7

hypothesis formation & testing, competing hypotheses

Doing an Analysis

This chapter explores a subset of the phonologies of a number of languages. The purpose of this chapter is to make explicit the reasoning typically applied to the task of solving a phonology problem. By studying models of problem solving, you not only better understand the logic of problem solving, you will also gain experience with rules and issues regarding underlying representations encountered in the languages of the world.

Analyzing a system of phonological alternations is not trivial: it requires practice, where you gain experience by solving phonological problems of increasing complexity, experience which facilitates subsequent problem solving. The wider your experience is with actual phonological processes and problem solving, the better able you will be to appreciate what processes are common in the languages of the world, and to understand the dynamics of hypothesis formation, testing and revision. The first analyses given here will be more explicit about the reasoning that goes into solving data sets of this nature, in some cases deliberately going down the wrong analytical path, so that you have the opportunity to recognise the wrong path, and see how to get back on the right path. In practice, many of the calculations that are involved here are done without explicitly thinking about it — once you have suitable experience with problem solving.

1. Yawelmani

Our first problem involves alternations in the verb paradigm in the Yawelmani dialect of Yokuts (California).

1.1. The data

Three phonological rules will be motivated by the following examples: vowel epenthesis, vowel shortening, and vowel harmony. It is not obvious what the underlying representation of verb roots is, so besides finding the rules we must make decisions about underlying forms.

(1)

nonfuture	*imperative*	*dubitative*	*passive aorist*	
xathin	xatk'a	xatal	xatit	eat
dubhun	dubk'a	dubal	dubut	lead by hand
xilhin	xilk'a	xilal	xilit	tangle
k'oʔhin	k'oʔk'o	k'oʔol	k'oʔit	throw
doshin	dosk'o	do:sol	do:sit	report
ṣaphin	ṣapk'a	ṣa:pal	ṣa:pit	burn
lanhin	lank'a	la:nal	la:nit	hear
mek'hin	mek'k'a	me:k'al	me:k'it	swallow
wonhin	wonk'o	wo:nol	wo:nit	hide
p'axathin	p'axatk'a	p'axa:tal	p'axa:tit	mourn
hiwethin	hiwetk'a	hiwe:tal	hiwe:tit	walk
ʔopothin	ʔopotk'o	ʔopo:tol	ʔopo:tit	arise from bed
yawalhin	yawalk'a	yawa:lal	yawa:lit	follow
paʔiṭhin	paʔiṭk'a	paʔṭal	paʔṭit	fight
ʔilikhin	ʔilikk'a	ʔilkal	ʔilkit	sing
logiwhin	logiwk'a	logwol	logwit	pulverize
ʔugunhun	ʔugunk'a	ʔugnal	ʔugnut	drink
lihimhin	lihimk'a	lihmal	lihmit	run
ʔayiyhin	ʔayiyk'a	ʔayyal	ʔayyit	pole a boat
t'oyixhin	t'oyixk'a	t'oyxol	t'oyxit	give medicine
luk'ulhun	luk'ulk'a	luk'lal	luk'lut	bury
so:nilhin	so:nilk'a	sonlol	sonlit	put on back
ʔa:milhin	ʔa:milk'a	ʔamlal	ʔamlit	help
mo:yinhin	mo:yink'a	moynol	moynit	become tired
ṣa:lik'hin	ṣa:lik'k'a	ṣalk'al	ṣalk'it	wake up

1.2. The first step: morphology

First we need a morphological analysis of the data. In a simple case, this involves looking at columns and rows of data, and figuring out which subparts of words are consistently present with one meaning, and which other subparts are consistently present with other meanings. This task is more complicated when the surface shape of roots and affixes changes due to phonological rules. We cannot provide a definitive morphological analysis of these data without knowing what the phonological system is, and certainty as to the phonological rules is impossible without knowing the morphological analysis. We break out of this seeming circle by adopting — and constantly revising — a preliminary and less precise analysis of the phonology and morphology. Improvement in the underlying representations should result in better rules, and as we refine the system of rules, the nature of the underlying distinctions hopefully becomes clearer.

In this case, four suffixes are added to roots, *-hin ~ -hun* 'nonfuture', *-k'a ~ -k'o* 'imperative', *-al ~ -ol* 'dubitative' and *-it ~ -ut* 'passive aorist'. The notation

-hin ~ *-hun* indicates that the suffix is pronounced either as *-hin* or as *-hun*. We need to discover when one form vs. the other is used, and express that relation in terms of an underlying form and a rule changing the underlying form.

Stem variants. Some stems have only one surface shape — *xat-* 'eat', *dub-* 'lead by hand', *xil-* 'tangle', and *k'oʔ-* 'throw', so the most natural assumption would be that these *are* the underlying forms for these particular stems (this assumption may turn out to be wrong, but it is a good starting assumption). Most stems in the data set have two surface manifestations. An important first step in understanding the rules of the language is to identify the alternations in the data, and one way to make the alternations explicit is to list the phonetic variants of each stem.

(2)

dos ~ do:s	'report'	ṣap ~ ṣa:p	'burn'
lan ~ la:n	'hear'	mek' ~ me:k'	'swallow'
won ~ wo:n	'hide'	p'axat ~ p'axa:t	'mourn'
hiwet ~ hiwe:t	'walk'	ʔopot ~ ʔopo:t	'arise from bed'
yawal ~ yawa:l	'follow'	paʔiṭ ~ paʔṭ	'fight'
ʔilik ~ ʔilk	'sing'	logiw ~ logw	'pulverize'
ʔugun ~ ʔugn	'drink'	lihim ~ lihm	'run'
ʔayiy ~ ʔayy	'pole a boat'	t'oyix ~ t'oyx	'give medicine'
luk'ul ~ luk'l	'bury'	so:nil ~ sonl	'put on back'
ʔa:mil ~ ʔaml	'help'	mo:yin ~ moyn	'become tired'
ṣa:lik' ~ ṣalk'	'wake up'		

In these cases, decisions must be made regarding the underlying forms.

Suffix variants. We must decide what the underlying form of each suffix is, and they all have two surface variants in terms of their vowel: either a nonrounded vowel, or a rounded vowel. For each suffix, we group the verbs in terms of which variant of the suffix is used with them.

(3)

-hin:	xat, xil, k'oʔ, dos, ṣap, lan, mek', won, p'axat, hiwet, ʔopot, yawal, paʔiṭ, ʔilik, logiw, lihim, ʔayiy, t'oyix, so:nil, ʔa:mil, mo:yin, ṣa:lik'
-hun:	dub, ʔugun, luk'ul
-k'a:	xat, dub, xil, ṣap, lan, mek', p'axat, hiwet, yawal, paʔiṭ, ʔilik, logiw, ʔugun, lihim, ʔayiy, t'oyix, luk'ul, so:nil, ʔa:mil, mo:yin, ṣa:lik'
-k'o:	k'oʔ, dos, won, ʔopot
-al:	xat, dub, xil, ṣa:p, la:n, me:k', p'axa:t, hiwe:t, yawa:l, paʔṭ, ʔilk, ʔugn, lihm, ʔayy, luk'l, ʔaml, ṣalk'

-ol:	k'oʔ, do:s, wo:n, ʔopo:t, logw, t'oyx, sonl, moyn
-it:	xat, xil, k'oʔ, do:s, ṣa:p, la:n, me:k', wo:n, p'axa:t, hiwe:t, ʔopo:t, yawa:l, paʔṭ, ʔilk, logw, lihm, ʔayy, t'oyx, sonl, ʔaml, moyn, ṣalk'
-ut:	dub, ʔugn, luk'l

1.3. Identifying phonological regularities

Vowel harmony. Having grouped the examples in this fashion, a phonological regularity can be detected. For the suffix *hin ~ hun*, the vowel *u* appears when the preceding vowel is *u*, and *i* appears in the suffix after any other vowel. The suffix *it ~ ut* obeys this same rule. The suffixes *k'a ~ k'o* and *al ~ ol* have the vowel *o* after *o*. This can be explained by positing a rule of vowel harmony between the suffix vowel and whatever vowel precedes it, where /a/ assimilates to /o/ and /i/ assimilates to /u/.

(4) $$\begin{bmatrix} V \\ \alpha hi \end{bmatrix} \rightarrow [+round] / \begin{bmatrix} V \\ \alpha hi \\ +\ round \end{bmatrix} C_0 \_\_\_$$

The variable notation — αhi.... αhi.... — expresses the condition that the vowels must have the same value of [hi], i.e. the harmonizing vowel must be [+hi] after a [+hi] round vowel, and [-hi] after a [-hi] round vowel, in order for the harmony rule to apply.

Vowel shortening. The next problem to tackle is the variation in the shape of the stem. A useful next step in trying to analyze that variation is to see whether the variants can be arranged into a small number of groups, organized according to the nature of the difference between the two stem shapes. In looking for such an organization, notice that some stems alternate in terms of having long versus short vowels, and in terms of having versus lacking a second vowel. Accordingly, we organize the data into the following classes of stem alternations (including the class of stems which have no alternation).

(5)

CVC —	xat, dub, xil, k'oʔ
CVC ~ CV:C —	dos ~ do:s, ṣap ~ ṣa:p, lan ~ la:n, mek' ~ me:k', won ~ wo:n
CVCVC ~ CVCV:C —	p'axat ~ p'axa:t, hiwet ~ hiwe:t, ʔopot ~ ʔopo:t, yawal ~ yawa:l

CVCVC ~ CVCC —	paʔiṭ ~ paʔṭ, ʔilik ~ ʔilk, logiw ~ logw, ʔugun ~ ʔugn, lihim ~ lihm, ʔayiy ~ ʔayy, t'oyix ~ t'oyx, luk'ul ~ luk'l
CV:CVC ~ CVCC —	so:nil ~ sonl, ʔa:mil ~ ʔaml, mo:yin ~ moyn, ṣa:lik' ~ ṣalk'

The initial hypothesis is that the invariant CVC stems have the underlying shape CVC. If there is no reason to make the underlying form be different from the surface form, the two forms should be assumed to be identical. Building on that decision, we will now set forth a hypothesis for stems which vary in shape between CVC and CV:C. It is highly unlikely that these stems also have the underlying shape CVC, since that would make it hard to account for stems such as /xat/ which are invariant CVC. We could not predict whether a stem vowel is supposed to have a length alternation or not, and the reasoning that leads to hypothesizing an underlying distinction /xat/ vs. /do:s/ which is contextually neutralized is exactly the same as that which leads to hypothesizing that in Russian (discussed in Chapter 4) "time" is underlyingly /raz/ and "forest" is /les/.

Given the conclusion that stems like *do:s ~ dos* have an underlying CV:C form, under what circumstance is the underlyingly long vowel of the stem shortened? Taking /do:s/ as a representative, and mechanically combining the assumed underlying stem with what we take to be the underlying form of the suffix, we arrive at the following underlying and surface relations.

(6)					
	underlying	do:s-hin	do:s-k'a	do:s-al	do:s-it
	surface	doshin	dosk'o	do:sol	do:sit

The change of /a/ to [o] is due to vowel harmony. There is also a change in vowel length before *k'a* and *hin*, and not before *-al* and *-it*. These suffixes are distinguished by whether they begin with a consonant or a vowel, thus whether combining the stem and suffix would result in the sequence V:CC. Scanning the entire data set reveals an important generalization, that a long vowel is always followed by CV, that is, a long vowel only occurs in an open syllable. The discovery of this generalization allows us to posit the following vowel shortening rule.

(7) V → [-long] / __ CC

This rule is all that is needed to explain both the invariant CVC stems and the alternating CV:C ~ CVC stems. Underingly /do:s-hin/ undergoes (7) and gives the surface form [doshin] — all other forms preserve the underlying length of the vowel. The existence of this rule also explains why we do not find the surface sequence V:CC — a long vowel before a cluster of two consonants — anywhere in the data, as such sequences undergo vowel shortening.

We turn next to the stems with the shape CVCVC ~ CVCV:C such as *p'axat ~ p'axa:t*. Since we have already encountered a rule which accounts for alternations in vowel length, we should immediately suspect that this length alternation is the same as the one just accounted for in CV:C ~ CVC stems. When we inspect the contexts where the long-vowel variant occurs, we see that there are long vowels when a vowel initial suffix is added, and short vowels when a consonant initial suffix is added. In other words, these stems are virtually the same as /CV:C/ stems, except that they have the underlying shape /CVCV:C/. We initially hypothesized that there was a rule of vowel shortening based on /CV:C/ stems, and that rule nicely handled those data. The way we formulated that rule was quite general, since it only said "shorten a long vowel before two consonants". Such a statement predicts that, if there are other stem shapes such as /CVCV:C/, they too will undergo that rule. We have now discovered that such stems do undergo the shortening rule, providing independent support for that rule.

Epenthesis. This reduces the unsolved part of the problem to two remaining classes of stems. In one of those, there is an alternation between presence versus absence of a vowel, and in the second group there is an alternation in vowel length as well as an alternation in the presence versus lack of a vowel in the second syllable; this should make us suspect that the vowel shortening rule applies to the second of these sets. Concentrating on the contexts where the stem has the shape CV(:)CVC as opposed to the shape CVCC, we notice that CV(:)CVC appears before consonant-initial suffixes and CVCC appears before vowel initial suffixes. We do not know at this point whether the second vowel is underlyingly part of the stem and is deleted in one context, or whether the vowel is inserted in a different context. Therefore, we will consider both possibilities: consideration of alternative hypotheses is an essential part of problem solving.

First suppose that the vowel is not part of the underlying representation of the stem. In that case, we assume the following representations

(8)					
	underlying	ʔilk-hin	ʔilk-k'a	ʔilk-al	ʔilk-it
	surface	ʔilik-hin	ʔilik-k'a	ʔilk-al	ʔilk-it
	underlying	ṣa:lk'-hin	ṣa:lk'-k'a	ṣa:lk'-al	ṣa:lk'-it
	surface	ṣa:lik'-hin	ṣa:lik'-k'a	ṣalk'-al	ṣalk'-it

Focusing on the hypothesized underlying representations where a vowel might be inserted, we notice that a vowel appears just in case the underlying representation has a sequence of three consonants. Looking at all of the data, we notice that there are no surface sequences of three or more consonants, making such an epenthesis approach plausible.

In order for an epenthesis solution to work, the actual quality of the inserted vowel must be completely predictable. If we were to discover that the qual-

ity of the second vowel is unpredictable, then it would necessarily be part of the underlying representation since unpredictble information must be in the underlying form. The vowel in the second syllable is always high, and is round just in case the preceding vowel is high and round. In other words, the vowel in question is a high vowel whose backness and roundness is predictable, given the rule of vowel harmony, and thus the vowel is fully predictable. Given the harmony rule, we can assume that the second vowel is *i*. It is then possible to account for these examples by applying the following rule of epenthesis.

(9) $\varnothing \rightarrow \underset{[+hi]}{V} \; / \; C \_\_ \; CC$

Given (9), the underlying form of the CVCiC ~ CVCC stems would be /CVCC/ and the underlying form of the CV:CiC ~ CVCC stems would be /CV:CC/. For stems like /ʔilk/, epenthesis applies to underlying /CVCC+CV(C)/ to give surface [CVCiC+CV(C)]: /ʔilk-hin/ → [ʔilikhin]. The alternant CVCC before VC suffixes — [ʔilkal] — directly reflects the underlying form.

For /CV:CC/ stems like /ṣa:lk'/, epenthesis will also apply to underlying /CV:CC+CV(C)/ giving the surface form [CV:CiC+CV(C)]: /ṣa:lk-hin/ → [ṣa:likhin]. When a VC suffix is added to such stems, there is no epenthesis, but we do find shortening of the underlyingly long vowel which stands before a consonant cluster. (/ṣa:lkal/ → [ṣalkal]). The rules of vowel harmony, epenthesis and vowel shortening, combined with our analyses of underlying representations, account for all aspects of the data in (1). We conclude that epenthesis is a *possible* account of these alternations.

The preceding analysis has assumed a rule of epenthesis based on underlying representations of the form /CVCC/ and /CV:CC/, but we should explore the competing hypothesis that the vowel found in these stems is not inserted, and is part of the underlying representation. Under that hypothesis, underlying representations of the relevant stems would be the following.

(10) paʔit, ʔilik, logiw, ʔugun, lihim, ʔayiy, t'oyix, luk'ul
so:nil, ʔa:mil, mo:yin, ṣa:lik'

Presuming that these are the underlying stems, a rule of vowel deletion is required to explain the discrepancy between surface and underlying forms, which can be seen in (11).

(11)

	underlying	luk'ul-hun	luk'ul-k'a	luk'ul-al	luk'ul-ut
	surface	luk'ul-hun	luk'ul-k'a	luk'l-al	luk'l-ut
	underlying	so:nil-hin	so:nil-k'a	so:nil-ol	so:nil-it
	surface	so:nil-hin	so:nil-k'a	sonl-ol	sonl-it

In forms which involve an alternation between a vowel and ∅, the context for vowel deletion would initially appear to be in an open syllable. This statement would produce too general a rule, since there are many vowels in open syllables, viz. *x**a**tal, k'**o**ʔit, d**o:**sit, p'**a**xathin* and *p'**a**x**a:**tal* among others. In some of these, deletion of a vowel would lead to a word-initial consonant cluster, i.e. we would predict **xtal, *k'ʔit, *dsit, *p'xathin* and **p'x**a:**tal*, and we see no word-initial clusters of consonants. If we are to have vowel deletion, the rule must be restricted from creating such clusters, so one way to enforce that requirement is to require the target of deletion to be preceded by the sequence VC. Thus, we might hypothesize the following syncope rule, one found in many languages.

(12) $V \rightarrow \varnothing$ / VC ___CV

This rule still makes incorrect predictions, since in fact there are vowels in the context VC___CV, as shown by forms such as *p'axa:tal, ʔopo:tit*, which according to (12) should be deleted. Since all such examples involve long vowels, it is a simple matter to restrict the assumed deletion rule to short vowels.

(13) $\underset{[-long]}{V} \rightarrow \varnothing$ / VC ___CV

With this rule of vowel syncope, the problem of vowel ~ ∅ alternations can also be accounted for. The remaining details of the analysis are exactly the same as they are under the assumption that there is a rule of vowel insertion.

