Explaining Analogy
Charles Reiss, Concordia University

> Since language is not, in its essence, a means for transmitting [cognitive] information—though no one denies that we constantly use language for this very purpose—then it is hardly surprising to find in languages much ambiguity and redundancy, as well as other properties that are obviously undesirable in a good communication code.
>
> —Morris Halle, "Confessio Grammatici"

I. The nature of parsing[*]

When one considers the phonology of a language as a formal computational device in isolation from the rest of the grammar, there is an obvious difference between the generative processes of producing surface forms from underlying representations and that of parsing surface forms to converge on a given underlying representation. This difference is that the former procedure is a one-to-one mapping[1] (a given underlying representation (UR) is mapped to a single surface form), whereas the latter can clearly be a one-to-many mapping (a single surface form can be mapped to a set of URs which are neutralized by the grammar). This can be illustrated by any number of well-known examples such as the following data from German:

(1) Production is *one-to-one*, parsing is *one-to-many*
    Production   Parsing
    /bunt/ > [bunt]   [bunt] > /bunt/ OR /bund/

Surface merger of underlyingly distinct forms is a well-known phenomenon, with examples available from a tremendous variety of human languages. Failure to recognize this phenomenon accounts for the inadequacy of the Optimality Theory (OT) parsing algorithm proposed by Smolensky (1996) which can be characterized as in (2):

(2) In parsing an OT-grammar selects the **most** harmonic *input/UR* for a given observed *output/surface form*

Only one UR can be **most** harmonic for a given mapping, so this algorithm fails in the basic task of providing a set of candidate parses for a given surface form (see Hale and Reiss 1997 for further discussion and alternatives to Smolensky's algorithm). A surface form [bunt] can only be parsed as /bunt/ by Smolensky's algorithm. Any appeal to top down processing to resolve the failure of Smolensky's parsing algorithm is inconsistent with well-established priming effects: "The general picture of lexical access during speech perception, then, is that initially it can discriminate only on phonological grounds. Only somewhat later in processing, after the syntactic and conceptual processors have gotten access to the list of possible candidates, can the ultimate choice of word be determined" (Jackendoff 1987: 103, cf. references therein).

II. Two kinds of 'analogy'

I will argue in this paper that the partial indeterminacy of parsing provides a straightforward explanation for the diachronic process commonly referred to as 'analogy', both within and across paradigms. I adopt the view of language change proposed in Hale (ms.) where change is seen as a mapping between successive grammars, rather than as change in a single grammar. With this distinction in mind, I

[1] Or many-to-one, since different URs can all map to the same PR. This can be due to lexical ambiguity, i.e. the lexical components of the different URs combine to produce phonologically identical strings as input to the phonology; or it can be due to structural ambiguity, whereby the phonology generates identical outputs from non-identical inputs (neutralization of contrast).

propose that analogical effects come about as a result of the ambiguity arising from the indeterminacy of parsing:

> (3) Parsing basis of analogical change: Given a grammar G_1, URs /a,b/ and a surface form Φ s.t. $G_1(a)=G_1(b)=\Phi$, Φ is ambiguous for a learner constructing G_2 using the output of G_1 as PLD.

One result of such ambiguity is 'analogy'. Note that this explanation of analogy is independent of the notion of the paradigm.[2] This is a strength of the proposal, since it is a well-known, though under-appreciated fact that analogical effects are attested both within and across paradigms. For example, in the history of Old Icelandic[3], there was a paradigm which included nom. sg. *mann-r / gen.sg. *mann-a* 'man'' Subsequently, these forms are attested as *maðr / manna*. The change of *mannr to *maðr* is clearly not a phonological one. In traditional terms, this change is based on 'analogy' to forms such as nom. sg. *guðr / gen. pl. *gunna* 'battle'.[4]

It is worth noting that this change in the history of Old Icelandic can be best explained without positing *any* change in the phonology of the successive stages. We need only posit changes in underlying representations. At each stage, both before and after the 'analogy', the relevant alternations can be generated by assuming, for example, a rule of cluster simplification which deletes *n* before *ðr* and a rule of cluster assimilation which assimilates *ð* to *n* between vowels:

