-II process of syncope, which is the mirror image of anaptyxis. This inverted syncope rule constitutes a genuine innovation. West Saxon already had a process deleting unchecked medial vowels, but this applied only to unfooted syllables preceded by a heavy root-syllable: e.g. nom.acc.sg. *hēafod* ~ dat.sg. *hēafde* ‘head’, from underlying /hæːɑ̯fod-/ (see e.g. Campbell 1959:§351). In contrast, the new inverted syncope rule is quantity-insensitive – like Orm’s – and applies after light syllables.

3.3 As trigger experiences shift, input optimization prompts restructuring

We have seen that late West Saxon underwent two simultaneous changes: the loss of apocope at level I, and the inversion of vowel epenthesis rules at level II. In the light of our discussion in §1 and §2, one is therefore led to surmise that, during the tenth century, children’s trigger experiences shifted, and input optimization caused underlying representations to be restructured accordingly. This section will show that this was indeed the case. The incidence of anaptyxis rose at an accelerated rate during the period. In a subset of cases, this development eliminated vowel~∅ alternations altogether, concealing the epenthetic nature of the vowels involved.

In §3.1 we pointed out that anaptyxis remained variable throughout late West Saxon. Unsurprisingly, insertion applied more frequently to C.r than to C.l or C.n clusters, as the formerincurred the most serious violations of CONTACT. In absolute terms, however, the rate of anaptyxis increased sharply during the tenth century. Bermúdez-Otero & Hogg (1999) sampled the prevalence of anaptyxis in the environment C__r at three different stages and obtained the results summarized in Table 6.

Table 6. Incidence of anaptyxis in the environment C__r *in tenth-century West Saxon*

Source	*Date*	*Incidence*
Alfred	*circa* 900	45%
Leechbook	mid 10^{th} century	54%
Ælfric	*circa* 1000	100%

Strikingly, by Ælfric's time anaptyxis ceased to be optional in C.r clusters.[17] This development had a dramatic effect upon *a*-stem nouns like *wæter*: the nom.acc.sg. continued to undergo parasiting as before, but overtly inflected forms were now obligatorily subject to anaptyxis. As a result, the underlying stem-final cluster /-Cr-/ now surfaced with an epenthetic vowel in every paradigmatic form without exception; see Table 7.

Table 7. Evolution of a*-stem nouns with underlying stem-final obstruent+sonorant clusters in West Saxon*

	Alfred		Ælfric	
nom.acc.sg.	*wæter*	*nægel*	*wæt**e**r*	*nægel*
nom.acc.pl.	*wætru~ wæteru*	*næglas*	*wæt**e**ra, -u*	*næglas~nægelas*
gen.pl.	*wætra~ wætera*	*nægla*	*wæt**e**ra*	*nægla~nægela*
	'water' (neut.)	'nail' (masc.)	'water'	'nail'

Towards the end of the tenth century, therefore, learners encountered no alternations in the paradigm of nouns such as *wæter*. Accordingly, they were led to posit /wæter-/ as the representation of the stem in the input to level II; cf. Figure 4. The PLD no longer provided direct evidence against this default option, whereas alternative /wætr-/ would have caused violations of DEP-V throughout the paradigm. The input to level II was accordingly restructured, and stem-final obstruent+sonorant clusters eliminated. Nouns with stem-final /-Cl-/ and /-Cn-/ clusters still showed alternations (often only variably); but this evidence was insufficient to prevent restructuring, probably because the restructured input representations coincided with the surface realization of the stem in the nom.acc.sg., that constituted the basic form in the paradigm (Lahiri 1982, Lahiri & Dresher 1983-84, Dresher 2000; see §1.2). Concomitantly, anaptyxis was replaced by an inverted process of syncope. In turn, the reanalysis of the input to level II was attended with constraint reranking at level I, with the hierarchy {NUC→V, SONPK→σ} » DEP-V climbing to the highest stratum in the grammar. Ultimately, restructuring reached underlying representations by input optimization with respect to the output of level I.

Interestingly, this development had a knock-on effect on level-I mappings. Following restructuring, the absence of apocope in nom.acc.pl. *wætera, -u* could no longer be imputed to blocking after obstruent+sonorant clusters. By the same token, the presence or absence of the nom.acc.pl. suffix [-u]~[-ɑ] became a lexical idiosyncrasy, as a new class of neuter *a*-stem nouns emerged in which the suffix failed to delete in an unfooted syllable; see Figure 6. Thus, apocope was lost as a phonological regularity, paving the way for the restoration of the suffix in environments that had previously induced deletion.

```
  μ  μ                  μ  μ
  |  |                  |  |
[.we.ro.].du.    but   [.wæ.te.].ru.
          ↓
          ∅
```

Figure 6. After restructuring, apocope becomes a lexical idiosyncrasy.

We should observe that apocope had survived for centuries despite considerable opacity. Notably, it underapplied in heavy neuter *ja*-stems as a result of the syncope of the thematic vowel **-i-*: e.g. **wī.ti.u* > *wī.tu* 'punishment' nom.acc.pl. (see e.g. Hogg 1997:121). In this sense, the restructuring of *a*-stems ending in obstruent+sonorant clusters appears to have been enough to tip the balance. Note, moreover, that with the loss of apocope all evidence of a phonological split among *a*-stem noun inflections disappears. For what it is worth, the evidence of the *Ormulum* shows that in Orm's early Middle English dialect all nominal inflection was assigned to level II (Bermúdez-Otero 1999:§4.2.3).

