Phonological Restructuring in Yidiɲ and its Theoretical Consequences

Bruce P. Hayes

UCLA

November 1997

Abstract

Detailed study of data in Dixon's (1977) grammar indicates that previous analyses of Yidiɲ have erred in supposing that the synchronic pattern of the language continues the historical pattern, whereby various nominal stems have their underlying final vowels deleted when no suffix follows. Instead, it appears that the system has undergone a radical reanalysis, whereby the suffixed forms are now projectable by general principles from the isolation forms. More precisely, a pattern of **multiple predictability** has developed: the form of suffixed allomorphs is largely predictable from the isolation allomorphs, but the older pattern, whereby isolation allomorphs can be predicted from the suffixed allomorphs, also persists.

From this descriptive result, three principal theoretical consequences are developed: (a) Yidiɲ possesses a fully-productive pattern of alternation that is not driven by markedness-faithfulness interactions; (b) the phonological constraints that are active in Yidiɲ likely include some that are quite unlikely to be members of a universal inventory; (c) there are more relations of predictability among surface forms in Yidiɲ than can be treated by the normal method, namely that of deriving all the surface allomorphs from a single underlying representation.

A tentative suggestion is made for how Optimality Theory might be extended to treat cases of this sort, by means of a class of "Anticorrespondence" constraints.

Phonological Restructuring in Yidiŋ and its Theoretical Consequences[1]

1. Introduction: Where Does Phonology Come From?

Optimality Theory (Prince and Smolensky 1993) has yielded a compelling picture of phonology, one in which quite intricate data patterns emerge from the relative ranking of markedness and faithfulness constraints. The markedness constraints are characteristically highly general, supported by cross-language typology, and in many instances explicable in terms of functional goals such as those involving articulation or perception. Faithfulness constraints are characteristically atomic, limited to bans on simple insertions, deletions, and feature changes. The theory is receiving increasing support from the study of acquisition (e.g. Gnanadesikan 1995, Pater 1996), from which it appears that many of the markedness constraints are spontaneously manifested in the speech of small children; thus the Stampean idea that phonology is to some degree internally-generated (Stampe 1973) has come again to the fore.

All of this looks like progress, especially when compared with the degree of ad hocness and language-specificity that prevailed in earlier work. The caution introduced here is this: does *all* phonology result from markedness-faithfulness interactions?

To begin, we will consider how exceptional phonology, not treatable under standard Optimality-theoretic assumptions, would be most likely to arise in the course of historical change.

It is a commonplace that phonologies are not transmitted directly from generation to generation; rather, it is *language data* that are transmitted, and the phonology must be constructed by each new generation on the basis of the input data, along with whatever is provided a priori by the human language faculty.

The process is not necessarily reliable, as language change attests. Consider in particular a case where the older generation comes to apply with ever-greater frequency a process of phonological deletion, plausibly the most dramatic of phonological changes. As the frequency of deletion comes to approach 100%, new language learners will ultimately be faced with a situation in which the recovery of underlying forms is not at all easy: the lost segments of non-alternating forms will not be recoverable at all, and even in alternating forms the character of lost segments might be difficult to obtain. Such situations are ripe for *restructuring*, the creation of a novel phonological system on the basis of the data pattern left behind by an older system.

---

[1] I would like to thank René Kager, Donca Steriade, the volume editors, and the participants in a UCLA seminar on Correspondence Constraints for helpful input on this research project.

Thus, there are plausibly two types of phonological change. One is gradual, system-internal change, in which markedness and faithfulness constraints are reranked; this is expected to yield coherent patterns amenable to treatment by standard OT approaches. The other is restructuring, which arises out of the attempt of a new generation to make sense out of a data pattern presented to it by an older generation. The data pattern may be a quite difficult one, due to the accidents of history, and it is an empirical question whether the language learners who carry out restructuring will necessarily limit themselves to systems expressible with only these formal resources.

What all this implies is that if we are to explore the full richness of human phonologies, clearly we should devote some attention to systems that are known to be restructured. It seems quite possible that such systems will put the mechanisms of the theory to the severest test, and suggest directions for revision.

This is what I have tried to do here. I will argue that the phonology of Yidiŋ, an Australian aboriginal language of North Queensland, has undergone a substantial restructuring, one which puts it beyond the reach of the central mechanisms of markedness and faithfulness in OT, and further seems mostly likely to be incompatible with the view of a universal constraint inventory. I will further speculate as to what modifications of the theory might be capable of producing workable OT analyses of Yidiŋ.

The research I report here extends an earlier tradition, whose important works include Vennemann (1972), Schuh (1972), Hale (1973), and Kenstowicz and Kisseberth (1977). Work in this earlier line likewise suggested restructuring as an important origin of phonological processes.

2. Yidiŋ

Yidiŋ phonology was worked out with considerable insight by R. M. W. Dixon (1977, hereafter D), in a intensively detailed description based on work with several of the last native speakers of the language. Dixon's data and generalizations have proven irresistible to phonological theorists, who have principally aimed to elucidate the interesting and unusual metrical structure of the language with particular versions of metrical stress theory.[2]

The focus here lies not in the metrical structures of Yidiŋ per se, but rather in the complex patterns (partly metrically governed) that are found throughout the Yidiŋ nominal and verbal

[2] See Nash (1979), Hayes (1980, 1982, 1995a), Halle and Vergnaud (1987), Kirchner (1993), Crowhurst and Hewitt (1995), and further references cited by Crowhurst and Hewitt.

paradigms. Almost every Yidiɲ stem shows alternations of vowel length, and many stems show vowel-zero alternations as well.

I will first review the Yidiɲ data from a historical viewpoint. This will have two advantages: it avoids an a priori expository prejudice for any particular analytical line, and it helps set the stage for a restructuring account, by showing what kind of data pattern must have faced the innovating generation of Yidiɲ speakers.

2.1 Historical Yidiɲ

The dramatic phonological alternations of Yidiɲ are largely the product of two historical changes. One of them is characterized by Dixon (D 43) as follows:

(1) **Penultimate Lengthening**

In every word with an odd number of syllables, the penultimate vowel is lengthened.

Penultimate Lengthening resulted in a huge number of alternations. For example (D 43), the underlying stem for 'mother', /mud^{j}am/, appears as [mud^{j}am] in the absolutive case, where no suffix is added; this is unaltered from its historical form. With the Purposive ending [-gu] added, the resulting form is trisyllabic (odd-syllabled), and thus was eligible for Penultimate Lengthening, which yielded [mud^{j}**a:**mgu]. The trisyllabic stem for 'dog', historically *[gudaga], appears in modern Yidiɲ as as [gud**a:**ga] in the absolutive case, since it is trisyllabic. But when the purposive suffix [-gu] is added to it, the form becomes even-syllabled, and therefore Penultimate Lengthening did not affect it: [gud**a**ga-gu]. A quadrisyllabic stem, [ŋunaŋgara] 'whale' (D 84) was unaltered in the absolutive, but received penultimate length in (for example) the dative, where a suffix renders the form pentasyllabic: [ŋunaŋgar**a:**-nda].

Naturally, one senses that there should be some connection between the interesting environment "penult of an odd-syllabled word" and the alternating stress pattern of Yidiɲ, laid out elsewhere in Dixon's grammar (D 40-41). To establish and formalize this connection is one goal of the many metrical accounts of Yidiɲ phonology.

The pattern expressed by Penultimate Lengthening continues to be highly productive in synchronic Yidiɲ, and is virtually exceptionless on the surface.

The other major phonological change that created modern Yidiɲ was as follows:

(2) **Final Syllable Deletion**

In a word ending in C_1 (C_2) V, delete C_2 and V, if:

a. The form that results ends in a legal word-final consonant (/l,r,ɽ,y,m,n,ɲ,ŋ/)
b. The form that results possesses an even number of syllables.

As one might expect, Final Syllable Deletion led to numerous phonological alternations, many of which persist into the synchronic state which Dixon describes. For example (D 45), if one takes the bare stem [buɲa] 'woman' and adds the basic postvocalic form of the ergative suffix [-ŋgu], one obtains [buɲaːŋ], which may be presumed to have been historically *[buɲaːŋgu] (and somewhat earlier, *[buɲaŋgu]). Similarly, the quadrisyllabic form [ŋunaŋgara] 'whale-absolutive' shows up in the ergative as [ŋunaŋgaraː-ŋ] (D 84), historically *[ŋunaŋgara-ŋgu]. The ergative ending can be seen in its unaltered historical form after a trisyllabic stem, e.g. in [mulari-ŋgu] 'initiated man-erg.' D 57.

It can be seen that Final Syllable Deletion, taken as a sound change, must have occurred after Penultimate Lengthening, since in [buɲaːŋ] and countless similar words, what was historically the penultimate vowel shows up as long. Indeed, Dixon's synchronic analysis recapitulates the historical ordering, placing Penultimate Lengthening first.

As a result of Final Syllable Deletion, a number of Yidiɲ suffixes have two dramatically different allomorphs, as was just shown for ergative [-ŋgu]/[-ːŋ]; for further cases see (3) below. Moreover, Final Syllable Deletion also yields alternations in stems. For example, the stem meaning 'moon' shows up as trisyllabic when suffixed: [gindanu-ŋgu] 'moon-erg.' D 57, since quadrisyllables never underwent Final Syllable Deletion. But alone, 'moon' was trisyllabic, and thus underwent Penultimate Lengthening and Final Syllable Deletion to yield modern Yidiɲ [gindaːn].

