Deriving the feature-filling / feature-changing contrast: An application to Hungarian vowel harmony*

Charles Reiss
Concordia University, Montréal

February 28, 2003

> Unfortunately, within linguistics it has not been generally recognized how important such formal, theoretical work is; instead there is a feeling that too much concern for theoretical detail is a waste of time...[T]he attitude that formal, theoretical work is bound to be both ad-hoc and sterile is, I am convinced, fundamentally mistaken ...
>
> Morris Halle (1975:530)

1 Distinctness and the interpretation of structural descriptions

In practice, as inspection of any introductory phonology book will show, it has been implicitly assumed in generative phonology that a rule will apply to any representation that contains a superset of the information contained in the rule's structural description (SD). In other words, if the SD of a rule R subsumes a representation Q, then Q is an input to R. Rules apply to natural classes of segments, and natural classes are represented by a representation that subsumes the representation of each of its members.

It turns out, however, that this is *not* the interpretive procedure developed in SPE (Chomsky & Halle, 1968), the foundational work in the field:

(1) Interpretive Procedure from SPE (Chapter 8, 337)

A rule of the form $A \rightarrow B/X\_\_\_\_Y$ applies to any string $Z = \ldots X'A'Y' \ldots$, where X', A', Y' are not distinct from X, A, Y, respectively; and it converts Z to $Z' = \ldots X'B'Y' \ldots$, where B' contains all specified features of B in addition to all features of A' not specified in B.

Distinctness is defined as follows:[1]

(2) Distinctness in SPE (336)

'Two units U1 and U2 are distinct if and only if there is at least one feature F such that U1 is specified [αF] and U2 is specified [βF] where α is plus and β is minus...'

The (typically implicit) appeal to subsumption in general phonological practice derives from the assumption of a logical equivalence between subsumption and non-distinctness, the idea that if x is non-distinct from y, then either x subsumes y or y subsumes x. This equivalence does not hold, however, except under the

*This paper has benefi tted from discussion with several audiences, including the *Acme Balkanica Conference* at Concordia University in 2001, and the Ohio State Linguistics Department. Especially helpful individuals have been Afton Lewis, Mark Hale, Sylvia Blaho, David Odden, Alain Thériault, Peter Liem, Morris Halle, Daniela Isac and anonymous reviewers. Finally, Péter Siptár deserves special thanks for vastly improving the discussion of Hungarian through careful reading and helpful criticism. This research was partially supported by the SSHRC-funded *Asymmetry Project* and several grants from Concordia University.

[1] I have omitted reference to non-binary feature values.

working assumption of the SPE era that representations are fully specified for all features. It *is* true that if either x subsumes y or y subsumes x, then x and y are non-distinct, but a simple example can illustrate that the converse is not valid if we allow for partially specified feature matrices in lexical entries.

Let x = [+round, -back] and let y = [+round, +hi]. The representations x and y do not disagree with respect to any features, and are thus non-distinct, but one clearly does not subsume the other. And we clearly do not expect, say, that x would satisfy a structural description specified as y.

As a further example, consider that by strict application of the SPE interpretive procedure, an underspecified vowel that had only the feature [-round] would satisfy the SD of a rule like (3), since the representation [-round] is not distinct from the representation [-nasal]:

(3) [-nasal] → [-voiced]

This is surely an undesirable result.

Non-distinct representations are, in the general case, what is called 'consistent' in unification-based frameworks—that is, they have no incompatible feature values. But non-distinctness, or consistency, does not reduce to subsumption.[2] The preceding discussion should make it clear that the interpretation of structural descriptions in generative phonology warrants reexamination.

2 Feature counting evaluation metrics

Perhaps all that is needed is to reject the SPE interpretive procedure in favor of one appealing to subsumption, since this is what the practice has been for the last several decades. Under this view any representation subsumed by (containing a superset of the information contained in) a rule's Structural Description is taken to be a licit input to the rule. This interpretive procedure has the desirable effect of allowing rules to apply to more than just single representations—they can apply to a natural class of representations whose description is subsumed by the rule's Structural Description.

This interpretive procedure entails that a rule that changed feature values, say from +F to -F for some feature F, would apply vacuously to representations that are already -F before the application of the rule.[3] For example, a straightforward statement of Polish or Polish coda devoicing might be written as in (4):

(4) [+cons, -son] → [-voiced] in CODA

This rule applies nonvacuously to [+voiced] inputs that are [+cons, -son]—in other words it makes them [-voiced]. However, according to the subsumption-based convention of rule interpretation, the rule also applies, albeit vacuously, to [-voiced] inputs that are [+cons, -son]. To reiterate, both [+voiced] and [-voiced] can satisfy the SD to be inputs to the rule.

This interpretation of SDs is related to the *SPE* feature-counting evaluation metric, the overarching goal of which is to minimize redundancy in the grammar, as seen in the Conciseness Condition formulated by Kenstowicz and Kisseberth (1979).

(5) The Conciseness Condition (one component of the SPE evaluation metric, from K&K:336)

> If there is more than one possible grammar that can be constructed for a given body of data, choose the grammar that is most concise in terms of the number of feature specifications.

With hindsight, it is now apparent that the Conciseness Condition is flawed by virtue of its parochialness—the model of grammar chosen by the analyst should take into account the models necessary to generate other languages as well as the one in question, and not just choose the most concise grammar that can generate

[2] I have found the same point made by Bayer and Johnson (1995: Section 2) in a discussion of Lambek Categorial Grammar: 'Interestingly, in cases where features are fully specified, these subsumption and consistency requirements are equivalent'. However, I do not think that the relevance of this observation to the application of phonological rules has been noted.

[3] I adopt without argument a binary-valued feature system. The paper is compatible with theories that allow various kinds of underspecification.

a given corpus. Thus we can see that the Conciseness Condition as stated here is in direct conflict with the search for Universal Grammar, the grammar of S_0, the initial state of the language faculty:[4]

(6) Choosing among extensionally equivalent grammars:
"Because evidence from Japanese can evidently bear on the correctness of a theory of S_0, it can have indirect—but very powerful– bearing on the choice of the grammar that attempts to characterize the I-language attained by a speaker of English." [Chomsky (1986:38)]

In other words, evidence from one language should bear on the best analysis of other languages. If two hypotheses, $\mathcal{A}$ and $\mathcal{B}$, concerning UG are empirically adequate to provide an explanatory account of English, but only one of the two, say $\mathcal{A}$, is adequate to provide an explanatory account of Japanese, then we should select $\mathcal{A}$ as the best available hypothesis for a theory of S_0 that can lead to acquisition of *both languages* .

The traditional interpretation of SDs such as (4) is not the only logical possibility—it could have been argued that a rule like Polish devoicing should be formulated so as not to apply vacuously, as in (7):

(7) $[+\text{cons, -son, +voiced}] \rightarrow$ [-voiced] in CODA

It seems that the decision to adopt the Conciseness Condition, and thus the rule format of (4), rather than (7), was motivated by the influence that engineering approaches to information theory had on the pioneers of generative phonology, an influence that has been described as leading to a dead end (Morris Halle, 1975:532 and p.c.). Formulation (4) was seen as the more efficient, and thus better, engineering solution since it was more concise than (7).[5]

In this paper, I explore another logical possibility for the interpretation of SDs and show that it solves longstanding problems in phonological theory—the question of how to allow rules to target unmarked or unspecified feature values and the intimately related issue of the distinction between feature-filling and feature-changing rules. I then apply the solution to the generation of the alternations seen in Hungarian vowel harmony.

3 Subsumption and structural descriptions—a problem

The SD in (4) subsumes various possible input representations. Crucially, all inputs must be specified for at least the features [+cons, -son]. For us to further understand the nature of the set of representations that can serve as inputs to the rule, we need to focus on features that are *absent* from the rule's SD. The traditional understanding of (4) depends on two distinct interpretations of the absence of a specification:

(8)
a. Absence of a feature value implies that the feature is *irrelevant* to the application of the rule.
b. Absence of a feature value implies that the feature does not need to be mentioned in the rule, because the rule neutralizes different values for the feature.

The absence of reference to features for place of articulation in the input of (4), for example, is interpreted as in (8a) to mean that the rule applies *regardless* of the place of articulation of the input consonant. In other words, features such as [cor] and [lab] do not appear in the rule because they are irrelevant to its application—Polish devoicing applies to obstruents at all places of articulation.

Assuming that Polish alternating stops are underlyingly [+voiced][6], the rule applies vacuously to underlying [-voiced] stops and it changes underlyingly [+voiced] stops to [-voiced]. Thus the rule's Structural

[4]An important question, addressed in another paper, is whether the the correct formulation of a rule is necessarily the most concise one that is consistent with the data and with the cross-linguistic (universal) demands discussed in this paper. I argue in this other work that learnability considerations provide yet another reason to favor less concise rules than we traditionally posit .