In sum, the accelerated rise in the incidence of anaptyxis in late West Saxon caused a crucial shift in the trigger experience of children. In stems containing a short root-vowel followed by a /Cr/ cluster, vowel~∅ alternations were altogether eliminated. As a result, epenthetic vowels were reanalyzed as underlying, in compliance with the requirements of input optimization. In turn, input restructuring disrupted the prosodic conditioning of apocope, which was accordingly lost.

4. PROSPECTS: THE LIFE CYCLE OF CONSTRAINT RANKINGS

During the 1980s and early 1990s, work in rule-based LP sought to effect a synthesis of disparate observations in synchronic and diachronic phonology. From a synchronic viewpoint, the mode of application of a phonological rule was seen as following from its position in a hierarchy of grammatical levels; notably, ascription to a particular level determined the morphosyntactic domain of the rule. Diachronically, rules were characterized as gradually percolating from lower to higher levels. This claim captured major properties of the life cycle of phonological processes, such as the rise of opacity, morphologically conditioned misapplication, and ultimate loss through lexicalization (Kiparsky 1988, 1995; Harris 1989; McMahon 1991, 2000a; Zec 1993).

Initially, the advent of OT appeared to compromise this synthesis of synchronic and diachronic evidence. In particular, Correspondence Theory (McCarthy & Prince 1995) has fostered a research program that rejects the interleaving of phonology with morphology and syntax, favoring instead a fully parallel treatment of the underlying-surface mapping. Nonetheless, an alternative strand of research has revealed major flaws in novel applications of correspondence such as OO-identity (e.g. Benua 1997) and Sympathy Theory (McCarthy 1999); this work demonstrates the need to reinstate cyclicity and level segregation as the major mechanisms governing the phonology-morphology interface (Orgun 1996a,b, Bermúdez-Otero 1999, Kiparsky 2000, forthcoming; see also Gess this volume, Jacobs this volume). Within the framework of interleaved OT, therefore, LP's insights into the life cycle of phonological patterns can be preserved and further developed.

In this connection, this chapter has shown that one of interleaved OT's major contributions to diachronic research lies in dissolving the conceptual difficulties that surround input restructuring in rule-based LP. As we have seen, the synchronic architecture of rule-based LP calls for a powerful learner pursuing a strategy of lexicon minimization. The evidence of analogical change, however, indicates that learners only countenance input-output disparity when presented with alternations. Interleaved OT resolves this tension by replacing lexical minimality with input optimization. As a result, the formal requirements of optimality and the functional demands of learnability converge.

University of Newcastle (Bermúdez-Otero), University of Manchester (Hogg)

5. NOTES

[1] Parts of the research reported in this chapter were presented at the Linguistics Seminar, Department of Linguistics, University of Manchester, May 5, 1998 (Bermúdez-Otero); at the XIV International Conference on Historical Linguistics, Vancouver, August 9-13, 1999 (Bermúdez-Otero & Hogg); and at the XI International Conference on English Historical Linguistics, Santiago de Compostela, September 7-11, 2000 (Bermúdez-Otero). We thank the audiences on all three occasions for their comments. We have also received helpful suggestions from Randall Gess, Bryan Gick, Eric Holt, April McMahon, and

Charles Reiss. We gratefully acknowledge the support of a British Academy Postdoctoral Fellowship to Ricardo Bermúdez-Otero and of the Leverhulme Trust to Richard Hogg.

[2] Prince & Smolensky's original term is 'lexicon optimization', which is perfectly adequate for strictly parallel implementations of OT. In interleaved or stratal models, however, the principle can be seen as regulating input-output relationships at each level of the grammar; see Bermúdez-Otero (1999:§3.3.3.2) and below. In such a context, therefore, the term 'input optimization' is preferable, for input forms coincide with underlying representations only at the highest grammatical level.

[3] Stress may be used to illustrate the distinction between the overt and covert components of a surface representation: one can assume that, for any given stimulus, the parser yields a representation of observable prominence relations (perhaps in the shape of a primitive grid; see Dresher 1996:253); in contrast, covert relationships such as foot constituency and foot headship will have to be assigned by the grammar. The task of learning such covert structure is a vital element in the logical problem of language acquisition: see e.g. Dresher (1994:72, 1996:251), Tesar & Smolensky (2000:6ff.). Note, incidentally, that Tesar & Smolensky (2000) use the term 'interpretive parsing' for the mapping of an overt form onto a full structural description, complete with covert structure. It is vital not to confuse this process with parsing understood – as here – as the mapping from an acoustic signal to an overt phonological representation. Similarly, Fitzpatrick & Wheeldon (2000) use the term 'feature extraction' for what we call 'parsing', and reserve the term 'parsing' for the computation of covert structure, involving grammatical derivations.

[4] A slightly different type of hypocorrection may affect the language-specific gradient patterns generated in the implementation module of the grammar (see Clements and Hertz 1996 for the nature of this module and Myers 2000 for the categorical/gradient distinction). The learner may successfully recover the output intended by the speaker, but may underestimate the speech rate, wrongly interpreting a fast-speech form as representing normal speech. This may initiate a feedback loop leading to categorical change in higher grammatical modules. See Gess (this volume) for germane discussion.