> (4) Two rules present at all stages Old Icelandic
> R1: n>∅ / __ ð C   ex. /gunð-r/ > guðr
> R2: ð> n/ n__V   ex. /gunð-a/ > gunna

Before the 'analogical' change the sequence *–nn-* is derivable either from underlying *–nn-* or from underlying *–nð-*, when occurring between vowels. Therefore, a learner who is depending on output such as *manna* generated by a pre-analogy grammar could 'mistakenly' posit /manð/ as the UR of the word meaning 'man'. Note that the *-nð-* cluster never surfaces in either form, since the *n* gets deleted by R1 or the *ð* assimilates by R2.

This kind of extension, or 'interparadigm analogy', has three features relevant to the present study. First, as just noted, the 'analogy' is not based upon any existing surface form within another paradigm, but rather on the neutralizing effect of the *grammar* which must be constructed once a learner acquires the *-nn-*/*-ð-* alternation for some morphemes. Second, the effect of the analogy is to *increase* surface allomorphy within paradigms. Third, as already mentioned, 'analogy' refers to a *diachronic* change and is thus not directly relevant to the study of grammars as static knowledge states, the object of inquiry recognized by the generative program.

Examples of (intra-)paradigm leveling are also well-known in the analogy literature. This type of change reduces allomorphy within paradigms. This kind of leveling will typically also have the effect of reducing opacity since the leveling of allomorphs makes the mapping of UR to PR more direct. In a later stage of Old Icelandic such changes are also attested. For example the paradigm *guðr / gunna* discussed above, was leveled to *gunnr /gunna* in later Old Icelandic. The 'man' word was not leveled, however, and survives with *–ð-* in the nom. sg. into Modern Icelandic. We will see below, however, that 'intra-paradigm leveling' is also a result of the acquisition process.

III. A unified account of analogy

In this section I propose a single account for the two kinds of analogy illustrated above. I will, however, first discuss paradigm leveling, which is claimed to be the more common process. The reduction of opacity through paradigm leveling occurs, not because of the learner's 'desire' to simplify the grammar,

---

[2] I use the term 'paradigm' in its traditional (and vague) sense of a group of words which can be viewed as inflected forms of a given base. I attribute no theoretical status to the paradigm since I believe it to be an epiphenomeon resulting from the generativity of the grammar.

[3] For fuller discussion of the Old Icelandic data see Reiss (1995, 1997).

[4] Parallel changes are also attested in adjectival and verbal paradigms, which demonstrates that 'four-part' or proportional analogy need not be based on a morphologically parallel relationship among participants forms, as Hock (1991:172) suggests is 'normally' the case.

as a functionalist might suggest, but just because, as Kiparsky and Menn (1977:73) point out "Opacity is a property of the relation between the grammar and the data. An opaque rule is not more complex, merely harder to discover." While it is intuitively obvious that it is 'hard to discover,' for example, that surface [nn] is derived from underlying /nð/, this intuition can be explained on the basis of a theory of phonological learnability. As Hale and Reiss (1997ab) argue, a child can acquire a lexicon only if the initial state of the grammar is such that surface forms and underlying representations are (a) identical to each other, and (b) identical to the child's parse of the PLD, the output of speakers of the target language. In OT terms this means that Faithfulness constraints must be ranked above well-formedness constraints initially (contra Smolensky 1996). In rule-based phonology, it means that the child's grammar initially contains no phonological rules (contra Natural Phonology theories of acquisition). Given that UG provides for maximally faithful mapping between UR and surface form as a default assumption, it is not surprising that earlier /gunða/ > [gunna] can be parsed by a learner as /gunna/. More precisely, given Hale and Reiss's assumptions, it is not only possible, but *necessary* that the 'direct' parse be the child's initial hypothesis. Only by failing to discover that which is 'hard to discover' does the learner continue to parse [gunna] as derived from /gunna/, and thus provide evidence for the attested 'analogical' change. The change is to be located in the new underlying representation which lacks an –ð-.