[5]The belief that the mind organized language in a maximally effi cient manner may have also motivated the Conciseness Condition. However, it could also have been argued that avoiding vacuous application would have constituted a more effi cient solution. Anderson's (1985:327) remarks on the topic are also telling: "Early concern for evaluation procedures . . . turned out to be something of a dead end." and "[T]he appeal of feature counting went away . . . not with a bang, but with a whimper."

[6]I will do so without argument here. The reader will notice that if Polish alternating stops are instead unmarked for voicing, the problem is to voice them in the appropriate contexts without targeting the non-alternating voiceless ones.

Description (SD) contains no reference to [voiced] since the rule neutralizes the distinction between [+voiced] and [-voiced]. This is interpretation (8b).

So, some features are absent from the SD because the rule does not affect them or depend on them in any way (8a), and others are absent because the rule neutralizes their two possible values (8b). To reiterate, any representation that is subsumed by the SD of the rule satisfies that SD. Thus (4) can apply to the following inputs:

(9) a. representations in which the absent features are irrelevant to rule application AND

b. representations in which the absent features are neutralized by the rule.

Consider, in contrast to the standard view of Polish-type patterns, a type of data which only became known much later in the history of generative phonology, a pattern which requires rules that fill in values on necessarily underspecified segments.

In Turkish, for example, Inkelas (1996; Inkelas & Orgun 1995) argues that there is necessarily a three-way contrast in voicing. Some stem-final stops show a t/d alternation (10a), with [t] appearing in codas and [d] appearing in onsets. Inkelas convincingly argues for an underlying segment that has all the features of a coronal stop, but is unspecified for [voiced]. She denotes this feature bundle as /D/. She states that the segment is assigned the value [-voiced] in codas, and [+voiced] elsewhere. Other stem-final stops consistently surface as [t] and thus are posited to be /t/ underlyingly (10b), and others surface as [d] consistently, and are thus posited to be underlying /d/ (10c).

(10) Turkish voicing alternations[7]

a. Alternating: [∅voiced] (unmarked for [voiced]) /D/
kanat 'wing' *kanatlar* 'wing-plural' *kanadım* 'wing-1sg.poss'

b. Non-alternating voiceless: [-voiced] /t/
sanat 'art' *sanatlar* 'art-plural' *sanatım* 'art-1sg.poss'

c. Non-alternating voiced: [+voiced] /d/
etüd 'etude' *etüdler* 'etude-plural' *etüdüm* 'etude-1sg.poss'

The rule responsible for making /D/ surface as [t] in codas would be identical to (4), but would have to be interpreted differently, since it crucially cannot apply to underlying /d/. In other words the representation of /D/ subsumes that of /d/ (and also that of /t/), but the rule that affects /D/ does not affect /d/. It is necessary to interpret the absence of [voiced] in the SD as in (11c), which completes the list of interpretations of absent features under discussion:

(11) a. Absence of a feature value implies that the feature is *irrelevant* to the application of the rule (=9a)

b. Absence of a feature value implies that the feature does not need to be mentioned in the rule, because the rule neutralizes different values for the feature (=9b)

[7]This data has been challenged in discussion on the grounds that some Turkish speakers do not pronounce the (c) forms with voiced obstruent in coda position. The irrelevance of such an objection is apparent, as long as the data represents a possible language. Since Orgun is a native speaker, I accept the data as given. The presence of inflectional morphology suggests that the forms should be treated as Turkish and not as French (the language they were borrowed from). I refer the reader to the cited works for further data showing that this Turkish case is not isolated—Inkelas discusses several cases with the same logical structure. She further shows that these data cannot be handled by labeling certain morphemes as 'exceptions'. For example, she provides examples of a single morpheme with both an obstruent that alternates in voicing and one that is consistently voiced (even in coda position): *edʒdat* 'ancestor', *edʒdatlar* 'ancestor-plural', *edʒdadɯ* 'ancestor-acc.' This example shows that failure to devoice obstruents in coda position cannot be a property of individual morphemes, since the stem-medial /dʒ/ remains voiced although it is always in coda position, whereas the stem-final /d/ alternates. Instead the two segments must be representationally distinct: the former is [+voiced] and the latter is underspecified for voicing.

c. Absence of a feature value implies that the feature *must* be absent from a potential input representation for the rule to apply

Without an intelligent homunculus, a mental grammar needs a solution to the problem of correctly selecting the relevant interpretation of a SD.

4 Earlier approaches

One way out of this dilemma would be to allow the grammar to refer to [∅voiced] as a possible specification:

(12) [+cons, -son, ∅voiced] → [-voiced] in CODA

I provide a principled argument against allowing [∅voiced] as a possible specification below. Traditionally, this move has been avoided by most researchers on the intuitive grounds that it represents an overly powerful enrichment of the representational apparatus of phonology.

Instead, however, the notational apparatus of rules has been enriched. Typically, a rule label, **Feature Filling** or, equivalently, **Structure Filling** is used, as in (13), to ensure that such a rule is not (over-)applied to fully specified segments and can only apply to provide feature values to underspecified segments. In the absence of such a label, the correct interpretation is left to the intelligence of the reader.

(13) **Feature Filling**: [+cons, -son] → [-voiced] in CODA

This is the solution proposed by Inkelas & Orgun (1995:777). I reproduce their rules exactly in (14).

(14) Feature-filling rules from Inkelas & Orgun (1995:777)

a. DEVOICING: Coda plosive → [-voiced] (structure-filling)

b. VOICING: Onset plosive → [+voiced] (structure-filling)

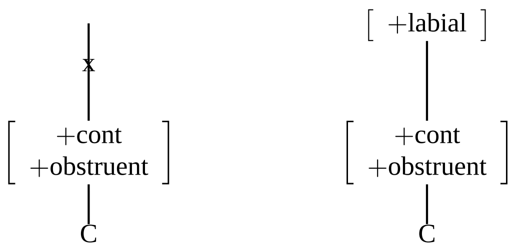
Their rule (14a) is basically equivalent to (13) and they provide a second feature filling rule (14b) to provide the alternating stops with [+voiced] in onsets.

Another example of the **Feature Filling** label can be found in McCarthy (1994:210), who formulates a rule spreading [pharyngeal] from a consonant to a following vowel. In addition to an autosegmental representation of spreading, McCarthy includes the following in the rule statement: ‘Condition: Feature-filling’. He notes that the ‘intent of the condition restricting [the spreading] to feature-filling is to block the rule from lowering any vowel other than the featureless vowel schwa’. Because it is featureless, consisting just of enough of a representation to identify it as a vowel, the representation of schwa will subsume that of every other vowel—less specification entails greater generality.

Kiparsky’s (1985) discussion of coronal underspecification briefly notes the problem treated here, stating that it is necessary ‘to work out some way of referring to unmarked segments’ when representations are not fully specified. However, his manner of distinguishing underspecified segments is not satisfactory since he introduces a new diacritic into representations just in places where, for example, [+coronal] (or a CORONAL class node) would be specified. Kiparsky proposes that (15a) be the representation of a coronal fricative such as /s/, where the x on the line to the missing node means ‘there is no specification on the tier of place features’. How do we know (or more importantly, how does the grammar ‘know’) that the x doesn’t denote underspecification for some other feature?

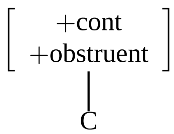
(15) Representations of /s/ and /f/ (Kiparsky 1985)

a. /s/ b. /f/



Kiparsky represents a non-coronal voiceless fricative, such as /f/ as in (15b) where the root node is associated to a labial place node. The natural class including both these voiceless fricatives would presumably be represented by Kiparsky as in (16), unspecified for features on the place tier:

(16) Representation of all fricatives in Kiparsky's (1985) system



Obviously, Kiparsky's system of representing /s/ in (15a) is the equivalent of specifying [+coronal], since the representation of /s/ contains information not present in (16) and can hardly be called underspecified.[8]

Similar use of a diacritic denoting the absence of an association to a given node can be found in Archangeli (1988). Where a 'melody unit or anchor' Z can be linked to a feature F *via* normal association lines between Z and F; obligatorily unlinked to F if Z is enclosed in a circle; or ambiguously linked or unlinked to F in the absence of an association line or circle.

(17) Linkage notation (adapted from Archangeli 1988): Z is a 'melody unit or anchor

a. (Z) unlinked to F

b. Z | F linked to F

c. Z ambiguously linked or unlinked to F

This notation presents a problem similar to that in Kiparsky's (1985) system, since the meaning of the circle around the Z in a given rule is 'unlinked to the feature F which the rule will provide'. Thus this notation of 'underspecification' requires reference to the very feature whose mention it is meant to avoid.

In the next section, I develop a Unified Interpretive Procedure (UIP) for Structural Descriptions that vitiates the need for explicit **Feature Filling** or **Feature Changing** diacritics, as well as the need to refer to features that are absent from a representation—I do not use [∅F] as a possible specification.

5 The Unified Interpretive Procedure

For a given rule $\mathcal{R}_a$, we can refer to its Structural Description as SD_a, its Structural Change as SC_a, and its Environment as Env_a, giving us the simple rule schema in (18):

[8] Kiparsky's suggestion leads to other problems as well—does the line in (15a) block spreading?