[5] Less compelling is Jakobson's (1929) appeal to universal implicational tendencies as engines of change. Essentially, Jakobson suggests that, given an implicational tendency $P \rightarrow Q$, then the introduction of P in a system having $\neg Q$ will create pressure in the direction of Q. The logic of this argument is fatally flawed; see Hawkins (1979:§5; 1983:§5.4.1) and McMahon (1994:§6.2) for a critique of similar proposals in diachronic syntax.

[6] The challenge is compounded by the fact that priming effects operate probabilistically: diachronic structure-preservation requirements formulated in absolute terms have failed to withstand empirical scrutiny. *Pace* de Chene & Anderson (1979) and Hayes (1989), for example, consonant deletion can trigger compensatory vowel lengthening without priming from phonological contrasts in either vowel length or syllable quantity (Hock 1986, Hayes 1989, Morin 1992, Lin 1997, Gess 1998).

[7] In a more speculative vein, Bermúdez-Otero (1999:199) suggests that, if the learning process is deterministic (Marcus 1980, Berwick 1985), then grammatical properties acquired relatively early will constrain the range of options available further down the learning path (Lightfoot 1989), potentially blocking certain types of phonologization.

[8] This suggestion is often accompanied by the claim that /r/-intrusion did not in fact arise by inversion, but this is by no means a necessary inference.

[9] 'Nonrandom' is the operative word in this statement. Charles Reiss (personal communication) points out that inversion may be supposed to affect a rather small percentage of coda deletion rules. If so, the failure of [ɗ] to be attested as a hiatus breaker need not have anything to do with markedness, but may simply reflect the fact that implosives (and implosive deletion rules) are statistically rare. However, in the light of our observation that most inverted hiatus rules appear to insert coronal sonorants, Reiss's argument entails the peculiar prediction that deletion rules targeting rhymal coronal sonorants are much more frequent than processes deleting any other type of coda consonant.

[10] Like 'hypocorrection' above, the term 'hypercorrection' here designates a type of misparsing (Ohala 1989, 1992, 1993). This should not be confused with hypercorrection in the sociolinguistic sense, which involves overgeneralization in the suppression of a stigmatized variant.

[11] Further research should determine how these sorting processes take place. At this stage, however, it is natural to surmise that sorting will be vulnerable to interference from frequency, proximity, and priming effects; see Reiss (this volume:§9.2). This could account for observed cases of sporadic 'four-part' analogy that increase allomorphy and input-output disparity (Vincent 1974, Reiss this volume).

[12] The literature on high vowel deletion is extensive: for handbook descriptions, see e.g. Brunner (1965), Campbell (1959), Hogg (1992), Luick (1964); for linear analyses in the style of SPE, see e.g. Dresher (1978), Kiparsky & O'Neil (1976), Peinovich (1979), Wagner (1969); for nonlinear treatments, see e.g. Bermúdez-Otero (2000, in preparation), Dresher & Lahiri (1991), Hutton (1998b), Idsardi (1994), Keyser & O'Neil (1985).

[13] We indicate ungrammaticality by means of a double asterisk and reserve single asterisks for reconstructed forms.

[14] As suggested to us by Randall Gess (personal communication), one could conceivably deal with the divergent behavior of nom.acc.pl. /-ɑ/ and gen.pl. /-ɑ/ by allowing faithfulness constraints to refer to morphological class, i.e. MAX-V(gen.pl.) » PARSE-σ » MAX-V(nom.acc.pl.), rather than postulating a stratal split. Fukazawa, Kitahara & Ota (1998) make a similar proposal in their analysis of Japanese, allowing faithfulness constraints to be indexed to different lexical classes. This proposal must however be rejected on general grounds. First, the cophonologies associated with different affixes can demonstrably differ in respect of the ranking of markedness constraints. Secondly, this strategy, unlike stratification, fails to yield a comprehensive solution to the opacity problem. See Bermúdez-Otero (1999) for discussion.

[15] This pattern of syllabification can in fact be traced back to the application of Sievers' Law in Indo-European. For an Optimality-Theoretic analysis, see Kiparsky (1998:§7) and Bermúdez-Otero (1999:§3.5.2.2, §3.5.2.5). For empirical corroboration from Old English manuscripts, see Wetzel (1981) and Lutz (1986).

[16] In stratal OT, Richness of the Base applies only to the input to the first stratum, i.e. to underlying representations. Systematic gaps in the input to noninitial strata are accounted for by constraints holding at higher levels. See Bermúdez-Otero (1999).

[17] In the pursuit of a standard *Schriftsprache*, Ælfric is known to have studiously avoided orthographic variation. Hence, sociolinguistic factors may partly account for the absolute regularity of Ælfric's data in Table 6. Recall, however, that the Ælfrician manuscripts do display variation in respect of anaptyxis in clusters other than C.r: e.g. *næglum* ~ *næg**e**lum* (see §3.1).

6. REFERENCES

Anderson, Stephen R. (1981). Why phonology isn't "natural". *Linguistic Inquiry* 12: 493-539.

Archangeli, Diana & Douglas Pulleyblank (1994). *Grounded phonology.* Cambridge, MA: The MIT Press.

Ash, Sharon (1982). The vocalization of intervocalic /l/ in Philadelphia. *SECOL Review* 6: 162-175.

Baković, Eric. (1999). Deletion, insertion, and symmetrical identity. *Harvard Working Papers in Linguistics* 7.

Bat-El, Outi (1996). Selecting the best of the worst: the grammar of Hebrew blends. *Phonology* 13: 283-328.