The diachronic restructuring of URs like *mann-* > *manð-* in interparadigm analogy is a change which *increases* allomorphy and is best understood by appeal to the non-unique solution to the parse of surface *VnnV* sequences by learners exposed to the output of grammars that neutralize underlying –nn- and –nð- under some conditions. Given this indeterminacy of the grammar, I propose that the child may be influenced by language specific — but *extragrammatical* — factors such as processing strategies in constructing UR's. [5] These factors will sometimes have the effect of over-riding the default grammatical analysis which assumes identity between UR and surface form. We can imagine, for example, the influence of 'priming effects'. Having recently heard a sequence [x], which is known to be derived from underlying /y/ (in both the learner's grammar and that of speakers of the target language), the learner might be led to assign the UR /y/ to another token of [x]. This process obviously cannot take into account whether the new token of [x] is merely derived from underlying /x/ for speakers of the target language. This misanalysis by the learner is exactly the process we can assume occurred in Old Icelandic. As we see, 'analogy' occurs independent of the existence of words which are morphologically related to the one which 'changes' diachronically.

It is now apparent that the simplest account of analogy, both a) interparadigm ('four-part') analogy and b) paradigm leveling, within a generative theory of grammar, is in the restructuring of the lexicon in the process of lexical acquisition. The shifts discussed above involved a) the acquisition of the morpheme meaning 'man' with an underlying *nð* instead of underlying *nn* and b) the acquisition of 'battle' with underlying *nn* instead of *nð*.

IV. Base Identity in Kenstowicz (1994)

We can turn now to an evaluation of the claim made by Kenstowicz (1994) that Optimality Theory, enriched by two forms of Output-Output Correspondence (OOC), namely Base-Identity (BI) and Uniform Exponence (UE), provides an account of apparent cases of analogy or leveling. We must first note that Kenstowicz does not discuss any cases of interparadigm (four-part) analogy, the type of change which increases allomorphy within a paradigm (such as that which gave rise to OI *maðr*). Below, I summarize two of Kenstowicz's examples, one for Base-Identity and one for Uniform Exponence, then evaluate the analyses.

The notion of Base-Identity, which motivates a set of OOC constraints, is defined by Kenstowicz as in (5).

(5) Base-Identity: Given an input structure [X Y] output candidates are evaluated for how well they match [X] and [Y] if the latter occur as independent words.

---

[5] This looks dangerously close to claiming that 'general problem solving skills' are invoked in language acquisition. As Morris Halle (p.c.) points out, these skills may not be part of the grammar, but they must be specific to language, since they operate on linguistic representations. Perhaps they can be compared to the task of providing pairs of rhyming words.

Kenstowicz goes on to propose that BI can explain some asymmetries between nouns and verbs in Korean. According to Kenstowicz, Korean allows no clusters in onsets or in codas, so stem-final /-ps-/, for example must simplify to [-p-] when a consonant-initial suffix (or no suffix follows). The stem /kaps-/ 'price', loses its /s/ before the conjunctive suffix /-kwa/, but not before the vowel-initial nominative suffix /-i/. In the citation form, there is no suffix, so the final cluster is again reduced, and the surface form is [kap]. The verbal stem /ēps/ 'not have' also loses its /s/ before a consonant-initial suffix, but retains it before a vowel-initial suffix. Korean verbs are bound morphemes, however, and so never appear without some kind of suffix. There is, therefore, no simple citation form for verb stems as there is for nouns. Kenstowicz derives the correct Korean output by ranking constraints against complex codas and onsets (*Complex) above a constraint requiring input consonants to appear in the output (Parse-C), as reproduced in (6) and (7).

(6) Forms in Standard Korean from Kenstowicz

/kaps/ 'price'		/ēps/ 'not have'	
kap	citation form	—	no citation form
kaps-i	nominative	ēps- ēss-ē	past-informal
kap-k'wa	'price and ... '	ēp-t'a	nonpast-formal

(7) Constraint ranking to generate cluster simplification

/kaps/	*Complex	Parse-C
☞ kap\<s\>		*
kaps	*!	