(18) Rule Schema[9]

$\mathcal{R}_a$: $SD_a \rightarrow SC_a$ in Env_a

We will employ Greek letter variables in the usual way: $\alpha \in \{+, -\}$.

I also assume that something like Principle 6 of Chomsky (1967) is valid: "Two successive lines of a derivation can differ by at most one feature specification".[10] Chomsky's Principle 6 helps restrict the notion of 'possible phonological rule', and thus it is desirable to conform to it. For our purposes, this principle means that SC_a will always contain a single feature specification, +F, -F or αF. We can now formulate the interpretive procedure for structural descriptions, replacing αF for SC_a .

(19) Unified Interpretive Procedure for Structural Descriptions

A representation Q is an input to a rule $\mathcal{R}_a$:

$SD_a \rightarrow \alpha$F in Env_a

if and only if SD_a subsumes Q and one of the following holds:

a. $-\alpha\text{F} \in SD_a$ (SD_a and thus each Q that satisfies SD_a is specified $-\alpha$F) OR

b. $-\alpha\text{F} \notin Q$ (no Q that satisfies SD_a is specified $-\alpha$F, and thus neither is SD_a specified $-\alpha$F)

First consider (19a): since (19a) requires that SD_a be specified $-\alpha$F, it follows that every representation Q subsumed by SD_a be thus specified. Since SD_a must subsume every input to the rule, each input must also be specified $-\alpha$F. A representation Q that satisfies this condition will undergo feature changing to αF.

Now consider (19b): the requirement of (19b) is that Q not be specified $-\alpha$F, so it can be either specified αF or *not specified at all* for feature F. Since Q is required not to be specified $-\alpha$F, any Q that satisfies this condition will not be subsumed by a SD_a which is $-\alpha$F. In other words, if Q is not $-\alpha$F, and Q is an input to $\mathcal{R}_a$, then SD_a is also not $-\alpha$F. (We thus see that (19a) and (19b) are mutually exclusive—they cannot be satisfied simultaneously.) If condition (19b) is fulfilled, the rule will either fill-in the value αF or vacuously 'change' αF to αF.

The two conditions thus require either (a) Q is $-\alpha$F or (b) Q is not $-\alpha$F. (Further conditions are imposed by Env_a, of course.) The existence of underspecification means that 'not $-\alpha$F' does not mean 'αF', but rather 'either αF or unspecified for F'.

Thus, if SC_a is -F, and SD_a is specified +F, then an input to $\mathcal{R}_a$ must contain +F in order to satisfy SD_a by condition a. However, if SC_a is -F but SD_a is not specified +F, then an input to $\mathcal{R}_a$ may not contain +F. It may be specified -F ($\mathcal{R}_a$ will apply vacuously in this case) or it may be unspecified for F ($\mathcal{R}_a$ will fill-in -F in this case) and thus satisfy SD_a by condition b.

Similarly, we can switch all the signs. If SC_a is +F, and SD_a is specified -F, then an input to $\mathcal{R}_a$ must contain -F in order to satisfy SD_a by condition a. However, if SC_a is +F, but SD_a is not specified -F then an input to $\mathcal{R}_a$ may not contain -F, but it may be specified +F ($\mathcal{R}_a$ will apply vacuously in this case) or it may be unspecified for F ($\mathcal{R}_a$ will fill-in +F in this case) and thus satisfy SD_a by condition b.

So, (19a) corresponds to traditional **Feature Changing** rules and (19b) to traditional **Feature Filling** rules. However, the UIP precludes the necessity of referring to unmarked values such as [∅voiced]. The crucial advance we have made is this: instead of having 'to work out some way of referring to unmarked segments' we have a way to ensure that they are treated as a class with representations that are vacuously affected by rules.

In a language like Polish, where we have a two-way voiced/voiceless contrast in obstruents we might be tempted to retain the traditional formulation of the devoicing rule and the traditional interpretation of structural descriptions. We would still generate Polish-type output. However, a truly explanatory approach to phonology allows us to see that Turkish can tell us something about Polish—the correct formulation of the

[9]More discussion is required for deletion, insertion and metathesis rules. These problems are addressed in work in progress. Also, the environement Env is, strictly speaking, part of the SD, but I will treat them separately for the sake of clarity.

[10]This notion can be adapted for more recent theories of representation. I do not make use of, for example, feature geometry, in this paper, and I have argued elsewhere that feature geometry is not a necessary or desirable part of phonological theory.

rule must be something closer to (7) than to (4). Since we are interested in Universal Grammar (UG), we are interested in a single interpretive procedure for all grammars. This was the point of the quotation from Chomsky in (6).

In this particular case of the representation of Polish devoicing, UG should be assumed to use the same interpretive procedure as is used in Turkish. More concise rules can be written for just the Polish data, but they would not be rules of human phonology, if the UIP is correct. Since the interpretive procedure for all languages is assumed to be identical, but the patterns to be accounted for are different, the rules themselves must differ as well. Polish uses (7), whereas Turkish uses (4).

In other words, the traditional account of Polish devoicing using the subsumption-based interpretive procedure would be extensionally equivalent to the account proposed here (using (7) and the UIP), but we now can choose between them in a principled fashion. Again, this is the type of argumentation suggested by (6).

A further implication of the UIP is that we now derive the intuitively valid result that rules do not treat representations that are +F and representations that are -F as a natural class to the exclusion of representations that are unmarked for F. With respect to the Turkish data discussed above this means that, for example, /t/ and /D/ constitute a natural class (they are *not* [+voiced]), and /d/ and /D/ do so as well (they are *not* [-voiced]), but that /t/ and /d/ do not, to the exclusion of /D/.

For the sake of explicitness, let us reiterate the difference between the simple subsumption-based interpretation of SDs and that given by the UIP. A rule like (4) should apply to [+voiced] stops according the traditional interpretive procedure. The SD of (4) is given as [+cons, -son], and since the representation of, say, /t/, /D/ and /d/ are subsumed by [+cons, -son], the rule should apply to all three. But his would not let us distinguish /d/ from /D/, which we need to do for Turkish.

However, using the UIP, a rule like (4), where the SC is [-voice], cannot apply to [+voiced] representations like /d/:

- (19a) is not satisfied since [+voice] is not in the SD of the rule as stated
- by (19b) any Q which is an input to the rule cannot be specified as [+voice]

Since neither condition is satisfied, the rule cannot apply to [+voiced] representations. However, both /t/ and /D/ satisfy condition (19a), since /t/ is [-voice] and the rule applies vacuously; and /D/ satisfies condition (19b), since /D/ is *not* [-voice] and thus the rule fills in this value. This would work perfectly for our model of Turkish.

Consider now the rule in (7), repeated here:

(20) [+cons, -son, +voiced] → [-voiced] in CODA

Since the SD of the rule contains [+voiced], the rule can obviously apply to [+voiced] representations. It cannot apply to representations that are either specified [-voice] or are unspecified for [voice]—the SD does not subsume such representations. This would work perfectly for our model of Polish.

6 Discussion

A reviewer of an abstract of this article objected to the UIP on the grounds that it is an extremely powerful device, since it is meant to be relevant to all rules in all languages. This objection reflects a misunderstanding of the notion of power in theory construction. In fact, a single interpretive procedure that holds for all rules in all languages provides a *less* powerful (and thus better) model than one that allows various devices on an *ad hoc* rule-by-rule and language-by-language basis. The objection can thus be dismissed.

Note that the result of the UIP in (19) can also be derived by allowing the use of logical negation in phonological representations. For example, if Turkish /t/ and /D/ were both specified [NOT+voiced], then the rule that fills in [-voiced] on /D/ in codas could refer to this specification:

(21) [+cons, -son, NOT+voiced] → [-voiced] in CODA

Allowing negation in representations to have scope over a single valued feature such as [+voiced] will not obviously create problems. However, allowing negation to have scope over sets of valued features would wreak havoc with the notion of natural class. It would allow us to treat the complement set of each natural class as a natural class. For example, the segments described by the set NOT[+voiced, +labial] would include both [d] and [p], but not [b].

What the UIP does is introduce logical negation into the interpretation of rules without enriching the set of primitives that can appear in lexical representations. The UIP does redefine the notion of natural class, in fact, but in a very restricted fashion: "[αF]" and "unspecified for F" constitute a natural class to the exclusion of "[-αF]" in the sense that the difference between otherwise identical members of a natural class "NOT[-αF]" is neutralized by the rule. Under this view natural classes are not defined by a feature matrix that subsumes a set of phonological representations, but instead by a set of phonological representations that are accepted as inputs to a rule, given the UIP. In other words, natural classes are derived from the nature of rule application, rather than constituting a primitive notion of phonological theory. I think this is a desirable result.