Benua, Laura (1997). *Transderivational identity: phonological relations between words*. PhD dissertation, University of Massachusetts, Amherst.

Bermúdez-Otero, Ricardo (1999). *Constraint interaction in language change.* [*Opacity and globality in phonological change.*] PhD dissertation, University of Manchester & Universidad de Santiago de Compostela.

Bermúdez-Otero, Ricardo (2000). The morphophonology of *a*-stem nouns in West Saxon. Handout of paper presented at the XI International Conference on English Historical Linguistics, Santiago de Compostela, 11 September 2000.

Bermúdez-Otero, Ricardo (in preparation). *The life cycle of constraint rankings: studies in early English morphophonology*. Ms., University of Newcastle upon Tyne.

Bermúdez-Otero, Ricardo & Kersti Börjars (in preparation). Phonology, syntax, and the problem of grounding. In Patrick Honeybone & Ricardo Bermúdez-Otero (eds). Linguistic knowledge: perspectives from phonology and from syntax. *Lingua*: Special Issue.

Bermúdez-Otero, Ricardo & Richard M. Hogg (1999). High Vowel Deletion in late Old English: optimality, opacity and learnability. Handout of paper presented at the XIV International Conference on Historical Linguistics, Vancouver, 9-13 August 1999.

Berwick, Robert C. (1985). *The acquisition of syntactic knowledge.* Cambridge, MA: MIT Press.

Blevins, Juliette (1997). Rules in Optimality Theory: two case studies. In Roca, Iggy (ed.) *Derivations and constraints in phonology*. Oxford: Clarendon Press. 227-260.

Boersma, Paul (1998). *Functional phonology: formalizing the interactions between articulatory and perceptual drives*. The Hague: Holland Academic Graphics.

Booij, Geert (1996). Cliticisation as prosodic integration: the case of Dutch. *The Linguistic Review* 13: 219-242.

Broadbent, Judith (1991). Linking and intrusive *r* in English. *University College London Working Papers in Linguistics* 3: 281-302.

Brunner, Karl (1942). *Altenglische Grammatik nach der angelsächsischen Grammatik von Eduard Sievers neubearbeitet.* 3rd edn (1965) Tübingen: Max Niemeyer.

Bybee, Joan L. (1998). The phonology of the lexicon: evidence from lexical diffusion. In Michael Barlow & Suzanne Kemmer (eds). *Usage-based models of language*. Stanford: CSLI. 65-85.

Campbell, Alistair (1959). *Old English grammar.* Oxford: Oxford University Press.

Clayton, Mary L. (1976). The redundance of underlying morpheme-structure conditions. *Language* 52: 295-313.

Clements, George N. (1990). The role of the sonority cycle in core syllabification. In John Kingston & Mary Beckman (eds). *Papers in laboratory phonology, 1: Between the grammar and physics of speech*. Cambridge: Cambridge University Press. 283-333.

Clements, George N. (1992). The sonority cycle and syllable organization. In Wolfgang U.Dressler, Hans C. Luschützky, Oskar E. Pfeiffer & John R. Rennison (eds) (1992). *Phonologica 1988. Proceedings of the Sixth International Phonology Meeting*. Cambridge: Cambridge University Press. 63-76.

Clements, George N. & Susan R. Hertz (1996). An integrated approach to phonology and phonetics. In Jacques Durand & Bernard Laks (eds). *Current trends in phonology: models and methods*. Salford: European Studies Research Institute. Vol. 1: 143-173.

Cole, Jennifer. (1995). The cycle in phonology. In Goldsmith, John A. (ed.) *The handbook of phonological theory*. Cambridge, MA: Blackwell Publishers. 70-113.

Cosijn, P. J. (1888). *Altwestsächsische Grammatik*. Vol. 1 (1883), Vol. 2: *Die Flexion* (1886). Den Haag: Martinus Nijhoff.

Croft, William (1995). Autonomy and functionalist linguistics. *Language* 71: 490-532.

Cutler, Anne & Jacques Mehler (1993). The periodicity bias. *Journal of Phonetics* 21. 103-108.

de Chene, E. Brent & Stephen R. Anderson (1979). Compensatory lengthening. *Language* 55: 505-535.

Donegan, Patricia (1993). On the phonetic basis of phonological change. In Jones, Charles (ed.) *Historical linguistics: problems and perspectives*. London: Longman 98-130.

Dresher, B. Elan (1978). *Old English and the theory of phonology.* PhD dissertation, University of Massachusetts, Amherst. Published (1985) New York: Garland.

Dresher, B. Elan (1993). The chronology and status of Anglian smoothing. In Sharon Hargus & Ellen M. Kaisse (eds). *Studies in Lexical Phonology*. (Phonetics and Phonology 4) San Diego, CA: Academic Press. 325-341.

Dresher, B. Elan (1994). Acquiring stress systems. In Ristad, Eric Sven (ed.) *Language computations. DIMACS Workshop on Human Language, March 20-22, 1992*. (DIMACS Series in Discrete Mathematics and Theoretical Computer Science 17). Providence, RI: American Mathematical Society. 71-92.

Dresher, B. Elan (1996). Learnability and phonological theory. In Jacques Durand & Bernard Laks (eds). *Current trends in phonology: models and methods*. Salford: European Studies Research Institute. Vol. 1: 245-266.

Dresher, B. Elan (1998). Paradigmatic levelling as a response to phonological incoherence. Handout of plenary paper presented at the X International Conference on English Historical Linguistics, Manchester, 21-26 August 1998.