/kaps+i/	*Complex	Parse-C
☞ kapsi		
kap\<s\>i		*!

Kenstowicz continues his discussion by noting that the above analysis does not generate the correct output for the "younger generation of Seoul speakers" who never have stem-final consonant clusters in nouns, yet do have clusters (like their elders) in verb forms. Kenstowicz provides the following data:

(10) Paradigms of "younger" Korean speakers

/kaps/	'price'	/ ēps/	'not have'
kap	citation form	—	no citation form
kap-i	nominative	ēps- ēss-ē	past-informal
kap-k'wa	'price and ... '	ēp-t'a	nonpast-formal

Kenstowicz's analysis of this dialectal difference is the following:

Thus, while /kaps+i/ surfaces as [kapi] with deletion of the /s/, / ēps+ēss+ē/ can never be realized as *[ēp+ēss+ē]. We can account for this asymmetry straightforwardly if the younger generation ranks Base-Identity above Parse-C. In other words, it is more important for the output of /kaps+i/ to resemble the output form of /kaps/ than to resemble the underlying input form.

The relevant tableaux are reproduced in (11). The independent word which serves as the Base is given at the bottom of the left-hand column, following Kenstowicz.

(11)

/kaps+i/	*Complex	Base-Identity	Parse-C
kapsi		*!	
☞ kap\<s\>i			*

Base: [kap]

4

Kenstowicz further explains that "Since Korean verbs always require an inflection there is no independently occurring output form of the stem to which the verb stem in / ēps+ ēss+ ē/ can be compared and so the identity constraint is vacuously satisfied. The Parse-C constraint demanding faithfulness to the underlying form will then choose the candidate that preserves the cluster (p.13)"

In a footnote, Kenstowicz mentions, but rejects, what is clearly the correct analysis: "While it is possible that younger speakers have restructured the input representation the Base-Identity constraint explains why verb stems may terminate in a CC cluster while noun stems systematically fail to do so." Even if true, this claim about the absence of CC clusters at the end of noun stems is not necessarily a linguistically significant generalization about the grammar of speakers of the relevant dialects, but may be merely a descriptive statement reflecting the history of the language. Besides, it is a well-known fact that analogical changes may affect individual lexical items, well-defined subsets of the lexicon, or the lexicon as a whole. The consistent 'choice', in nouns, of the non-cluster-containing parse by younger speakers will be discussed below. [6]

This application of Output-Output Correspondence must be evaluated on empirical, theoretical and general grounds. On the most superficial level, we see that the proposed Correspondence relations are strictly speaking outside of the domain of theoretical synchronic linguistics. In the Korean example, the grammar of younger speakers is analyzed with respect to the grammar of older speakers. The two grammars differ in output and Kenstowicz assumes that this difference is located in the ranking of relevant constraints and not in the form of lexical entries. Since younger speakers never exhibit consonant clusters in any forms of the word for 'price', the null hypothesis should be that there is no cluster underlyingly (see below). The grammars of their parents are irrelevant to an analysis of their grammars.

V. 'Regular' Analogy in Lexical Categories

I turn now to the question of how analogy could possibly occur regularly within a definable subcomponent of the lexicon: as Kenstowicz states, the historical cluster simplification in Korean nouns is regular, whereas verbs are unaffected. An explanation for this again can be sought in the acquisition process. I assume that the only goal children have, if, indeed, 'goal' is an appropriate term for the unconscious pattern recognition process of language acquisition,[7] is to acquire the ambient language. Given the non-uniqueness of the solutions provided by the parser (discussed above), we were forced to assume that non-grammatical factors can come into play in the construction of URs. Besides the cases which Kenstowicz discusses, it is also true that Korean has nouns which never contain stem-final clusters, for either older or younger speakers: /cip/ 'house (Martin 1954).