Unlike Penultimate Lengthening, Final Syllable Deletion has not left a clean, across-the-board data pattern in contemporary Yidiɲ. Rather, there are many exceptions. These arose perhaps as analogical restorations, or perhaps the original process was lexically irregular in the first place.

Among the suffixes, we find that the majority of forms that could in principle alternate actually do ((3a)). Exceptions, however, are non-negligible ((3b)).

(3)a. **Suffixes which alternate by Final Syllable Deletion**

[-ŋgu] ~ [-ːŋ]	ergative
[-ɲa] ~ [-ːɲ]	accusative
[-yi] ~ [-ːy]	nominal comitative
[-ɲu] ~ [-ːɲ]	past ([-n-] conjugation)
[-l-ɲu] ~ [-ːl]	past ([-l-] conjugation)
[-ɽ-ɲu] ~ [-ːɽ]	past ([-ɽ-] conjugation)
[-ɲu-nda] ~ [-ɲuː-n]	dative subordinate (verbal)
[-ŋa] ~ [-ːŋ]	verbal comitative
[-ŋa] ~ [-ːŋ]	verbal causative

b. **Suffixes which do not alternate by Final Syllable Deletion**

[-nda]	dative (nominal)
[-na]	purposive
[-n-dʲi], [-l-dʲi], [-ɽ-dʲi]	'lest' inflection (in [-n-], [-l-], and [-ɽ-] conjugations)

Note that this list is incomplete, as there are three other suffixes that alternate by syllable count, but in ways that cannot be dealt with by Final Syllable Deletion alone: locative [-la] ~ [-ː], ablative [-mu] ~ [-m], and genitive [-ni] ~ [-ːn] ~ [-ni] ~ [-nə].

Looking at the inventory of roots, we find a similar bifurcation. Dixon (D 58) counts 80 roots that alternate by Final Syllable Deletion (e.g. [gindanu-ŋgu] ~ [gindaːn]), and 52 roots that do not undergo the process, even though they are phonologically eligible. An example of the latter is the stem for 'initiated man', which undergoes only Penultimate Lengthening where applicable: [mulaːri] (not *[mulaːr]) ~ [mulari-ŋgu].

Due to all of this lexical idiosyncrasy, the synchronic version of Final Syllable Deletion as stated by Dixon ends up incorporating various mechanisms to make reference to individual lexical items possible:

(4) **Final Syllable Deletion (from D 48)**

$XV_1C_1(C_2)V_2\# \rightarrow XV_1C_1$

if (a) $XV_1C_1(C_2)V_2\#$ is an odd-syllabled word;
and (b) C_1 is one of the set of allowable word-final consonants;

and (c) EITHER (i) there is a morpheme boundary between V_1 and C_1
OR (ii) V_2 is a "morphophoneme": A, I or U

In this formulation, a "morphophoneme" is understood to be the final vowel of one of the stems like /gindanu/ that is (more or less idiosyncratically) eligible for deletion; Dixon writes /gindanU/. Moreover, it is understood that case (4c.i), the suffix truncation case, permits of lexical exceptions, namely the suffixes listed in (3b).

Plainly, if Final Syllable Deletion was once a fully-productive, across-the-board process of Yidiŋ, its current status is rather attenuated, with considerable exceptionality and lexical idiosyncrasy.

2.2 A Traditional Analysis

Dixon provides a cogent traditional phonological analysis of the Yidiŋ alternations, using a fairly standard post-*SPE* approach. This account has served as the basis of almost all subsequent treatments of Yidiŋ. As with many phonological analyses, the synchronic description recapitulates history.

Dixon's basic assumptions are as follows.

- Underlying representations of stems include all the vowels that a stem displays anywhere in its paradigm. Thus the UR of surface [ginda:n] 'moon' is /gindanu/, since the /u/ shows up in suffixed forms like [gindanu-ŋgu].

- Vowels are assumed to be short underlyingly, except in the rare cases where they show up long across the board. Thus the /a/ of [ginda:n] is underlying short /a/, whereas the nonalternating /u:/ of [durgu:] 'mopoke owl' (D 84) is underlyingly long.

- The historical sound changes of Penultimate Lengthening and Final Syllable Deletion are continued in synchronic Yidiŋ as phonological rules, applied in the order just given, thus: /gindanu/ → ginda:nu → [ginda:n]. Clearly, rule ordering is required to make this solution work, since the odd-syllabled structural description of Penultimate Lengthening is met only at the deep level of representation, before Final Syllable Deletion has rendered the form even-syllabled.

- Since Final Syllable Deletion is quite irregular in its application, both stems and suffixes are lexically marked for whether they may undergo it.[3]

[3] Dixon uses "morphophonemes" for stems and exception features for suffixes, but the basic distinction is the same.

To my knowledge, all of the many subsequent analyses of Yidiɲ, my own included, have followed Dixon on these basic points. However, the traditional approach has two crucial defects which, to my knowledge, have not been previously noticed. Both indicate that it is not sufficient as a true characterization of the Yidiɲ speaker's knowledge of her language.

2.2.1 Defects of the Traditional Analysis I: Distribution of Invariant /V:/

The first problem arises from the distribution of vowel length in Yidiɲ, which is worth examining in detail.

Aside from Penultimate Lengthening, already discussed, there are three sources of long vowels:

I. Certain suffixes idiosyncratically cause the vowel of the preceding syllable to surface as long. For example, when the antipassive suffix [-ːdʲi], a prelengthener, is attached to the stem meaning 'see, look' (and the past tense ending [-ɲu] is added to the result), we get [waw**aː**-dʲi-ɲu], with length on the second syllable (D 218). Dixon plausibly traces the prelengthening property of these suffixes to historical sources in which the lengthening was compensatory, arising to fill the length slot of the deleting segments. However, as he shows, the synchronic situation appears fairly clearly to involve an arbitrary, morphologically-triggered lengthening.

II. Dixon suggests (D 77-83) that some instances of [iː] result from monophthongization of underlying /iy/.[4]

III. Finally, in just a few stems, vowel length is an invariant property of the stem; i.e. is phonemic under any analysis. Dixon does not speculate on where these long vowels come from historically. They are unlikely to be ancient, as they do not correspond with the length reconstructed from the evidence of neighboring languages (D 70); and indeed this reconstructed length has been lost in Yidiɲ.

Dixon collected sixteen morphemes with underlying length. Thirteen are disyllables, with the long vowel in final position, e.g. [durguː] 'mopoke owl', [giŋaː] 'vine species'; and the remaining three are are quadrisyllables, with length in second or fourth position: [galambaɽaː] 'march fly', [waɲinbaraː] 'what's the matter?' and [waɽaːbuga] 'white apple

[4] It is not a foregone conclusion that this suggestion should be accepted; in particular, it leads to a puzzling asymmetry in the phonemic long vowel inventory, whereby only two of the three Yidiɲ vowels (/uː/, /aː/) are permitted to occur in underlying forms. If we take [iː] at face value this asymmetry disappears. The cost is a rather peculiar arrangement in suffix allomorphy whereby some stems ending in /i/ (arguably the lowest-sonority vowel) must be lexically marked to take the suffix allomorphs that are otherwise used for consonant-final stems.

tree'. It can be determined that vowel length in these forms has nothing to do with Penultimate Lengthening, because it is invariant throughout the paradigm: thus [durguː] 'mopoke owl-absolutive', [durguː-n] 'genitive', [durguː-nda] 'dative'; [durguː-nu-la] 'locative of genitive'; etc. (D 84, 137).

By inspecting the forms in Dixon's list of invariant long vowels (D 85-86), one can easily determine the following:

- There are no invariant long vowels found in **odd syllables**. This gap is explainable on metrical grounds, [5] and it is not crucial to present concerns.

- Invariant long vowels may never occur in **closed syllables**. This gap will be quite crucial to the discussion, and should be borne in mind.

- Finally, invariant long vowels do not occur in **trisyllabic stems**.[6] A hypothetical case, to show what such a long vowel would look like if it existed, is /nulaːri/, which would show up as [nulaːri] in the unsuffixed absolutive case, and (for example) as [nulaːri-ŋgu] in the ergative.

Note that there are no *phonological* reasons why *[nulaːri-ŋgu] could not exist: its length pattern is completely legal. It arises, for instance, when a disyllabic stem with a final long vowel is followed by two suffixes, as in [durguː-nu-la] 'mopoke owl-genitive-locative', given above. The same length pattern can also be created when one takes a disyllabic stem and adds two suffixes of which the first is a pre-lengthener: a form given earlier, [wawaː-dʲi-ɲu] 'see-antipassive-past', is an example. Finally, the same length pattern also appears in the monomorphemic form [waɽaːbuga] 'white apple tree'.

Could these gaps in the Yidiɲ stem inventory be accidental? I have calculated the expected number of aberrant stems on the following basis. A rough check of Dixon's Yidiɲ glossary[7] yielded 437 disyllabic stems, of which 13 have a long vowel. There are 55 quadrisyllabic stems, of which 3 have a long vowel. Now, there are 206 trisyllabic stems. Following the percentages observed for the other lengths, we should find somewhere between 6 and 11 trisyllabic stems with a long vowel, but in fact there are none (a fact confirmed by Dixon,

[5] See the downloadable Appendix to this article, section 11.3.3; cited in fn. 17.