Dresher, B. Elan (2000). Analogical levelling of vowel length in West Germanic. In Lahiri, Aditi (ed.) *Analogy, levelling, markedness: principles of change in phonology and morphology*. Berlin: Mouton de Gruyter.

Dresher, B. Elan & Aditi Lahiri (1991). The Germanic foot: metrical coherence in Old English. *Linguistic Inquiry* 22: 251-286.

Elugbe, Ben Ohi (1982). The vowels of Proto-Ẹdoid. *Journal of the Linguistic Association of Nigeria* 1: 107-115.

Elugbe, Ben Ohi (1989). Edoid. In Bendor-Samuel, John (ed.) *The Niger-Congo languages: a classification and description of Africa's largest language family*. Lanham: University Press of America. 291-304.

Fitzpatrick, Jennifer & Linda R. Wheeldon (2000). Phonology and phonetics in psycholinguistic models of speech perception. In Noel Burton-Roberts, Philip Carr & Gerard Docherty (eds). *Phonological knowledge: conceptual and empirical issues*. Oxford: Oxford University Press. 131-160.

Fodor, Jerry A. (1983). *The modularity of mind: an essay on faculty psychology*. Cambridge, MA: The MIT Press.

Fukazawa, Haruka, Mafuyu Kitahara & Mitshiko Ota (1998). Lexical stratification and ranking invariance in constraint-based grammars. *Papers from the Annual Regional Meeting, Chicago Linguistics Society* 34. (ROA-267-0698, Rutgers Optimality Archive, http://roa.rutgers.edu.)

Gerfen, Chip (2001). A critical view of licensing by cue: codas and obstruents in Eastern Andalusian Spanish. In Lombardi, Linda (ed.) *Segmental phonology in Optimality Theory: constraints and representations*. Cambridge: Cambridge University Press. 183-205.

Gess, Randall (1998). Compensatory lengthening and structure preservation revisited. *Phonology* 15: 353-366.

Gess, Randall (this volume). On re-ranking and explanatory adequacy in a constraint-based theory of phonological change.

Gick, Bryan (1999). A gesture-based account of intrusive consonants in English. *Phonology* 16: 29-54.

Gick, Bryan (2002). An X-ray investigation of pharyngeal constriction in American English schwa. *Phonetica* 59: 38-48.

Gick, Bryan, A. Min Kang & D. H. Whalen (2002). MRI evidence for commonality in the post-oral articulations of English vowels and liquids. *Journal of Phonetics* 30: 357-372.

Giegerich, Heinz J. (1999). *Lexical strata in English: morphological causes, phonological effects*. Cambridge: Cambridge University Press.
Hale, Mark & Charles Reiss (1998). Formal and empirical arguments concerning phonological acquisition. *Linguistic Inquiry* 29: 656-683.
Hale, Mark & Charles Reiss (2000). Phonology as cognition. In Noel Burton-Roberts, Philip Carr & Gerard Docherty (eds). *Phonological knowledge: conceptual and empirical issues*. Oxford: Oxford University Press. 161-184.
Halle, Morris & William J. Idsardi (1997). *r*, hypercorrection, and the Elsewhere Condition. In Roca, Iggy (ed.) *Derivations and constraints in phonology*. Oxford: Clarendon Press. 331-348.
Harris, John (1989). Towards a lexical analysis of sound change in progress. *Journal of Linguistics* 25: 35-56.
Harris, John (1994). *English sound structure*. Oxford: Blackwell.
Harrison, K. David & Abigail Kaun (2000). Pattern-responsive lexicon optimization. *Papers from the Annual Meeting of the North East Linguistic Society* 30.
Haspelmath, Martin (1999). Optimality and diachronic adaptation. *Zeitschrift für Sprachwissenschaft* 18: 180-205.
Hawkins, John A. (1979). Implicational universals as predictors of word order change. *Language* 55: 618-648.
Hawkins, John A. (1983). *Word order universals*. New York: Academic Press.
Hayes, Bruce (1989). Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20: 253-306.
Hayes, Bruce (1999). Phonetically-driven phonology: the role of Optimality Theory and Inductive Grounding. In Michael Darnell, Edith Moravcsik, Frederick J. Newmeyer, Michael Noonan & Kathleen Wheatley (eds). *Functionalism and formalism in linguistics*. Vol. 1: *General papers*. Amsterdam: John Benjamins. 243-285.
Hock, Hans Heinrich (1986). Compensatory lengthening: in defense of the concept 'mora'. *Folia Linguistica Historica* 20: 431-460.
Hogg, Richard M. (1992). *A grammar of Old English*. Vol. 1: *Phonology*. Oxford: Basil Blackwell.
Hogg, Richard M. (1997). The morphology and dialect of Old English disyllabic nouns. In Raymond Hickey & Stanisław Puppel (eds). *Language history and linguistic modelling: a festschrift for Jacek Fisiak on his 60th birthday*. Berlin: Mouton de Gruyter. 113-126.
Hogg, Richard M. (2000). On the (non-)existence of High Vowel Deletion. In Aditi Lahiri (ed.) *Analogy, levelling,markedness: principles of change in phonology and morphology*. Berlin: Mouton de Gruyter. 353-376.
Hume, Elizabeth & Keith Johnson (2001). A model of the interplay of speech perception and phonology. In Elizabeth Hume & Keith Johnson (eds). *The role of speech perception in phonology*. New York: Academic Press. 3-26.
Hurford, James R. (1990). Nativist and functional explanations in language acquisition. In Roca, Iggy M. (ed.). *Logical issues in language acquisition*. Dordrecht: Foris Publications. 85-136.
Hutton, John (1998a). The development of secondary stress in Old English. In Hogg, Richard M. & Linda van Bergen (eds). *Historical linguistics 1995: selected papers from the 12th International Conference on Historical Linguistics, Manchester, August 1995*. Vol. 2: *Germanic linguistics*. Amsterdam: John Benjamins Publishing Company. 115-130.
Hutton, John (1998b). Stress in Old English, *giet ongean*. *Linguistics* 36: 847-885.