A possible solution relies on a form of staged, diachronic lexical diffusion and hypercorrection. Given any number of 'misacquired' URs, even a single one, the dialectal discrepancy between such forms and forms with the historically 'correct' clusters can be extended in sociolinguistic dialect borrowing. For example, a child may 'mistakenly' adopt the hypothesis that the citation form of a noun and the UR are identical in a given case, since there are stems like *cip*. In the appropriate sociolinguistic context, this child's 'missing' cluster in a form like [kapi] can serve as a model of imitation and over-generalization for others. This will even have the effect of other speakers replacing in usage forms with clusters by forms without clusters. To be precise, such speakers will have two different forms of a single etymological root, and the selection between the two will depend on sociolinguistic factors.[8] We are safe in assuming that such sociolinguistic diffusion took place, since it is implausible that a whole generation of Koreans spontaneously 'misacquired' their parents' language in exactly the same way, *viz.* by failing to produce exactly the same set of consonant clusters in nouns. This sociolinguistic phenomenon is necessary to an explanation of diachronic change, but irrelevant to the analysis of synchronic grammatical states.

---

[6] For individual lexical items, consider the change of earlier English *reach / raught* to modern *reach / reached* as opposed to *teach / taught*. For an 'analogical' change affecting the whole lexicon, including nouns, adjectives, verbs and pronominal forms, see my fuller discussion of the Old Icelandic facts in Reiss 1997.

[7] "Language acquisition is something that happens to a child placed in a certain environment, not something that the child does" (Chomsky 1993:29).

[8] There are two noteworthy aspects of this discussion: 1)it is not necessarily the case that the hypercorrected or borrowed form of the root 'replace' one which was acquired earlier; 2)this account does not rely on indeterminacy in the functioning of the grammar, since the choice of root-form is sociolinguistically conditioned. See Hale (ms.) for fuller discussion.

Finally, it is worth pointing out that the *kaps/kap* alternation is not merely triggered by the presence/absence of a following vowel. As Martin (1954: 20) explains, Standard Korean presents many cases in which final stem-final clusters are simplified before a vowel:

> Before a vowel which does not begin a particle, the copula, or an inflectional ending, the usual treatment is to reduce the excess [that is, stem-final material which cannot be syllabified in the coda or in a following onset when the syntactic juncture is not 'strong enough' to allow resyllabification—cr]: *kap olumyen* 'when the price rises', *kap ālki elyewe* 'it's hard to find out the price'.

The existence of such surface forms with simplified clusters, despite the existence of a following vowel, is surely relevant to the opacity of underlying forms. Nouns in Standard Korean thus show two variants with a very complex distribution: clusters surface before vowel initial morphemes within some phonological domain in which resyllabification occurs, say the phonological word or the clitic group. Cluster simplification occurs before consonant initial morphemes within such domains, or before vowels which lie outside of such domains. Since, as Kenstowicz points out, verbs are always inflected, the conditions on cluster simplification in verb stems are less opaque (and thus 'easier to discover')—the following morpheme is always within the domain of potential resyllabification, so clusters survive before a vowel and are simplified before a consonant. Such differences of opacity between the conditioning of cluster simplification in verbs and nouns, may have contributed to a consistent reanalysis of underlying representations in only one of these categories. Also, since noun stems can occur in uninflected form, whereas verb stems cannot, the former can occur prepausally, whereas the latter never do (Martin 1954:20). In citation form, noun stems are clearly prepausal, and thus subject to cluster simplification in Standard Korean.

VI. Uniform Exponence in Kenstowicz (1995)

Kenstowicz adopts a second type of OOC which can be invoked in cases where there is no isolation form of a morpheme to which other forms can be compared. The effect of such a constraint, which is dubbed Uniform Exponence (UE) is to "minimize allomorphic differences":

> (12) Uniform Exponence: minimize the differences in the realization of a lexical item (morpheme, stem, affix, word.)