[6] Or more generally, in odd-syllabled stems. There are so few pentasyllabic stems that the absence of invariant long vowels in them could easily be an accident. No stem in Yidiɲ is longer than pentasyllabic.

[7] All stems counted except verbs and adverbs, which never have invariant length.

D 86). Checking with a chi square test, we find this there is only a 1.2% chance that this situation could arise by accident.

The absence of long vowels from closed syllables likewise appears not to be accidental. Consider that of 398 total stems in the glossary that end in an open syllable, 15 have a final long vowel. There are 301 stems that end in a consonant. Of these, we would expect about 11 to have a long vowel, but none do. A chi square test indicates that the probability that this could arise by chance is about 0.07%.

One can actually defend the view that Yidiŋ speakers were *tacitly aware* of these data patterns (though not necessarily using the scheme that the traditional analysis posits). As will be seen below, there is good evidence that Yidiŋ speakers often had to concoct inflected forms of stems that they were only familiar with in the bare absolutive form. When they did this, they generally treated surface [V:] as (in Dixon's terms) a derived [V:], not an underlying one; thus [CVC**V:**C] normally appears suffixed as [CVC**V**CV$_i$-CV], where V$_i$ is some vowel, and [CVC**V:**CV] as [CVC**V**CV-CV].

The asymmetry in the distribution of long vowels is an unexplained oddity for the traditional approach. In derivational theories, such gaps must be stated as constraints on underlying representations, as follows:

(5) **Yidiŋ Deep Phonotactics**

In underlying representations:

a. There are no trisyllabic stems with a long vowel.
b. There are no long vowels in closed syllables.

These underlying phonotactic constraints will guarantee that there are no alternations of the type *[CVCV:C] ~ [CVCV:C-CV] or *[CVCV:CV] ~ [CVCV:CV-CV]. I take them to be descriptively accurate, but explanatorily deficient: why should there be a ban on long vowels in specifically trisyllabic stems? Further, why should such a ban occur in (of all languages), Yidiŋ, where because of Penultimate Lengthening, on the surface there is a long vowel in *every* trisyllabic stem? Why, in addition, should long vowels be banned underlyingly in closed syllables, where on the surface it is in fact quite normal for a long vowel to occur in a closed syllable?

Intuitively, there is an answer to these questions. Yidiŋ, in the traditional account, is a language with many derived long vowels, but few underlying ones. In contexts where derived long vowels abound, apparently Yidiŋ speakers have made the assumption that every long

vowel is a derived one. This seems plausible, but it has no straightforward translation in the traditional framework.

For accounts using Optimality Theory, the gaps in the invariant long vowel inventory are more embarrassing still, as under the doctrine of the Richness of the Base (Prince and Smolensky 1993, Smolensky 1996), OT has aspired (rightly, in my opinion) to avoid constraints on underlying forms entirely.

2.2.2 Defects of the Traditional Analysis II: Patterns of Vowel Restoration

Putting this problem to the side for the moment, let us consider another difficulty for the traditional analysis. One claim that the traditional approach makes quite explicitly is this: for stems that alternate by Final Syllable Deletion, the allomorph of a stem that packs the greatest amount of phonological information is the suffixed allomorph. This is the allomorph that preserves the crucial final vowel (any of the three Yidiŋ vowels /i/, /a/, or /u/) that cannot be determined by inspecting the truncated isolation stem. The basic pattern of predictability is claimed to be "from the context stem, you can predict the isolation stem", but not vice versa.

This is the claim, but the actual facts of Yidiŋ are fascinatingly different. In the great bulk of the cases, the form of the full stem is predictable (up to free variation) from the form of the isolation stem. This view is adumbrated in a fascinating section (D 59-65) of Dixon's grammar.

During the Yidiŋ elicitations, once Dixon had figured out the basic pattern of Final Vowel Deletion, he took care to obtain a suffixed form for every stem he had previously encountered only in the plain absolutive form. In a number of cases, the very same isolation form yielded more than one suffixed form, either from different speakers or from the same speaker on different occasions. The pattern of free variation is complex; and for the moment I will discuss only the statistically predominant case, which I will call the "standard" outcomes.

There are about 79 stems to consider; namely, those which are truncated in their isolation form but show up with an extra vowel in their suffixed form. As Dixon (D 60) notes, there is a strong regularity governing what vowel must be added to the absolutive base to obtain the inflected stem: in 57 out of 79 cases, it is a *copy of the second stem vowel*; thus [CVC$\mathbf{V_i}$:C] ~ [CVC$\mathbf{V_i}$C$\mathbf{V_i}$-CV].[8] Here are six examples (D 65-68) culled from the 57 total cases; [-CV] represents any suffix. The vowel copy relation is emphasized here with boldfacing.

[8] My counts differ slightly from Dixon's due to a different procedure: I omit cases with free variation and cases where the vowel is unknown, but include forms from Dyalŋuy ("mother-in-law language").

(6) [bab**a:**l] ~ [bab**a**l**a**-CV] 'bone'
[band^{j}**a:**ɽ] ~ [band^{j}**a**ɽ**a**-CV] 'madness'
[d^{j}ig**u:**r] ~ [d^{j}ig**u**r**u**-CV] 'thunderstorm'
[gab**u:**l] ~ [gab**u**l**u**-CV] 'stick for carrying fish'
[gaw**i:**ɽ] ~ [gaw**i**ɽ**i**-CV] 'crescent shaped'
[mind**i:**r] ~ [mind**i**r**i**-CV] 'salt-water centipede'

This is an interesting observation, and becomes more so if we add a slight emendation to it: when the consonant that appears at the end of the isolation allomorph is a *nasal*, the vowel that is added is virtually always /u/. This provision brings an additional 14 cases into the realm of predictability. Here are five representative examples, taken from the 14:

(7) [baɽi:**ɲ**] ~ [baɽi**ɲ** **u**-CV] 'tree used for handles'
[d^{j}ala:**m**] ~ [d^{j}ala**m** **u**-CV] 'fresh, young'
[d^{j}uri:**n**] ~ [d^{j}uri**n** **u**-CV] 'leech'
[ginda:**n**] ~ [ginda**n** **u**-CV] 'moon'

As these cases show, the requirement for /u/ after nasals appears to outrank the requirement for the last two stem vowels to be the same; there are no cases at all like "[wug**a:m**] ~ [wug**ama**-CV]" where the vowel identity requirement overrides the post-nasal /u/ requirement. Naturally, there are a number of examples in which both vowel identity and postnasal position lead to the same result.

Putting Dixon's vowel-copy pattern together with the post-nasal principle, we get a striking outcome: the inflected stem is to a fair degree predictable from the truncated absolutive stem. I state this principle in preliminary form below:

(8) **Stem Vowel Restoration**

a. If the truncated absolutive stem is [CVC$_0$V:**N**], where N is a nasal consonant, then the full inflected stem must be augmented by /-u/.

b. Otherwise, the full inflected stem of [CVC$_0$**V$_i$:**C] must be augmented by /-**V$_i$**/.

Of the 79 relevant stems, 71 have their missing vowels filled in correctly by Stem Vowel Restoration (57 by vowel copying, and 14 by post-nasal /u/). Seven additional cases show free variation, all between the vowel predicted by Stem Restoration and some other vowel.

(9)	[gadʲ**u:**l]	~	[gadʲ**ula**-CV], [gadʲ**ulu**-CV]	‘dirty (e.g. water)’
	[gaw**u:**l]	~	[gaw**ula**-CV], [gaw**ulu**-CV]	‘blue gum tree’
	[gunb**u:**l]	~	[gunb**ula**-CV], [gunb**ulu**-CV]	‘Dyalnguy ?: billy-can’
	[mag**u:**l]	~	[mag**ula**-CV], [mag**ulu**-CV]	‘a root food’
	[gangu:**n**]	~	[gangu**na**-CV], [gangu**nu**-CV]	‘bushes arranged as trap’
	[ɲag**i:**l]	~	[ɲag**il****a**-CV], [ɲag**il****i**-CV]	‘warm’
	[wag**a:**ɽ]	~	[wag**a**ɽ**i**-CV], [wag**a**ɽ**a**-CV]	‘wide’

In three cases, Dixon could only obtain forms in which the consultants followed the alternate vowel restoration strategy of suffix vowel copying (discussed below in 7.1), so that the inherent stem vowel could not be determined. These cases are listed and discussed in (16) below. Finally, there were only eight cases that are outright exceptions; that is, show an invariant vowel contrary to that predicted by (8). These exceptions are the following:

(10)	[dʲamb**u:**l]	~	[dʲamb**ula**-CV]	‘two’
	[gang**u:**l]	~	[gang**ula**-CV]	‘grey wallaby’
	[gamb**i:**ɽ]	~	[gamb**i**ɽ**a**-CV]	‘tablelands’
	[gaɲdʲ**i:**l]	~	[gaɲdʲ**il****a**-CV]	‘crab’
	[gubu:**m**]	~	[gubu**ma**-CV]	‘black pine’
	[gul**a:**ɽ]	~	[gul**a**ɽ**i**-CV]	‘big-leaved fig tree’
	[guŋg**a:**ɽ]	~	[guŋg**a**ɽ**i**-CV]	‘north’
	[waŋ**a:**ɽ]	~	[waŋ**a**ɽ**i**-CV]	‘pre-pubescent boy’

Despite these exceptions, it can be argued that Stem Vowel Restoration is productive in Yidiɲ; that is, “psychologically real”. In particular, for some forms Dixon located evidence (D 65-68) from cognate stems of neighboring languages which indicate what the originally truncated final vowel must have been. In most of these, it appears that Yidiɲ has altered the original historical pattern of alternation so as to conform to the Stem Restoration principle.