Idsardi, William J. (1994). Open and closed feet in Old English. *Linguistic Inquiry* 25: 522-533.
Inkelas, Sharon (1995). The consequences of optimization for underspecification. *Proceedings of the North East Linguistic Society* 25: 287-302.
Inkelas, Sharon (2000). Phonotactic blocking through structural immunity. In Stiebels, Barbara & Dieter Wunderlich (eds). *Lexicon in focus.* (Studia Grammatica 45). Berlin: Akademie Verlag. 7-40.
Itô, Junko, R. Armin Mester & Jaye Padgett (1995). Licensing and underspecification in Optimality Theory. *Linguistic Inquiry* 26: 571-613.
Jacobs, Haike (this volume). The emergence of quantity-sensitivity in Latin: secondary stress, Iambic Shortening and theoretical implications for 'mixed' stress systems.
Jakobson, Roman (1929). Remarques sur l'évolution phonologique du russe comparée à celle des autres langues slaves. *Travaux du Cercle Linguistique de Prague* 2. Reprinted (1962) in *Selected writings*. Vol. I: *Phonological studies*. The Hague: Mouton & Co. 7-116.
Kager, René (1999). *Optimality theory.* Cambridge: Cambridge University Press.
Kaisse, Ellen M. & Patricia A. Shaw (1985). On the theory of Lexical Phonology. *Phonology Yearbook* 2: 1-30.
Kean, Mary-Louise (1974). The strict cycle in phonology. *Linguistic Inquiry* 5: 179-203.
Kenstowicz, Michael & Charles W. Kisseberth (1977). *Topics in phonological theory*. New York: Academic Press.
Keyser, Samuel Jay & Wayne O'Neil (1985). *Rule generalization and opacity in language change.* Dordrecht: Foris.
King, Robert D. (1969). *Historical linguistics and generative grammar.* Englewood Cliffs, NJ: Prentice-Hall.
King, Robert D. (1973). Rule insertion. *Language* 49: 551-578.
King, Robert D. (1976). The history of final devoicing in Yiddish. Bloomington, IN: Indiana University Linguistics Club.
Kiparsky, Paul (1965). *Phonological change*. PhD dissertation, MIT.
Kiparsky, Paul (1968). Linguistic universals and language change. In Bach, E. & R. Harms (eds). *Universals in linguistic theory.* New York: Holt, Rinehart and Winston. Reprinted in Kiparsky, P. (1982). *Explanation in phonology.* Dordrecht: Foris. 13-43.
Kiparsky, Paul (1973). Abstractness, opacity and global rules. In Fujimura, Osamu (ed.) *Three dimensions of linguistic theory*. Tokyo: TEC Company. 57-86.
Kiparsky, Paul (1982a). From Cyclic Phonology to Lexical Phonology. In van der Hulst, Harry & Norval Smith (eds). *The structure of phonological representations.* Dordrecht: Foris. Part I: 131-75.
Kiparsky, Paul (1982b). Lexical morphology and phonology. In In-Seok Yang (ed.) *Linguistics in the morning calm: selected papers from SICOL-1981*. Seoul: Hanshin. 3-91.
Kiparsky, Paul (1984). On the lexical phonology of Icelandic. In Claes-Christian Elert, Irene Johansson & Eva Strangert (eds.) *Nordic prosody III: papers from a symposium*. (Acta Universitatis Umensis: Umeå Studies in the Humanities 59). Stockholm: Almqvist & Wiksell. 135-64.
Kiparsky, Paul (1988). Phonological change. In Newmeyer, Frederick J. (ed.) *Linguistics: the Cambridge survey*. Cambridge: Cambridge University Press. 363-415.
Kiparsky, Paul (1993). Blocking in nonderived environments. In Hargus, Sharon & Ellen M. Kaisse (eds). *Studies in Lexical Phonology.* (Phonetics and Phonology 4) San Diego: Academic Press. 277-313.

Kiparsky, Paul (1995). The phonological basis of sound change. In Goldsmith, John A. (ed.) *The handbook of phonological theory*. Cambridge, MA: Blackwell Publishers. 640-670.

Kiparsky, P. (1998). Sievers' Law as prosodic optimization. In Jay Jasanoff, H. Craig Melchert & Lisi Oliver (eds). *Mír Curad. Studies in honor of Calvert Watkins.* Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft.

Kiparsky, Paul (2000). Opacity and cyclicity. *The Linguistic Review* 17: 351-365.

Kiparsky, Paul (forthcoming). *Paradigm effects and opacity*. Stanford: CSLI Publications.

Kiparsky, Paul & Wayne O'Neil (1976). The phonology of Old English inflections. *Linguistic Inquiry* 7: 527-557.