Kenstowicz proposes using this constraint to account for the behavior in some (unidentified) Spanish dialects of the morpheme written *des-*. This morpheme is realized as *de*[h]- invariantly, whereas other tokens of [h] are synchronically derivable from /s/:

(13)	/mes/		/des-/	
	meh	'month'	deh-calzar	'unshoe'
	mes-eh	plural	de.h-e.cho	'refuse'

Rather than assuming that there is a difference in underlying form, that is, that the relevant UR's are /mes-/ and /deh-/, Kenstowicz assumes that both types of morpheme have underlying /s/. In order to assure that *des-* is realized invariantly as [deh-], Kenstowicz posits a Uniform Exponence constraint which is apparently specific to this one morpheme (since he gives no other examples). This is instead of assuming that the language has both underlying and derived [h].

What we must ask ourselves is this: Given a non-alternating morpheme of the shape [deh-] what is a learner most likely to posit as the underlying form? Kenstowicz states that, in general, underlying /s/ is the only source for [h], but we must ask whether the more salient linguistically significant generalization is this historical fact, or the fact that the morpheme in question has one, and only one, realization. Since this morpheme must be lexically marked to ensure that it is always realized with [h], we can consider two ways of achieving this. Kenstowicz's solution is to posit an abstract diacritic which applies to a single morpheme and causes it to surface with [h] in onsets. An alternative is to differentiate this morpheme by assigning it a UR which contains /h/. In coming to a decision on this question, I tend to agree with a comment by Alan Prince posted to the OT discussion list (Nov.21, 1996):

> [W]hy, in a grammar G such that G(a)=G(b) for potential input elements /a/,/b/, [is] a nonalternating observed element [a] … not (sometimes, always, freely) lexically /b/[?] The correct answer is surely 'why bother?' — i.e. to set up /b/ for [a] when /a/ will do.

However, Kenstowicz's solution raises another problem which is more serious than Prince's concerns of elegance. The notion of richness of the base, discussed by Prince and Smolensky (1993) and Smolensky (1996), which claims that the nature of an OT grammar assures that all inputs will surface as well-formed outputs, loses all its force if it can be over-ridden by arbitrarily marking individual morphemes as exceptions. Obviously, all theories must come to grips with exceptions, but Kenstowicz's solution is a blatant violation of the arbitrary appeal to cophonologies discussed by Inkelas, Orgun and Zoll (1996). These authors argue convincingly against unprincipled appeals to cophonologies to account for apparent exceptional behavior. In general, they license appeal to a cophonology only in cases where the 'exceptional' morphemes constitute a well-defined morphological, syntactic (or even semantic) category. Otherwise, apparent exceptions should be handled by positing distinct UR's for morphemes which display different alternation patterns. For example, instead of positing an arbitrary class of morphemes which exceptionally fail to show the effects of coda devoicing in Turkish, they posit an underlying three-way contrast in feature specification. The root meaning 'wing' ends in an alternating stop — [t] in codas and [d] in onsets — whereas the other two roots end in stops which do not alternate, as in (14) below..

(14) a. Alternating (unspecified for voicing):
　　　 kanat 'wing'　kanatlar 'wing-plural'　kanadɪm 'wing-1sg.poss'
　　 b. Non-alternating voiceless (specified [-voiced]):
　　　 sanat 'art'　sanatlar 'art-plural'　sanatɪm 'art-1sg.poss'
　　 c. Non-alternating voiced (specified [+voiced]):
　　　 etüd 'etude'　etüdler 'etude-plural'　etüdüm 'etude-1sg.poss'

In order to generate the attested patterns, Inkelas, Orgun and Zoll assume a feature-filling rule (or rather the OT equivalent[9]) which assigns [+voice] to onset obstruents and [-voice] to coda obstruents, in case they have no underlying voicing specification.