1.4. Evaluating alternatives

In terms of simply generating the data, both the syncope and epenthesis analyses work. The question then becomes, is there a reason to chose one of these hypotheses over the other? It is entirely possible that we will not be able to come up with any compelling reasons for selecting one analysis over the other, in which case we must simply accept the fact that there are two equally plausible ways to account for the facts. As far as the simplicity, naturalness and generality of the two analyses is concerned, neither theory is superior to the other. Processes inserting vowels to break up CCC clusters are very common, as are rules of syncope which delete short vowels in the context VC___CV.

By sheer counting of symbols, the epenthesis rule might be slightly superior since it only requires reference to five entities and syncope requires reference to seven entities. Such literal symbol counting, practiced in the early era of generative phonology, is misguided.

We should also consider the factual predictions of the two analyses. The epenthesis analysis predicts that there should be no CCC sequences in the language, and this appears to be correct. On the other hand, the syncope analysis predicts that there should be no short vowels in the context VC___CV, which also appears to be correct. Interestingly, neither account actually makes the prediction of the competing analysis — so, the epenthesis analysis does not preclude the exis-

tence of short vowels in the VC__CV context, and the syncope analysis does not preclude the existence of CCC sequences. If it turns out that there *are* CCC sequences in the language, the epenthesis solution will probably have to be rejected; whereas is there *are* VCVCV sequences in the language, the syncope analysis will probably have to be rejected. This would motivate further research into the language, to determine if one of these analyses makes a bad prediction.

A related issue to consider is the question of 'coincidence', in terms of assumed underlying representations. In lieu of a specific rule which restricts the occurrence of phonemes in some environment, we expect phonemes to combine without any constraints. Clearly there must be some constraints on underlying representations in Yawelmani, since for example we do not find underlying representations such as /ioate/ with sequences of vowels. In this case, there is no motivation from phonological alternations to suspect that there might be underlying forms such as /ioate/. As far as logical possibilities in underlying forms are concerned for the issue at hand — epenthesis versus deletion — both analyses result in systematic gaps in the logically possible underlying forms. Under the epenthesis analysis, there are apparently no stems of the underlying form /CVCVC/, although there are stems of the form /CVCV:C/. Under the syncope analysis, we notice that all short second-syllable vowels in disyllabic stems are in fact /i/ (surface [u] in some cases, in accordance with vowel harmony).

At this point, it is impossible to give strong arguments in favor of one analysis over another, so we accept this indeterminacy for now. The fundamental point is that each analysis implies a set of predictions about possible and impossible forms in the language, and these predictions need to be tested against the available data. In this case, we have not been able to determine that one theory is clearly superior to the other. The main research problem which we face is that the corpus of data from Yawelmani available to us at this point is restricted, so we cannot know whether generalizations which we extract about the language based on this particular corpus are representative of the language as a whole. Even if we had access to a reference grammar for the language, there is some chance that our empirical generalizations based on the data from that grammar do not hold for the whole language, if the author of the grammar is not be aware of all examples.

2. Kihehe

The following data illustrate phonological processes of Kihehe (Tanzania). Each noun is in one of 15 numbered noun classes, like genders in French or German. The class of a noun is marked by a prefix. The goal is to determine the underlying form of stems and prefixes, and explain the processes at work in these data.

2.1. The data

Here are the relevant data from nouns.

(14)

Class 1

mutesi	‘trapper’	mulagusi	‘sorcerer’
mutelesi	‘cook’	muñwi	‘drinker’
mwiimbi	‘singer’	mweendi	‘one who likes people’
mwaasi	‘builder’	moogofi	‘one who is afraid’
moofusi	‘one who washes’	muuci	‘one who comes’

Class 2

vatesi	‘trappers’	valagusi	‘sorcerers’
vatelesi	‘cooks’	vañwi	‘drinkers’
viimbi	‘singers’	veendi	‘ones who like people’
vaasi	‘builders’	woogofi	‘ones who are afraid’
woofusi	‘one who washes’	wuuci	‘one who comes’

Class 3

muhoomi	‘cow hump’	muhogo	‘cassava’
mufuniko	‘cover’	muvili	‘body’
mwiina	‘hole’	mwiigiigi	‘shadow’
mweenda	‘cloth’	mooto	‘fire’
muuñu	‘salt’		

Class 4

mihoomi	‘cow humps’	mihogo	‘cassavas’
mifuniko	‘covers’	mivili	‘bodys’
miina	‘holes’	miigiigi	‘shadows’
myeenda	‘cloths’	myooto	‘fires’
myuuñu	‘salts’		

Class 6

mavafi	‘hairy caterpillars’	masaasi	‘bullets’
maboga	‘pumpkins’	mayayi	‘legs’
miino	‘teeth’	miiho	‘eyes’

Class 7

kigidi	‘waist’	kingaamba	‘sweet potato’
kisogo	‘back of head’	čuula	‘frog’
čuunga	‘wet lowland’	čaanga	‘grave’
kifiniko	‘dinky cover’	kivili	‘dinky body’
kihoomi	‘dinky cow hump’	kivafi	‘dinky hairy caterpillar’
čooto	‘dinky fire’	čeenda	‘dinky cloth’
čuuñu	‘dinky salt’	kiiho	‘dinky eye’
kiina	‘dinky hole’	kiigiigi	‘dinky shadow’

Class 8

figidi	'waists'	fingaamba	'sweet potatos'
fisogo	'backs of head'	fyuula	'frogs'
fyuunga	'wet lowlands'	fyaanga	'graves'
fifiniko	'dinky covers'	fivili	'dinky bodys'
fihoomi	'dinky cow humps'	fivafi	'dinky hairy caterpillars'
fyooto	'dinky fires'	fyeenda	'dinky cloths'
fyuuñu	'dinky salts'	fiiho	'dinky eyes'
fiina	'dinky holes'	fiigiigi	'dinky shadows'

Class 11

luteefu	'reed mat'	lupava	'stirring stick'
lutego	'trap'	ludali	'power
luhaanga	'sand'	lwiimbo	'song'
lweendo	'loving'	lwaaniko	'dry stuff'
lwiifwi	'chameleon'		

Class 12

kateefu	'small mat'	kakoongo	'small wound'
kafuniko	'small cover'	kangaamba	'small sweet potato'
kaasi	'small builder'	kiimbi	'small singer'
kaanga	'small grave'	kooto	'small file'
kuula	'small frog'	kuunga	'small wet lowland'

Class 13

tuteefu	'small mats'	tukoongo	'small wounds'
tufuniko	'small covers'	tungaamba	'small sweet potatos'
twaasi	'small builders'	twiimbi	'small singers'
twaanga	'small graves'	tooto	'small files'
tuula	'small frogs'	tuunga	'small wet lowlands'

Class 14

wuvaso	'sleeping place'	wulime	'cultivating'
wugali	'porridge'	wutiitu	'blackness'
weelu	'whiteness'	wuumi	'life'
woogofu	'fear'	wiiyooga	'mushroom'
waangufu	'speed'		

2.2. Morphological analysis

As always, a preliminary morphological analysis is the first step in solving this phonology problem. Each noun has some prefix that marks noun class, fol-

lowed by a stem. We also see, comparing nouns in various classes, that the same stems can appear in different classes, so for example class 3 *mu-hoomi* 'cow hump' is clearly related to class 4 *mi-hoomi* 'cow humps' — singulars and plurals are marked by changes in class; class 11 *lu-teefu* 'reed mat' is clearly related to *ka-teefu* 'small mat' and *tu-teefu* 'small mats'. The class prefixes have a number of phonetic manifestations, so we find *mu-*, *mw-* and *m-* for classes 1 and 3, *va*, *v-* and *w-* for class 2, *mi-*, *my-* and *m-* for class 4, *ma-* and *m-* for cl. 6, *ki-* and *č-* for class 7, *fi-* and *fy-* for class 8, *lu-* and *lw-* for class 11, *ka-* and *k-* for class 12, *tu-* and *tw-* for class 13, and *wu-*, *w-* for class 14.

2.3. Phonological alternations

Noun stems fall in two groups in terms of phonological processes: those which begin with a consonant, and those beginning with a vowel. Examples of stems which begin with a consonant are *-tesi* (cf. *mu-tesi, va-tesi*) and *-lagusi* (cf. *mu-lagusi, va-lagusi*); examples of stems which begin with vowels are *-iimbi* (cf. *mw-iimbi*, *v-iimbi*) and *-eendi* (*mw-eendi, v-eendi*). The best phonological information about the nature of the prefix is available from its form before a consonant, so our working hypothesis is that the underlying form of the noun prefix is that found before a consonant — it preserves more information.

As we try to understand the phonological changes found with vowel-initial stems, it is helpful to look for a general unity behind these changes. One important generalization about the language, judging from the data, is that there are no vowel sequences in the language (what may seem to be sequences such as *ii*, *ee* are not sequences, but are the orthographic representation of single long vowel segments). Given the assumption that the prefixes for classes 1 and 2 are respectively /mu/ and /va/, the expected underlying forms of 'singer' and 'singers' would be /mu-iimbi/ and /va-iimbi/. These differ from the surface forms [mw-iimbi] and [v-iimbi]: in the case of /mu-iimbi/, underlying /u/ has become [w], and in the case of underlying /va-iimbi/, underlying [a] has been deleted. In both cases, the end result is that an underlying cluster of vowels has been eliminated.

Glide formation vs. vowel deletion. Now we should ask, why is a vowel deleted in one case but turned into a glide in another case? The answer lies in the nature of the prefix vowel. The vowel /u/ becomes the glide [w], and the only difference between *u* and *w* is that the former is syllabic (a vowel) where the latter is nonsyllabic. The low vowel /a/, on the other hand, does not have a corresponding glide in this language (or in any language). In other words, a rule of glide-formation simply could not apply to /a/ and result in a segment of the language.

To make progress in solving the problem, we need to advance hypotheses and test them against the data. We therefore assume the following rules of glide formation and vowel deletion.

(15) $\begin{matrix} V \\ [+hi] \end{matrix} \rightarrow$ [-syll] / __ V *Glide Formation*

(16) V → ∅ / __ V *a-Deletion*

By ordering (16) after (15), we can make (16) very general, since (15) will have already eliminated other vowel sequences. At this point, we can simply go through the data from top to bottom, seeing whether we are able to account for the examples with no further rules — or, we may find that other rules become necessary.

For nouns in class 1, the examples *mw-iimbi*, *mw-eendi* and *mw-aasi* are straightforward, deriving from /mu-iimbi/, /mu-eendi/ and /mu-aasi/. The forms *m-oogofi*, *m-oofusi* and *m-uuci* presumably derive from /mu-oogofi/ and /mu-oofusi/ and /mu-uuci/. The vowel /u/ has been deleted, which seems to run counter to our hypothesis that high vowels become glides before vowels. It is possible that there is another rule that deletes /u/ before a round vowel.

(17) $u \rightarrow \varnothing / \_\_\_ \begin{matrix} V \\ [+round] \end{matrix}$ *u-Deletion*

We could also consider letting the glide formation rule apply and then explain the difference /mu-aasi/ → *mw-aasi* vs. /mu-oofusi/ → *m-oofusi* by subjecting derived *mw-oofusi* to a rule deleting *w* before a round vowel.

(18) w → ∅ / ___ [+round] *w-Deletion*

Thus we must keep in mind two hypotheses regarding /u+o/ and /u+u/ sequences.

v-rounding. Now consider class 2. In stems beginning with a vowel, we easily explain *v-iimbi, v-eendi* and *v-aasi* from *va-iimbi, va-eendi* and *va-aasi*, where *a*-deletion applies. Something else seems to be happening in *w-oogofi, w-oofusi* and *w-uuci* from *va-oogofi, va-oofusi* and *va-uuci*. Application of *a*-Deletion would yield *v-oogofi, v-oofusi* and *v-uuci*, which differ from the surface forms only in the replacement of *v* by *w*. Since this process takes place before a round vowel, we conjecture that there may be an assimilation rule such as the following.

(19) v → w / ___ [+round] *v-Rounding*

If there is such a rule in the language, it would eliminate any sequences *vu*, *vo*: and the data contains no such sequences. There is still a problem to address, that *w*-Deletion (18) should apply to *woogofi* but it does not — the surface form is not *[oogofi]. Two explanations come to mind. One is that *v*-Rounding is ordered after *w*-Deletion, so at the stage where *w*-Deletion would apply, this word has the shape

voogofi and not *woogofi* (so *w*-deletion cannot apply). The other is that (18) needs to be revised, so that it only deletes a post-consonantal *w* before a round vowel.

(20) $w \rightarrow \varnothing / C \_\_\_ [+round]$

Our decision-making criteria are not stringent enough that we can definitively chose between these solutions, so we will leave this question open for the time.

Moving to other classes, the nouns in class 3 present no problems. Glide formation applies to this prefix, so /mu-iina/ → [mw-iina], and before a round vowel derived *w* deletes, so /mu-ooto/ → *mw-ooto* which then becomes [m-ooto].

Front vowels and glides. The nouns in class 4 generally conform to the predictions of our analysis. Note in particular that underlying /mi-uuñu/ and /mi-ooto/ undergo Glide Formation before a round vowel. Such examples show that it was correct to state glide formation rule in a more general way, so that all high vowels (and not just /u/) become glides before any vowel (not just non-round vowels).

We cannot yet fully explain what happens with noun stems beginning with the vowel *i*, as in *m-iina, m-iigiigi*. Given /mi-iina/, /mi-iigiigi/, we predict surface **my-iina,* **my-iigiigi*. This is reminiscent of the problem of /mu-oogofi/ and /mu-uuci/ and we might want to generalize the rule deleting a glide, to include deleting a front glide before a front vowel (analogous to deleting a round glide before a round vowel). What prevents us from doing this is that while *w* deletes before both *u* and *o*, *y* only deletes before *i* and not *e*, as we can see from *my-eenda*. It might be more elegant or symmetrical for round glides to delete before round vowels of any height *and* front glides to delete before front vowels of any height, but the facts say otherwise: a front glide only deletes before a front *high* vowel.

(21) $\begin{bmatrix} + \text{hi} \\ - \text{back} \\ - \text{syl} \end{bmatrix} \rightarrow \varnothing / \_\_\_ \begin{bmatrix} + \text{hi} \\ - \text{back} \end{bmatrix}$ *y-Deletion*

Checking other classes: discovering a palatalization rule. The class 6 prefix *ma-* presents no surprises at all: it appears as *ma-* before a consonant, and its vowel deletes before another vowel, as in *m-iino* from *ma-iino*. The class 7 prefix, on the other hand, is more complex. Before a consonant it appears as *ki-*, and also appears as *k*(*i*)- before *i*. Before other vowels, it appears as *č*, as in *č-uula, č-aanga, č-ooto* and *č-eenda*. Again, we continue the procedure of comparing the underlying and predicted surface forms (predicted by mechanically applying the rules which we have already postulated to the underlying forms we have committed ourselves to), to see exactly what governs this discrepancy. From underlying *ki-uula, ki-aanga, ki-ooto* and *ki-eenda* we would expect *ky-uula, ky-aanga, ky-ooto* and *ky-eenda*, given Glide Formation. The discrepancy lies in the fact that the predicted sequence

ky has been fused into *č*, a process of palatalization found in many languages. Since *ky* is nowhere found in the data, we can confidently posit the following rule.

(22) ky → č *Palatalization*

Since /ki/ surfaces as [č] when attached to a vowel-initial noun stem, the question arises as to what has happened in *k-iiho, k-iina* and *k-iigiigi*. The glide formation rule should apply to /ki-iiho/, /ki-iina/ and /ki-iigiigi/ giving *ky-iiho, ky-iina* and *ky-iigiigi*, which we would expect to undergo (22). But there is a rule deleting *y* before *i*. If *y* is deleted by that rule, it could not condition the change of *k* to *č*, so all that is required is the ordering statement that *y*-Deletion precedes palatalization (22). Thus /ki-iina/ becomes *ky-iina* by Glide Formation, and before the Palatalization rule can apply, *y*-Deletion rule (21) deletes the glide that is crucial for (22).

Deciding on the form of w-deletion; degemination. At this point, we can quickly check the examples in classes 8, 11, 12 and 13 and verify that our analysis explains all of these forms as well. The final set of examples are those in class 14, which has the prefix /wu/. This prefix raises a question in terms of our analysis: why do we have the sequence [wu], which is eliminated by a rule elsewhere? One explanation is the statement of the rule itself: if (20) is the correct rule, then this *w* could not delete because it is not preceded by a consonant. The other possibility is that [wu] actually comes from /vu/ by applying *v*-Rounding (19), which we assumed applies after *w*-Deletion. While both explanations work, the analysis where [wu] is underlying /vu/ has the disadvantage of being rather abstract, in positing an underlying segment in the prefix which never appears as such. This issue was presaged in chapter 4 and is discussed in more detail in chapter 9: for the moment we will simply say that given a choice between a concrete analysis where the underlying form of a morpheme is composed only of segments which actually appear as such in some surface manifestation of the morpheme, and an abstract form with a segment that never appears on the surface, the concrete analysis is preferable to the abstract one, all other things being comparable. On that basis, we decide that the underlying for of the class 14 prefix is /wu/, which means that the proper explanation for failure of *w*-deletion lies in the statement of *w*-deletion itself, as (20).

Still analysing this class of nouns, we now focus on examples where the prefix precedes a vowel-initial stem, e.g. *w-eelu, w-uumi, w-oogofu, w-iiyooga* and *w-aangufu* from underlying /wu-eelu/, /wu-uumi/, /wu-oogofu/, /wu-iiyooga/ and /wu-aangufu/. Applying Glide Formation would give the surface forms **ww-eelu,* **ww-uumi,* **ww-oogofu,* **ww-iiyooga* and **ww-aangufu*, which differ from the surface form in a simple way, that they have a geminate *w* where the actual form has only a single *w* (in fact, there do not seem to be any geminate consonants in the language), which allows us to posit the following degemination rule.

(23) $C_iC_i \rightarrow C_i$ *Degemination*

2.4. Extending the data

Verbs are subject to these same rules, as some additional data will show, and an analysis of verbs will provide additional support for aspects of this analysis. Kihehe is a tone language, and while we have not been concerned with accounting for tone (and have not marked tones), in the following data, tones are marked, and can be predicted by rule. In analyzing these data, we want to account for the placement of the High tone (H), which is marked with an acute accent.