We can now return to the rejection of the use of $\emptyset$ as a coefficient value for features. Obviously, we could introduce this value, allowing the set of values to range over $\{+, -, \emptyset\}$. However, this move would have implications for the behavior of natural classes that appear to be undesirable. If phonological rules could refer to [$\emptyset$F] in structural descriptions, then it would be possible to apply rules to segments so specified without affecting other segments. For example, it should be possible to affect Turkish /D/ to the exclusion of both /d/ and /t/, say by rounding it before round vowels:

(22) A hypothetical rule

[$\emptyset$voiced] $\rightarrow$ [+labial] before [+labial]

My intuition is that we won't find such processes, but introducing $\emptyset$ as a feature value allows such possibilities since [$\emptyset$voiced] describes a natural class to the exclusion of /d/ and /t/. In contrast, such a process cannot be modeled using the UIP approach. Underspecified segments can never be referred to without referring to the segments with which they neutralize on the surface.

Either the introduction of negation into the set of lexical representational primitives ([NOT+voiced]) or the use of $\emptyset$ as a possible feature coefficient ([$\emptyset$voiced]) can be used to correctly model data with the logical structure of the Turkish stop alternations. However, I have provided arguments that introducing logical negation into the *interpretive procedure for rules* is empirically preferable to both of these alternatives. We turn now to discussion of a more complex rule type.

7 Rules containing variables

In the previous discussion α was used as a metalanguage variable to refer to either '+' or '-' within rules. In this section, I show that UIP applies without modification also in cases where α denotes a variable within the rules themselves.

If SC_a contains the variable αF, then inputs to $\mathcal{R}_a$ must either be specified for $-\alpha$F or else representations specified $-\alpha$F will not satisfy the SD (following UIP). Since α ranges over $+$ and $-$, $-\alpha$ also ranges over $+$ and $-$. This means that if the SC_a is specified for αF, and SD_a is *not* specified $-\alpha$F, then inputs to $\mathcal{R}_a$ cannot be specified for either +F or -F. Thus $\mathcal{R}_a$ must be feature-filling—its inputs must be unspecified for F.

Immediately we see that this forces us to revise a type of rule that is commonly invoked—the type that involves feature changing assimilation to either -F or +F. For example, the voicing assimilation seen in Russian prepositions in (23a) is represented in standard generative phonology as in (23b) (Halle & Clements 1983). We see that /t/ voices to [d] and that /z/ devoices to [s], when preceding a voiced or voiceless obstruent, respectively.

(23) Traditional statement of Russian voicing assimilation (to be rejected in accordance with (19))

a.

	'from'	'without'	'next to'
'rose'	at rózɨ	b^{y}iz rózɨ	u rózɨ
'Ala' (name)	at álɨ	b^{y}iz álɨ	u álɨ
'cow'	at karóvɨ	b^{y}is karóvɨ	u karóvɨ
'beard'	ad baradɨ́	b^{y}iz baradɨ́	u baradɨ́
'sister'	at s^{y}istrɨ́	b^{y}is s^{y}istrɨ́	u s^{y}istrɨ́

b. [+cons, -son] $\rightarrow$ [αvoiced] / __ [-son, αvoiced]

According to UIP (19) the presence of [αvoiced] in the SC and the lack of reference to [voiced] in the SD, forces the interpretation that inputs lack [$-\alpha$voiced]. Since $\alpha \in \{+,-\}$ this rules out any value for [voiced] in the input. Thus, in accordance with UIP, this rule would only be licit as a feature-filling rule, applying to consonants that lacked a value for [voiced].

Assuming, as is traditional, that Russian obstruents are all specified as either [+voiced] or [-voiced] underlyingly, the analysis could, in accordance with UIP, be split into two rules:

(24) Two valued assimilation

a. [+voiced, +cons, -son] $\rightarrow$ [-voiced] / __ [-voiced]

b. [-voiced, +cons, -son] $\rightarrow$ [+voiced] / __ [+voiced]

The first rule (24a) would apply only to [+voiced] stops. The second rule (24b) would apply only to [-voiced] stops.

However, it now becomes apparent that we can collapse the two processes and still honor the UIP.

(25) [+cons, -son, $-\alpha$voiced] $\rightarrow$ [αvoiced] / __ [-son, αvoiced]

The rule, like Polish (and Russian) coda devoicing, is thus purely feature-changing: whenever there is feature changing, there is no vacuous application. If we compare (25) to (23b) we see that once again, our UIP drives us to accept a less concise representation of the rule. The new formulation contains three mentions of α, as compared with two in the original.[11]

I will now apply UIP to the generation of a more complex body of data, the trigger-target relations found in Hungarian vowel harmony.

8 Regular Patterns in Hungarian Vowel Harmony

In the following discussion, I treat only the fully productive patterns of Hungarian vowel harmony in regular stems. For example, I leave aside the representational issues associated with front vowel stems that take back vowel suffixes, rule iterativity, and also the treatment of transparent vowels. There are four main patterns of harmonic alternations, which I first present with the orthographic vowels, since the orthography has typically guided (excessively, I believe) other treatments of this data. I then turn to the phonetic values of the vowels in section 9, which shows some of the deficiencies of relying too heavily on the orthography, and then to the lexical representation of alternating and non-alternating vowels in section 10. A set of rules for generating the surface foms from the proposed lexical representations is developed in section 11 and schematic derivations are given in 12.

[11] Of course, it is common to represent such a process as a rule autosegmentally spreading the [voice] specification from an obstruent to the one preceding it. In recent work (Reiss, 2003ab) I argue that much of autosegmental phonology has to be replaced by an algebraic notation using variables and indices. The more powerful algebraic notation is necessary, since phonological computation appears to require the use of quantifiers in order to express non-identity conditions. In other words, some version of the old-fashioned alpha notation may be the best way to express such processes.

Nothing crucial relies on deriving all of these alternations from rules that correspond to 'structure-filling' rules, however, the complex patterns and the relative familiarity of the data make Hungarian vowel harmony an attractive testing ground for a new proposal.

8.1 Harmony involving *a/e*

Suffixes such as inessive *-ban/-ben* surface as *-ban* after back vowels and *-ben* after front vowels: *dobban* 'in a drum', *szemben* 'in an eye'. These suffixes are never rounded and thus surface with the unrounded front variant also after a front round vowel: *tökben* 'in a pumpkin'.

8.2 Harmony involving *á/é*

There are also alternations of the so-called 'long' versions of these vowels, as in the translative suffix *-vá/-vé* (the *v* assimilates to a preceding consonant): : *dobbá* '(turn) into a drum', *szemmé* '(turn) into an eye', *tökké* '(turn) into a pumpkin'.

8.3 Harmony involving *u/ü*, *ú/ű* and *ó/ő*

The other alternations with two forms always have round vowels. They agree with the backness of the preceding vowel. Some consistently have a short, high rounded vowel: *angolul* 'in English', *törökül* 'in Turkish', *lengyelül* 'in Polish'. Others have the long counterparts of these high vowels: *lábú* '-legged', *fejű* '-headed'. Finally, one set consistently has a long, mid rounded vowel: *dobtól* 'from a drum', *szemtől* 'from an eye', *töktől* 'from a pumpkin'. Thus, the height of the preceding vowel never affects the harmonic vowel: *tűztől* 'from a fire'.

8.4 Harmony involving *e/ö/o*

Finally we must consider harmony patterns with three alternants. These are restricted to a single pattern involving mid vowels: *e/ö/o*. This pattern is seen in the superessive suffix *-en/-ön/-on*: *szemen* 'on an eye', *tökön* 'on a pumpkin', *dobon* 'on a drum'.[12] The vowel of this suffix is present underlyingly, and it can be accounted for straightforwardly.

However, there are other vowels that participate in a vowel-zero alternation and may be analyzed as epenthetic. These vowels generally show the same *e/ö/o* three-way contrast. Such is the case of the accusative suffix: *szemet* 'eye-ACC', *tököt 'pumpkin-*ACC*'*, dobot *'drum-*ACC*'*. These 'unstable' vowels also surface as *a/e* after so-called 'lowering stems'. I leave a discussion of such stems aside for the purposes of this paper.

9 Vowel Phonetics

In the preceding section I referred to the vowels using standard orthography and occasional reference to their phonetic features. In this section I provide the IPA symbols that represent the vowels in standard pronunciation, along with a featural representation. The umlaut marks front rounded vowels. The lengthened umlaut marks long front rounded vowels, and an acute accent marks all the other long vowels. Thus, one of the two orthographic marks of length is always present on a phonetically long vowel, and never present (aside from a few frozen spellings) on short vowels. However, these vowels do not always differ from their 'short' counterparts in length alone—they may also differ featurally, with respect to height, rounding or tenseness.

The IPA symbols and featural descriptions I have assigned are based on examination of a variety of sources, most importantly consultation with native speaker linguists Péter Siptár and Sylvia Blaho.

[12] The superessive does also surface as just *-n* when the stem ends in a vowel, as in *kapu-n* 'on a gate'. I assume this results from a straightforward vowel deletion rule.