(11)	**Form**			**Gloss**	**Cognate**
	[gaw**u:**l]	~	[gaw**ul**(**a**/**u**)-CV]	‘blue gum tree’	[gawul**a**] (Dya:bugay)
	[mag**u:**l]	~	[mag**ul**(**a**/**u**)-CV]	‘a root food’	[magul**a**] (Dya:bugay)
	[ɲag**i:**l]	~	[ɲag**il**(**a**/**i**)-CV]	‘warm’	[ɲagil**a**] (Gunggay), [ɲigal**a**] (Mamu Dyirbal)
	[yag**u:**ɲ]	~	[yag**uɲu**-CV]	‘echidna’	[yuguɲ**a**ŋ] (Gunggay)
	[muɽi:**n**]	~	[muɽi**nu**-CV]	‘ashes’	[murin**i**] (Dya:bugay)
	[bandʲ**a:**ɽ]	~	[bandʲ**a**ɽ**a**-CV]	‘madness’ in head	[bandʲ**a**ɽ] (Dyirbal)

[dʲuŋg**uːm**]	~	[dʲuŋg**umu**-CV]	'worm'	[dʲuŋgum] (Dya:bugay)
[gabaː**n**]	~	[gaba**nu**-CV]	'rain'	[gabaːn] (Dya:bugay)
[gindaː**n**]	~	[ginda**nu**-CV]	'moon'	[gindaːn] (Dya:bugay)
[gugiː**ɲ**]	~	[gugi**ɲu**-CV]	'flying fox'	[gugi] (Mamu Dyirbal), [gugiːn] (Dya:bugay)
[malaː**n**]	~	[mala**nu**-CV]	'right hand'	[malaːn] (Dya:bugay)

There are, however, three troublesome forms that work the wrong way:

(12)

Form		**Gloss**		**Cognate**
[gub**uː**l]	~	[gubul(**a**/**u**)-CV]	D(?): 'billy-can'	[gunbul**u**] (Ngadjan Dyirbal)
[wag**aː**ɽ]	~	[wagaɽ(**a**/**i**)-CV]	'wide'	[wagar**a**] (Dya:bugay)
[gamb**iː**ɽ]	~	[gambiɽ**a-CV**]	'tablelands'	[gamb**il**] (Dyirbal, Mamu Dyirbal)

Assuming that some explanation is possible for these cases,[9] the main force of the comparative data is that Stem Restoration has been an active force in reshaping the inventory of alternating stems in Yidiɲ . Note that several forms in (11) affirm the greater strength of the post-nasal subcase of Stem Vowel Restoration over the vowel copying subcase.

Consider further the cases of free variation seen in (9). Since these forms always involve a vowel of the quality predicted by Stem Restoration, it seems very likely that these cases illustrate change in progress, with new vowels derived by Stem Restoration displacing the older, etymologically correct vowels.

This, then, is a novel phenomenon to be accounted for. Previously, we have imagined that the truncated allomorph is largely predictable from the full allomorph; this is just what one would expect from a historical process that deleted all three stem-final vowels. But now we can see that *the full allomorph is also largely predictable from truncated allomorph*. It will be recalled from above that 52 of the 132 eligible stems simply don't apocopate; an example is [mulaːri] ~ [mulari-ŋgu] 'initiated man.' Because of these cases, it is not particularly more effective to try to predict the absolutive forms from the inflected forms than vice versa.

2.2.3 An Attempt to Save the Traditional Analysis: Deep Constraints

The traditional analysis, as it has always been stated, has nothing to say about the data patterns involved in Stem Vowel Restoration. Could it be beefed up to include these crucial and productive generalizations? Earlier, we rather awkwardly covered a similar problem (the

[9] For discussion of possible accounts, see the downloadable Appendix to this article, cited in fn. 17.

distribution of non-alternating long vowels) by adding constraints on underlying forms. Let us consider the same strategy here:

(13) **Deep Constraints as an Attempted Substitute For Stem Vowel Restoration**

a. In an underlying stem of shape /CVC_0VNV_i/, where N is a nasal, V_i must be /u/.
b. Otherwise, in an underlying stem of shape /$CVC_0V_jCV_i$/, V_i and V_j must be identical.

Interestingly, these proposed constraints *fail empirically*. To see this, recall that about 30% of the eligible trisyllabic stems idiosyncratically fail to alternate by Final Vowel Deletion, instead having trisyllabic allomorphs across the board. In the traditional approach, these stems have just the same kinds of underlying representation as the alternating stems, except that they possess an exception feature that blocks Final Vowel Deletion.[10] Thus, any constraint on underlying forms that holds true of the alternating stems should hold true just as well of the non-alternating ones.

In fact, this is not so. Among non-alternating stems listed in Dixon's glossary, only 8 out of 21, or 38%, of the stems obey the constraints proposed in (14); this may be compared with the 83-91 % obedience rate (depending on how the free variants are counted) in the alternating stems.

(14)

Obey (8): 8		**Disobey (8): 13**	
[baŋga**mu**]	'English potato'	[bib**iya**]	**'**coconut tree'
[bud^j**a**l**a**]	'fine, finely ground'	[dam**a**r**i**]	'silly'
[dal**iyi**]	'hunger'	[d^{j}iy**uya**]	'catbird'
[dig**a**r**a**]	'coast'	[galg**a**l**i**]	**'**curlew'
[d^{j}ud**ulu**]	'brown pigeon'	[gaɽa**na**]	'black cockatoo'
[gud^j**a**r**a**]	'broom-like implement'	[gum**a**ɽ**i**]	'red'
[gug**ulu**]	'recitative style'	[gurg**iya**]	'khaki bream'
[ŋaw**uyu**]	'salt-water turtle'	[mug**a**ɽ**u**]	'fish net'
		[mul**a**r**i**]	'initiated man'
		[mulɲ**a**r**i**]	'blanket'
		[ɲirij**ɲi**]	'long peppery fruit'
		[waɲ**ir**a]	'what kind of'
		[yuɽ**iya**]	'saltwater snake species'

[10] Alternatively, as in Dixon's account, the undergoers bear a special diacritic that triggers the rule.

There is another comparison possible which makes the same point. In Dixon's glossary, there are 33 stems that in principle would be eligible for Final Vowel Deletion, but cannot undergo it because the final vowel is preceded by an obstruent or by a consonant cluster (both of these conditions reliably block the process). Now, of these, only 14, or 42%, obey the deep constraints (for example, [gud**a**g**a**] 'dog-absolutive'), whereas 19 are like [bind**u**b**a**] 'crayfish-absolutive' and disobey them.

What seems to be the correct generalization is this: the final vowel of a trisyllabic stem is strongly constrained to obey the generalizations of Stem Vowel Restoration only when this vowel alternates with zero. It would seem that Stem Vowel Restoration is a principle governing *the relationships between forms*, not *the shape of underlying forms*. Thus, even with the additional device of constraints on underlying forms, the traditional analysis has no grip on the crucial data. The facts suggest rather forcefully, I think, that this analysis should be reconsidered. Before doing so, we must consider further data that also point in this direction.

2.3 More on Stem Vowel Restoration: The Alternative Strategies

The data pattern of Stem Vowel Restoration is actually more intricate than the previous section suggests. Dixon's careful description (D 59-65) actually records three *additional* patterns, which apparently occur largely in free variation with the statistically dominant pattern just noted.

2.3.1 The Suffix Vowel Copy Pattern

On some occasions, Dixon's consultants volunteered forms in which the vowel added to the isolation stem was a copy of the vowel of the following suffix. Thus, one consultant gave as the inflected forms of absolutive [gambiːn] 'top-knot pigeon' the following (D 61):

(15)	[gambin**u**-ŋg**u**]	ergative
	[gambin**a**-l**a**]	locative
	[gambin**i**-y**i**]	comitative

In a later elicitation, this consultant provided only forms of the type [gambin**u**-CV], which follow the primary pattern, specifying /u/ post-nasally.[11]

There were three stems Dixon collected which always alternated by this pattern:

(16)	[mugiːɽ]	~	[mugiɽ**V**-C**V**]	'small mussels'
	[wubuːl]	~	[wubul**V**-C**V**]	'lucky (at hunting, etc.)'
	[wurguːl]	~	[wurgul**V**-C**V**]	'pelican'

[11] It is unknown whether this vowel is the etymologically correct one.

There is no evidence in Dixon's grammar to support the view that the suffix vowel copy pattern is **obligatory** for these stems, and indeed for the case of [mugiɽ**V**] Dixon explicitly states that other outcomes also were found (D 62).

2.3.2 The Schwa Pattern

There also occurred cases (D 62) in which the "restored" vowel was a schwa. Thus a consultant offered as the genitive of [gubuːm] 'black pine' the form [gubum**ə**ni], instead of the primary (and probably, etymologically correct) form [gubum**a**ni]. The appearance of schwa is extraordinary, since schwa does not otherwise occur as a vowel of Yidiŋ.