Koefoed, G. (1974). On formal and functional explanation: some notes on Kiparsky's 'Explanation in phonology'. In Anderson, John M. & Charles Jones (eds). *Historical Linguistics II: theory and description in phonology. Proceedings of the First International Conference on Historical Linguistics, Edinburgh 2nd - 7th September 1973.* Amsterdam: North-Holland. 276-293.

Kuryłowicz, Jerzy (1949). La nature des procès dites 'analogiques'. *Acta Linguistica* 5: 15-37.

Labov, William (1994). *Principles of language change.* Vol 1: *Internal factors.* Oxford: Basil Blackwell.

Lahiri, Aditi (1982). *Theoretical implications of analogical change: evidence from Germanic languages.* PhD dissertation, Brown University.

Lahiri, Aditi & B. Elan Dresher (1983-84). Diachronic and synchronic implications of declension shifts. *The Linguistic Review* 3: 141-163.

Lahiri, Aditi & B. Elan Dresher (1999). Open syllable lengthening in West Germanic. *Language* 75: 678-719.

Lass, Roger & John M. Anderson (1975). *Old English phonology*. Cambridge: Cambridge University Press.

Lightfoot, David (1989). The child's trigger experience: degree-0 learnability. *Behavioural and Brain Sciences* 12: 321-375.

Lightfoot, David (1998). The development of grammars. *Glot International* 3: 3-8.

Lightfoot, David (1999). *The development of language: acquisition, change, and evolution.* Oxford: Blackwell.

Lin, Yen-Hwei (1997). Syllabic and moraic structures in Piro. *Phonology* 14: 403-436.

Lloret-Romanyach, Maria-Rosa (1988). *Gemination and vowel length in Oromo morphophonology.* PhD dissertation, Indiana University.

Löfqvist, Anders, Thomas Baer, Nancy S. McGarr & Robin Seider Story (1989). The cricothyroid muscle in voicing control. *Journal of the Acoustical Society of America* 50: 1166-79.

Luick, Karl (1964). *Historische Grammatik der englischen Sprache.* Edited by F. Wild & H. Koziol. Part 1 (1914-1921), Part 2 (1929-1940). Stuttgart: Bernhard Tauchnitz / Oxford: Basil Blackwell.

Lutz, Angelika (1986). The syllabic basis of word division in Old English manuscripts. *English Studies* 3: 193-210.

Mańczak, Witold (1958). Tendences générales des changements analogiques. *Lingua* 7: 298-325, 387-420.

Marcus, Mitchell P. (1980). *A theory of syntactic recognition for natural language.* Cambridge, MA: MIT Press.

Mascaró, Joan (1976). *Catalan phonology and the phonological cycle*. PhD dissertation, MIT.

Matisoff, James A. (1973). Tonogenesis in Southeast Asia. In Larry M. Hyman (ed.) *Consonant types and tone.* Los Angeles: University of Southern California. 97-114.

McCarthy, John J. (1993). A case of surface constraint violation. *Canadian Journal of Linguistics* 38: 169-195.
McCarthy, John J. (1999). Sympathy and phonological opacity. *Phonology* 16: 331-399.
McCarthy, John J. & Alan Prince (1995). Faithfulness and reduplicative identity. In Jill Beckman, Laura Walsh-Dickey & Suzanne Urbanczyk (eds). *Papers in Optimality Theory.* (University of Massachusetts Occasional Papers in Linguistics 18) Amherst: Graduate Linguistic Student Association. 249-384.
McMahon, April M. S. (1991). Lexical phonology and sound change: the case of the Scottish Vowel Length Rule. *Journal of Linguistics* 27: 29-53.
McMahon, April M. S. (1994). *Understanding language change.* Cambridge: Cambridge University Press.
McMahon, April M. S. (2000a). *Lexical phonology and the history of English.* Cambridge: Cambridge University Press.
McMahon, April M. S. (2000b). *Change, chance, and optimality.* Oxford: Oxford University Press.
McMahon, April M. S., Paul Foulkes & Laura Tollfree (1994). Gestural representation and Lexical Phonology. *Phonology* 11: 277-316.
Morin, Yves-Charles (1992). Phonological interpretations of historical lengthening. Paper presented at the Seventh International Phonology Meeting, Krems.
Myers, Scott (2000). Boundary disputes: the distinction between phonetic and phonological sound patterns. In Noel Burton-Roberts, Philip Carr & Gerard Docherty (eds). *Phonological knowledge: conceptual and empirical issues*. Oxford: Oxford University Press. 245-272.
Nigoyi, Partha & Robert C. Berwick (1997). Evolutionary consequences of language learning. *Linguistics and Philosophy* 20: 697-719.
Ohala, John J. (1989). Sound change is drawn from a pool of synchronic variation. In Leiv Egil Breivik & Ernst Håkon Jahr (eds). *Language change: contributions to the study of its causes.* Berlin: Mouton de Gruyter. 173-198.183.
Ohala, John J. (1992). What's cognitive, what's not, in sound change. In Günter Kellermann & Michael D. Morrissey (eds). *Diachrony within synchrony: language history and cognition. Papers from the International Symposium at the University of Duisburg, 26-28 March 1990.* Frankfurt am Main: Peter Lang. 309-355.
Ohala, John J. (1993). The phonetics of sound change. In Jones, Charles (ed.) *Historical linguistics: problems and perspectives.* London/New York: Longman. 237-278.
Orgun, Cemil Orhan (1996a). *Sign-Based Morphology and Phonology, with special attention to Optimality Theory*. PhD dissertation, University of California, Berkeley.
Orgun, Cemil Orhan (1996b). Sign-Based Morphology: a declarative theory of phonology-morphology interleaving. Ms., University of California, Berkeley. (ROA-122-0496, Rutgers Optimality Archive, http://ruccs.rutgers.edu/roa.html). Condensed version published in Ben Hermans & Marc van Oostendorp (eds) (1999). *The derivational residue in phonological Optimality Theory*. Amsterdam: John Benjamins Publishing Company. 247-267.
Peinovich, M. P. (1979). *Old English noun morphology: a diachronic study.* (North-Holland Linguistic Series 41). Amsterdam: North-Holland Publishing Company.
Pope, John C. (1967-8). *Homilies of Ælfric: a supplementary collection*. 2 vols. (Early English Text Society 259-260.) London: Oxford University Press.