Many discussions of Hungarian vowels abstract away from predictable and non-distinctive features. For example, Siptár (1994) says of the *á* vowel that the height (extra low, in comparison with *a/e*, which are often treated as low) and backness (fairly central) is "a matter of phonetic implementation since in the (morpho)phonological pattern of Hungarian [this vowel] behaves as a low back vowel (*e.g.* with respect to vowel harmony, long/short alternations, *etc.*)" (175). I will assume a minimal amount of underspecification (following Inkelas 1996), and thus include such features. Assuming an inventory of binary vowel features that consists of {hi, lo, bk, rd, ATR} I prefer to remain faithful to the phonetic facts wherever possible and to show how doing so helps us to account for a complex set of interactions.

(26) Surface vowels of Hungarian

orthography	IPA	features	length
i	[i]	[+hi, -lo, -bk, -rd, +ATR]	SHORT
í	[iː]	[+hi, -lo, -bk, -rd, +ATR]	LONG
ü	[y]	[+hi, -lo, -bk, +rd, +ATR]	SHORT
ű	[yː]	[+hi, -lo, -bk, +rd, +ATR]	LONG
e	[ɛ]	[-hi, -lo, -bk, -rd, -ATR]	SHORT
é	[eː]	[-hi, -lo, -bk, -rd, +ATR]	LONG
ö	[ø]	[-hi, -lo, -bk, +rd, +ATR]	SHORT
ő	[øː]	[-hi, -lo, -bk, +rd, +ATR]	LONG
u	[u]	[+hi, -lo, +bk, +rd, +ATR]	SHORT
ú	[uː]	[+hi, -lo, +bk, +rd, +ATR]	LONG
o	[o]	[-hi, -lo, +bk, +rd, +ATR]	SHORT
ó	[oː]	[-hi, -lo, +bk, +rd, +ATR]	LONG
a	[ɔ]	[-hi, -lo, +bk, +rd, -ATR]	SHORT
á	[aː]	[-hi, +lo, +bk, -rd, +ATR]	LONG

Note that all the high vowels are [+ATR]. For the mid, front, unrounded vowels the short one is lax ([-ATR]), the long one tense ([+ATR]). However, the other four mid vowels, all the round ones, are all tense. Finally, note that orthographic *a* represents [ɔ], a mid, back, round, lax vowel, whereas *á* is low and unrounded and tense.

Most analyses of Hungarian abstract away from all of these surface distinctions, for example, treating *a* as an unrounded vowel for the purposes of harmony. In the following, I follow a different course.

10 Vowels in Underlying Representation

Following recent work such as Inkelas (1996) and others, I assume that underspecification is only posited by a learner, when it is forced to do so by alternations in the data. In other words, it is logically possible that Polish has /t/ *vs.* /D/, rather than /t/ *vs.* /d/, but we assume that /D/ is only posited if /t/ and /d/ are 'already assigned'. The child's default parse is that underlying segments are identical to surface segments, and this default is rejected only if alternations force a new analysis[13], as in Turkish. Polish does not present the child with data that will force such a change. Since the transparent parse works, the learner sticks with it and does not consider other extensionally equivalent grammars.[14]

(27) Inkelas on underspecification

> underlying representation is determined solely by optimization with respect to the grammar, not by imposing any type of constraints directly on underlying representation… [this] results in the use of underspecification only when there are alternant surface forms… (Inkelas 1996:1).

[13] See Hale & Reiss (1998) for arguments concerning the transparency of the child's initial parse.

[14] An additional source of underspecifi cation may arise in order to capture distributional patterns, but the point here is that there is no obvious need to minimize the amount of information stored in lexical representations, since this represents a longer learning path—see Hale & Reiss (2000) for discussion.

Inkelas specifically rejects philosophical arguments for underspecification:

(28) 'Grammar-blind' approaches to underspecification rejected by Inkelas 1996, q.v. for references

- **Markedness** (universal, language-specific, or contextual); unmarked material is underspecified
- **Redundancy**; redundant feature values (determined on the basis of the segment inventory) are underspecified
- **Predictability**: predictable material is underspecified

According to Inkelas, '[t]he only motivation for underspecification is to capture alternations in the optimal way' (1996: 2). The underspecification posited for the harmonic vowels of Hungarian will be guided by this principle.

10.1 Nonalternating vowels

In the case of non-alternating vowels, I assume that they enter and exit the phonology with the same featural representations. There are no redundancy rules, for example. Thus, both nonalternating and alternating high vowels are [+ATR] in the lexicon, and they surface as [+ATR]. The fact that [+ATR] is predictable from [+hi] is not encoded in the grammar.

10.1.1 Stem vowels

Aside from a few length alternations and vowel-zero alternations that will not concern us, all stem (and prefix) vowels in Hungarian have a constant value.[15] All of the vowels in (26) appear in stems, and many in prefixes. I assume that they all have the underlying and surface featural representations given above.

10.1.2 Suffix vowels

Vowel harmony in Hungarian only affects suffixes. However, not all suffixes are subject to harmony. The following suffixes are among those that do not alternate: *-ig* 'until', *-ként* 'as', *-kor* 'at a time': *hatkor* 'at six o'clock', *hétkor* 'at seven o'clock'.

10.2 Alternating vowels

I now propose underlying representations for the alternating harmonic vowels. I will assume that each of the alternations mentioned in section 8 can be captured *via* a partially underspecified representation and a set of phonological rules that **fill in** some feature values—no feature changing rules are needed. Length is mentioned in the prose, but not represented featurally, since I assume it is represented through multiple linking to a timing tier.[16] .

10.2.1 The lexical form of the *a/e* vowel

Suffixes such as the inessive *-ban/-ben*, illative *-ba/-be* and sublative *-ra/re* are all stored with the following feature values for their alternating vowel: [-hi, -lo, -ATR]. This should not be controversial, since such suffixes do in fact alternate with respect to [bk] and [rd] (*a* is [+round]). The specification [-lo] is maintained, since both realizations of the vowel, [ɔ, ɛ], are mid vowels.

[15] Of course, a full evaluation of my proposal will ultimately require accounting for the length alternations as well. I leave this for future research.

[16] The careful reader will have noted that I earlier (fn.11) rejected some uses of autosegmental phonology, but am here appealing to an autosegmental representation of length. This will have to be dealt with.

10.2.2 The lexical form of the *á/é* vowel

These vowels are never high or round, but they alternate between front mid and low back. I assume that both forms are [+ATR], although this may be debatable. Accordingly, the underlying form is [-hi, -rd, +ATR].

10.2.3 The lexical form of the consistently rounded vowels

These alternations can all be generated by the same rule—the only alternating feature is [bk].

- *u/ü*
  This alternation is between a front and back lax high round vowel. It is stored as [+hi, -lo, +rd, +ATR], with no specification for [bk]. This feature is filled in by context.
- *ú/ű*
  This alternation is between long vowels that are consistently high, round and tense. These vowels are represented as [+hi, -lo, +rd, +ATR].
- *ó/ő*
  This alternation is between long vowels that are consistently mid, round and tense. These vowels are represented as [-hi, -lo, +rd, +ATR].

10.2.4 The lexical form of the *e/ö/o* vowel

The most complex alternation to deal with is that which shows three surface forms. As mentioned above, the presence of a vocalic feature bundle is underlying in some cases of three-way alternations and epenthetic in others. Of course, in the latter case, it is misleading to refer to their lexical representation. However, I will assume that both underlying and epenthetic three-form suffixes enter the phonology as [-hi, -lo] 'Feature-filling' rules will provide the values for [bk] and [rd] and [ATR].

10.3 Personal pronoun forms

Corresponding to the case endings discussed above are a set of stems that inflect for person. For example, the inessive *-ban/-ben* corresponds to *bennem, benned, benne, bennünk, bennetek, bennük* 'in me, in you ...'. The stem appears with the front vowel [ɛ] with which the person endings harmonize. Similarly, the delative *-ról/-/ről* corresponds to *rólam, rólad, róla, rólunk, rólatok, róluk* 'from off of me, from off of you ...', all with back vowel suffixes. It is thus tempting to see these case stems as providing evidence for the underlying vowel quality of the case suffixes.

However, in several categories, the case stem and case suffix do *not* match phonologically. The superessive ending *-en, -on, -ön* corresponds to the case stem in *rajtam, rajtad, rajta, rajtunk, rajtatok, rajtuk* 'on me, on you ...'. The elative *-ból/-/ből* corresponds to the case stem in *belőlem, belőled, belőle, belőlünk, belőletek, belőlük* 'from in me, from in you ...'. Thus I will assume that the case endings and the case stems are listed independently in the lexicon. In other words, *contra* Ringen & Vago (1998), I conclude that there is a set of suffixes, such as /-bVn/, with underspecified, alternating vowels, and a set of stems such as /bɛn-/ with fully specified, nonalternating vowels. In some instances (*e. g.* , inessive /-bVn/ and /bɛn-/), the forms are clearly related etymologically, whereas in others (*e. g.* , elative/-bVl/ and /belől-/, or superessive /-Vn/ and /rajt-/) the relationship is opaque. The case stems contain fully specified vowels that trigger harmony in the personal suffixes; the case endings harmonize with the stems they attach to—there is no need to posit feature-changing processes in those instances where the case stem and case suffix just happen to look alike.