(24)

V	**V for**	**V for each**	**make V**
kúkama	kúkamíla	kúkamilána	kúkamyá
kúsana	kúsaníla	kúsanilána	kúsanyá
kútova	kútovéla	kútovelána	kútovyá
kúlava	kúlavíla	kúlavilána	kúlavyá
kúfwiíma	kúfwiimíla	kúfwiimilána	kúfwiimyá
kúkalaánga	kúkalaangíla	kúkalaangilána	kúkalaangyá
kúkaláva	kúkalavíla	kúkalavilána	kúkalavyá
kwéenda	kwéendéla	kwéendelána	kwéendyá
kwíimba	kwíimbíla	kwíimbilána	kwíimbyá
kóogópa	kóogopéla	kóogopelána	kóogopyá

be V'd	**V us**	**V them**	**meaning**
kúkamwá	kútukáma	kúvakáma	milk
kúsanwá	kútusána	kúvasána	comb
kútowá	kútutóva	kúvatóva	beat
kúlawá	kútuláva	kúvaláva	look at
kúfwiimwá	kútufwiíma	kúvafwiíma	hunt
kúkalaangwá	kútukalaánga	kúvakalaánga	fry
kúkalawá	—	—	take bath
kwéendwá	kútweénda	kúveénda	love
kwíimbwá	kútwiímba	kúviímba	sing
kóogopwá	kútoogópa	kúwoogópa	fear

The morphology. These data indicate that all verbs begin with *kú* or something derivable from /kú/ by the rules already motivated, thus we assume that *kú-* is an inflectional prefix. In addition, all verbs end with the vowel *a*, which is probably a morpheme since it is unlikely that every root would end in exactly the same vowel. The stem 'milk' is probably *-kam-*. Various grammatical relations are expressed by suffixes standing between the stem and the suffix *-a*, such as *-il-* 'for', *-an-* 'each other', *-y-* 'make', *-w-* passive: the objects 'us' and 'them' are marked by the prefixes *-tu-* and *-va-* between the prefix *kú-* and the verb stem.

Phonological rules. Looking at the last three roots, which are vowel-initial, the prefixes *kú-*, *tu-* and *va-* are subject to the rules motivated on the basis of nouns, where /u/ becomes [w] before a vowel, but deletes after a consonant and before a round vowel (so, /ku-oogopa/ → *kwoogopa* → [koogopa]); the sequence *vo* becomes *wo* (/ku-va-oogopa/ → *kuvoogopa* → [kuwoogopa]). The change of /v/ to *w* is also seen in examples such as *kútowá* and *kúlawá*, coming (apparently) from /ku-tov-w-a/ and /ku-lav-w-a/. The rule of *v*-Rounding would derive *kútowwá* and *kúlawwá*, and the actual phonetic forms can be accounted for based on that intermediate form by Degemination.

One additional segmental process of vowel harmony is motivated by the above examples. The benefactive suffix retains its underlying high vowel in forms such as *kúkam-íl-a, kúsan-íl-a* and *kúfwiim-íl-a*, but that vowel assimilates in height to a preceding mid vowel in examples such as *kútov-él-a, kwéend-él-a* and *kóogop-él-a*. This motivates the following vowel harmony rule:

(25) $V \rightarrow [-hi] / \underset{\begin{bmatrix}-hi\\-low\end{bmatrix}}{V\ C_0}\ \_\_\_$ *Vowel Harmony*

Regarding tone, most examples have a H tone on the second-to-last vowel of the word (this may be the second part of a long vowel in the penultimate syllable, or the only vowel of a short penultimate syllable), which can be accounted for by the following rule.

(26) $V \rightarrow [+H] / \_\_\ C_0\ V\#$ *Tone Assignment*

In some verbs, this H is missing — cf. *kúkama, kúsana, kútova*. Applying this tone assignment rule to these forms would result in outputs such as **kúkáma, *kúsána, *kútóva*, with H tones on adjacent vowels. Since our examples contain no cases of consecutive toned vowels, we may assume a rule along the following lines.

(27) $V \rightarrow [-H] / \underset{[+H]}{V\ C_0}\ \_\_\_$

What about the columns with the suffixes *-y-* 'make' and *-w-* 'passive', which have word-final H, not penult H ? We expect **kúkalaángwa*. But if these two suffixes are underlyingly *i* and *u*, then the underlying form of *kúkalaangwá* would be /kú-kalaang-u-a/. H tone would be assigned to the penultimate vowel under that assumption, giving *kúkalaangúa*. However, we already know that there is a rule of Glide Formation which would turn *u* and *i* into *w* and *y* before vowels, a rule which has obviously applied in these forms. Since only syllabic elements can bear tones, the tone on the penultimate vowel apparently shifts to the final syllable, where it can be pronounced.

Such tone shift, where the tone of a vowel shifts to another vowel when the original vowel deletes or desyllabifies is common in tone languages and is discussed in the last chapter

3. Icelandic

Our next example is alternations in noun inflection in Modern Icelandic.

3.1. The data

The relevant data are in (28). The task is to provide a unique underlying representation for each stem and case suffix, state what phonological rules are required to account for these data, and indicate what order they apply in, when the ordering of rules matters.

(28)

hestür	hattür	heimür	gröütür	skougür	*nom. sg*
hest	hatt	heim	gröüt	skoug	*acc. sg.*
hesti	hatti	heimi	gröüti	skougi	*dat. sg.*
hests	hatts	heims	gröüts	skougs	*gen. sg.*
hestar	hattar	heimar	gröütar	skougar	*nom. pl.*
hesta	hatta	heima	gröüta	skouga	*acc. pl.*
hestüm	höttüm	heimüm	gröütüm	skougüm	*dat. pl*
‘horse’	‘hat’	‘home’	‘porridge’	‘forest’	
garðür	laiknir	hirðir	himinn	morgünn	*nom. sg.*
garð	laikni	hirði	himin	morgün	*acc. sg.*
garði	laikni	hirði	himni	morgni	*dat. sg.*
garðs	laiknis	hirðis	himins	morgüns	*gen. sg.*
garðar	laiknar	hirðar	himnar	morgnar	*nom. pl.*
garða	laikna	hirða	himna	morgna	*acc. pl.*
görðüm	laiknüm	hirðüm	himnüm	morgnüm	*dat. pl.*
‘garden’	‘doctor’	‘herd’	‘heaven’	‘morning’	
stoull	magaull	yöküll	θümall	mour	*nom. sg.*
stoul	magaul	yökül	θümal	mou	*acc. sg.*
stouli	magauli	yökli	θümli	mou	*dat. sg.*
stouls	magauls	yöküls	θümals	mous	*gen. sg.*
stoular	magaular	yöklar	θümlar	mouar	*nom. pl.*
stoula	magaula	yökla	θümla	moua	*acc. pl.*
stoulüm	magaulüm	yöklüm	θümlüm	mouüm	*dat. pl.*
‘chair’	‘flank’	‘glacier’	‘thumb’	‘peat’	
akür	hamar	hver	galdür	byour	*nom. sg.*
akür	hamar	hver	galdür	byour	*acc. sg.*
akri	hamri	hver	galdri	byour	*dat. sg.*
akürs	hamars	hvers	galdurs	byours	*gen. sg.*

akrar	hamrar	hverar	galdrar	byourar	*nom. pl.*
akra	hamra	hvera	galdra	byoura	*acc. pl.*
ökrüm	hömrüm	hverüm	göldrüm	byourüm	*dat. pl.*
'field'	'hammer'	'geyser'	'magic'	'beer'	

3.2. Morphological analysis

It is not immediately clear what are appropriate underlying representations for some case suffixes. It would appear that the nom. sg. ending is something like *-ür* or maybe *-ir*, although sometimes you just find lengthening of a stem final consonant. We start by assuming that the acc. sg. has no case suffix, the dat. sg. is *-i*, the gen. sg. is *-s*, the nom. pl. is *-ar*, the acc. pl. is *-a* and the dat. pl. is *-üm*, since in almost all stems, that is how these suffixes are actually manifested. It would similarly not be unreasonable to assume that the acc. sg. form is a close approximation of the underlying form of the stem.

3.3. Phonological alternations

On the basis of these assumptions about underlying forms, we can identify some phonological alternations which need to be explained. First and foremost, we need to explain the consonantal variation in the nom. sg. Second, we need to explain the alternation between [a] and [ö] in examples such as [hatta] ~ [höttüm]. Third, there is a vowel ~ ∅ alternation as in [himin] ~ [himna] and [morgün] ~ [morgna]. Fourth, the dat. sg. form generally appears as [i], but in some cases does not surface. We will try to solve one of these problems, selecting at random, since at this point we have no reason to think that finding a solution to one of these problems is dependent on finding a solution to any other of these problems.

The vowel of the nominative singular. We will begin with the problem of the nominative singular. The first step in taming this problem is to state exactly what the problem is. There are many apparent realizations of this suffix, depending on the noun stem to which it is attached, we find *-ür*, *-ir*, *-r*, *-∅* , *-l* and *-n*. Constructing this list of surface realizations alone is enough to allow us to make an intelligent initial guess about the underlying form, which is that the nom. sg. is -(V)*r*, since half of the variants of this affix actually contain *-r* (of course, this assumption could be wrong, since numerical counts are not arguments for underlying forms, only suggestions, but again we need to start somewhere). The next step is to extract generalizations about the contexts where each variant is used. We would start by listing the presumably underlying stems themselves, noting that we have *-ur* with /hest/, /hatt/, /gröüt/, /heim/, /garð/ and /skoug/, *-r* with /mou/, /laikni/ and /hirði/, ∅ with /akür/, /hamar/, /hver/ and /galdür/, *-l* with /stoul/, /magaul/ and /yökül/, and *-n* with /himin/ and /morgün/. At this point, generalizations about the underlying form become easier to see: we find [l] after /l/, [n] after /n/, ∅ after /r/,

You may wonder, why assume that *garðür* illustrates the variant *-ur* and *galdür* does not? We assumed that the acc. sg. best reflects the underlying form, and since [ür] is present in acc. sg. *galdür* but not *garð*, [ür] must be part of the stem in *galdür* and not in *garðür*.

[r] after a vowel and [ür] after any other consonant. We can conclude that the most likely underlying forms for this suffix are /ür/ and /r/.

Having identified the nature of the conditioning environment and armed with two hypotheses about the underlying form, it is time to transform this information into specific rules. Since underlying representations and rules go hand in hand, we need to determine whether one of the assumed underlying representations for the suffix results in more plausible rules. Let of consider the entailments of these underlying forms in terms of the rules that they commit us to.

(29) **Hypothesis: /ür/**

No change:
hest-ür → hestür, hatt-ür → hattür, skoug-ür → skougür, heim-ür → heimür, garð-ür → garðür, gröüt-ür → gröütür

Deletion of a vowel
mou-ür → mour, laikni-ür → laiknir, hirði-ür → hirðir

Deletion of a vowel and deletion of r
akür-ür → akür, hver-ür → hver, byour-ür → byour, galdür-ür → galdür, hamar-ür → hamar

Deletion of a vowel and assimilation
stoul-ür → stoull, yökul-ür → yökül-ür, θümal-ür → θümall, himin-ür → himinn, magaul-ür → magaull, morgün-ür → morgünn

Hypothesis: /r/

Insertion of a vowel
hest-r → hestür, hatt-r → hattür, skoug-r → skougür, heim-r → heimür, garð-r → garðür, gröüt-r → gröütür

No change
mou-r → mour, laikni-r → laiknir, hirði-r → hirðir

Deletion of r
akür-r → akür, hver-r → hver, byour-r → byour, galdür-r → galdür, hamar-r → hamar

Assimilation
stoul-r → stoull, yökül-r → yökül-ür, θümal-r → θümall, himin-r → himinn, magaul-r → magaull, morgün-r → morgünn

We will start with the assumption that the suffix is underlyingly /ür/. Given that, a deletion rule is required to eliminate *ü* from this suffix when it is preceded by one of /r,l,n/.

(30) $$\begin{bmatrix} +\text{syl} \\ +\text{hi} \\ +\text{rd} \end{bmatrix} \rightarrow \varnothing \,/\, \text{V} \begin{bmatrix} +\text{coronal} \\ +\text{sonorant} \end{bmatrix} \_\_\ \text{r \#}$$

Applying this rule to underlying *himinür, stoulür, hverür* would yield forms such as *himinr, stoulr, hverr*, and these outputs would be subject to other rules.

In addition, given the assumption of underlyingly /ür/, we would require a rule to delete the vowel from the suffix when the preceding stem ends in a vowel. This rule would allow us to account for forms such as *mour, laiknir*, and *hirðir*, from assumed *mou-ür, laikni-ür*, and *hirði-ür*. Deletion of a vowel after another vowel is not implausible, so we might postulate the following rule.

(31) $V \rightarrow \varnothing / V \_\_$

However, this is too general, since *u* can be preceded by other vowels — cf. *mour, skougür, magaull*. This particular statement of the rule makes a prediction that certain kinds of phonetic sequences should not occur, and that prediction is wrong.

Our rule went wrong in that it does not distinguish supposed vowel combinations which would be created by concatenation of morphemes (these sequences do undergo reduction) from diphthongs which are contained wholely within a single morpheme (which do not undergo reduction). We could attempt to overcome this shortcoming by specifically requiring that the two vowels be in separate morphemes, as indicated in the following rule.

(32) $V \rightarrow \varnothing / V + \_\_$

Even this restriction is insufficient, since it does not explain why the supposed suffix vowel in /laikni-ür/ and /hirði-ür/ deletes, but in the nom. pl, acc. pl and dat pl., the vowels of the suffix *-ar*, *-a* and *-üm* are not deleted (cf. *laiknar, hirðar, laikna, hirða, laiknüm, hirðüm*): rather, the vowel of the stem deletes. Particularly troublesome for the hypothesis that the nom. sg suffix is /ür/ is the fact that the dative plural suffix *-üm* acts so different. These problems could be remedied by requiring the vowel which deletes to be *ü*, and by deleting *ü* only before *r*.

(33) $ü \rightarrow \varnothing / V + \_\_r$

In lieu of a competing hypothesis, it is difficult to judge the correctness of this rule, but given the very specific information needed in this rule to make it work, you should be suspicious of the rule. The general idea of vowel deletion after a vowel or before a vowel is reasonable, but if you pay attention to which vowel deletes and when it deletes, the vowel deletion approach is not promising.

At this point, we are so thoroughly suspicious of the hypothesis of underlying /ür/ that we have a reason to look for an alternative hypothesis, to see if different assumptions about underlying forms simplify the description. The competing hypothesis that the suffix is /r/ requires a rule to insert *ü* before *r* just in case a consonant precedes.

(34) $\varnothing \rightarrow$ ü / C __ r# (=) $\varnothing \rightarrow \begin{bmatrix} +syl \\ +hi \\ +rd \\ -bk \end{bmatrix} / \text{ C } \_\_\_ \begin{bmatrix} -syl \\ +son \end{bmatrix} \#$

This rule does not apply to /morgün-r/, since we have [morgünn] and not *[morgünür], but that fact does not have to be directly stated in the epenthesis rule. The explanation is straightforward: another rule eliminates underlying /nr#/, giving [nn] instead: rule ordering matters. Given the generality of the epenthesis rule versus the highly specific nature of the *ü*-deletion rule, we reject the *ü*-deletion hypothesis: therefore the underlying form of the nominative singular must be /r/.

Vowel deletion. Now that we understand that the nom. sg. suffix is /r/ and *ü* which appears before it is actually epenthetic, we turn to vowel-plus-vowel sequences. The stems *lakni, hirði* and *mou* end in vowels or diphthongs, and when a vowel initial suffix comes after the stem, a vowel is deleted. Examples are repeated below, this time including in parentheses the underlying vowel which is deleted.

(35)				
	laikn(i)-i	hirð(i)-i	mou-(i)	*dat. sg.*
	laikn(i)-ar	hirð(i)-ar	mou-ar	*nom. pl.*
	laikn(i)-a	hirð(i)-a	mou-a	*acc. pl.*
	laikn(i)-üm	hirð(i)-üm	mou-üm	*dat. pl*
	‘doctor’	‘herd’	‘peat’	

The simple generalization is that the vowel *i* deletes before or after another vowel between morphemes (in an example such as [laikni] from /laikni+i/, we cannot tell which *i* is deleted). Thus we may posit the following rule.

(36) $i \rightarrow \varnothing / \left\{ \begin{matrix} V + \_ \\ \_ + V \end{matrix} \right\}$

This can be written *i* → ⌀ % __+V. The symbol ‘%’ is the mirror-image notation meaning “before or after”.

Sonorant clusters with r. Two other rules are required which affect C+r sequences. One assimilates /r/ to a preceding /l/ or /n/. The question arises, are /l/ and /n/, excluing /r/, a natural class? The consonants /l/ and /n/ have in common the properties of being coronal sonorants, but so does /r/. The consonant /r/ is [-nasal], but so is /l/; /r/ is [-lateral], but so is /n/. Thus, feature theory says that it would be impossible to refer to the class of consonants /l,n/ excluding /r/. But it is not *necessary* to explicitly state the assimilation rule so that it only applies after /l,n/, since /r/ deletes after another /r/ anyhow. In other words, we need the following rule:

(37) r → ∅ / r___

The sonorant-assimilation rule can therefore be stated generally as:

Recall the use of Greek-letter variable for formulating assimilation rules, discussed in chapter 6. This rule states that /r/ takes on the same values for lateral and nasal as found in the preceding consonant.

(38) $$r \rightarrow \begin{bmatrix} \alpha\text{lateral} \\ \beta\text{nasal} \end{bmatrix} / \begin{bmatrix} +\text{sonorant} \\ +\text{coronal} \\ \alpha\text{lateral} \\ \beta\text{nasal} \end{bmatrix} \_\_$$

This rule change /lr/ into [ll], /nr/ into [nn] and vacuously changes /rr/ into [rr]: the independent process of r-deletion will still simplify the resulting sequence of *r*'s.

Syncope. The next problem which we will take on is the vowel ~ ∅ alternation found for example in [himin] ~ [himni]. Not all stems participate in this alternation, so we do find the alternation in the stems *akür, galdür, himin, hamar, morgun, yökül* and *θümal* but not *hest, hatt, heim, garð, gröüt, skoug, mou, stoul, hver, byour, magaul* (we will consider /laikni/ and /hirði/ later). A simple generalization determines which stems alternate: only single vowels outside the initial syllable are subject to the alternation.