2.3.3 Nonalternation

The last of the three supplementary patterns is the rarest of all: it is simply the attachment of the suffix to the truncated isolation stem itself, as in [ginda:n] ~ [ginda:n-ɖʲi] 'moon-absolutive, comitative' (D 64). If the suffix involved has a special allomorph for consonant-final stems, such as comitative [-ɖʲi], that is the allomorph is used (the postvocalic allomorph of the comitative can be seen in the normal variant [gindanu-yi]).

The nonalternation pattern is only doubtfully well-formed: when Dixon played tapes to his speakers of themselves saying such forms, they sometimes felt that these were errors. Thus whatever analysis is adopted for them should not treat them as fully integrated into the language.

Now, consider again the traditional analysis of the vowel ~ zero alternations. The outlook for this analysis clearly becomes worse when these cases of free variation are considered. Systematic free variation is typically felt to diagnose optional or competing rule systems, but the traditional account has no alternative but to place the variation in the underlying forms themselves, with massive loss of generalization.

2.4 The Genitive Ending

The Yidiŋ genitive marker [-ːn] ~ [-ni] ~ [-nu] ~ [-nə] provides one further argument against the traditional analysis. Genitive in Yidiŋ is treated by Dixon as a derivational category. The genitive stem created by affixation of [-ːn] ~ [-ni] ~ [-nu] ~ [-nə] is inflected with the normal nominal cases (absolutive, ergative, dative, etc.). Let us consider the distribution of the allomorphs of the genitive.

(a) [-ːn] shows up in absolutives (i.e., "bare genitives") when two conditions are met: that its use should result in a word that is even-syllabled, in accord with the general Yidiŋ preference, and that the stem be vowel final, so that the Yidiŋ requirements for segment sequencing may be respected.

(c) The third allomorph, [-nu] shows up when further suffixes are added to plain genitives that end in [-ːn]. Thus paradigms like those of (17), which show [-nː] in the plain genitive, may be amplified as follows:

(19) [buɲa] 'woman-abs.' [buɲa-**ːn**] 'genitive' [buɲa-**nu**-nda] 'gen.+dative' D 53
[dʲadʲa] 'child-abs.' [dʲadʲa-**ːn**] 'genitive' [dʲadʲa-**nu**-ŋgu] 'gen.+ergative' D 136

(d) The fourth and final allomorph of the genitive is [-nə]. It occurs in free variation with /-nu/.

What are we to make of this pattern? First, the distribution of [-ni] vs. [-ːn] plainly follows the normal pattern in Yidiɲ: allomorphy is guided by an even-syllable target. The remaining, more puzzling, cases of (c) and (d) above follow under the assumption that the [u] and [ə] of [-nu] and [-nə] are *restored vowels*, just like the restored vowels of monomorphemic stems. Thus, [buɲaː-n] restores /u/ postnasally in [buɲa-n**u**-ŋgu], but may also restore schwa in [buɲa-n**ə**-ŋgu] (form implied in D 54).[12]

[12] Dixon does not report cases in which the genitive suffix vowel is restored by Suffix Vowel Copy (section 2.3.1), or is not restored (non-alternation; section 2.3.3). I conjecture that this is an accidental data gap: suffix-copy forms, and especially non-vowel-restored forms, are rare, and it

The genitive data are of interest because they illustrate the relative strength of two generalizations that compete. Had Yidiŋ speakers opted for an *invariant underlying form* for the genitive, then the abundant instances of [-ni] (that is, with any odd-syllabled or consonant-final stem) would imply that this allomorph should also appear when Final Syllable Deletion is blocked by the presence of an additional suffix. In actual fact, the inserted [u] and [ə] vowels show that the extra vowel is being projected on the basis of the phonological shape of the *bare genitive stem form*. Thus, Stem Vowel Restoration takes place even when there is substantial evidence available for what the underlying stem vowel is supposed to be.

I take this to be one further reason to hold the traditional account in grave doubt: it must treat the genitive suffix with allomorphy, replicating the pattern that occurs independently with simple noun stems.

3. Theoretical Consequences

All the above discussion has as its general conclusion that Yidiŋ is not fully or properly analyzed under the traditional account. The traditional account, because it merely recapitulates history, does not capture the new patterns that arose as further generations of Yidiŋ speakers reanalyzed the system.

Is this conclusion of importance just to Yidiŋists, or does it have more general theoretical consequences? I would suggest the latter, on several grounds.

3.1 The "Inside-Out" Preference in Phonology

In Hayes (1995b) I suggested, following much earlier work, that phonological systems tend to organize themselves in ways that permit derived forms (such as the suffixed case forms of Yidiŋ) to be predicted from the base forms (usually, as in Yidiŋ, isolation forms). In that article, I presented a couple of examples suggesting that languages often rearrange their phonologies so that this will be true. Further support for this position is presented by Kenstowicz (1996).

Yidiŋ provides the most intricate case of this sort I have yet seen. It would appear that a serious reorganization of the phonology has taken place, permitting the suffixed allomorphs of stems to be predicted from their isolation allomorphs. In other words, the "direction of predictability" in the system has to a fair degree reversed. In the older system, the isolation form of a stem like [ginda:n] could be predicted from its suffixed form [gindanu-CV], by means of Final Syllable Deletion. But following restructuring, it became largely possible to predict the suffixed form on the basis of the isolation form, using general and productive principles. Thus

---

could easily have happened that within the relative rare morphological category of suffixed genitives, Dixon did not find any cases.

Yidiŋ phonology went from having an "outside-in" character to the (I claim, less marked)

It is easy to imagine when this change must have happened: "outside-in" phonology would have been quite stable while Final Syllable Deletion remained as an *optional* pattern of alternation. At one point there would have occurred archaizing free variants like [ginda:nu], alongside innovating [ginda:n]. Archaic [ginda:nu] would have provided ample evidence to learners for the underlying form of the stem. It is when Final Syllable Deletion became entirely obligatory that the restructuring process would likely have begun.

3.2 Rule Inversion and the Basis of Phonological Constraints

It has been suggested, e.g. by Tesar and Smolensky (1996), that phonology is learned simply by using the input data to rank a set of constraints given *a priori*; that is, as part of Universal Grammar. Under this approach, one would expect all phonologies to be highly principled in character, given that all of their ingredients are universal and only their ranking is language-specific.

The facts considered here do not necessarily refute this view. But they do encourage one to think of the acquisition problem in broader terms. Consider the pattern of nominal inflection in Yidiŋ, at the stage in the language's history just after Final Syllable Deletion became obligatory. I will use the term "Pre-Yidiŋ" to refer to this crucial stage. As Dixon suggests (D 59, 64-65), the Pre-Yidiŋ data pattern must have been quite hard to learn, because the unsuffixed absolutive stems occur in text more often than all other inflected forms put together. The restructuring that took place, converting Pre-Yidiŋ to Modern Yidiŋ, presumably was a direct response to this difficulty.

Let us hypothesize that a major goal of the language-acquiring child is to produce *correct and accurate reproductions* of the inflected forms spoken around her. Given that phonology is often cleanly-patterned and based on markedness constraints of wide applicability, the Tesar/Smolensky strategy of ranking only principled constraints will often suffice, or come close to it, in attaining accurate reproduction. But as we have seen, sometimes historical change deals the child a difficult hand. My contention is that in such cases the child does what is necessary, which includes the creation of relatively ad hoc, language-specific constraints. Yidiŋ seems to be an example of this, especially in its requirement that [u] be the restored vowel after nasals.

The specific scheme I suggest is as follows: in cases of great difficulty, such as was found in Pre-Yidiŋ, the language learner will seize upon generalizations that are *statistically useful, albeit imperfect*, and make use of them to improve her ability to guess unknown inflected forms. There are several cases where this appears to have happened here.

Consider first the observation made earlier that (under the traditional analysis) underlying long vowels are excluded from trisyllabic stems. Here, the crucial statistical pattern was that (as a matter of historical accident) the great majority of long vowels in Yidiŋ nouns were due historically to Penultimate Lengthening. Thus, if a language learner heard [CVC**V:**CV] in an isolation stem, it was a truly excellent bet that the second stem vowel would appear as short under suffixation: [CVC**V**CV-CV]. I believe that this is the origin of (what has been called) the ban on underlying long vowels in trisyllables. Whatever invariant long vowels may have existed in trisyllabic stems in Pre-Yidiŋ (these would have had paradigms like [CVC**V:**CV] ~ [CVC**V:**CV-CV]) were drowned in the statistically predominant pattern of alternating long vowels.[13] The need for language learners to be able to project the vowel lengths of suffixed forms from isolation forms thus rendered it impossible to sustain a vowel length contrast in trisyllabic stems.

The second case to consider is the principles of Stem Vowel Restoration, covered in sections 2.2.2 and 2.3 above. I believe it plausible that these principles also originated as exaggerations of statistical patterns already present in the Yidiŋ lexicon. In attempting to provide the missing vowel, Yidiŋ learners relied on a slight pre-existing tendency toward vowel harmony in the final syllables of stems, as well as a slight pre-existing preference for /u/ after nasals (see (14) and immediately following discussion). These slight tendencies, which may well have been true *by sheerest accident* in Pre-Yidiŋ, were exaggerated and employed as the best available means of predicting the quality of the inserted vowel in the reanalyzed modern language.