11 The Rules

In this paper, I am concerned only with capturing the interactions between triggers and targets of harmony. Therefore, I will not address the issue of transparent vowels or of roots with surface front vowels that trigger back harmony. I will also not propose an explicit mechanism for harmony in terms of syllable structure or higher-level structures. I will adopt a simplified view in which harmony occurs between a vowel and the vowel that precedes it. The following seven rules are posited with crucial ordering noted. All seven are what would be called feature-filling or structure-filling rules in other accounts.[17] This does not need to be stated for each rule, since it follows from the UIP.

$\mathcal{R}_1$: V → [αbk] / [αbk] _
This rule applies first. According to the UIP, it can only apply to vowels that have no specification for [bk]. Thus it applies to all of the alternating vowels and none of the non-alternating ones.

$\mathcal{R}_2$: [-lo, +bk] → [+rd]
In accordance with the UIP, this rule applies to [-lo, +bk] that are NOT [-rd]. The rule applies vacuously to already rounded vowels that are also [-lo, +bk], and in feature filling fashion to [-lo, +bk] vowels that are unspecified for [rd]. Since $\mathcal{R}_2$ applies to vowels which became [+bk] *via* $\mathcal{R}_1$, $\mathcal{R}_2$ must follow $\mathcal{R}_1$. For example, a vowel that enters the phonology as [-hi, -lo] could become [+bk] by $\mathcal{R}_1$ if it follows a back vowel, and then [+rd] by $\mathcal{R}_2$. It would then receive and ATR value from $\mathcal{R}_6$ and at this point it would be fully specified for all the vowel features.

$\mathcal{R}_3$: [+bk, -rd] → [+lo]
This rule ensures that *á* is realized as a low vowel, the only one in the language. Crucially, the vowel of the *á/é* alternation is underlyingly specified as [-rd]. The rule refers to [+bk], and since all the alternating vowels are underlyingly unspecified for [bk], $\mathcal{R}_3$ must follow $\mathcal{R}_1$.

$\mathcal{R}_4$: [-hi, -bk, -ATR] → [-rd]
This rule ensures that the *e* of the *e/a* alternation surfaces as unround. It cannot effect this feature insertion for the *e* of the *e/ö/o* alternation, since the latter is not specified as [-ATR] until $\mathcal{R}_6$ applies.

$\mathcal{R}_5$: V → [αrd] / [αrd] _
By the UIP, this rule cannot apply to any vowel that is already specified for rounding. It turns out that it only ends up affecting the front vowels of the *e/ö/o* alternation. All other vowels that receive rounding by rule get [+rd] from $\mathcal{R}_2$ or $\mathcal{R}_4$. Thus, $\mathcal{R}_5$ follows these two.

$\mathcal{R}_6$: [-hi, -lo, αrd] → [αATR]
This rule makes mid vowels that are unspecified for [ATR] receive an [ATR] value that agrees with their rounding value. The UIP requires that inputs to this rule not have an [ATR] value, thus there is no chance that long *é* will be affected by this rule. This rule applies after $\mathcal{R}_2$ and also after $\mathcal{R}_5$, since these fill in the values of [rd] required for the rule to apply.

$\mathcal{R}_7$: [-hi, -bk] → [-lo]
This provides the [-lo] value to the front member of the *á/é* alternation. Since the SD refers to [bk], $\mathcal{R}_7$ must follow $\mathcal{R}_1$. The rule will apply vacuously to a derived [-hi, -bk, -lo] *e*, as well.

12 Derivations

The following tables provide derivations of each of the alternations. Input and output are given with orthographic vowels. The only features of the trigger of harmony that are relevant are [bk] and [round], so it is not necessary to provide all trigger-target pairs. For example, the vowel *u* will trigger the exact same harmony patterns as *o*. I have provided schematic derivations for four trigger vowels representing each of the four possible combinations of these features: [+bk, -rd], [+bk, +rd], [-bk, -rd] and [-bk, -rd].

In the first table (29), I have annotated the cases where the rule does not affect the input, whereas in the subsequent tables I merely note 'N.A.' for 'not affected'. If a box is marked NS (no subsumption), it

[17]Some also happen to be what are traditionally called fi ll-in rules—rules that supply default values only with reference to context within the segment itself.

means that the SD of the rule does not subsume the potential input representation. For example, $\mathcal{R}_3$ does not apply to any forms in table (29), since the structural description of $\mathcal{R}_3$ does not subsume any of the input representations in this table. Thus, such cases conform to standard practice—the rule does not apply because it does not match the SD.

However, some cells are marked UIP to show that the current rule's SD would be met under traditional assumptions, but it is not met under the UIP. For example, $\mathcal{R}_5$ does not apply to any forms in (29) since all the inputs are specified for [rd] by this point in the derivation, and $\mathcal{R}_5$ must be feature-filling by the UIP.

Finally, some cells are marked VAC, where the SD is satisfied under the UIP, but the effect of the rule is not visible (it applies vacuously). This is the case for $\mathcal{R}_7$ in the second and fourth data columns in (29), where it applies vacuously to vowels that are already [-hi, -lo, -bk]. The only rules that affect the output are $\mathcal{R}_1$ (for all forms in this table), $\mathcal{R}_2$ (forms preceded by a back vowel trigger), and $\mathcal{R}_4$ (forms preceded by a front vowel trigger).

(29) Generating the *a/e* alternation

	á __	e __	o __	ö __
In.	[−hi, +lo, +bk, −rd, +ATR] [−hi, −lo, −ATR]	[−hi, −lo, −bk, −rd, −ATR] [−hi, −lo, −ATR]	[−hi, −lo, +bk, +rd, +ATR] [−hi, −lo, −ATR]	[−hi, −lo, −bk, +rd, +ATR] [−hi, −lo, −ATR]
$\mathcal{R}_1$	——— [−hi, −lo, +bk, −ATR]	——— [−hi, −lo, −bk, −ATR]	——— [−hi, −lo, +bk, −ATR]	——— [−hi, −lo, −bk, −ATR]
$\mathcal{R}_2$	——— [−hi, −lo, +bk, +rd, −ATR]	NS	——— [−hi, −lo, +bk, +rd, −ATR]	NS
$\mathcal{R}_3$	NS	NS	NS	NS
$\mathcal{R}_4$	NS	——— [−hi, −lo, −bk, −rd, −ATR]	NS	——— [−hi, −lo, −bk, −rd, −ATR]
$\mathcal{R}_5$	UIP	UIP	UIP	UIP
$\mathcal{R}_6$	UIP	UIP	UIP	UIP
$\mathcal{R}_7$	NS	VAC	NS	VAC
Out.	*á-a*	*e-e*	*o-a*	*ö-e*

In (30) we derive the *á/é* alternation. We see that rules $\mathcal{R}_1$ and $\mathcal{R}_3$ affect the forms triggered by preceding back vowels, whereas rules $\mathcal{R}_1$ and $\mathcal{R}_7$ affect the forms triggered by preceding front vowels.

(30) Generating the *á/é* alternation

	á __	e __	o __	ö __
In.	$\begin{bmatrix}-\text{hi}\\+\text{lo}\\+\text{bk}\\-\text{rd}\\+\text{ATR}\end{bmatrix}\begin{bmatrix}-\text{hi}\\-\text{rd}\\+\text{ATR}\end{bmatrix}$	$\begin{bmatrix}-\text{hi}\\-\text{lo}\\-\text{bk}\\-\text{rd}\\-\text{ATR}\end{bmatrix}\begin{bmatrix}-\text{hi}\\-\text{rd}\\+\text{ATR}\end{bmatrix}$	$\begin{bmatrix}-\text{hi}\\-\text{lo}\\+\text{bk}\\+\text{rd}\\+\text{ATR}\end{bmatrix}\begin{bmatrix}-\text{hi}\\-\text{rd}\\+\text{ATR}\end{bmatrix}$	$\begin{bmatrix}-\text{hi}\\-\text{lo}\\-\text{bk}\\+\text{rd}\\+\text{ATR}\end{bmatrix}\begin{bmatrix}-\text{hi}\\-\text{rd}\\+\text{ATR}\end{bmatrix}$
$\mathcal{R}_1$	——— $\begin{bmatrix}-\text{hi}\\-\text{rd}\\+\text{bk}\\+\text{ATR}\end{bmatrix}$	——— $\begin{bmatrix}-\text{hi}\\-\text{rd}\\-\text{bk}\\+\text{ATR}\end{bmatrix}$	——— $\begin{bmatrix}-\text{hi}\\-\text{rd}\\+\text{bk}\\+\text{ATR}\end{bmatrix}$	——— $\begin{bmatrix}-\text{hi}\\-\text{rd}\\-\text{bk}\\+\text{ATR}\end{bmatrix}$
$\mathcal{R}_2$	N.A.	N.A.	N.A.	N.A.
$\mathcal{R}_3$	——— $\begin{bmatrix}-\text{hi}\\+\text{lo}\\-\text{rd}\\+\text{bk}\\+\text{ATR}\end{bmatrix}$	N.A.	——— $\begin{bmatrix}-\text{hi}\\+\text{lo}\\-\text{rd}\\+\text{bk}\\+\text{ATR}\end{bmatrix}$	N.A.
$\mathcal{R}_4$	N.A.	N.A.	N.A.	N.A.
$\mathcal{R}_5$	N.A.	N.A.	N.A.	N.A.
$\mathcal{R}_6$	N.A.	N.A.	N.A.	N.A.
$\mathcal{R}_7$	N.A.	——— $\begin{bmatrix}-\text{hi}\\-\text{lo}\\-\text{rd}\\-\text{bk}\\+\text{ATR}\end{bmatrix}$	N.A.	——— $\begin{bmatrix}-\text{hi}\\-\text{lo}\\-\text{rd}\\-\text{bk}\\+\text{ATR}\end{bmatrix}$
Out.	*á-á*	*e-é*	*o-á*	*ö-é*

The next two tables (31, 32) show the derivation of front/back alternations of consistently round vowels. In both cases the only relevant rule is $\mathcal{R}_1$. Since the long high round vowels *ű, ú* behave in completely parallel fashion to the short ones *ü, u*, we show only the latter.