Now we must ask under what circumstances the vowel deletes. Taking /himin/ as representative, we can list the contexts:

(39) *CVCVC stem*
himin-n (nom. sg.), himin (acc. sg.), himin-s (gen. sg.)

CVCC stem
himn-i (dat. sg.), himn-ar (nom. pl.), himn-a (acc. pl.), himn-üm (dat. pl)

In other words, there is no vowel before a vowel-initial suffix.

Having isolated the context in which a vowel is deleted, we can offer a phonological rule of vowel syncope.

(40) $$\underset{[-\text{long}]}{V} \rightarrow \varnothing \,/\, VC_0 \_\_\_ CV$$

ü-umlaut. This now leaves us with the problem of the alternation between [a] and [ö]. In looking for a context where this vowel change happens, we note that it takes place before the dative plural suffix *-üm*, which underlyingly has the front round vowel [ü], and thus the rule involves an assimilation in roundness and backness.

(41) $a \rightarrow ö \,/\, \_\_ C_0\, ü$ $\qquad$ $(a \rightarrow \begin{bmatrix} -\text{back} \\ +\text{rd} \end{bmatrix} / \_\_\_ C_0 \begin{bmatrix} -\text{back} \\ +\text{rd} \end{bmatrix})$

It is evident, given examples such as [hattür] from /hatt-r/, that the vowel [ü] inserted by ü-epenthesis does not trigger this rule, which can be explained by ordering the rule of round-harmony (41) before u-insertion (34).

i-deletion. The final fact to be explained is that while the dative singular suffix is *-i*, the dative singular of the stems *akür, hver, byour, galdür* is identical to the stem — the vowel *i* is missing. This can be accounted for by a rule deleting *i* after *r*.

(42) $i \rightarrow \varnothing$ / r __ #

Reconsidering /akür/. We are nearly finished with our analysis of Icelandic phonology, but one area of data needs further consideration. We assumed the underlying representations of 'field' and 'magic' to be /akür/ and /galdür/, based on the fact that that is how they appear phonetically in the accusative singular, and this form has generally been a good diagnostic of the underlying representation. However, there is a problem with assuming underlying /akür/ and /galdür/, that the rule of rounding assimilation (41) would be expected to apply in these forms, giving incorrect **ökür, *ökürs*. This problem can be resolved by modifying our assumption about the underlying form, since we already have a rule which inserts *ü* before *r* — a rule which applies after rounding assimilation (epenthetic *ü* does not trigger rounding). Therefore, we change our assumption about underlying forms, to /akr/ and /galdr/. This entails a small change in the way that we formalize the rule of epenthesis, since that rule as presently stated only inserts *ü* before *r* which is in word final position, and yet we also want to be able to insert *ü* before *r* which stands before another consonant, in order to explain /akr-s/ → [akürs].

(43) $\varnothing \rightarrow$ ü / C __ r {C,#}

Commonly, the expression {C,#} indicates syllable structure: the rule prevents Cr at the end of a syllable.

4. Modern Hebrew

The next case-study comes from a set of alternations in the conjugation of verbs in a certain derivational class in Modern Hebrew.

These data are from a nonstandard dialect that has pharyngeals which were lost in the standard dialect, either being deleted (in the case of ʕ) or changed to *x* (in the case of *ḥ*).

4.1. The data

The goal of this problem is to determine the underlying representations of the verbal prefix and the stems, as well as whatever rules are needed to account for these phonological alternations. In some cases, a related word is provided in order to clarify aspects of the underlying stem. The data to be accounted for are in (44).

(44)

1 sg.	*2 sg. msc*	*3 sg. fem*	*gloss*	*related word*	*gloss*
itparnasti	itparnes	itparnesu	earn		
itparsamti	itparsem	itparsemu	become famous		
idbalbalti	idbalbel	idbalbelu	be confused		
idgalgalti	idgalgel	idgalgelu	revolve		
itħamakti	itħamek	itħamku	turn away		
itlabašti	itlabeš	itlapšu	get dressed		
idbadarti	idbader	idbadru	make fun		
idgarašti	idgareš	idgaršu	divorce		
itpalalti	itpalel	itpalelu	pray		
itxamamti	itxamem	itxamemu	warm		
itmotati	itmotet	itmotetu	quake		
itʔošašti	itʔošeš	itʔošešu	recover		
idbodati	idboded	idbodedu	seclude oneself		
istaparti	istaper	istapru	get a haircut	sapar	barber
istarakti	istarek	istarku	comb hair	ma-srek	comb
ištaparti	ištaper	ištapru	improve	šipur	improvement
itˢtalamti	itˢtalem	itˢtalmu	have photo taken	tˢalem	photographer
izdakanti	izdaken	izdaknu	age	zaken	old
izdarasti	izdarez	izdarzu	hurry	zariz	alert
itamamti	itamem	itamemu	feign innocence	tamim	innocent
idardarti	idarder	idarderu	decline	dirdur	rolling
itpataħti	itpateaħ	itpatħu	develop		
idgalaħti	idgaleaħ	idgalħu	shave		
itnatˢaħti	itnatˢeaħ	itnatˢħu	argue		
ištagati	ištagea	ištagʕu	become mad		
itparati	itparea	itparʕu	cause disorder		
itmaleti	itmale	itmalʔu	become full		
itpaleti	itpale	itpalʔu	become surprised		
itnaseti	itnase	itnasʔu	feel superior		

4.2. Morphological analysis

Each of these verbs verb has a prefix which is either /it/ or /id/, and the prefix transparently surfaces as one of these two variants in most examples. The 1 sg. form is marked with a suffix *-ti*, the 3 sg. fem. has the suffix *-u*, and the 2 sg. masc. has no suffix. The vowel in the second stem syllable is underlyingly the same for all verbs: this fact is not entirely obvious from these data but is made obvious by a more extensive analysis of the morphological structure of words in the language. An analysis of the phonological factors surrounding the second vowel will show that these surface variants can be derived from one particular underlying

vowel. Derivationally related words, such as the root underlying *ištaparti* 'improve' and *šipur* 'improvement', have in common a set of consonants, but their vowels differ (vowel changes are a means of indicating derivational relations in Semitic languages, which we will not be concerned with).

4.3. Phonological alternations

Voicing assimilation. As for the choice between an underlying voiced or voiceless consonant, scanning the data reveals that a voiced consonant appears before voiced obstruents and a voiceless consonant appears before voiceless obstruents and sonorants. Since sonorants are phonetically voiced, it is clear that there is no natural context for deriving the voiceless consonant [t], so we assume that the prefix is underlyingly /it/. Before a voiced obstruent, a voiceless obstruent becomes voiced.

(45) $[\text{-sonorant}] \rightarrow [\text{+voice}] \; / \; \underline{\quad\quad} \; \begin{bmatrix}\text{C} \\ \text{-son} \\ \text{+voi}\end{bmatrix}$

Alternations in V_2. The second vowel of the stem has three phonetic variants: [a] as in *itparn**a**sti*, [e] as in *itparn**e**s*, and ∅ as in *idbad**e**ru* (cf. *idbad**e**r*). Deletion of the second stem vowel only takes place before the suffix *-u*, so we will first attempt to decide when the vowel is deleted. A partial specification of the context for vowel deletion is before C+V, which explains why the 1 sg. and 2 sg. masc. forms (with the suffixes *-ti* and *-∅*) do not undergo vowel deletion. The next step in determining when a vowel is deleted is to sort the examples into two groups: those with vowel deletion, and those with no vowel deletion. In the following examples, the site of vowel deletion (or its lack) is marked with an underscore.

(46) *Vowel deletion*

itħam_ku	itlap_šu	idbad_ru
idgar_šu	istap_ru	istar_ku
ištap_ru	itˢtal_mu	izdak_nu
izdar_zu	itmal_ʔu	itpal_ʔu
itnas_ʔu	itpat_ħu	idgal_ħu
itnatˢ_ħu	ištag_ʕu	itpar_ʕu

No vowel deletion

itparn<u>e</u>su	itpars<u>e</u>mu	idbalb<u>e</u>lu
idgalg<u>e</u>lu	idarderu	itpal<u>e</u>lu
itxam<u>e</u>mu	itmot<u>e</u>tu	itʔoš<u>e</u>šu
idbod<u>e</u>du	itam<u>e</u>mu	

Based on this grouping, we discover a vowel is deleted when it is preceded by just a single consonant; if two consonants precede the vowel, there is no deletion.

However, it is not always the case that a vowel deletes after a single consonant, so our rule cannot simply look for one versus two consonants. There are cases such as *itʔošešu* where there is no vowel deletion, despite the fact that there is only a single consonant before the vowel. Inspecting all of those examples, we discover that the consonants preceding and following the vowel are the same, and in every case where a vowel is deleted, the preceding and following consonants are different. Thus, a vowel deletes only if it is preceded by a single consonant, and that consonant must be different from the consonant that follows the vowel (which is indicated informally as '$C_i...C_j$' in the rule).

(47) $e \rightarrow \varnothing$ / V C_i ___ C_j V

At this point, we now clearly recognize this process as a kind of syncope, a phonological rule which we have encountered many times before.

Closed syllable lowering. Now we turn to the alternation between [a] and [e]. Concentrating on the first set of examples in the dataset, we find [a] before CC (*itparnasti*), and [e] before C# or CV (*itparnes, itparnesu*). Assuming that this distribution is generally valid, we would therefore posit the following rule to derive [a] from /e/.

(48) $e \rightarrow a$ / __ CC

An attempt to derive [e] from underlying /a/ runs into the difficulty that the context "when followed by C# or CV" is not a coherent context, but is just a set of two partially related contexts. This motivates the decision to select underlying /e/.

In four examples, the second stem vowel /e/ appears as [a] before a single consonant, namely the 1 sg. forms *itmotati, idbodati, ištagati* and *itparati*. These examples fall into two distinct subgroups, as shown by looking at their underlying stems, which is revealed in the 3 sg. feminine forms (*itmotet-u, idboded-u* and *ištagʕu, itparʕu*). In the first two examples the stems underlyingly end in a coronal stop *t* or *d*, and in the second two examples the stems underlyingly end in the voiced pharyngeal ʕ. At the underlying level, the second stem vowel is followed by two consonants (/itmotetti/, /itbodedti/, /ištageʕti/ and /itpareʕti/). Surface [a] is explained on the basis of the underlying consonant cluster — it must simply be assured that the rules simplifying these clusters apply after (48).

In the first two examples (*itmotati* and *idbodati* from /itmotat-ti/ and /idbodad-ti/) combination of the 1 sg. suffix with the root would (after assimilation of voicing) be expected to result in **itmotatti* and **idbodatti*. In fact, the data provide no examples of geminate consonants, and where geminates might have been created by vowel syncope in *idbodedu*, syncope is blocked. Thus, the language

seems to be pursuing a strategy of avoiding the creation of geminate consonants. We can account for this simplification of consonant clusters by the following rule.

(49) $\begin{Bmatrix} t \\ d \end{Bmatrix} \rightarrow \varnothing \; / \; \_\_\_ \begin{Bmatrix} t \\ d \end{Bmatrix}$ or $C_iC_i \rightarrow C_i$

This rule also explains *itamem* and *idarder*, where the stem begins with /t/ or /d/. The underlying forms would be /it-tamem/ and /it-darder/: the surface form with a single consonant reflects the application of this consonant-degemination process.

Stems with final pharyngeals and laryngeals. The vowel quality of /šageʕ/ and /pareʕ/ will be left aside temporarily. We thus turn to the stems represented in *it-patahti, idgalahti* and *itnatˢahti*. What is problematic about these stems is the appearance of [ea] when no suffix is added, viz. *itpateah, idgaleah* and *itnatˢeah*. Assuming the underlying forms to be *itpateh, idgaleh* and *itnatˢeh* (selecting /e/ as the second vowel, analogous to *itparnes, itlabeš* and *idboded*) we would need a rule inserting the vowel [a]. These stems have in common that their final consonant is the pharyngeal [ħ], suggesting a rule along the following lines.

(50) $\varnothing \rightarrow$ a / e __ ħ

Why does this rule only apply in the suffixless 2 sg. masc. form? When the stem is followed by *-u* (/itpatehu/ → [itpathu]) the vowel /e/ is deleted by the syncope rule, so there is no vowel before *h*. Syncope does not apply before the suffix *-ti* in /itpatehti/ → [itpatahti] but there is still no epenthetic vowel. The reason is that underlying /e/ changes to [a] by rule (48), before a cluster of consonants. Since that rule changes /e/ to *a* but (50) applies after *e*, prior application of (50) deprives vowel insertion of a chance to apply.

Now returning to the stems *šageʕ* and *pareʕ*, we can see that this same process of vowel insertion applies in these stems in the 2 sg. masc. Starting from /ištageʕ/ and /itpareʕ/, vowel epenthesis obviously applies to give intermediate *ištageaʕ* and *itpareaʕ*. This argues that the epenthesis rule should be generalized so that both of the pharyngeal consonants trigger the process.

(51) $\varnothing \rightarrow$ V / e __ C
[+low] [+low]

The forms derived by (51) are close to the actual forms, which lack the consonant ʕ, and with an appropriate consonant deletion rule we can finish the derivation of these forms. To formalize this rule, we need to determine where the consonant ʕ appears in the language: our data indicates that it appears only before a vowel, never before a consonant or at the end of a word which is to say it never appears at the end of a syllable). Knowing this generalization, we posit the following rule.

(52) ʕ → ∅ / ___ {C,#} (=) ʕ → ∅ / ___ .

No further rules are needed to account for this set of examples. In *ištagati* and *itparati*, from *ištageʕti* and *itpareʕti*, there is no epenthetic vowel. This is predicted by our analysis, since these verbs must undergo the rule lowering /e/ to [a] before CC, and as we have just argued, vowel lowering precedes vowel epenthesis (thus preventing epenthesis from applying). In this respect, *ištagati* and *itparati* are parallel to *itpateah, idgaleah* and *itnatˢeah*. The non-parallelism derives from the fact that syllable-final ʕ is deleted, so predicted *ištagaʕti* and *itparaʕti* are realized as *ištagati* and *itparati* thanks to this deletion.

The final set of verb stems typified by the verb *itmaleti ~ itmale ~ itmalʔu* exbibits a glottal stop in some contexts and ∅ in other contexts. The two most obvious hypotheses regarding underlying form are that the stem is /male/, or else /maleʔ/. It is difficult to decide between these possibilities; we will explore both possibilities. Suppose, first, that these stems end in glottal stop. In that case, we need a rule deleting glottal stop syllable finally — a similar rule was required to delete the consonant ʕ. A crucial difference between stems ending in ʕ and stems presumably ending in ʔ is that the stem vowel /e/ does not lower to [a] before *-ti* in the latter set. Thus, deletion of ʔ would have to be governed by a different rule than deletion of ʕ, since ʔ-deletion precedes lowering and ʕ-deletion follows lowering.

An alternative possibility that we want to consider is that these stems really end in a vowel, not glottal stop. Assuming this, surface [itpaleti] would simply reflect concatenation of the stem /pale/ with the suffix, and no phonological rule would apply. The problem is that we would also need to explain why the rule of syncope does not apply to [itpaleti], since the phonetic context for that rule is found here. The glottal-final hypothesis can explain failure of syncope rather easily, by ordering glottal stop deletion after syncope — when syncope applies, the form is /itpaleʔti/, where the consonant cluster blocks syncope.

Metathesis. The last point regarding the Hebrew data is the position of *t* in the prefix. The consonant of the prefix actually appears after the first consonant of the stem in the following examples.

(53)	istaparti	'get a haircut'	istarakti	'comb hair'
	ištaparti	'improve'	itˢtalamti	'have photo taken'
	izdakanti	'age'	izdarasti	'hurry'

We would have expected forms such as [itsaparti], [itšaparti], [ittˢalamti] by just prefixing *it-* to the stem. A metathesis rule is therefore needed which moves *t* to be after these consonants. What makes this group of consonants — [s,š,tˢ,z] — a natural class is that they are all and only the strident coronals. We can thus formalize this rule as follows: a coronal stop followed by a coronal strident switch order.

(54) $\begin{bmatrix} + \text{cor} \\ - \text{cont} \end{bmatrix} \begin{bmatrix} + \text{cor} \\ + \text{strid} \end{bmatrix} \rightarrow \begin{bmatrix} + \text{cor} \\ + \text{strid} \end{bmatrix} \begin{bmatrix} + \text{cor} \\ - \text{cont} \end{bmatrix}$

The ordering of this metathesis rule with respect to the voicing assimilation rule is crucial. Given underlying /it-zakanti/, you might attempt to apply metathesis first, which would yield *iztakanti*, where voiceless *t* is placed after stem-initial *z*. The voicing assimilation rule (in a general form, applying between all obstruents) might apply to yield **istakanti*. So if metathesis applies before voicing assimilation, we will derive an incorrect result, either **iztakanti* if there is no voicing assimilation (assuming that the rule only turns voiceless consonants into voiced ones) or **istakanti* if there is voicing assimilation. However, we will derive the correct output if we apply voicing assimilation first: /itzakanti/ becomes *idzakanti*, which surfaces as [izdakanti] by metathesis. With this ordering, we have completed our analysis of Modern Hebrew phonology.

5. Japanese

The analysis of phonological alternations found in connection with the conjugation of verbs in Japanese provides our final illustration of the kinds of issues that must be considered in coming up with appropriate rules and underlying representations. In solving this problem, it is particularly important to make the correct assumptions about underlying representations, since the selection of underlying forms goes hand in hand with stating the rules correctly.

5.1. The data

The relevant data are given in (55).