At the time Dixon collected his data, the principles of Stem Vowel Restoration had not yet stabilized; they still competed with rival strategies for stem vowel restoration. These competing strategies, laid out in section 2.3, were presumably created out of whole cloth by Yidiŋ learners, and indeed they seem to have a more principled, markedness-driven character. The tendency towards vocalic harmony in epenthetic vowels is widespread in languages, and vocalic harmony also exists as a typical constraint in child phonology. Schwa is likewise a typical quality for epenthetic vowels.[14]

---

[13] There is every reason to believe that such long vowels once existed, since the historical processes that create long vowels (loss of C from VCV, compensatory lengthening) are not ordinarily sensitive to syllable count.

[14] Before concluding that these strategies necessarily constitute "UG in action," however, one would want to rule out an alternative suggested by Dixon (D 62): that the restored vowels are meant to be inconspicuous in their context, thus helping to avoid embarrassment at being unable to remember the "correct" vowel. Schwa is a good candidate here, since it is roughly equidistant perceptually from the three full-vowel possibilities that could otherwise occur. Copying of a neighbor vowel is also a good choice: given the pervasive existence of vowel-to-vowel coarticulation (Öhman 1966), any vowel in the crucial position will be shifted somewhat in the

Consider finally the vowel length alternation in earlier stems of the type [CVCV:C] ~ [CVCVCV-CV]. Here the projection of suffixed form from isolation form worked with a combination of the previous principles. The exceptionless pattern of inserting a vowel after /...CV:C/ arose from the fact that all, or virtually all isolation stems ending in /...CV:C/ were derived historically by Penultimate Lengthening and Final Syllable Deletion. Any historical stems that had once alternated as [CVCV:C] ~ [CVCV:C-CV] would have been drowned in the vastly larger statistical pattern, and would have undergone readjustment to [CVCV:C] ~ [CVCVC**V**-CV], using the same principles for choice of epenthetic vowel just discussed.

The upshot of all this is, perhaps, an extension of one's conceptions of "where phonology comes from". Rather than constituting just a language-particular prioritization of general, a priori principles, *some* phonology seems to represent the relatively ad hoc response of learners to conundrums presented to them by historical change. Faced with such a conundrum, learners are capable of fabricating entirely new phonology, which has no direct connection with a language's earlier, more "motivated" form. This conclusion follows earlier research on "rule inversion", cited above in the Introduction.

It is true that *most* phonology does seem to have the principled character that arises from well-motivated constraints. The reason for this statistical predominance is that the rather more ad hoc principles of the type discussed here arise only in the context of restructuring, where peculiar historical evolution forces the language learner to develop odd constraints to deal with the resulting data conundrums.

Moreover, we can expect that even the constraints that arise from restructuring will not be utterly unprincipled; they too must fall at least loosely within the set of constraints possible within phonological theory in general.

3.3 Alternation Not Driven by Markedness-Faithfulness Interactions

In the Introduction to this article I mentioned that it is a great and novel virtue of Optimality Theory that it ties the pattern of alternation in a language to basic principles of markedness. At the same time, it does seem that there are cases of productive alternation that cannot be reduced to the simple interaction of markedness and faithfulness. Yidiŋ provides a fairly clear example.

Consider the vowel ~ zero alternations we have been examining. They may usefully compared with another type of alternation, which is extremely common: the appearance of derived vowel length whenever a [-CV] suffix is attached to a [CVCVC] stem (cf. (1),

---

direction of its neighbors. Thus a vowel copying a neighbor is a good bet for resembling the unknown "correct" vowel.

Penultimate Lengthening). An example given earlier is [mud^{j}am] 'mother-absolutive' ~ [mud^{j}a:m-gu] 'mother-purposive'. Curiously, in this pattern of alternation, Yidiɲ *creates* a word shape ([CVCV:CCV]) which it elsewhere tries to *avoid*. Specifically, where [CVCV:CCV] would arise simply from the concatenation of [CVCV:C] and [-CV], Yidiɲ productively inserts vowels to avoid this outcome, in the cases of Stem Vowel Restoration, discussed at length above.

In brief, we have two patterns of alternation:

(20) a. [CVC**V:**C] ~ [CVC**V**C**V**-CV] (Stem Vowel Restoration cases)
b. [CVC**V**C] ~ [CVC**V:**C-CV] (ordinary suffixation to /CVCVC/)

The pattern in (20a) cannot be driven by orthodox markedness-faithfulness interactions, as the following reasoning shows. Assume for purposes of argument that there exist highly-ranked markedness constraints that force [CVCV:C] + [CV] in (20a) to surface as [CVCVCV-CV]. The penalty in faithfulness for doing this is (as stated in the system of McCarthy and Prince 1995) a violation of IDENT([+LONG]), due to the loss of vowel length, and a violation of DEP(V), due to the insertion of a vowel. Yet, in (20b), where our starting point is [CVCVC] plus [-CV], the very same markedness constraints are satisfied by the candidate *[CVCVCV-CV], which is not the winner; and the penalty in faithfulness is *less*, namely just a violation of DEP(V). Thus it is not at all clear how the real winner [CVCV:C-CV] could ever be made to defeat its ill-formed rival, at least without serious amplifications of the theory.

The reader might object at this point that the [CVC**V**CV-CV] output in forms of the type (20a) is not the result of GEN altering the underlying representation, but simply of the underlying representation surfacing unaltered. But the whole point of the above discussion was to show that this point of view is wrong: the traditional analysis fails to capture the facts in precisely this area. It is more plausible to suppose that alternations like [CVCV:C] ~ [CVCVC**V-CV**] really are due to processes of epenthesis and vowel shortening, somehow expressed appropriately within Optimality Theory. It is these processes that cannot be treated as straightforward markedness-faithfulness interaction.

The Yidiɲ case differs from other examples of this type in a crucial way. In many languages, there is problematic phonology that involves lexical exceptions or special vocabulary strata. It is easy to imagine a theory that sets up a treatment for exceptions (e.g. Pater (ms.)) or strata (Ito and Mester 1995), so that such alternations do indeed follow from markedness-faithfulness interaction. The Yidiɲ case is different because the problem is entirely general (involving essentially all forms of the relevant shape), and not limited to any kind of stratum or exception class.

Thus, classical OT needs amplification; a means of driving alternations that cannot be the result of markedness-faithfulness interaction. This is not to deny that this interaction is the central and normal cause of alternation in phonology: cases like Yidiŋ are the exception, with an origin in historical restructuring.

3.4 Underlying Representations

Consider the task of taking the "Wug" test in Yidiŋ. In this test, named after the seminal work of Berko (1958), an experimental subject is given one inflected form of a novel stem, and is asked to say what the other inflected forms would be; thus, in the classic case for English, one is asked "What is the plural of [wʌg]?" and replies "[wʌgz]."

For a hypothetical Yidiŋ form like "[bagaːn]", we already know that Yidiŋ speakers would pass the Wug test: as shown above, in the past various forms along the lines of [bagan**u**-nda] have already been invented, "correctly", by Yidiŋ speakers. The speakers succeeded in passing the Wug test by inventing the principles of Stem Vowel Restoration, which redefined what counts as the correct answer.

What, then, of the *other* direction for Wug testing, namely suffixed form to plain stem? The sort of question asked here is: "If there were a particular thing that in the dative was called a [baganu-nda], what might be its absolutive form be?"

Here, alas, we lack data on Yidiŋ speakers' Wug-testing performance, but it seems fairly likely (based on my own informal experience in Wug-testing speakers of other languages) that they would be able to name the two most likely candidates, namely [bagaːn] and [bagaːnu]. Let us suppose tentatively that this is the case, and see what would follow.

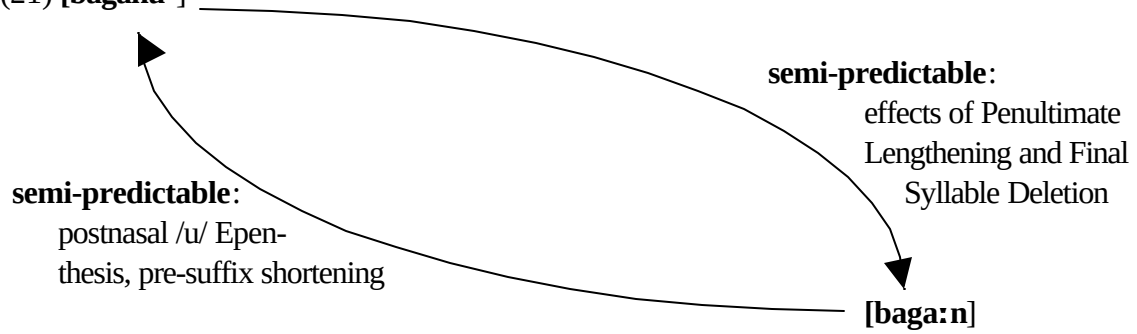
Here is the crucial empirical fact. With roughly 90% certainty, it is predictable in Yidiŋ that the dative form of a hypothetical absolutive [bagaːn] will be [baganu-nda], at least as a free variant. This is because the principles of Stem Vowel Restoration work correctly about 90% of the time. But simultaneously, by Dixon's figures (D 58) it is about 70% predictable that the absolutive form of a hypothetical dative [baganu-nda] will be [bagaːn] (and not [bagaːnu]).

With this in mind, we can consider the question: How do speakers pass the Wug test? Under traditional accounts of phonology (both Optimality-theoretic and earlier generative ones), the crucial mediating element is the *underlying representation*. The taker of the Wug test uses her phonological and morphological knowledge in the "backwards" direction to deduce what could be the underlying form or forms of what she is hearing, then applies the morphology and phonology in the forward direction to figure out the predicted surface form (or forms) that answer the original question.