(31) Generating the *u/ü* alternation

	á __	e __	o __	ö __
In.	$\begin{bmatrix}-\text{hi}\\+\text{lo}\\+\text{bk}\\-\text{rd}\\+\text{ATR}\end{bmatrix}\begin{bmatrix}+\text{hi}\\-\text{lo}\\+\text{rd}\\+\text{ATR}\end{bmatrix}$	$\begin{bmatrix}-\text{hi}\\-\text{lo}\\-\text{bk}\\-\text{rd}\\+\text{ATR}\end{bmatrix}\begin{bmatrix}+\text{hi}\\-\text{lo}\\+\text{rd}\\+\text{ATR}\end{bmatrix}$	$\begin{bmatrix}-\text{hi}\\-\text{lo}\\+\text{bk}\\+\text{rd}\\+\text{ATR}\end{bmatrix}\begin{bmatrix}+\text{hi}\\-\text{lo}\\+\text{rd}\\-\text{ATR}\end{bmatrix}$	$\begin{bmatrix}-\text{hi}\\-\text{lo}\\-\text{bk}\\+\text{rd}\\+\text{ATR}\end{bmatrix}\begin{bmatrix}+\text{hi}\\-\text{lo}\\+\text{rd}\\+\text{ATR}\end{bmatrix}$
$\mathcal{R}_1$	——— $\begin{bmatrix}+\text{hi}\\-\text{lo}\\+\text{bk}\\+\text{rd}\\+\text{ATR}\end{bmatrix}$	——— $\begin{bmatrix}+\text{hi}\\-\text{lo}\\-\text{bk}\\+\text{rd}\\+\text{ATR}\end{bmatrix}$	——— $\begin{bmatrix}+\text{hi}\\-\text{lo}\\+\text{bk}\\+\text{rd}\\+\text{ATR}\end{bmatrix}$	——— $\begin{bmatrix}+\text{hi}\\-\text{lo}\\-\text{bk}\\+\text{rd}\\+\text{ATR}\end{bmatrix}$
$\mathcal{R}_2$	N.A.	N.A.	N.A.	N.A.
$\mathcal{R}_3$	N.A.	N.A.	N.A.	N.A.
$\mathcal{R}_4$	N.A.	N.A.	N.A.	N.A.
$\mathcal{R}_5$	N.A.	N.A.	N.A.	N.A.
$\mathcal{R}_6$	N.A.	N.A.	N.A.	N.A.
$\mathcal{R}_7$	N.A.	N.A.	N.A.	N.A.
Out.	*á-u*	*e-ü*	*o-u*	*ö-ü*

(32) Generating the *ó/ő* alternation

	á __	e __	o __	ö __
In.	$\begin{bmatrix}-\text{hi}\\+\text{lo}\\+\text{bk}\\-\text{rd}\\+\text{ATR}\end{bmatrix}\begin{bmatrix}-\text{hi}\\-\text{lo}\\+\text{rd}\\+\text{ATR}\end{bmatrix}$	$\begin{bmatrix}-\text{hi}\\-\text{lo}\\-\text{bk}\\-\text{rd}\\-\text{ATR}\end{bmatrix}\begin{bmatrix}-\text{hi}\\-\text{lo}\\+\text{rd}\\+\text{ATR}\end{bmatrix}$	$\begin{bmatrix}-\text{hi}\\-\text{lo}\\+\text{bk}\\+\text{rd}\\+\text{ATR}\end{bmatrix}\begin{bmatrix}-\text{hi}\\-\text{lo}\\+\text{rd}\\+\text{ATR}\end{bmatrix}$	$\begin{bmatrix}-\text{hi}\\-\text{lo}\\-\text{bk}\\+\text{rd}\\+\text{ATR}\end{bmatrix}\begin{bmatrix}-\text{hi}\\-\text{lo}\\+\text{rd}\\+\text{ATR}\end{bmatrix}$
$\mathcal{R}_1$	——— $\begin{bmatrix}-\text{hi}\\-\text{lo}\\+\text{bk}\\+\text{rd}\\+\text{ATR}\end{bmatrix}$	——— $\begin{bmatrix}-\text{hi}\\-\text{lo}\\-\text{bk}\\+\text{rd}\\+\text{ATR}\end{bmatrix}$	——— $\begin{bmatrix}-\text{hi}\\-\text{lo}\\+\text{bk}\\+\text{rd}\\+\text{ATR}\end{bmatrix}$	——— $\begin{bmatrix}-\text{hi}\\-\text{lo}\\-\text{bk}\\+\text{rd}\\+\text{ATR}\end{bmatrix}$
$\mathcal{R}_2$	N.A.	N.A.	N.A.	N.A.
$\mathcal{R}_3$	N.A.	N.A.	N.A.	N.A.
$\mathcal{R}_4$	N.A.	N.A.	N.A.	N.A.
$\mathcal{R}_5$	N.A.	N.A.	N.A.	N.A.
$\mathcal{R}_6$	N.A.	N.A.	N.A.	N.A.
$\mathcal{R}_7$	N.A.	N.A.	N.A.	N.A.
Out.	*á-ó*	*e-ő*	*o-ó*	*ö-ő*

In (33) rules $\mathcal{R}_1$ and $\mathcal{R}_6$ affect all forms. $\mathcal{R}_2$ only affects forms triggered by back vowels, and $\mathcal{R}_5$ only affects forms triggered by front vowels. The input vowel is merely specified [-hi, -lo], but a three way *e/ö/o* alternation is generated.

(33) Generating the *e/ö/o* alternation

	á __	e __	o __	ö __
In.	$\begin{bmatrix} -\text{hi} \\ +\text{lo} \\ +\text{bk} \\ -\text{rd} \\ +\text{ATR} \end{bmatrix} \begin{bmatrix} -\text{hi} \\ -\text{lo} \end{bmatrix}$	$\begin{bmatrix} -\text{hi} \\ -\text{lo} \\ -\text{bk} \\ -\text{rd} \\ -\text{ATR} \end{bmatrix} \begin{bmatrix} -\text{hi} \\ -\text{lo} \end{bmatrix}$	$\begin{bmatrix} -\text{hi} \\ -\text{lo} \\ +\text{bk} \\ +\text{rd} \\ +\text{ATR} \end{bmatrix} \begin{bmatrix} -\text{hi} \\ -\text{lo} \end{bmatrix}$	$\begin{bmatrix} -\text{hi} \\ -\text{lo} \\ -\text{bk} \\ +\text{rd} \\ +\text{ATR} \end{bmatrix} \begin{bmatrix} -\text{hi} \\ -\text{lo} \end{bmatrix}$
$\mathcal{R}_1$	——— $\begin{bmatrix} -\text{hi} \\ -\text{lo} \\ +\text{bk} \end{bmatrix}$	——— $\begin{bmatrix} -\text{hi} \\ -\text{lo} \\ -\text{bk} \end{bmatrix}$	——— $\begin{bmatrix} -\text{hi} \\ -\text{lo} \\ +\text{bk} \end{bmatrix}$	——— $\begin{bmatrix} -\text{hi} \\ -\text{lo} \\ -\text{bk} \end{bmatrix}$
$\mathcal{R}_2$	——— $\begin{bmatrix} -\text{hi} \\ -\text{lo} \\ +\text{bk} \\ +\text{rd} \end{bmatrix}$	N.A.	——— $\begin{bmatrix} -\text{hi} \\ -\text{lo} \\ +\text{bk} \\ +\text{rd} \end{bmatrix}$	N.A.
$\mathcal{R}_3$	N.A.	N.A.	N.A.	N.A.
$\mathcal{R}_4$	N.A.	N.A.	N.A.	N.A.
$\mathcal{R}_5$	N.A.	——— $\begin{bmatrix} -\text{hi} \\ -\text{lo} \\ -\text{bk} \\ -\text{rd} \end{bmatrix}$	N.A.	——— $\begin{bmatrix} -\text{hi} \\ -\text{lo} \\ -\text{bk} \\ +\text{rd} \end{bmatrix}$
$\mathcal{R}_6$	——— $\begin{bmatrix} -\text{hi} \\ -\text{lo} \\ +\text{bk} \\ +\text{rd} \\ +\text{ATR} \end{bmatrix}$	——— $\begin{bmatrix} -\text{hi} \\ -\text{lo} \\ -\text{bk} \\ -\text{rd} \\ -\text{ATR} \end{bmatrix}$	——— $\begin{bmatrix} -\text{hi} \\ -\text{lo} \\ +\text{bk} \\ +\text{rd} \\ +\text{ATR} \end{bmatrix}$	——— $\begin{bmatrix} -\text{hi} \\ -\text{lo} \\ -\text{bk} \\ +\text{rd} \\ +\text{ATR} \end{bmatrix}$
$\mathcal{R}_7$	N.A.	N.A.	N.A.	N.A.
Out.	*á-o*	*e-e*	*o-o*	*ö-ö*