(55)

present	*negative*	*volitional*	*past*	*inchoative*	*gloss*
neru	nenai	netai	neta	neyo:	sleep
miru	minai	mitai	mita	miyo:	see
šinu	šinanai	šinitai	šinda	šino:	die
yomu	yomanai	yomitai	yonda	yomo:	read
yobu	yobanai	yobitai	yonda	yobo:	call
kat^{s}u	katanai	kačitai	katta	kato:	win
kasu	kasanai	kašitai	kašita	kaso:	lend
waku	wakanai	wakitai	waita	wako:	boil
t^{s}ugu	t^{s}uganai	t^{s}ugitai	t^{s}uida	t^{s}ugo:	pour
karu	karanai	karitai	katta	karo:	shear
kau	kawanai	kaitai	katta	kao:	buy

5.2. Morphological analysis

We could make an initial guess regarding suffixes, which leads to the following hypotheses: *-u* = 'present', *-nai* = 'negative', *-tai* = 'volitional', *-ta* = 'past' and *-yo:* = 'inchoative': that analysis seems reasonable given he first two verbs in the data. We might also surmise that the root is whatever the present tense form is without the present ending, i.e. underlying *ner, mir, šin, yom, yob, kat^s, kas, wak, t^{s}ug, kar*, and *ka*. In lieu of the application of a phonological rule, the surface form of a word should simply be whatever we hypothesize the underlying form of the root to be, plus the underlying form of added affixes. Therefore, given our preliminary theory of roots and suffixes in Japanese, we predict the following surface forms, with hyphens inserted between morphemes to make the division of words into roots and suffixes clear: it is important to understand the literal predictions of your analysis, and to compare them with the observed facts.

(56) **Predicted surface forms**

present	*negative*	*volitional*	*past*	*inchoative*
ner-u	ner-nai	ner-tai	ner-ta	ner-yo:
mir-u	mir-nai	mir-tai	mir-ta	mir-yo:
šin-u	šin-nai	šin-tai	šin-ta	šin-yo:
yom-u	yom-nai	yom-tai	yom-ta	yom-yo:
yob-u	yob-nai	yob-tai	yob-ta	yob-yo:
kat^s-u	kat^s-nai	kat^s-tai	kat^s-ta	kat^s-yo:
kas-u	kas-nai	kas-tai	kas-ta	kas-yo:
wak-u	wak-nai	wak-tai	wak-ta	wak-yo:
t^{s}ug-u	t^{s}ug-nai	t^{s}ug-tai	t^{s}ug-ta	t^{s}ug-yo:
kar-u	kar-nai	kar-tai	kar-ta	kar-yo:
ka-u	ka-nai	ka-tai	ka-ta	ka-yo:

The forms which are correct as is are underlined: as we can see, all of the present tense forms are correct, and none of the others are. It is no surprise that the present tense forms are correct, since we decided that the underlying form of the root is whatever we find in the present tense minus the vowel *-u*. It is possible, but unlikely, that every other word undergoes some phonological rule.

Changing our hypothesis. Since our first guess about underlying forms is highly suspect, we should consider alternative hypotheses. Quite often, the cause of analytic problems is incorrect underlying forms. One place to consider revising the assumptions about underlying representations would be ithe underlying forms of the affixes. It was assumed — largely on the basis of the first two forms *nenai* and *minai* — that the negative suffix is underlyingly *-nai*. However, in most of the examples, this apparent suffix is preceded by the vowel *a* (*šinanai, yomanai, yobanai*

and so on), which suggests the alternative possibility that the negative suffix is really *-anai*. Similarly, the decision that the volitional suffix is underlyingly *-tai* was justified based on the fact that it appears as *-tai* in the first two examples; however, the suffix is otherwise always preceded by the vowel *i* (*šinitai, yomitai, yobitai* and so on), so this vowel might analogously be part of the suffix.

A fact strongly suggests that the initial hypothesis about the underlying forms of suffixes was incorrect. The past tense suffix, which we also assumed to be *-ta*, behaves very differently from the volitional suffix, and thus we have *šinitai* versus *šinda, yomitai* versus *yonda, kačitai* versus *katta, karitai* versus *katta* (there *are* similarities such as *kašitai* and *kašita* which must also be accounted for). It is quite unlikely that we can account for these very different phonological patterns by reasonable phonological rules if we assumes that the volitional and past tense suffixes differ solely by the presence of final *i*.

It is this realization, that there is a thorough divergence between the past tense and volitional suffixes in terms of how they act phonologically, that provides the key to identifying the right underlying forms. Given how similar these two suffixes are in surface forms, -(*i*)*tai* vs. -(*i*)*ta*, but how differently they behave phonologically, they must have quite different underlying forms. Since the past tense suffix rarely has a vowel and the volitional suffix usually does, we modify our hypothesis so that the volitional is /-itai/ and the past tense is /ta/. Because the negative acts very muct like the volitional in terms of where it has a vowel, we also adopt the alternative that the negative is /anai/.

These changed assumptions about underlying representations of suffixes yield a significant improvement in the accuracy of our predicted surface forms, as indicated in (57), with correct surface forms underlined.

(57) **Modified predicted surface forms**

present	*negative*	*volitional*	*past*	*inchoative*
ner-u	ner-anai	ner-itai	ner-ta	ner-yo:
mir-u	mir-anai	mir-itai	mir-ta	mir-yo:
šin-u	šin-anai	šin-itai	šin-ta	šin-yo:
yom-u	yom-anai	yom-itai	yom-ta	yom-yo:
yob-u	yob-anai	yob-itai	yob-ta	yob-yo:
katˢ-u	katˢ-anai	katˢ-itai	katˢ-ta	katˢ-yo:
kas-u	kas-anai	kas-itai	kas-ta	kas-yo:
wak-u	wak-anai	wak-itai	wak-ta	wak-yo:
tˢug-u	tˢug-anai	tˢug-itai	tˢug-ta	tˢug-yo:
kar-u	kar-anai	kar-itai	kar-ta	kar-yo:
ka-u	ka-anai	ka-itai	ka-ta	ka-yo:

Implicitly, we know that forms such as predicted *[katˢanai] (for [katanai]) and *[kas-itai] (for [kašitai]) must be explained, either with other changes in underying forms, or by hypothesizing rules.

We will consider one further significant modification of the underlying representations, inspired by the success that resulted from changing our assumptions about *-itai* and *-anai*, in reducing the degree to which underlying and surface forms differ. The original and dubious decision to treat these suffixes as *tai* and *nai* was influenced by the fact that that is how they appear with the first two verbs. It is also possible that our initial hypothesis about the underlying form of these two verb roots was incorrect. There is good reason to believe that those assumptions were indeed also incorrect. Compare the surface form of the three verbs in our dataset which, by hypothesis, have roots ending in *r*.

(58)

present	*negative*	*volitional*	*past*	*inchoative*	*gloss*
ner-u	ne-nai	ne-tai	ne-ta	ne-yo:	sleep
mir-u	mi-nai	mi-tai	mi-ta	mi-yo:	see
kar-u	kar-anai	kar-itai	katt-a	kar-o:	shear

Clearly, the supposed roots /ner/ and /mir/ act quite differently from /kar/. The consonant *r* surfaces in most of the surface forms of the verb 'shear', whereas *r* only appears in 'sleep' and 'see' in the present tense. In other words, there is little reason to believe that the first two roots are really /ner/ and /mir/, rather than /ne/ and /mi/: in contrast, there seems to be a much stronger basis for saying that 'shear' is underlyingly /kar/. Now suppose we change our assumption about these two verbs, and assume that /ne/ and /mi/ end in vowels.

(59) **Modified predicted surface forms**

present	*negative*	*volitional*	*past*	*inchoative*
ne-u	ne-anai	ne-itai	ne-ta	ne-yo:
mi-u	mi-anai	mi-itai	mi-ta	mi-yo:
šin-u	šin-anai	šin-itai	šin-ta	šin-yo:
yom-u	yom-anai	yom-itai	yom-ta	yom-yo:
yob-u	yob-anai	yob-itai	yob-ta	yob-yo:
kat^{s}-u	kat^{s}-anai	kat^{s}-itai	kat^{s}-ta	kat^{s}-yo:
kas-u	kas-anai	kas-itai	kas-ta	kas-yo:
wak-u	wak-anai	wak-itai	wak-ta	wak-yo:
t^{s}ug-u	t^{s}ug-anai	t^{s}ug-itai	t^{s}ug-ta	t^{s}ug-yo:
kar-u	kar-anai	kar-itai	kar-ta	kar-yo:
ka-u	ka-anai	ka-itai	ka-ta	ka-yo:

In terms of being able to predict the surface forms of verbs without phonological rules, this has resulted in a slight improvement of predictive power (sometimes involving a shuffling of correct and incorrect columns, where under the current hypothesis we no longer directly predict the form of the present tense, but we now can generate the past and inchoative forms without requiring any further rules). More important is the fact that we now have a principled basis, in terms of differ-

ent types of underlying forms, for predicting the different behavior of the verbs which have the present tense *neru, miru* versus *karu*, which are in the first two cases actually vowel-final roots, in contrast to a consonant-final root.

5.3. Phonological rules

Since we have made reasonable progress in solving the problem of underlying forms, we will attempt to discover phonological rules which explain remaining differences between underlying and surface forms — though it always remains possible that we will need to change our assumed underlying forms, as our analysis progresses. The approach to take is to look at forms which are still not completely explained, and construct hypotheses to account for these forms: what new rules are needed to get from the underlying to surface forms. One useful way to approach this is to look for columns or rows of data where similar things seem to be happening. The incorrectly predicted forms are re-listed below, this time excluding the forms which are already explained, with information about the nature of the problem added. If a segment is predicted but does not actually surface, that segment is placed in parentheses; if there is a segment which appears in the surface form but which does not appear to be present in the underlying form, the segments is placed in square brackets; segments whose phonetic quality differs from the predicted quality are italicized.

(60)	*present*	*negative*	*volitional*	*past*	*inchoative*	
	ne[r]u	ne(a)nai	ne(i)tai			sleep
	mi[r]u	mi(a)nai	mi(i)tai			see
				šin*t*a	šin(y)o:	die
				yo*mt*a	yom(y)o:	read
				yo*bt*a	yob(y)o:	call
		ka*tˢ*anai		ka*tˢ*ta	ka*tˢ*(y)o:	win
			kas*i*tai	kas[i]ta	kas(y)o:	lend
				wa*k*ta	wak(y)o:	boil
				t*ˢ*ug*t*a	t*ˢ*ug(y)o:	pour
				ka*r*ta	kar(y)o:	shear
		ka(w)anai		ka[t]ta	ka(y)o:	buy

The glide in the inchoative. In order to explain most of the problems which arise with the inchoative form, we will consider the possibility that there is a rule deleting consonants after consonants, since that is the nature of the problem with the inchoative colum. Such a consonant deletion cannot be totally general, i.e. deleting any consonant after any other consonant, since as is evident in the past tense column, the consonant clusters [tt] and [nd] are possible in the language. Nevertheless, these two clusters are a rather restricted subset of the imaginable two-consonant combinations which can be formed from the consonants of the lan-

guage, and this is a good indication that there may be some process deleting a consonant after another consonant. Thus we might assume a rule deleting the glide *y* after a consonant.

(61) $\begin{bmatrix} \text{- cons} \\ \text{- back} \end{bmatrix} \rightarrow \varnothing \,/\, \text{C}\_\_$

If you know Japanese, you may know of words with *y* after a consonant, e.g. [To:kyo:]. We restrict ourselves to the specific dataset given here, but a restriction on the rule that the deleted consonant must be suffix-initial solves this problem

The postulation of any such rule immediately makes a prediction about possible surface forms: there should be no sequences of consonant plus glide in the data. Since there are none in the data at hand, our hypothesis has passed an important test. Armed with this rule, we have accounted for a very large chunk of otherwise problematic examples in (60) — all of the inchoative forms except for *kao:* 'buy', where the glide deletes but there seems to be no consonant which would condition deletion of the glide.

Vowel deletion. Another area where some success is possible in reconciling underlying and surface forms by focusing on possible segment sequences is with the verbs 'sleep' and 'see'. The difference between the predicted (*neanai, mianai; neitai, miitai*) and actual forms (*nenai, minai; netai, mitai*) of the negative and volitional forms is that the actual forms lack the suffix vowel. In the predicted forms, we find a sequence of vowels, whereas in the actual form, only the first of those vowels is found. This raises the question whether we might postulate a rule deleting a vowel after another vowel. In positing such a rule, we want to consider what V-V sequences are found in the data. The sequence [ai] exists in the volitional and negative suffixes, and in past tense *waita*; also [ui] in the past of 'pour'; also the sequences [ao:] and [au] in the verb 'buy'. We do not find sequences of vowels with the front vowels [e] or [i] plus a vowel ([ia], [ii], [ea] and [ei]). Therefore, we posit the following rule of vowel deletion.

(62) $\text{V} \rightarrow \varnothing \,/ \begin{bmatrix} \text{+ syl} \\ \text{- back} \end{bmatrix} \_\_$

This resolves many problematic forms of 'sleep' and 'die', such as the change /ne-itai/ → [netai], but there are still examples that we still cannot explain. In the present tense, we find [neru] and [miru], which we presume derives from /ne-u/ and /mi-u/. The vowel deletion rule (62) should apply to these underlying forms, resulting in incorrect *[ne] and *[mi]. We might try to resolve this by assuming that the vowel [u] cannot be deleted by (62) — we would then need to restricting the rule to exclude round vowels from deletion. If /u/ fails to be deleted in /ne-u/, perhaps a consonant is inserted thereby eliminating the cluster of vowels.

We will consider another possibility later, that the present suffix is /ru/, so rather than inserting it in *neru*, we delete it in [yomu].

(63) $\varnothing \rightarrow r/\begin{bmatrix} +\text{syl} \\ -\text{back} \end{bmatrix}$ __ V

Armed with these new rules, we will have actually accounted for all forms of the verbs 'sleep' and 'see'.

Nasal + consonant. The remaining problems have been reduced to a very small set. A comparison of presumed underlying and surface past forms is given below.

(64)

/šinta/	[šinda]	/yomta/	[yonda]
/yobta/	[yonda]	/katˢta/	[katta]
/kasta/	[kašita]	/wakta/	[waita]
/tˢugta/	[tˢuida]	/karta/	[katta]
/kata/	[katta]		

The problem posed by the past tense form is that by combining the root with the suffix *-ta*, underlying clusters of consonants would be created, but there are very severe restrictions on what consonant clusters exist in Japanese. The simplest problem is that presented by [šinda] from /šinta/, where /t/ becomes voiced after a nasal. A process of post-nasal voicing is rather common in the languages of the world, so we may hypothesize that there is such a process in Japanese.

(65) C → [+voice] / [+nasal] ___

The data further suggest that the rule applies in other examples, since we see that in the past tense [yonda] of the roots /yom/ and /yob/, the final consonant of the root is a nasal on the surface, and /t/ becomes voiced.

We account for the stems /yob/ and /yom/ by noting that the final consonant in these roots becomes [n], which is part of the change from the nonexistent sequences /mt/ and /bt/ to the actually occurring [nd]. Thus, these consonants become [n] before /t/ (and subsequently, /t/ voices after the derived [n]).

(66) $[-\text{coronal}] \rightarrow \begin{bmatrix} +\text{coronal} \\ +\text{nasal} \end{bmatrix} / \_\_\_ \text{C}$

Although the data only illustrate nasalization before /t/, (66) is stated as generally as possible, predicting that /k/ or /d/ would nasalize as well.

Watching for contexts where a phenomenon seems to be relevant to more than one form, we also notice that the surface forms [waita] and [tˢuida] differ from their underlying forms /wakta/ and /tˢugta/ by replacing the preconsonantal velar with the vowel [i], suggesting a vocalization rule such as the following.

(67) C V
[+hi] → [-back] / ___ C

This rule accounts for [waita], and almost accounts for [tˢuida]: but we still need to explain why the suffix consonant is voiced. The underlying representation itself provides a reason for this voicing, since underlyingly, /t/ is preceded by a voiced consonant in /tˢugta/. We know that /t/ voices in another context, after a nasal, so we could account for voicing in [tˢuida] by re-stating the rule so that it applies not just after nasals (which are voiced), but after all voiced consonants. By applying the voicing rule which is sensitive to underlying consonant voicing before the velar-vocalization rule, we can explain the opaque surface difference [waita] versus [tˢuida], as deriving from the voicing of the consonant which precedes it underlyingly. We also want to be sure to apply rule (67) before rule (66), given the way we have formulated these rules. We did not explicitly restrict (66), which changes non-coronals to [n] before a consonant, to applying only to labials. Therefore, the more specific rule (67) must apply first, otherwise velars would also be incorrectly turned into [n] before a consonant.

5.4. Taking stock

We should review the analysis to be sure there are no loose ends. We have six rules — y-deletion, vowel-deletion, r-insertion, consonant voicing, velar vocalization, and labial nasalization — which, given our assumptions regarding roots and suffixes, account for most of the forms in the dataset. It is important to recheck the full dataset against our rules, to be certain that our analysis does handle all of the data. A few forms remain which we cannot fully explain.

The forms which we have not yet explained are the following. First, we have not explained the variation in the root-final consonant seen in the verb 'win' (*katˢ-u, kat-anai, kač-itai, kat-ta, kat-o:*). Second, we have not accounted for the variation between *s* and *š* in the verb 'shear', nor have we explained the presence of the vowel [i] in the past tense of this verb. Finally, in the verb 'buy' we have not explained the presence of [w] in the negative, the appearance of a second [t] in the past tense form, and we have not explained why in the inchoative form [kao:] the suffix consonant *y* deletes.

Correcting the final consonant. The first problem to tackle is the variation in the final consonant of 'win'. Looking at the correlation between the phonetic realization of the consonant and the following segment, we see that [tˢ] appears before [u], [č] appears before [i], and [t] appears elsewhere. It was a mistake to assume that the underlying form of this root contains the consonant /tˢ/; instead, we will assume that the underlying consonant is /t/ (so nothing more needs to be said about the surface forms *kat-anai, kat-ta*, and *kat-o:*). Looking more generally at the dis-

tribution of [č] and [t^s] in the data, [č] only appears before [i], and [t^s] only appears before [u], allowing us to posit the following rules.

(68) $t \rightarrow$ [+delayed release] / __ u

(69) $t \rightarrow \begin{bmatrix} + \text{del. rel} \\ - \text{anterior} \end{bmatrix}$ / __ i

Moving to 'lend', we find a related problem that /s/ appears as [š] before [i]. This is reminiscent of the process which we assumed turning *t* into *č* before *i*. In fact, we can decompose the process *t* → *č* into two more basic steps: /t/ becomes an affricate before [i], and *s* and *t^s* become alveopalatal [š] and [č] before the vowel [i].

i-epenthesis. All that remains to be explained about 'lend' is why [i] appears in the past tense, i.e. why does /kasta/ become *kasita* (whence [kašita])? This is simple: we see that [st] does not exist in the language, and no assimilations turn it into an existing cluster, so [i] is inserted to separate these two consonants.