Returning to the Spanish example, we might consider that a class of morphemes containing just one member, *des*, would constitute a coherent, well-defined class. However, it turns out that in at least some of the Spanish dialects which have invariant [deh] there are other morphemes with invariant [h], regardless of whether a vowel or consonant follows. These include the plural forms of the article, orthographic *las, los*, which are pronounced [lah], [loh], not only in *La*[h] *Palmas*, but even in a form like *Lo*[h] *Angeles*, as well as uninflected forms like *ma*[h] 'more'. This set of morphemes clearly does not form a coherent class. Therefore, we are forced to adopt a 'prespecification' analysis (to adopt the terminology of Inkelas, Orgun and Zoll), namely, that surface [h] is derived from /s/ only in morphemes which alternate.

It has been pointed out that such an analysis fails to capture the fact that very many Spanish [h]'s *are* derived from /s/. In response, consider that the situation in Spanish is formally identical to German devoicing, mentioned above. Consider how one might apply the use of BI or UE to an analysis of the German coda devoicing facts. Rather than positing a difference in UR between what is traditionally assumed to be the two roots /bunt/ and /bund/, we might assume that the two roots are both /bund/ underlyingly. The well-known alternations exhibited by these roots ([bunt]/[bundəs] vs. [bunt]/[buntəs]) could then be accounted for in the following fashion: the alternating paradigm is due to standard phonological processes, coda devoicing or its OT equivalent; the non-alternating paradigm is subject to the same phonology, but a (set of) morpheme specific UE constraint(s) guarantees that alternations are suppressed for some morphemes. If one sought to be truly perverse, it could be claimed that such an account allows for the expression of a newly discovered generalization: no German roots end with a voiceless obstruent.

---

[9] See Reiss 1997 for discussion of potential difficulties with generating feature filling in OT. In brief, since a common interpretation of markedness is that it corresponds to greater representational complexity, it is unclear how underspecified structures could ever violate well-formedness constraints and require 'repair' by the grammar.

VII. Discussion

We have seen that analyzing the Korean and Spanish data by means of Output-Output Correspondence flies in the face of common-sense and also is used by Kenstowicz to set up Correspondences that hold *between different grammars.* We must also reiterate that we would not expect OOC to provide an account of analogy, since OT is a theory of grammar, and analogy is a diachronic process, a relationship between grammars.

OOC runs into some further problems as an account for analogy. Note that both forms of OOC considered in this paper demand Correspondence between 'related' forms, forms that at least share certain morphological material. The discussion of Base-Identity in Korean made reference to paradigms which are leveled in the direction of a base form which, crucially, occurs as an independent word. The discussion of Uniform Exponence in Spanish, relied crucially on a single morpheme being assigned a unique phonetic realization at the cost of violating (by hypothesis) some otherwise general patterns in the distribution of sounds. If we refer back to the 'extension' of –ð- into the 'man' word in Old Icelandic, sketched above, or indeed to any of the well-attested cases of interparadigm analogy, we find that OOC cannot even begin to provide a motivation for such diachronic processes. The diachronic restructuring of URs like /mann-/ > /manð-/ actually *decreases* the uniformity of exponence (i.e. *increases* allomorphy).

The theory proposed here, one which posits lexical restructuring, accounts for both inter- and intra-paradigm analogy in a unified fashion: both result from restructuring of underlying representations by an acquirer *vis-à-vis* the target language. The theory provides a diachronic solution to a diachronic phenomenon. It does not confuse the generative notion of language as mental grammar with the socio-political notion of language as speech community, in which context, for example, one might describe relationships between the speech of younger and older speakers. The theory proposed here thus better explains of the diachronic nature of what is traditionally called analogical change by maintaining an explicit generative theory of grammar. By removing the burden of explanation for 'analogy' from the theory of grammar (in the generative tradition), and locating the source of 'analogy' in acquisition and sociolinguistic borrowing, we end up with a more constrained theory of grammar.

Once the data discussed by Kenstowicz are seen in this light, it becomes clear that at least some of the puzzles for which OOC was proposed disappear. The empirical basis for such powerful constraint types is thus significantly weakened. See Hale and Kissock (1997) and Hale, Kissock and Reiss (1997ab) for further critique of OOC.