We thus see that a complex set of vowel-vowel interactions can be derived using a set of simple rules, appropriately applied in accordance with the UIP. As mentioned in passing, there clearly remain other problems to solve in the domain of Hungarian vowels, and it will be interesting to see if the present proposal can shed new light on these classic problems. While there is no logical necessity to derive these patterns exclusively with the equivalent of feature-filling rules, doing so allows us to simplify the account of vowel harmony in a straightforward fashion, a topic we now turn to.

13 Mixed stems

By generating Hungarian vowel harmony in this way, we find it is unnecessary explain why harmony does not occur between stems in a compound (e.g., *balta+nyél* 'hatchet handle') or between a prefix and a stem (e.g., *meg+lát* 'catch sight of'), despite the standard view on this matter as expressed by authors like Spencer (1996:178) and Siptár & Törkenczy (2000: Chapter 6). These authors attribute the failure of harmony to apply in such cases to the assignment of prefixes and first compound members to separate phonological words from the stem that follows them. While such a division may be reasonable on other grounds, vowel harmony, or rather its failure to apply, provides no evidence for a phonological word break between such constituents. Under the model presented here, all prefix and stem vowels are fully specified[18] and the vowel harmony rules I propose only affect underspecified vowels, when interpreted using the UIP.

14 Conclusions

We have seen that the goal of conciseness, or maximal generality, in rule formulation was inconsistent with the search for UG. Rules were formulated in order to generate a given corpus of data, not with the aim of

[18]Aside from some epenthetic, 'unstable' vowels, which receive their full specification from within their stem.

modeling a single human phonological component. It is desirable to have a single interpretive procedure for determining whether a representation satisfies the Structural Description of a rule, and thus serves as an input to the rule. This single procedure should work for all rules in all languages. In this paper, I have explored a logical alternative to the interpretive procedure adopted in SPE, and argued that this Unified Interpretive Procedure solves longstanding problems by making it unnecessary to refer to a third value of binary features [∅F], to introduce negation into lexical representations (*e.g.* [NOT+rd]) or to introduce a FEATURE-FILLING/FEATURE-CHANGING diacritic on rules.

I have argued that a small set of rules, provided with the UIP can generate a set of data as complex as the target-trigger relations of Hungarian vowel harmony. The UIP provides us with a principled distinction between feature filling and feature changing rules, one that is read off of the rule's representation mechanically. We do not need labels (or a clever homunculus) to tell us if a rule is feature-filling or feature changing, and we do not need diacritics in representations to refer to features that are not present. If the purely feature filling analysis of Hungarian presented here[19] is in fact valid, then we have shown that this conclusion leads to an overall simplification in our understanding of Hungarian phonology—there is less evidence for positing phonological domains within words to account for the lack of harmony in mixed stems and compounds. Not surprisingly, rather than being "ad-hoc and sterile", attention to theoretical detail can lead to improved analyses.

One might wonder whether the UIP is just a notational variant of the rule labels 'feature-filling' and 'feature changing', since it appears to replace these two labels with a disjunctive condition on rule application. There is a good argument that it is not. Labels could be used with great latitude, in principle. For example, using the mechanism of rule labeling, a rule could be stated that filled in values for [rd] only on segments *lacking* specification for [hi]. The notion of a rule label does not preclude this possibility, though I believe it is not attested, and the UIP tells us why it is not—feature-filling and feature-changing differences derive from the UIP. The UIP uses the SC of a rule to determine which features must be absent from a representation for it to serve as an input to that same rule. In other words, the UIP explains why just the feature that appears in the SC is the one for which a rule can require obligatory absence.

The UIP allows representations in the lexicon to be non-distinct, and it even makes it possible for one representation to subsume another, since the UIP will allow the two to be treated differently by the grammar in a principled fashion. Specifically, a rule can affect a representation A, yet fail to affect another representation B, even when A is more general than (subsumes) B. As the reference to earlier work by Inkelas & Orgun, Kiparsky, Archangeli and McCarthy demonstrates, this has been a longstanding problem in phonological theory.

By developing a workable (albeit incomplete) underspecification analysis for Hungarian, we are working towards a simpler model of Hungarian phonology, in general. If the contrast between suffix alternations and prefix constancy, for example, does not call for a stratal phonology, the question arises of how general this result is—both for Hungarian and for universal phonology. Can the acceptance of the UIP as formulated here be used to simplify other analyses that drew on the arsenal of traditional generative phonology, such as the Derived Environment Condition, the Strict Cycle Condition, Structure Preservation and the ElsewhereCondition? This important issue will provide fertile terrain for future research.

At this point we can compare the quotation from Morris Halle with which we began to a competing view of the proper goals of phonology:

> We urge a reassessment of this essentially formalist position. If phonology is separated from the principles of well-formedness (the 'laws') that drive it, the resulting loss of constraint and theoretical depth will mark a major defeat for the enterprise [Prince & Smolensky (1993: 198, see also p.3)]

This paper is an attempt to support Halle's view. We should develop the necessary formal tools to express phonological processes, and we should understand how our notation works, whether it is constraint- or rule-based. I find this to be a more interesting and pressing problem, one that can lead to greater 'theoretical depth', than a taxonomic encoding of dubious 'principles of wellformedness'.

[19] Or another feature fi lling analysis.

References

Anderson, Stephen. 1985. *Phonology in the Twentieth Century*. University of Chicago.

Archangeli, Diana. 1988. *Underspecification in Yawelmani Phonology and Morphology*. (revision of 1984 MIT dissertation). New York: Garland.

Bayer, Sam & Mark Johnson. 1995. Features and Agreement. *Proceedings of the 33rd Annual Meeting of the Association for Computational Linguistics.* San Francisco: Morgan Kaufmann.

Chomsky, Noam. 1967. Some General Properties of Phonological Rules. *Language* 43:102-128.

Chomsky, Noam and Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper and Row.

Hale, Mark & Charles Reiss. 2000. The subset principle in phonology: Why the *tabula* can't be *rasa*. Under revision for *Journal of Linguistics*.

Hale, M. & C. Reiss. 1998. Formal and empirical arguments concerning phonological acquisition. *Linguistic Inquiry* 29: 656-683.

Halle, Morris. 1975. Confessio Grammatici. *Language* 51:525-35.

Halle, Morris & G.N. Clements. *Problem Book in Phonology*. Cambridge, MA: MIT Press.

Inkelas, Sharon. 1996. Archiphonemic Underspecifi cation. Ms. UC Berkeley.

Inkelas, Sharon and Orhan Orgun. 1995. Level ordering and economy in the lexical phonology of Turkish. *Language 71*. 763-793.

Kiparsky, Paul. 1985. Some Consequences of Lexical Phonology. *Phonology Yearbook* 2:85-138.

McCarthy, John. 1994. The phonetics and phonology of Semitic pharyngeals. In P. Keating, ed., *Phonological Structure and Phonetic Form: Papers in Laboratory Phonology III*. 191-233. Cambridge University Press.

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

Reiss, Charles. 2003a. Towards a theory of fundamental phonological relations. To appear in A.M. di Sciullo, ed., *Papers on Linguistic Asymmetry.* John Benjamins Press..

Reiss, Charles. 2003b. Quantifi cation in phonological rules. To appear in *The Linguistic Review 20: Special Issue on Typology in Phonology*, Spring 2003.

Ringen, Catherine & Robert Vago. 1998. Hungarian vowel harmony in Optimality Theory. *Phonology* 15.3: 393-416.

Siptár, Péter. 1994. The vowel inventory of Hungarian: Its size and structure. *The Even Yearbook 1994*:175–184. Budapest: ELTE SEAS Working Papers in Linguistics.

Siptár, Péter & Miklós Törkenczy. 2000. *The Phonology of Hungarian.* Oxford: Oxford University Press.

Spencer, Andrew. 1996. *Phonology*. Oxford: Blackwell.