Prince, Alan & Paul Smolensky (1993). *Optimality Theory: constraint interaction in generative grammar*. Ms., Rutgers University Center for Cognitive Science Technical Report #2.

Pulleyblank, Douglas (1997). Optimality Theory and features. In Diana Archangeli & D. Terence Langendoen (eds). *Optimality Theory: an overview*. Oxford: Blackwell. 59-101.

Pulleyblank, Douglas & William J. Turkel (1996). Optimality Theory and learning algorithms: the representation of recurrent featural asymmetries. In Jacques Durand & Bernard Laks (eds.) *Current trends in phonology: models and methods*. Salford: European Studies Research Institute. 655-686.

Pylyshyn, Zenon W. (1984). *Computation and cognition: toward a foundation for cognitive science*. Cambridge, MA: The MIT Press.

Reiss, Charles (this volume). Language change without constraint reranking.

Selkirk, Elizabeth O. (1984). On the major class features and syllable theory. In Mark Aronoff & Richard T. Oehrle (eds). *Language sound structure: Studies in phonology presented to Morris Halle by his teacher and students*. Cambridge, MA: MIT Press. 107-136.

Sproat, Richard & Osamu Fujimura (1993). Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics* 21: 291-311.

Steriade, Donca (1995). Underspecification and markedness. In Goldsmith, John A. (ed.) *The handbook of phonological theory*. Cambridge, MA: Blackwell Publishers. 114-174.

Steriade, Donca (1997). Phonetics in phonology: the case of laryngeal neutralization. Ms., University of California, Los Angeles.

Suzuki, Seiichi (1988). On the 1 sg. pres. ind. ending *-u* and High Vowel Deletion in Anglian and other West Germanic languages. *Indogermanische Forschungen* 93: 210-224.

Svantesson, Jan-Olof (1983). *Kammu phonology and morphology*. (Travaux de l'institut de linguistique de Lund 18). Lund: Gleerup.

Svantesson, Jan-Olof (1989). Tonogenetic mechanisms in northern Mon-Khmer. *Phonetica* 46: 60-79.

Sweet, Henry (1871). *King Alfred's West-Saxon version of Gregory's Pastoral care*. (Early English Text Society Original Series 45, 50). Reprinted (1958). London: Oxford University Press.

Tesar, Bruce & Paul Smolensky (2000). *Learnability in Optimality Theory*. Cambridge, MA: The MIT Press.

Vennemann, Theo (1972). Rule inversion. *Lingua* 29: 209-242.

Vennemann, Theo (1988). *Preference laws for syllable structure and the explanation of sound change, with special reference to German, Germanic, Italian and Latin*. Berlin: Mouton de Gruyter.

Vincent, Nigel (1974). Analogy reconsidered. In Anderson, John M. & Charles Jones (eds). *Historical Linguistics II: theory and description in phonology. Proceedings of the First International Conference on Historical Linguistics, Edinburgh 2nd - 7th September 1973.* Amsterdam: North-Holland. 427-445.

Wagner, Karl Heinz (1969). *Generative grammatical studies in the Old English language*. Heidelberg: Julius Groos.

Weinreich, Uriel, William Labov & Marvin I. Herzog (1968). Empirical foundations for a theory of language change. In Lehmann, Winfred P. & Yakov Malkiel (eds.) *Directions for historical linguistics: a symposium.* Austin: University of Texas Press. 95-195.

Wetzel, Claus-Dieter (1981). *Die Worttrennung am Zeilenende in altenglischen Handschriften.* Frankfurt: Lang.

Wilbur, Ronnie (1973). *The phonology of reduplication.* PhD dissertation, University of Illinois, Urbana-Champaign.
Winitz, Harris, M. E. Scheib & J. A. Reeds (1972). Identification of stops and vowels for the burst portion of /p,t,k/ isolated from conversational speech. *Journal of the Acoustical Society of America* 51: 1309-1317.
Wurzel, Wolfgang U. (1989). *Inflectional morphology and naturalness.* Dordrecht: Kluwer.
Yip, Moira (1996). Lexicon optimization in languages without alternations. In Jacques Durand & Bernard Laks (eds). *Current trends in phonology: models and methods*. Vol. 2. Salford: European Studies Research Institute. 759-790.
Zec, Draga (1993). Rule domains and phonological change. In Sharon Hargus & Ellen M. Kaisse (eds). *Studies in Lexical Phonology*. (Phonetics and Phonology 4). San Diego, CA: Academic Press. 365-405.