VIII. Conclusions

The appeal to OOC is part of a general resurgence of functionalist thinking in linguistics: it is intuitively satisfying to imagine that a principle like Uniform Exponence which would minimize allomorphy is relevant to grammar. However, as we have seen there are serious problems with this view. First, lexical restructuring can account for all the cases of 'analogy' in which OOC has been invoked, as well as for cases where OOC is not only irrelevant, but is also violated outright. Basic scientific methodology forces us to reject the OOC explanation. Second, even if one were willing to accept functionalist arguments for why leveling of paradigms occurs (to reduce allomorphy), it would still be misguided to build OOC into the grammar: 'analogy' occurs in the course of transmission/acquisition of language. Since aspects of acquisition are responsible for 'analogy', it would be redundant to posit that OOC (grammar) was responsible for analogy as well. Again, basic scientific methodology forces us to reject the OOC explanation. Finally, the view of opacity expounded by Kiparsky and Menn (1977) and referred to above has leveling as a natural consequence. No new grammatical machinery is posited under such a view. Once again, we must reject the OOC analysis. Functional considerations may or may not be relevant to a full understanding of language, including the nature of change and acquisition, but strictly speaking, these notions are neither useful for understanding analogy, nor are they part of grammar *per se.*

References

Chomsky, N. 1993. *Language and Mind.* Moyer Bell: Wakefield, Rhode Island.

Hale, M. *Theory and Method in Historical Linguistics.* Ms. Concordia University.

Hale, M. and M. Kissock. 1996. *The Phonology-Syntax Interface in Rotuman.* Austronesian Formal Linguistics Association meeting, UCLA (to appear in proceedings of same).

-------. 1997. *The Incomplete Phase in Rotuman: a synchronic analysis*, ms., Concordia University and Oakland University.

Hale, M., M. Kissock and C. Reiss. 1997a. Evaluating the Evidence for Output-Output Correspondence. Proceedings of *BLS 23*. Berkeley Linguistics Society.

--------. 1997b. What is Output? Output-Output Correspondence in OT. Proceedings of *WCCFL 16*. CSLI Publications: Stanford.

Hale, M. and C. Reiss. 1997a. The Comprehension/Production Dilemma in Child Language: A Response to Smolensky. ROA 170. (submitted to *LI*).

--------. 1997b. How to Parse (and How Not To) in OT Phonology. *Proceedings of NELS 27*. Graduate Linguistics Association, UMass: Amherst.

Inkelas, S., O. Orgun and C. Zoll. 1996. Exceptions and static phonological patterns: cophonologies vs. prespecification. ROA 124.

Jackendoff, Ray. 1987. Consciousness and the Computational Mind. MIT Press: Cambridge, MA.

Kenstowicz, M. 1994. Base-Identity and Uniform Exponence: Alternatives to Cyclicity. ROA 103. To appear in Jacques Durand and Bernard Laks, eds. (1995) *Current Trends in Phonology: Models and Methods*. CNRS, Paris-X and University of Salford Publications.

Kiparsky, P. 1968. Linguistic Universals and Linguistic Change. In E. Bach and R.T. Harms, eds., *Universals in Linguistic Theory*. New York: Holt, Rinehart and Winston, Inc.

Kiparsky, P. and L. Menn. 1977. On the Acquisition of Phonology. In *Language, Learning and Thought*, ed. J. Macnamara. New York: Academic Press.

Martin, S. 1954. *Korean Morphophonemics*. William Dwight Whitney Linguistics Series. Baltimore: LSA.

Prince, A. and P. Smolensky. 1993. Optimality Theory: Constraint Interaction in Generative Grammar.Technical Report Rutgers Center for Cognitive Science, Rutgers University, New Brunswick, N.J.

Reiss, C. 1997. The origins of the nn/ð alternation in Old Icelandic. *NOWELE* vol. 30, pp. 135-158.

Reiss, C. 1995. *A theory of assimilation with special reference to Old Icelandic*. Ph.D. dissertation, Harvard.

Smolensky, P. 1996. On the Comprehension/Production Dilemma in Child Language. *Linguistic Inquiry* 27, 720-731.