In the Wug test for Yidiŋ considered here, the two reasonable choices for an underlying representation of a surface absolutive [bagaːn] are /bagaːn/, if the alternation is due to epenthesis and vowel shortening process, or /baganu/, if the alternation is actually due to the lengthening and Final Syllable Deletion posited in the traditional account. But *neither* choice does real justice to the pattern of predictability. The underlying representation /bagaːn/ claims that the isolation stem allomorph [bagaːn] is arbitrary from the viewpoint of the inflected form [baganunda] (it is not; it is predictable on about a 70% basis); while /baganu/ claims that the inflected form [baganunda] is arbitrary from the viewpoint of the plain stem [bagaːn] (it is not; it is predictable on about a 90% basis).

This situation, in principle, has major theoretical consequences. As textbooks like Kenstowicz and Kisseberth (1979) ably lay out, the central role of phonological underlying representations is to serve as a repository of unpredictable information concerning the pronunciation of a given morpheme. The patterns of predictability among the members of the paradigm of that morpheme are supposed to follow from the phonological rules (or other principles, such as GEN and constraints in OT), which act to derive all surface allomorphs of a stem from the unique underlying form. What is special about Yidiŋ is that there is *more predictability* present in the system than can be accounted for under the usual assumption of derivation from a single underlying form. The following diagram illustrates this point.

(21) **[baganu-]**



This situation is not unique to Yidiŋ. In Hayes (1995b), I laid out the evidence for a comparable situation in English. In English, it is largely predictable that forms ending in [...ənt] will alternate with [...ɛnt] upon suffixation, as in numerous pairs like ˈ*elem*[ə]*nt* ~ *ele*ˈ*m*[ɛ]*ntal*, ˈ*accid*[ə]*nt* ~ *acci*ˈ*d*[ɛ]*ntal*, etc. (Here is the Wug test: if X = [ˈpɛɹədənt], what is [Xəl]? Answer: [ˌpɛɹəˈdɛntəl]). But it is *also* predictable that stems which end in [...ɛnt] when affixed will show [...ənt] in their isolation forms (Wug test: if X+əl = [ˌpɛɹəˈdɛntəl], what is X? Answer: [ˈpɛɹədənt]). Derivation from a single underlying form does not suffice to account for both relations.[15]

[15] Some further details, to nail down the argument:

(a) Suppose we use a rule (or equivalent) of Vowel Reduction to guarantee the [...ɛnt] → [...ənt] mapping. Then, to get the [...ənt] → [...ˈɛnt] mapping, we need to restrict underlying forms so that final /nt/ is always preceded by /ɛ/. But this can't be right, because English has many words like *st*[ʌ]*nt*, [æ]*nt*, *st*[ɪ]*nt*, ˈ*galliˌv*[æ]*nt*, etc. It is only where the principles of English stress happen to make the final syllable stressless (i.e., in polysyllabic nouns) that the underlying vowel must be limited to /ɛ/. So the underlying structure condition would have to duplicate the stress rules in its structural description, a highly undesirable move.

(b) Alternatively, we could suppose that the underlying form of the relevant words ends in /...ənt/, and posit a /ə/ → ɛ
here again, we must keep vowels other than schwa from occurring in this underlying position, else they would show up on the surface when assigned stress. This forces the need for an underlying constraint, just as before, only this time requiring final /nt/ to be preceded by /ə/. Obviously, this constraint, too, is falsified by *stunt*, *ant*, *stint*, etc., unless it duplicates the stress rules.

(c) A final option is to make Vowel Reduction cyclic, forcing [...ənt] on the first cycle, with Full Vowel Restoration applying on the second cycle. This fails, because Vowel Reduction cannot in general be cyclic (compare ˈ*at*[ə]*m* ~ *a*ˈ*t*[ɑ]*mic*, ˈ*syst*[ə]*m* ~ *sys*ˈ*t*[ɛ]*mic*.

The conclusion is that, just as in Yidiŋ, the crucial generalization relates surface forms to each other, not surface forms to underlying forms.

3.5 Summary

Optimality Theory, at least as it is usually practiced today, is in one sense a conventional theory of derivational phonology, in which which surface forms are related to one another by deriving them from a common underlying form. As such, it suffers from the problems that beset all such theories. These are: that no means is provided to account for the characteristic "inside-out" bias of phonology, and that no means is provided to account for hypercharacterized phonological systems like Yidiŋ in which there is more interparadigm predictability than can be handled by deriving all allomorphs of a morpheme from a single underlying form.

I have also suggested two lessons from Yidiŋ that pertain specically to OT. First, it is unrealistic to expect all constraints to be utterly principled manifestations of UG: some constraints result from the efforts of language learners to pass the Wug test (that is, to project novel members of paradigms) in cases where historical change has served them up a difficult data pattern. It seems that in such cases learners will take advantage of regularities present in the data, even relatively arbitrary ones. Second, it is unrealistic to view all alternation in phonology as driven by the relative ranking of faithfulness and markedness constraints (though obviously, much alternation is). Here again, the need for language learners to deal with the Wug test for difficult data patterns dealt them by the accidents of history leads the learners to create fully-productive and widespread patterns of alternation that go beyond what can be treated with the resources of contemporary OT.

4. Where to Go From Here?

A clear picture of how phonological theory should respond to cases like Yidiŋ is not yet in view. But a few speculative remarks might be helpful.

First, in line with the title of this book, I would assert that while the Yidiŋ data involve a kind of "residue" (set of unaccountable phenomena) for standard Optimality Theory, it is hardly a *derivational* residue. Nothing in traditional derivational phonology seems to promise any better account of the data here, and indeed there are phenomena in Yidiŋ for which non-derivational OT seems very well suited: Yidiŋ phonology is quite "conspiratorial," in the sense of Kisseberth (1970), and conspiracies are just what OT was designed to handle. Some notable conspiracies in Yidiŋ are the "look-ahead" property of Final Syllable Deletion (2), which only applies when a legitimate final consonant will remain after truncation, and the pervasive phonotactic target of even-syllabledness.

Second, as a way of shedding light on the problems raised here, it might be worthwhile briefly to situate Optimality Theory in its historical context. The generative phonology of the 1960's and early 1970's was very much focused on "deep" phenomena: patterns of alternation that (it was claimed) could be treated only with quite complex and intricate mappings from underlying to surface representations. This led to a strong emphasis on language-particular

derivations as the central element of the theory; and this in turn led to two long-term crises that were ultimately resolved only with the introduction of OT. These crises were (i) the conspiracy problem and (ii) an inability to relate the content of language-particular analyses to general principles of markedness.

The approach that OT used to solve these two problems was to turn the research strategy on its head: instead of positing rules to account for alternation, with the hope of somehow finessing the problems of conspiracies and markedness, classical OT let the crucial principles that govern conspiracies and markedness serve as the heart of the theory, and hoped instead to finesse the problem of complex patterns of alternation.

Specifically, the device that OT provides to permit alternation is the ranking of particular faithfulness constraints sufficiently low in the grammar. Faithfulness constraints are characteristically atomistic, banning simple feature switches, insertions, and deletions. Given that a earlier research program felt it necessary to devote massively complex resources to the description of alternation, we should not be surprised if the much more impoverished capacities of contemporary OT are not up to the task of handling the more complex alternation patterns.[16]

Is this cause for despair? My own judgment is that it is certainly not. Arguably, we are very much better off under OT than we were before. We now have a theory that promises an adequate account of the very general and pervasive phenomena of conspiracies and markedness; earlier, we had a theory that had very little to say about these matters. The payoff for returning to earlier approaches would be quite small, and would primarily concern only the residual class of phonological phenomena that arise from restructuring.

4.1 Anticorrespondence

As for what is needed in OT to treat facts like those given in this article, I tentatively suggest that we should address the question of patterns of alternation rather more directly than they are addressed with simple faithfulness constraints. I conjecture that what is needed is contraints of "Anticorrespondence," which would *actively require* morphemes to alternate in particular ways. An Anticorrespondence constraint would say: "if morpheme μ appears with shape X in a particular context C, it must appear with shape X' in a distinct context C'". One might call the allomorph of context C the *base allomorph* and the allomorphy of context C' the *projected allomorph*.

Here is a brief Yidiŋ example. As noted above, all isolation [CVCV:C] stems in Yidiŋ are amplified by a vowel when a suffix is added to them, as in [bab**a:**l] ~ [bab**ala**-CV] 'bone' from

[16] A large further class of cases that should be considered in this context are those previously treated with opaque rule ordering. These are now the subject of a growing literature, which attempts to reconcile opacity with OT.

(6). As was observed in section 3.3, it does not appear to be possible to treat this as a simple assertion of markedness over faithfulness, since parallel cases where the isolation form is [CVCVC] do not undergo a similar shift. An Anticorrespondence constraint to treat this is stated in (22), where "X" designates some non-null string:

(22) VOWEL RESTORATION

$$\varnothing \; / \; \text{V:C} \; \_\_\_ \;] \; // \; \_\_\_ \;]_{\text{Word}} \rightarrow \text{V} \; // \; \_\_\_ \; \text{X}$$

The notation here involves both single and double environment slashes.

- The single slash refers to the applicable context *within the base allomorph*. Thus, in (22) the single slash indicates that the constraint is focusing on the right edge (expressed as the null string) of a base allomorph that ends in /...CV:C/.