(70) $\varnothing \rightarrow \begin{bmatrix} + \text{syl} \\ + \text{hi} \\ - \text{back} \end{bmatrix} / \begin{bmatrix} + \text{cont} \\ - \text{son} \end{bmatrix} \_\_ \begin{bmatrix} + \text{cor} \\ - \text{cont} \end{bmatrix}$

r-assimilation and final w. Turning now to the form [katta] 'shear (past)' from /kar-ta/, a simple assimilation is needed to explain this form:

(71) $r \rightarrow C_i$ / __ C_i

The last remaining problems are in the verb 'buy', where we must explain the extra [t] in [katta], the presence of [w] in [kawanai], and the loss of /y/ in the inchoative form [kao:]. We might explain the form [kawanai] by a rule of *w*-insertion inserting *w* between two occurrences of the vowel [a]; more puzzling is the form [katta], which we presume derives from /ka-ta/. It would be very unusual for a consonant to spontaneously double between vowels. Since there are so many problems associated with this one root, perhaps the problem lies in our assumptions about the underlying form of this root. Perhaps the *w* in [kawanai] is part of the root itself. What would be the benefit of assuming that this root is really /kaw/? First, it explains the presence of *w* in [kawanai]. Second, it provides a basis for the extra [t] in [katta]: /w/ assimilates to following [t]. Such an assimilation is implicit in our analysis, namely rule (71) assimilating /r/ to /t/. We can generalize this rule to applying to both /r/ and /w/, which are oral sonorant. Finally, positing underlying /kaw/ helps to resolve the mystery of why /y/ deletes in the inchoative form

[kao:], when otherwise /y/ only deletes when it is preceded by a consonant. If we start with /ka-yo:/ there is no reason for /y/ to delete, but if we start with /kaw-yo:/, /y/ is underlyingly preceded by a consonant /w/, which causes deletion of *y*, and then /w/ itself is deleted.

The cost of this analysis — a small cost — is that we must explain why [w] does not appear more widely in the root, specifically, why do we not find surface [w] in *ka-u*, *ka-itai* and *ka-o:*. The answer lies in the context where [w] appears: [w] only appears before a low vowel, suggesting the following rule.

(72) $w \rightarrow \varnothing / \_ \begin{bmatrix} V \\ -\text{low} \end{bmatrix}$

At this point, we have a complete analysis of the data. The rules (in shorthand versions) and underlying forms are recapitulated below.

(73) *Roots*:
/ne/ ‘sleep’, /mi/ ‘see’, /šin/ ‘die’, /yom/ ‘read’, /yob/ ‘call’, /kat/ ‘win’, /kas/ ‘lend’, /wak/ ‘boil’, /t^{s}ug/ ‘pour’, /kar/ ‘shear’, /kaw/ ‘buy’

Suffixes:
-u ‘present’, *-anai* ‘negative’, *-itai* ‘volitional’, *-ta* ‘past’, *-yo:* ‘inchoative’

Rules:

$y \rightarrow \varnothing$ / C__
$\varnothing \rightarrow r$ / e,i __ V
b,m $\rightarrow$ n / __ t
t $\rightarrow$ t^s / __ u,i
$\varnothing \rightarrow$ i / s __t
w $\rightarrow \varnothing$ /__ V
[-lo]

V $\rightarrow \varnothing$ / e,i ___
[-round]
k,g $\rightarrow$ i / ___ t
t^s,s $\rightarrow$ č,š / __ i
r,w $\rightarrow$ t / __ t
t $\rightarrow$ d / C ___
[+voi]

Progress by hypothesis forming and testing. Three important points have emerged as our analysis developed. First, analysis procedes step-by-step, by forming specific hypotheses which we then check against the data, revising those hypotheses should they prove to be wrong. Second, it is vital to consider more than one hypothesis: if we had only pursued the first hypothesis that the roots /ne/, /mi/, /kar/ and /kaw/ were really underlying /ner/, /mir/, /kar/ and /ka/, we would never have been able to make sense of the data. The most important skill that you can bring to the task of problem-solving is the ability to create and evaluate competing hypotheses intended to explain some fact. Finally, it is particularly important to remember that assumptions about underlying representations go hand-in-hand with the phonological rules which you postulate for a language. When you check your

solution, the problem may not be that your rules are wrong, but that your underlying forms are wrong. By continuously reviewing the analysis, and making sure that the rules works and your assumptions about underlying forms are consistent, you should arrive at the stage that no further improvements to the analysis are possible, given the data available to you.

It might occur to you that there are aspects of the underlying representation which could still be questioned. Consider the present tense form, which we assumed was /u/. An alternative may be considered: the suffix might be /ru/. The presence of underlying /r/ in this suffix is made plausible by the fact that *r* actually appears in the forms *miru*, *neru*. We assumed that *r* is epenthetic, but perhaps it is part of the present suffix. That would allow us to eliminate the rule of *r*-epenthesis which is needed only to account for [neru] and [miru]. At the same time, we can also simplify the rule of vowel deletion, by removing the restriction that only non-round vowels delete after [e] and [i]: we made that assumption only because /ne-u/ and /mi-u/ apparently did not undergo the process of vowel deletion.

Any change in assumptioned underlying forms requires a reconsideration of those parts of the analysis relevant to that morpheme. We would then assume the underlying forms /šin-ru/, /yom-ru/, /kat-ru/ and so on, with the root final consonant being followed by /r/. This /r/ must be deleted: but notice that we already have a rule which, stated in a more general form, would delete this /r/, namely the rule deleting /y/ after a consonant.

(74) [+sonor] → ∅ / C ___

If we generalize that rule to apply to any sonorant consonant after a consonant, we eliminate the rule of *r*-insertion, and generalize the rules *y*-deletion and vowel-deletion, which results in a better analysis.

Summary

Analyzing a complex set of data into a consistent system of underlying representations and rules requires you to pay attention to details. A solution to a problem requires that you formulate reasoned hypotheses and test them against the data. The most important skill needed to test a hypothesis is that you must apply your rules completely literally. Do what the rule says must be done, and if that does not give you the correct result, you must change your underlying representations, rules, or rule ordering. The ability to conceive of and evaluate multiple hypotheses is one of the most important skills in problem solving.

Exercises

1: Serbo-Croatian

These data from Serbo-Croatian have been simplified in two ways, to make the problem more manageable. Vowel length is omitted, and the only accent that is included is the predictable accent. Invariant lexical acent is not marked, and your analysis should explain how accent is assigned in the predictable-accent class, where accent is marked. You cannot write rules which predict accent for those words in the unpredictable accent class, and you cannot (and should not try to) write a rule which somehow predicts *whether* a word receives a predictable accent. Ignore the accent of words with no accent mark (other parts of the phonology of such words must be accounted for). Past tense verbs all have the same general past tense suffix, and that the difference between masculine, feminine and neuter past tense involves the same suffixes as are used to mark gender in adjectives.

Adjectives

Masc.	*Fem.*	*Neut.*	*Pl.*	
mlád	mladá	mladó	mladí	young
bogat	bogata	bogato	bogati	rich
béo	belá	beló	belí	white
veseo	vesela	veselo	veseli	gay
debéo	debelá	debeló	debelí	fat
mío	milá	miló	milí	dear
zelén	zelená	zelenó	zelení	green
kradén	kradená	kradenó	kradení	stolen
dalék	daleká	dalekó	dalekí	far
visók	visoká	visokó	visokí	high
dubók	duboká	dubokó	dubokí	deep
križan	križana	križano	križani	cross
sunčan	sunčana	sunčano	sunčani	sunny
svečan	svečana	svečano	svečani	formal
bogat	bogata	bogato	bogati	rich
rapav	rapava	rapavo	rapavi	rough
yásan	yasná	yasnó	yasní	clear
vážan	važná	važnó	važní	important
sítan	sitná	sitnó	sitní	tiny
ledan	ledna	ledno	ledni	frozen
tának	tanká	tankó	tankí	slim
krátak	kratká	kratkó	kratkí	short
blízak	bliská	bliskó	bliskí	close
úzak	uská	uskó	uskí	narrow
dóbar	dobrá	dobró	dobrí	kind
óštar	oštrá	oštró	oštrí	sharp

bodar	bodra	bodro	bodri	alert
ustao	ustala	ustalo	ustali	tired
múkao	muklá	mukló	muklí	hoarse
óbao	oblá	obló	oblí	plump
pódao	podlá	podló	podlí	base

Verbs

1sg pres	*masc. past*	*fem. past*	*neut. past*	
tepém	tépao	teplá	tepló	wander
skubém	skúbao	skublá	skubló	tear
tresém	trésao	treslá	tresló	shake
vezém	vézao	vezlá	vezló	lead

2: Standard Ukrainian

Standard Ukrainian has palatalized and non-palatalized consonants, but only non-palatalized consonants before *e*. Consonants are generally palatalized before *i*, with some apparent exceptions such as *bil*^{y} 'ache', which need not be seen as exceptions, given the right analysis. Give ordered rules to account for the alternations of the following nouns. The alternation between *o* and *e* is limited to suffixes. Also for masculine nouns referring to persons, *ov/ev* is inserted between the root and the case suffix in the locative singular (see 'son-in-law', 'grandfather'). The data are initially ambiguous as to whether or not the alternations between *o* and *i* and between *e* and *i* are to be implemented by the same rule. Consider both possibilities; give an argument for selecting one of these solutions.

Masculine nouns

Nom. sg.	*Dat. pl.*	*Dat. sg.*	*Loc. sg.*	*Gloss*
zub	zubam	zubov^{y}i	zub^{y}i	tooth
sv^{y}it	sv^{y}itam	sv^{y}itov^{y}i	sv^{y}it^{y}i	light
z^{y}at^{y}	z^{y}at^{y}am	z^{y}atev^{y}i	z^{y}atev^{y}i	son-in-law
koš^{y}il^{y}	košel^{y}am	košelev^{y}i	košel^{y}i	basket
zlod^{y}iy	zlod^{y}iyam	zlod^{y}iyev^{y}i	zlod^{y}iyev^{y}i	thief
m^{y}is^{y}at^{sy}	m^{y}is^{y}at^{sy}am	m^{y}is^{y}at^{s}ev^{y}i	m^{y}is^{y}at^{sy}i	month
korovay	korovayam	korovayev^{y}i	korovayi	round loaf
kam^{y}in^{y}	kamen^{y}am	kamenev^{y}i	kamen^{y}i	stone
m^{y}id^{y}	m^{y}id^{y}am	m^{y}idev^{y}i	m^{y}id^{y}i	copper
xl^{y}iw	xl^{y}ivam	xl^{y}ivov^{y}i	xl^{y}iv^{y}i	stable
holub	holubam	holubov^{y}i	holub^{y}i	dove
s^{y}in	s^{y}inam	s^{y}inov^{y}i	s^{y}in^{y}i	son
leb^{y}id^{y}	lebed^{y}am	lebedev^{y}i	lebed^{y}i	swan
sus^{y}id	sus^{y}idam	sus^{y}idov^{y}i	sus^{y}idov^{y}i	neighbor
čolov^{y}ik	čolov^{y}ikam	čolov^{y}ikov^{y}i	čolov^{y}ik^{y}i	man
l^{y}id	ledam	ledov^{y}i	led^{y}i	ice

bil^{y}	bol^{y}am	bolev^{y}i	bol^{y}i	ache
riw	rovam	rovov^{y}i	rov^{y}i	ditch
stiw	stolam	stolov^{y}i	stol^{y}i	table
d^{y}id	d^{y}idam	d^{y}idov^{y}i	d^{y}idov^{y}i	grandfather
l^{y}it	l^{y}otam	l^{y}otov^{y}i	l^{y}ot^{y}i	flight
mist	mostam	mostov^{y}i	most^{y}i	bridge

Neuter nouns

Nom. sg.	*Gen. sg.*	*Dat. sg.*	*Loc. sg.*	*Gen. pl.*	*Gloss*
t^{y}ilo	t^{y}ila	t^{y}ilu	t^{y}il^{y}i	t^{y}iw	body
koleso	kolesa	kolesu	koles^{y}i	kol^{y}is	wheel
ozero	ozera	ozeru	ozer^{y}i	oz^{y}ir	lake
selo	sela	selu	sel^{y}i	s^{y}iw	village
pole	pol^{y}a	pol^{y}u	pol^{y}i	pil^{y}	field
slovo	slova	slovu	slov^{y}i	sliw	word
more	mor^{y}a	mor^{y}u	mor^{y}i	mir^{y}	sea

3: Somali

In the following Somali data, [ɖ] is a voiced retroflex stop and [ɽ] is a voiced retroflex spirant. Account for all phonological alternations in these data. In your discussion of these forms, be sure to make it clear what you assume the underlying representations of relevant morphemes are. Your discussion should also make it clear what motivates your underlying representations and rules. For instance if you could analyse some alternation by assuming underlying X and rule Y, say why (or whether) that choice is preferable to the alternative of assuming underlying P and rule Q.

Singular	*Sing. Definite*	*Plural*	*Gloss*
daar	daarta	daaro	house
gees	geesta	geeso	side
laf	lafta	lafo	bone
lug	lugta	luɣo	leg
naag	naagta	naaɣo	woman
tib	tibta	tiβo	pestle
sab	sabta	saβo	outcast
bad	bada	baðo	sea
ǰid	ǰida	ǰiðo	person
feeɖ	feeɖa	feeɽo	rib
ʕiir	ʕiirta	ʕiiro	buttermilk
ʔul	ʔuša	ʔulo	stick
bil	biša	bilo	month
meel	meeša	meelo	place
kaliil	kaliiša	kaliilo	summer

nayl	nayša	naylo	female lamb
sun	sunta	sumo	poison
laan	laanta	laamo	branch
sin	sinta	simo	hip
dan	danta	dano	affair
daan	daanta	daano	river bank
saan	saanta	saano	hide
nirig	nirigta	nirgo	baby female camel
gaβaḍ	gaβaḍa	gabḍo	girl
gaǰan	gaǰanta	gaǰmo	arm
hoɣol	hoɣoša	hoglo	downpour
baɣal	baɣaša	baglo	mule
waħar	waħarta	waħaro	female kid
irbad	irbada	irbaðo	needle
kefed	kefeda	kefeðo	pan
ǰilin	ǰilinta	ǰilino	female dwarf
bohol	bohoša	boholo	hole
jirid	jirida	jirdo	trunk
ʔaayad	ʔaayada	ʔaayaðo	miracle
gaʕan	gaʕanta	gaʕmo	hand
ʔinan	ʔinanta	ʔinano	daughter

3sg msc. pst	*3sg fem. past*	*1pl past*	*Gloss*
suɣay	sugtay	sugnay	wait
kaβay	kabtay	kabnay	fix
siðay	siday	sidnay	carry
dilay	dišay	dillay	kill
ganay	gantay	gannay	aim
tumay	tuntay	tunnay	hammer
argay	aragtay	aragnay	see
gudbay	guðubtay	guðubnay	cross a river
qoslay	qosošay	qosollay	laugh
hadlay	haðašay	haðallay	talk

4: Latin

Provide a complete account of the following phonological alternations in Latin, including underlying forms for nouns stems.

Nominative	*Genitive*	*Gloss*
arks	arkis	fortress
duks	dukis	leader
daps	dapis	feast
re:ks	re:gis	king

falanks	falangis	phalanx
filiks	filikis	fern
lapis	lapidis	stone
li:s	li:tis	strife
fraws	frawdis	deceit
noks	noktis	night
frons	frontis	brow
frons	frondis	leaf
inku:s	inku:dis	anvil
sors	sortis	lot
fu:r	fu:ris	thief
murmur	murmuris	murmur
augur	auguris	augur
arbor	arboris	tree
pugil	pugilis	boxer
sal	salis	salt
adeps	adipis	fat
apeks	apikis	top
pri:nkeps	pri:nkipis	chief
ekwes	ekwitis	horseman
miles	militis	soldier
no:men	no:minis	name
karmen	ka:rminis	song
lu:men	lu:minis	light
wenter	wentris	belly
pater	patris	father
kada:wer	kada:weris	corpse
tu:ber	tu:beris	swelling
piper	piperis	pepper
karker	karkeris	prison

The following five nouns and adjectives select a different genitive suffix, *-i:* as opposed to *is*. You cannot predict on phonological grounds what nouns take this suffix, but otherwise these words follow the rules motivated in the language.

li:ber	li:beri:	free
miser	miseri:	wretched
ager	agri:	field
sinister	sinistri:	left
liber	libri:	book

What other phonological rule or rules are needed to account for the following data?

as	assis	whole
os	ossis	bone
far	farris	spell
mel	mellis	honey
o:s	o:ris	mouth
flo:s	flo:ris	flower
mu:s	mu:ris	mouse
cru:s	cru:ris	leg
kinis	kineris	ash
pulvis	pulveris	dust

5: Turkish

Provide a phonological analysis of the following data from Turkish.

nom	*poss*	*dat*	*abl*	*nom. pl*	
oda	odas**ɨ**	odaya	odadan	odalar	room
dere	deresi	dereye	dereden	dereler	river
ütü	ütüsü	ütüye	ütüden	ütüler	iron
balo	balosu	baloya	balodan	balolar	ball
ar**ɨ**	ar**ɨ**s**ɨ**	ar**ɨ**ya	ar**ɨ**dan	ar**ɨ**lar	bee
la:	la:s**ɨ**	la:ya	la:dan	la:lar	la (note)
bina:	bina:s**ɨ**	bina:ya	bina:dan	bina:lar	building
imla:	imla:s**ɨ**	imla:ya	imla:dan	imla:lar	spelling
be:	be:si	be:ye	be:den	be:ler	B (letter)
kep	kepi	kepe	kepten	kepler	cap
at	at**ɨ**	ata	attan	atlar	horse
ek	eki	eke	ekten	ekler	affix
ok	oku	oka	oktan	oklar	arrow
güč	gŭ̈jü	gŭ̈je	güčten	güčler	power
ahmet	ahmedi	ahmede	ahmetten	ahmetler	Ahmed
kurt	kurdu	kurda	kurttan	kurtlar	worm
türk	türkü	türke	türkten	türkler	Turk
genč	genči	genče	genčten	genčler	young
halk	halk**ɨ**	halka	halktan	halklar	folk
üst	üstü	üste	üstten	üstler	upper plane
sarp	sarp**ɨ**	sarpa	sarptan	sarplar	steep
harp	harb**ɨ**	harba	harptan	harplar	war
alt	alt**ɨ**	alta	alttan	altlar	bottom
renk	rengi	renge	renkten	renkler	color
his	hissi	hisse	histen	hisler	feeling
hür	hürrü	hürre	hürden	hürler	free
mahal	mahall**ɨ**	mahalla	mahaldan	mahallar	place
hak	hakk**ɨ**	hakka	haktan	haklar	right

zam	zamm**ɨ**	zamma	zamdan	zamlar	inflation
af	aff**ɨ**	affa	aftan	aflar	excuse
arap	arab**ɨ**	araba	araptan	araplar	Arab
koyun	koyunu	koyuna	koyundan	koyunlar	sheep
pilot	pilotu	pilota	pilottan	pilotlar	pilot
kitap	kitab**ɨ**	kitaba	kitaptan	kitaplar	book
domuz	domuzu	domuza	domuzdan	domuzlar	pig
davul	davulu	davula	davuldan	davullar	drum
bay**ɨ**r	bay**ɨ**r**ɨ**	bay**ɨ**ra	bay**ɨ**rdan	bay**ɨ**rlar	slope
somun	somunu	somuna	somundan	somunlar	loaf
fikir	fikri	fikre	fikirden	fikirler	idea
isim	ismi	isme	isimden	isimler	name
boyun	boynu	boyna	boyundan	boyunlar	neck
čevir	čevri	čevre	čevirden	čevirler	injustice
devir	devri	devre	devirden	devirler	transfer
koyun	koynu	koyna	koyundan	koyunlar	bosom
kar**ɨ**n	karn**ɨ**	karna	kar**ɨ**ndan	kar**ɨ**nlar	thorax
burun	burnu	burna	burundan	burunlar	nose
ak**ɨ**l	akl**ɨ**	akla	ak**ɨ**ldan	ak**ɨ**llar	intelligence
šehir	šehri	šehre	šehirden	šehirler	city
namaz	namaz**ɨ**	namaza	namazdan	namazlar	worship
zaman	zama:n**ɨ**	zama:na	zamandan	zamanlar	time
harap	hara:b**ɨ**	hara:ba	haraptan	haraplar	ruined
i:kaz	i:ka:z**ɨ**	i:ka:za	i:kazdan	i:kazlar	warning
hayat	haya:t**ɨ**	haya:ta	hayattan	hayatlar	life
ispat	ispa:t**ɨ**	ispa:ta	ispattan	ispatlar	proof
inek	inei	inee	inekten	inekler	cow
mant**ɨ**k	mant**ɨɨ**	mant**ɨ**a	mant**ɨ**ktan	mant**ɨ**klar	logic
ayak	aya**ɨ**	ayaa	ayaktan	ayaklar	foot
čabuk	čabuu	čabua	čabuktan	čabuklar	quick
dakik	dakii	dakie	dakikten	dakikler	punctual
merak	mera:k**ɨ**	mera:ka	meraktan	meraklar	curiosity
tebrik	tebri:ki	tebri:ke	tebrikten	tebrikler	greetings
hukuk	huku:ku	huku:ka	hukuktan	hukuklar	law

6: Kera

Propose rules to account for the following alternations. It will prove useful to think about Kera vowels in terms of high versus non-high vowels. Also, in this language it would be convenient to assume that [h] and [ʔ] are specified as [+low]. Pay attention to both verbs like *b**ɨ**lan* 'want me', *balnan* 'wanted me' and *balla* 'you must want!', i.e. there are present, past and imperative forms involved, certain tenses being marked by suffixes. Finally, pay attention to what might look like

a coincidence in the distribution of vowels in the underlying forms of verb roots: there are no coincidences.

haman	‘eat me’	se:nen	‘my brother’
hamam	‘eat you m.’	se:nem	‘your masc. brother’
hɨmi	‘eat you f.’	si:ni	‘your fem. brother’
hɨmu	‘eat him’	si:nu	‘his brother’
hama	‘eat her’	se:na	‘her brother’
hamaŋ	‘eat you pl.’	se:neŋ	‘your pl. brother’
kolon	‘change me’	gi:din	‘my belly’
kolom	‘change you masc.’	gi:dim	‘your masc. belly’
kuli	‘change you fem.’	gi:di	‘your fem. belly’
kulu	‘change him’	gi:du	‘his belly’
kola	‘change her’	gi:dɨ	‘her belly’
koloŋ	‘change you pl.’	gi:diŋ	‘your pl. belly’
cɨ:rɨn	‘my head’	gunun	‘wake me’
cɨ:rɨm	‘your masc. head’	gunum	‘wake you masc.’
ci:ri	‘your fem. head’	guni	‘wake you fem.’
cu:ru	‘his head’	gunu	‘wake him’
cɨ:rɨ	‘her head’	gunɨ	‘wake her’
cɨ:rɨŋ	‘your pl. head’	gunuŋ	‘wake you pl.’
bɨlan	‘want me’	ŋɨfan	‘meet me’
bɨlam	‘want you masc.’	ŋɨfam	‘meet you masc.’
bɨli	‘want you fem.’	ŋɨfi	‘meet you fem.’
bɨlu	‘want him’	ŋɨfu	‘meet him’
bɨla	‘want her’	ŋɨfa	‘meet her’
bɨlaŋ	‘want you pl.’	ŋɨfaŋ	‘meet you pl.’
ʔasan	‘know me’	ʔapan	‘find me’
ʔasam	‘know you masc.’	ʔapam	‘find you masc.’
ʔɨsi	‘know you fem.’	ʔɨpi	‘find you fem.’
ʔɨsu	‘know him’	ʔɨpu	‘find him’
ʔasa	‘know her’	ʔapa	‘find her’
ʔasaŋ	‘know you pl.’	ʔapaŋ	‘find you pl.’
haran	‘give me back’		
haram	‘give you masc. back’		
hɨri	‘give you fem. back’		
hɨru	‘give him back’		
hara	‘give her back’		

haraŋ	‘give you pl. back’

balnan	‘wanted me’	ŋafnan	‘met me’
balnam	‘wanted you masc.’	ŋafnam	‘met you masc.’
bɨlni	‘wanted you fem.’	ŋɨfni	‘met you fem.’
bɨlnu	‘wanted him’	ŋɨfnu	‘met him’
balna	‘wanted her’	ŋafna	‘met her’
balnaŋ	‘wanted you pl.’	ŋafnaŋ	‘met you pl.’
balla	‘you must want!’	ŋafla	‘you must meet!’

ba	‘not’	pa	‘again’	bɨpa	‘no more’

7: Keley-i

Account for the alternations in the following verbs. The different forms relate to whether the action is in the past or future, and which element in the sentence is emphasised (subject, object, instrument). Roots underlyingly have the shape CVC(C)VC, and certain forms such as the subject focus future require changes in the stem that result in a CVCCVC shape. This may be accomplished by reduplicating the initial CV- for stems whose first vowel is [e] (*ʔum-bebhat* ← *behat*) or doubling the middle consonant (*ʔum-buŋŋet* ← *buŋet*). The contrastive identification imperfective form conditions lengthening of the consonant in the middle of the stem, when the first vowel is not [e] (*memayyuʔ* ← *bayuʔ*). These changes are part of the morphology, so do not attempt to write phonological rules to double consonants or reduplicate syllables. Be sure to explicitly state the underlying form of each root and affix. Understanding the status of [s] and [h] in this language is important in solving this problem. It is also important to consider exactly what underlying nasal consonant is present in these various prefixes and infixes — there is evidence in the data which shows that the underlying nature of the nasal explains certain observed differences in phonological behavior.

subject focus future	*direct object focus past*	*instrumental focus past*	
ʔumduntuk	dinuntuk	ʔinduntuk	punch
ʔumbayyuʔ	binayuʔ	ʔimbayuʔ	pound rice
ʔumdillag	dinilag	ʔindilag	light lamp
ʔumgubbat	ginubat	ʔiŋgubat	fight
ʔumhullat	hinulat	ʔinhulat	cover
ʔumbuŋŋet	binuŋet	ʔimbuŋet	scold
ʔumgalgal	ginalgal	ʔiŋgalgal	chew
ʔumʔagtuʔ	ʔinagtuʔ	ʔinʔagtuʔ	carry on head
ʔumʔehneŋ	ʔinehneŋ	ʔinʔehneŋ	stand
ʔumbebhat	binhat	ʔimbehat	cut rattan
ʔumdedʔek	dinʔek	ʔindeʔek	accuse

ʔumtuggun	sinugun	ʔintugun	advise
ʔumtetpen	simpen	ʔintepen	measure
ʔumpeptut	pintut	ʔimpetut	dam
ʔumhehpuŋ	himpuŋ	ʔinhepuŋ	break a stick
ʔumtetkuk	siŋkuk	ʔintekuk	shout
ʔumkekbet	kimbet	ʔiŋkebet	scratch
ʔumbebdad	bindad	ʔimbedad	untie
ʔumdedgeh	diŋgeh	ʔindegeh	sick

instrumental past focus	*contrastive id. imperfective*	*contrastive id. perfective*	
ʔinduntuk	menuntuk	nenuntuk	punch
ʔimbayuʔ	memayyuʔ	nemayuʔ	pound rice
ʔindilag	menillag	nenilag	light lamp
ʔiŋgubat	meŋubbat	neŋubat	fight
ʔinhulat	menullat	nenulat	cover
ʔintanem	menannem	nenanem	plant
ʔimpedug	memdug	nemdug	chase
ʔimbedad	memdad	nemdad	untie
ʔiŋkebet	meŋbet	neŋbet	scratch
ʔimbekaʔ	memkaʔ	nemkaʔ	dig
ʔintepen	mempen	nempen	measure
ʔintebaʔ	membaʔ	nembaʔ	kill a pig
ʔintekuk	meŋkuk	neŋkuk	shout
ʔindegeh	meŋgeh	neŋgeh	sick
ʔinhepaw	mempaw	nempaw	possess
ʔinteled	menled	nenled	sting
ʔindeʔek	menʔek	nenʔek	accuse
ʔinʔebaʔ	meŋbaʔ	neŋbaʔ	carry on back
ʔinʔinum	meŋinnum	neŋinum	drink
ʔinʔagtuʔ	meŋagtuʔ	neŋagtuʔ	carry on head
ʔinʔalaʔ	meŋallaʔ	neŋalaʔ	get
ʔinʔawit	meŋawwit	neŋawit	get

The following past subject clausal focus forms involve a different prefix, using some of the roots found above. A number of roots require reduplication of the first root syllable.

nandunduntuk	‘punch’	nampepedug	‘chase’
naŋkekebet	‘scratch’	nambebekaʔ	‘dig’
nantetekuk	‘shout’	nandedeʔek	‘accuse
nanʔeʔebaʔ	‘carry on back’	nanʔiʔinum	‘drink’
nantanem	‘plant’		

8: Kikuria

In some (but not all) of the examples below, morphemes boundaries have been been introduced to assist in the analysis. Every noun is assigned to a grammatical class conventionally given a number (1-20), which is indicated by a particular prefix on the nouns (e.g. *omo-* for cl. 1); there are also pronoun prefixes on verbs marking subject and object for each class. Tones may be disregarded (however, it is predictable in the infinitive). It is important to pay attention to interaction between processes in this problem.

ogo-táángá	'to begin'	oko-gɛ́sa	'to harvest'
oko-rɔ́ga	'to witch'	oko-réma	'to plow'
oko-hóórá	'to thresh'	ugu-sííká	'to close a door'
ugu-súraangá	'to sing praise'	uku-gíingá	'to shave'
ugu-túúhá	'to be blunt'		

ogo-kó-báră	'to count you (sg)'	uku-gú-súraánga	'to praise you (sg)'
oko-mó-báră	'to count him'	uku-mú-súraánga	'to praise him'
ogo-tó-báră	'to count us'	ugu-tú-súraánga	'to praise us'
oko-gé-báră	'to count them (4)'	uku-gí-súraánga	'to praise it (4)'
oko-ré-báră	'to count it (5)'	uku-rí-súraánga	'to praise it (5)'
uku-bí-báră	'to count it (8)'	uku-bí-súraánga	'to praise it (8)'
uku-chí-báră	'to count it (10)'	ugu-chí-súraánga	'to praise it (10)'

oko-mó-gó-gɛsɛ́ra	'to harvest it (3) for him'
uku-mú-gú-siíkya	'to make him close it (3)'
uku-mú-gú-siíndya	'to make him win it (3)'
oko-bá-súraánga	'to praise them'
oko-mó-bá-suráángéra	'to praise them for him'
oko-bá-mú-suráángéra	'to praise him for them'

to V	*to make to V*	*to V for*	*to make V for*	
okoréma	ukurímyá	okoréméră	ukurímíryá	'weed'
okoróma	ukurúmyá	okoróméră	ukurúmíryá	'bite'
okohóórá	ukuhúúryá	okohóórérá	ukuhúúríryá	'thresh'
okohéétóká	ukuhíítúkyá	okohéétókerá	ukuhíítúkiryá	'remember'
okogéémbá	ukugíimbyá	okogéémbérá	ukugíimbíryá	'make rain'
ogosóóká	ugusúúkyá	ogosóókérá	ugusúúkíryá	'respect'
ogotégétă	ugutígityă	ogotégéterá	ugutígitiryá	'be late'
okorɔ́ga	okorógyá	okorɔ́gɛ́ră	okorógéryá	'bewitch'
okogɔ́ɔ́gá	okogóógyá	okogɔ́ɔ́gɛ́rá	okogóógéryá	'slaughter'
okogɔ́ɔ́tá	okogóótyá	okogɔ́ɔ́tɛ́rá	okogóótéryá	'hold'
ogosɔ́ka	ogosókyá	ogosɔ́kɛ́ră	ogosókéryá	'poke'

ogotɛ́rɛ́kă	ogotɛ́rɛ́kyá	ogotɛ́rɛ́kɛrá	ogotɛ́rɛ́keryá	‘brew’
okogɛ́sa	okogɛ́syá	okogɛ́sɛ́ră	okogɛ́séryá	‘harvest’
ogosɛ́ɛ́nsá	ogosɛ́ɛ́nsyá	ogosɛ́ɛ́nsɛ́rá	ogosɛ́ɛ́nséryá	‘winnow’

to V	*to make to V*	*to V for*	*to make V for*	
ugusííká	ugusííkyá	ogosɛ́ɛ́kɛ́rá	ugusííkíryá	‘to close’
ukurúga	ukurúgyá	okorógɛ́ră	ukurúgíryá	‘to cook’
ugusúka	ugusúkyá	ogosókɛ́ră	ugusúkíryá	‘to plait’
ukurííngá	ukurííngyá	okoréɛ́ngɛ́rá	ukurííngíryá	‘to fold’
ugusííndá	ugusííndyá	ogosɛ́ɛ́ndɛ́rá	ugusííndíryá	‘to win’

imperative	*infinitive*	*they will V*	*then will V for*	
remă	okoréma	mbareréma	mbareréméra	‘cultivate’
bară	okobára	mbarebára	mbarebáréra	‘count’
ată	ogɔɔ́ta	mbarɛɛ́ta	mbarɛɛ́tɛ́ra	‘be split’
ahă	okɔɔ́ha	mbarɛɛ́ha	mbarɛɛ́hɛ́ra	‘pick greens’
agă	okɔɔ́ga	mbarɛɛ́ga	mbarɛɛ́gɛ́ra	‘weed’
aangá	okɔɔ́nga	mbarɛɛ́nga	mbarɛɛ́ngɛ́ra	‘refuse’
andeká	okɔɔ́ndɛ́kă	mbarɛɛ́ndɛ́ka	mbarɛɛ́ndɛ́kɛra	‘write’

imperative	*3s subjunctive*	*3s subjunctive for*	
remă	aremɛ̆	aremerɛ́	‘cultivate’
tɛrɛká	atɛrɛkɛ́	atɛrɛkɛ́rɛ	‘brew’
ebă	ɛɛbɛ̆	ɛɛbɛrɛ́	‘forget’
egă	ɛɛgɛ̆	ɛɛgɛrɛ́	‘learn’
ogă	ɔɔgɛ̆	ɔɔgɛrɛ́	‘be sharp’
ɛyă	ɛɛyɛ̆	ɛɛyɛrɛ́	‘sweep’
ɔrɔká	ɔɔrɔkɛ́	ɔɔrɔkɛ́rɛ	‘come out’

9: Lardil

Account for the phonological alternations seen in the data below.