- The first double slash refers to the *context in which the base form appears*, which in this case is isolation (i.e., no suffix is present). As noted earlier, the use of isolation allomorphs as base allomorphs is quite characteristic.

- The second double slash specifies the context of the projected allomorph. In this case, the variable X indicates that the projected allomorph occupies nonfinal position.

- Lastly, the material at the right side of the arrow indicates what must occur in the projected allomorph as a replacement for what occurs as the initial symbol of the constraint. Here, zero must correspond to some vowel.

Thus, we can read the VOWEL RESTORATION constraint (22) as follows: "zero at the right edge of an isolation base allomorph ending in [...V:C] must correspond with a vowel for allomorphs occurring with a suffix." The actual quality of the inserted vowel would be determined by additional constraints, ranked freely in order to account for free variation (see section 2.3).

A constraint like VOWEL RESTORATION has a rather crudely descriptive character. It should be remembered, however, that Anticorrespondence constraints arise, according to the present view, only in acquisitional conundrums, where historical change has created a tough data pattern. In such cases, the ordinary analytical approach of ranking faithfulness constraints against markedness constraints has already been tried by the learner and has failed. In such an instance, the simple task of collecting stem allomorphs, and the contexts where they occur, at least has the advantage of straightforwardness and directness.

Anticorrespondence constraints, quite naturally, resemble rules. A crucial difference is that they refer to surface members of a paradigm, rather than to underlying forms. In the present case, this would seem to be a clear virtue. As argued in section 3.4, Yidiŋ shows more

relations of predictability among surface forms than can be accounted for by the method of deriving all surface forms from a single underlying representation. In contrast, Anticorrespondence constraints may be stated for *all* the relations of mutual predictability among surface allomorphs, and thus are capable in principle of characterizing such relations completely and coherently.

At this point, I am out of space, and must refer the interested reader to a longish downloadable Appendix[17] for a tentative analysis of Yidiŋ that makes use of Anticorrespondence constraints.

4.2 Conclusions

In terms of "good news" and "bad news," the inspection of Yidiŋ data carried out here has yielded rather mixed results. While Yidiŋ emerges as an even more interesting language than has previously been thought (or so I believe), *phonology* comes out looking rather less principled and harder to do research in. If I am right in supposing that something like Anticorrespondence constraints are needed for Yidiŋ phonology, then it would seem that phonology simply cannot be done on the fully-principled basis laid out in classical Optimality Theory, and must indeed on occasion make use of a certain amount of brute-force descriptive power. Is phonology still worth doing in this case?

My own opinion is that it most certainly is, but we will have to be shrewder in the research strategies we adopt. One possibility is to locate particular empirical domains in which "pure" phonology of the type identified by classical OT can be isolated and studied; for discussion of such cases, see Hayes (in press). Another possibility is to let considerations of *learnability* serve as a guide to research. Suppose we developed the ability to write computational algorithms that could learn entire phonological systems. Any proposed model of this type faces a serious danger, namely that the number of hypotheses it must consider and test against data might be impossibly large. If it were possible to develop an algorithm that could learn, say, the Yidiŋ data pattern using only limited time and computational resources, then whatever the model assumed about the nature of phonological constraints would receive a kind of semi-empirical support. It seems likely that only a very thoughtful, highly principled theory of constraints could constrain the hypothesis set to the point that learning could occur in a feasible amount of time.

The crucial point here is that the "brute force" objection that can be made to constraints like Anticorrespondence is fundamentally an objection about learnability. It is only direct research on learning that will determine to what extent such an objection is legitimate.

---

[17] Available at the author's Web site: http://www.humnet.ucla.edu/humnet/linguistics/people/hayes/hayes.htm.

The philosophical conclusion here, which is hardly a novel one, is that languages are often more interesting and challenging than our current theories allow for. The proper response when one discovers this is not (usually) to abandon the theory in despair, but to exercise some imagination and initiative: what modifications to the theory, and what novel research strategies, would bring the new phenomena into line? Ultimately, taking on such challenges seems likely to be a very fruitful approach.

References

Berko, Jean (1958) "The child's learning of English morphology," *Word* 14, 150-177.

Crowhurst, Megan and Mark Hewitt (1995) "Prosodic overlay and headless feet in Yidiŋ," *Phonology* 12, 39-84.

Dixon, Robert M. W. (1977) *A Grammar of Yidiŋ*, Cambridge University Press.

Gnanadesikan, Amalia. 1995. "Markedness and Faithfulness constraints in child phonology." ROA-67, Rutgers Optimality Archive, http://ruccs.rutgers.edu/roa.html.

Hale, Kenneth (1973) "Deep-surface canonical disparities in relation to analysis and change," in Thomas Sebeok, ed., *Current Trends in Linguistics* 11, 401-458.

Halle, Morris and Jean-Roger Vergnaud (1987) *An Essay on Stress*, MIT Press, Cambridge, MA.

Hayes, Bruce (1980) *A Metrical Theory of Stress Rules*, Ph.D. dissertation, MIT, Cambridge, MA. [Published 1985: Garland Press, New York.]

Hayes, Bruce (1982) "Metrical Structure as the Organizing Principle of Yidin^{y} Phonology," in H. van der Hulst and N. Smith, eds., *The Structure of Phonological Representations, Part I*, Foris Publications, Dordrecht, The Netherlands.

Hayes, Bruce (1995a) *Metrical Stress Theory: Principles and Case Studies*, University of Chicago Press, Chicago.

Hayes, Bruce (1995b) "On what to teach the undergraduates: some changing orthodoxies in phonological theory," in Ik-Hwan Lee, ed., *Linguistics in the Morning Calm 3*, Hanshin, Seoul, pp. 59-77.

Hayes, Bruce (in press) "Phonetically-Driven Phonology: The Role of Optimality Theory and Inductive Grounding," to appear in the proceedings volume of the 1996 Milwaukee Conference on Formalism and Functionalism in Linguistics.

Itô, Junko and Armin Mester (1995) "Japanese phonology," in John Goldsmith, eds., *The Handbook of Phonological Theory*, Basil Blackwell, Oxford, pp. 817-838.

Kenstowicz, Michael (1996) "Base-Identity and Uniform Exponence: alternatives to cyclicity," in J. Duran and B. Laks, eds., *Current Trends in Phonology*, European Studies Research Institute, University of Salford Publications, 363-393. [Also: Rutgers Optimality Archive 103, http://ruccs.rutgers.edu/roa.html].

Kenstowicz, Michael and Charles Kisseberth (1977) *Topics in Phonological Theory*, Academic Press, New York.

Kenstowicz, Michael and Charles Kisseberth (1979) *Generative Phonology: Description and Theory*, Academic Press, New York.

Kisseberth, Charles (1970) "On the functional unity of phonological rules," *Linguistic Inquiry* 1, 291-306.

Kirchner, Robert (1993) "Optimizing Yidiŋ phonology," ms., UCLA.

McCarthy, John and Alan S. Prince (1995) "Faithfulness and reduplicative identity," in Jill N. Beckman, Laura Walsh Dickey, and Suzanne Urbanczyk, eds., *Papers in Optimality Theory, University of Massachusetts Occasional Papers* 18, 249-384. [Also: Rutgers Optimality Archive 60, http://ruccs.rutgers.edu/roa.html].

Nash, David (1979) "Yidiny stress: a metrical account," *North Eastern Linguistic Society (NELS) 9, CUNYForum* 7-8, 112-130.

Öhman, Sven (1966) "Coarticulation in VCV utterances: spectrographic measurements," *Journal of the Acoustical Society of America* 39, 151-168.

Pater, Joe. 1996. "Minimal violation and phonological development." Ms., University of British Columbia, Vancouver.

Pater, Joe (ms.) "On the nonuniformity of weight-to-stress and stress preservation effects in English," ms., McGill University. [Rutgers Optimality Archive 107, http://ruccs.rutgers.edu/roa.html].

Prince, Alan S. and Paul Smolensky (1993) *Optimality Theory: Constraint Interaction in Generative Grammar*, ms., Dept. of Linguistics, Rutgers University, New Brunswick, NJ and Program in Cognitive Science, Johns Hopkins University, Baltimore.

Schuh, Russell (1972) "Rule inversion in Chadic," *Studies in African Linguistics* 3, 379-397.

Smolensky, Paul (1993) "Harmony, markedness and phonological activity," Rutgers Optimality Archive 87, http://ruccs.rutgers.edu/roa.html.

Smolensky, Paul (1996) "The initial state and 'Richness of the Base' in Optimality Theory," http://www.cogsci.jhu.edu/TechReports. [Also: Rutgers Optimality Archive 154, http://ruccs.rutgers.edu/roa.html].

Stampe, David (1973) *A Dissertation on Natural Phonology*. Ph.D. dissertation, University of Chicago. Distributed 1979 by Indiana University Linguistics Club, Bloomington.

Tesar, Bruce B. and Paul Smolensky (1993) "The learnability of Optimality Theory: an algorithm and some basic complexity results," Rutgers Optimality Archive 2, http://ruccs.rutgers.edu/roa.html.

Tesar, Bruce B. and Paul Smolensky (1996) "Learnability in Optimality Theory," Rutgers Optimality Archive 110, http://ruccs.rutgers.edu/roa.html.

Vennemann, Theo (1972) "Rule inversion," *Lingua* 29, 209-42.

Department of Linguistics
UCLA
Los Angeles, CA 90095-1543
USA
bhayes@humnet.ucla.edu