Bare N	*Accusative*	*Nonfuture*	*Future*	*Gloss*
kentapal	kentapalin	kentapalŋar	kentapaluṛ	dugong
keṯ̪ar	keṯ̪arin	keṯ̪arŋar	keṯ̪aruṛ	river
miyaṛ	miyaṛin	miyaṛŋar	miyaṛuṛ	spear
yupur	yupurin	yupurŋar	yupuruṛ	red rock cod
taŋur	taŋurin	taŋurŋar	taŋuruṛ	crab sp.
yaraman	yaramanin	yaramanar	yaramankuṛ	horse
maaṇ	maaṇin	maaṇar	maaṇkuṛ	spear
pirŋen	pirŋenin	pirŋenar	pirŋenkuṛ	woman
mela	melan	melaŋar	melaṛ	sea
ṯ̪awa	ṯ̪awan	ṯ̪awaŋ	ṯ̪awaṛ	rat

wanka	wankan	wankaŋar	wankaṛ	arm
kuŋka	kuŋkan	kuŋkaŋar	kuŋkaṛ	groin
tarŋka	tarŋkan	tarŋkaŋar	tarŋkaṛ	barracuda
ŋuka	ŋukun	ŋukuŋar	ŋukuṛ	water
ŋuṛa	ŋuṛun	ŋuṛuŋar	ŋuṛuṛ	forehead
kaṭa	kaṭun	kaṭuŋar	kaṭuṛ	child
muṇa	muṇun	muṇuŋar	muṇuṛ	elbow
ŋawa	ŋawun	ŋawuŋar	ŋawuṛ	wife
keṇṭe	keṇṭin	keṇṭiŋar	keṇṭiwuṛ	wife
t^{y}impe	t^{y}impin	t^{y}impiŋar	t^{y}impiwuṛ	tail
ŋiṇe	ŋiṇin	ŋiṇiŋar	ŋiṇiwuṛ	skin
pape	papin	papiŋar	papiwuṛ	father's mother
t^{y}empe	t^{y}empen	t^{y}empeŋar	t^{y}empeṛ	mother's father
wiṭe	wiṭen	wiṭeŋar	wiṭeṛ	interior
waŋal	waŋalkin	waŋalkar	waŋalkuṛ	boomerang
men^{y}el	men^{y}elkin	men^{y}elkar	men^{y}elkuṛ	dogfish sp.
makar	makarkin	makarkar	makarkuṛ	anthill
yalul	yalulun	yaluluŋar	yaluluṛ	flame
mayar	mayaran	mayaraŋar	mayaraṛ	rainbow
ṯalkur	ṯalkuran	ṯalkuraŋar	ṯalkuraṛ	kookaburra
wiwal	wiwalan	wiwalaŋar	wiwalaṛ	bush mango
karikar	karikarin	karikariŋar	karikariwuṛ	butter-fish
yiliyil	yiliyilin	yiliyiliŋar	yiliyiliwuṛ	oyster sp
yukar	yukarpan	yukarpaŋar	yukarpaṛ	husband
pulŋar	pulŋarpa	pulŋarpaŋar	pulŋarpaṛ	huge
wulun	wulunkan	wulunkaŋar	wulunkaṛ	fruit sp.
wuṭal	wuṭalt^{y}in	wuṭalt^{y}iŋar	wuṭalt^{y}iwur	meat
kantukan	kantukantun	kantukantuŋar	kantukantuṛ	red
karwakar	karwakarwan	karwakarwaŋar	karwakarwaṛ	wattle sp.
ṯurara	ṯuraraŋin	ṯuraraŋar	ṯuraraŋkuṛ	shark
ŋalu	ŋalukin	ŋalukar	ŋalukuṛ	story
kurka	kurkaŋin	kurkaŋar	kurkaŋkuṛ	pandja
taŋku	taŋkuŋin	taŋkuŋar	taŋkuŋkuṛ	oyster sp.
kurpuṛu	kurpuṛuŋin	kurpuṛuŋar	kurpuṛuŋkuṛ	lancewood
putu	putukan	putukaŋar	putukaṛ	short
maali	maaliyan	maaliyaŋar	maaliyaṛ	swamp turtle
t^{y}iṇtirpu	t^{y}iṇtirpuwan	t^{y}iṇtirpuwaŋar	t^{y}iṇtirpuwaṛ	willie wagtail
pukat^{y}i	pukat^{y}iyan	pukat^{y}iyaŋar	pukat^{y}iyaṛ	hawk sp.
murkuni	murkuniman	murkunimaŋar	murkunimaṛ	nullah
ŋawuŋa	ŋawuŋawun	ŋawuŋawuŋar	ŋawuŋawuṛ	termite
tipiti	tipitipin	tipitipiŋar	tipitipiwuṛ	rock-cod sp.
ṯapu	ṯaput^{y}in	ṯaput^{y}iŋar	ṯaput^{y}iwuṛ	older brother
muŋkumu	muŋkumuŋkun	muŋkumuŋkuŋar	muŋkumuŋkuṛ	wooden axe

tʸumputʸu tʸumputʸumpun tʸumputʸumpuŋar tʸumputʸumpuŗ dragonfly

10: Sakha (Yakut)

Give a phonological analysis to the following case-marking paradigms of nouns in Sakha.

noun	*plural*	*associative*	*gloss*
aɣa	aɣalar	aɣaliin	father
paarta	paartalar	paartaliin	school desk
tia	tialar	tialiin	forest
kinige	kinigeler	kinigeliin	book
ǰie	ǰieler	ǰieliin	house
iye	iyeler	iyeliin	mother
kini	kiniler	kiniliin	3rd person
bie	bieler	bieliin	mare
oɣo	oɣolor	oɣoluun	child
χopto	χoptolor	χoptoluun	gull
börö	börölör	börölüün	wolf
tial	tiallar	tialliin	wind
ial	iallar	ialliin	neighbor
kuul	kuullar	kuulluun	sack
at	attar	attiin	horse
balik	baliktar	baliktiin	fish
iskaap	iskaaptar	iskaaptiin	cabinet
oɣus	oɣustar	oɣustuun	bull
kus	kustar	kustuun	duck
tünnük	tünnükter	tünnüktüün	window
sep	septer	septiin	tool
et	etter	ettiin	meat
örüs	örüster	örüstüün	river
tiis	tiister	tiistiin	tooth
soroχ	soroχtor	soroχtuun	some person
oχ	oχtor	oχtuun	arrow
oloppos	oloppostor	oloppostuun	chair
ötöχ	ötöχtör	ötöχtüün	abandoned farm
ubay	ubaydar	ubaydiin	elder brother
saray	saraydar	saraydiin	barn
tiy	tiydar	tiydiin	foal
atiir	atiirdar	atiirdiin	stallion
oyuur	oyuurdar	oyuurduun	forest
üčügey	üčügeyder	üčügeydiin	good person
eǰiiy	eǰiiyder	eǰiiydiin	elder sister
tomtor	tomtordor	tomtorduun	knob

moɣotoy	moɣotoydor	moɣotoyduun		chipmunk
kötör	kötördör	kötördüün		bird
bölköy	bölköydör	bölköydüün		islet
χatɨŋ	χatɨŋnar	χatɨŋnɨɨn		birch
aan	aannar	aannɨɨn		door
tiiŋ	tiiŋner	tiiŋniin		squirrel
sordoŋ	sordoŋnor	sordoŋnuun		pike
olom	olomnor	olomnuun		ford
oron	oronnor	oronnuun		bed
bödöŋ	bödöŋnör	bödöŋnüün		strong one

noun	*partitive*	*comparative*	*ablative*	*gloss*
aɣa	aɣata	aɣataaɣar	aɣattan	father
paarta	paartata	paartataaɣar	paartattan	school desk
tɨa	tɨata	tɨataaɣar	tɨattan	forest
kinige	kinigete	kinigeteeɣer	kinigetten	book
ǰie	ǰiete	ǰieteeɣer	ǰietten	house
iye	iyete	iyeteeɣer	iyetten	mother
kini	kinite	kiniteeɣer	kinitten	3rd person
bie	biete	bieteeɣer	bietten	mare
oɣo	oɣoto	oɣotooɣor	oɣotton	child
χopto	χoptoto	χoptotooɣor	χoptotton	gull
börö	börötö	börötööɣör	böröttön	wolf
tɨal	tɨalla	tɨallaaɣar	tɨaltan	wind
ɨal	ɨalla	ɨallaaɣar	ɨaltan	neighbor
kuul	kuulla	kuullaaɣar	kuultan	sack
moχsoɣol	moχsoɣollo	moχsoɣollooɣor	moχsoɣolton	falcon
at	atta	attaaɣar	attan	horse
balɨk	balɨkta	balɨktaaɣar	balɨktan	fish
ɨskaap	ɨskaapta	ɨskaaptaaɣar	ɨskaaptan	cabinet
oɣus	oɣusta	oɣustaaɣar	oɣustan	bull
kus	kusta	kustaaɣar	kustan	duck
tünnük	tünnükte	tünnükteeɣer	tünnükten	window
sep	septe	septeeɣer	septen	tool
et	ette	etteeɣer	etten	meat
örüs	örüste	örüsteeɣer	örüsten	river
tiis	tiiste	tiisteeɣer	tiisten	tooth
soroχ	soroχto	soroχtooɣor	soroχton	some person
ötöχ	ötöχtö	ötöχtööɣör	ötöχtön	abandoned farm
ubay	ubayda	ubaydaaɣar	ubaytan	elder brother
saray	sarayda	saraydaaɣar	saraytan	barn
tɨy	tɨyda	tɨydaaɣar	tɨytan	foal
atɨɨr	atɨɨrda	atɨɨrdaaɣar	atɨɨrtan	stallion

χirur	χirurda	χirurdaaɣar	χirurtan	surgeon
üčügey	üčügeyde	üčügeydeeɣer	üčügeyten	good person
tomtor	tomtordo	tomtordooɣor	tomtorton	knob
moɣotoy	moɣotoydo	moɣotoydooɣor	moɣotoyton	chipmunk
kötör	kötördö	kötördööɣör	kötörtön	bird
suorɣan	suorɣanna	suorɣannaaɣar	suorɣantan	blanket
χatiŋ	χatiŋna	χatiŋnaaɣar	χatiŋtan	birch
aan	aanna	aannaaɣar	aantan	door
tiiŋ	tiiŋne	tiiŋneeɣer	tiiŋten	squirrel
sordoŋ	sordoŋno	sordoŋnooɣor	sordoŋton	pike
olom	olomno	olomnooɣor	olomton	ford
bödöŋ	bödöŋnö	bödöŋnööɣör	bödöŋtön	strong one

noun	*dative*	*accusative*	*gloss*
aɣa	aɣaɣa	aɣani	father
ǰie	ǰieɣe	ǰieni	house
iye	iyeɣe	iyeni	mother
oɣo	oɣoɣo	oɣonu	child
börö	böröɣö	börönü	wolf
tial	tialga	tiali	wind
kuul	kuulga	kuulu	sack
at	akka	ati	horse
balik	balikka	baligi	fish
iskaap	iskaapka	iskaabi	cabinet
oɣus	oɣuska	oɣuhu	bull
kus	kuska	kuhu	duck
sep	sepke	sebi	tool
et	ekke	eti	meat
tiis	tiiske	tiihi	tooth
ot	okko	otu	grass
soroχ	soroχχo	soroɣu	some person
ötöχ	ötöχχö	ötöɣü	abandoned farm
oχ	oχχo	oɣu	arrow
saray	sarayga	sarayi	barn
tiy	tiyga	tiyi	foal
kötör	kötörgö	kötörü	bird

oyuun	oyuuŋŋa	oyuunu	shaman
χatiŋ	χatiŋŋa	χatiŋi	birch
aan	aaŋŋa	aani	door
olom	olomŋo	olomu	ford

noun	*ourN*	*gloss*	*noun*	*our N*	*gloss*
aɣa	aɣabɨt	father	iye	iyebit	mother
uol	uolbut	son	kötör	kötörbüt	bird
kɨlaas	kɨlaaspɨt	classroom	ɨskaap	ɨskaappɨt	cabinet
kuorat	kuorappɨt	town	tiis	tiispit	tooth
ohoχ	ohoχput	stove	tünnük	tünnükpüt	window
aan	aammɨt	door	kapitan	kapitammɨt	capitain
tiiŋ	tiiŋmit	squirrel	oron	orommut	bed
kün	kümmüt	day			

11: Sadžava Ukrainian

Give a phonological analysis of the following data. Assume that all surface occurrences of k^y and g^y in this language are derived by rule. Also assume that stress is located on the proper vowel in the underlying representation: the rules for shifting stress are too complex to be considered here. Nouns in declension II depalatalizes a consonant before the locative suffix, and nouns in declension III depalatalize in the genitive. The variation in the genitive and locative sg. suffix in declension I (*-i* or *-a* versus *-u*) is lexically governed: do not write rules which select between these suffixes. Concentrate on extablishing the correct underlying representations for the noun stem.

Declension I

Nom. sg.	*Gen. sg.*	*Loc. sg.*	*Gloss*
plást	plastá	plas^{y}k$^{y\acute{}}$i	layer
skorúx	skoruxá	skorus^{y}í	mountain ash
ɣ^{y}r$^{y\acute{}}$ix	ɣ^{y}r^{y}ixá	ɣ^{y}r^{y}is$^{y\acute{}}$i	sin
pastúx	pastuxá	pastus^{y}í	herdsman
m^{y}n^{y}úx	m^{y}n^{y}úxa	m^{y}n^{y}ús^{y}i	fish sp.
plúɣ	plúɣa	plúz^{y}i	plow
s^{y}t$^{y\acute{}}$íɣ	stóɣa	stóz^{y}i	stack
sák	sáka	sát^{sy}i	fishnet
bék	bəká	bət$^{sy\acute{}}$i	bull
lést	ləstá	ləs^{y}k$^{y\acute{}}$i	letter
lést	lésta	lés^{y}k^{y}i	leaf
p^{y}l$^{y\acute{}}$ít	plóta	plók^{y}i	wicker fence
s^{y}m^{y}r$^{y\acute{}}$íd	smróda	smróg^{y}i	stench
f$^{y\acute{}}$íst	fostá	fos^{y}k$^{y\acute{}}$i	tail
m$^{y\acute{}}$íst	mósta	mós^{y}k^{y}i	bridge
l$^{y\acute{}}$íd	lǽdu	lədú	ice
d^{y}r$^{y\acute{}}$ít	dróta	drók^{y}i	thick wire
m$^{y\acute{}}$íd	mǽdu	mədú	honey
v$^{y\acute{}}$íl	volá	vol$^{y\acute{}}$í	ox
v$^{y\acute{}}$íz	vóza	vóz^{y}i	cart

sér	séra	sér^{y}i	cottage cheese
s^{y}n$^{y\acute{}}$ip	snopá	snop$^{y\acute{}}$i	sheaf
ɣréb	ɣrəbá	ɣrəb^{y}í	mushroom
læb^{y}id	læbəda	læbəg^{y}i	swan
bær^{y}iɣ	bærəɣa	bærəz^{y}i	shore
pər$^{y\acute{}}$iɣ	pəróɣa	pəróz^{y}i	dumpling
por^{y}íɣ	poróɣa	poróz^{y}i	threshhold
bol^{y}ék	bol^{y}əká	bol^{y}ət^{sy}í	abcess
vór^{y}iɣ	vóroɣa	vóroz^{y}i	enemy
kónək	kónəka	kónət^{sy}i	grasshopper
pót^{y}ik	potóka	potót^{sy}i	stream
t$^{y\acute{}}$ík	tóka	tót^{sy}i	current
k$^{y\acute{}}$íl	kolá	kol^{y}í	stake

Declension II

Nom. sg.	*Gen. sg.*	*Loc. sg.*	*Gloss*
kovál^{y}	koval^{y}é	kovalé	blacksmith
ǰm$^{y\acute{}}$íl^{y}	ǰm^{y}il^{y}é	ǰm^{y}ilé	bumblebee
k^{y}r^{y}íl^{y}	k^{y}r^{y}il^{y}é	k^{y}r^{y}ilé	rabbit
učétəl^{y}	učétəl^{y}ə	učétələ	teacher
græb^{y}in^{y}	græbən^{y}ə	græbənə	comb
ólən^{y}	ólən^{y}ə	ólənə	deer
yač^{y}m$^{y\acute{}}$ín^{y}	yačmæn^{y}ə	yačmænə	barley
yás^{y}in^{y}	yásən^{y}ə	yásənə	ash tree
z^{y}ék^{y}	z^{y}ék^{y}ə	z^{y}étə	son-in-law

Declension III

Nom. sg.	*Gen. sg.*	*Gloss*
más^{y}k^{y}	mástə	fat
s^{y}m$^{y\acute{}}$ír^{y}k^{y}	smærtə	death
v^{y}ís^{y}k^{y}	v^{y}ístə	news
rág^{y}is^{y}k^{y}	rádostə	joy
s^{y}íl^{y}	sólə	salt
póš^{y}is^{y}k^{y}	póšəstə	epidemic
zám^{y}ik^{y}	zámətə	snowstorm
skátər^{y}k^{y}	skátərtə	tablecloth
k^{y}ís^{y}k^{y}	kóstə	bone

12: Koromfe

Koromfe has two kinds of vowels, [-ATR] ɪʊɛɔa and [+ATR] iueoʌ. Provide an analysis of the alternations in the following data, which involve singular and plural forms of nouns and different tense-inflections for verbs.

Singular	*Plural*	*gloss*
gɪbrɛ	gɪba	hatchet
hubre	hubʌ	ditch
nɛbrɛ	nɛba	pea
diŋgre	diŋgʌ	bush type
zoŋgre	zoŋgʌ	wing
lɔ̃ŋgrɛ	lɔ̃ŋga	shoe
hullre	hullʌ	gutter
sɛkrɛ	sɛka	half
tɛfrɛ	tɛfa	cotton fiber
dabɛɛrɛ	dabɛɛya	camp
dɔɔrɛ	dɔɔya	long
gɪgaarɛ	gɪgaaya	vulture
pʊpaarɛ	pʊpaaya	grass type
koire	koyʌ	bracelet
dʊmdɛ	dʊma	lion
hulomde	hulomʌ	marrow
tɛmdɛ	tɛma	beard
logomde	logomʌ	camel
bɪndɛ	bɪna	heart
hɔ̃ndɛ	hɔ̃na	hoe
honde	honʌ	bean
geŋde	geŋʌ	pebble
zɛŋdɛ	zɛŋa	upper arm
bɛllɛ	bɛla	back
yɪllɛ	yɪla	horn
selle	selʌ	space
pallɛ	pala	stretcher
deŋgele	deŋgelʌ	open area
sembele	sembelʌ	piece
dãĩnɛ	dãỹã	wood
hʊ̃ĩnɛ	hʊ̃ỹã	caterpillar
kɔ̃ĩnɛ	kɔ̃ỹã	squirrel
kɔ̃ɔ̃nɛ	kɔ̃ɔ̃ỹã	old
sɔ̃ɔ̃nɛ	sɔ̃ɔ̃ỹã	period
bɛtɛ	bɛra	male animal
datɛ	dara	chest
gete	gerʌ	forked stick
gote	gorʌ	stream
bɪtɛ	bɪra	frog
dɔtɛ	dɔra	cloud

neutral	*past*	*progressive*	*gloss*
ta	taɛ	taraa	shoot
gɔ	gɔɛ	gɔraa	go back
kʊ	kɔɛ	kʊraa	kill
tu	toe	turʌʌ	coat
li	lee	lirʌʌ	forget
dɪ	dɛ	dɪraa	eat
tã	tãɛ̃	tãnaa	contradict
nɛ̃	nɛ̃	nɛ̃naa	defecate
saɪ	sayɛ	saɪraa	separate
yɛɪ	yɛyɛ	yɛɪraa	waste
sɔɪ	sɔyɛ	sɔɪraa	split
ỹɛ̃ĩ	ỹɛ̃ỹɛ̃	ỹɛ̃ĩnaa	catch
dɔ̃ĩ	dɔ̃ỹɛ̃	dɔ̃ĩnaa	dream
kɛndɪ	kɛndɛ	kɛndraa	finish
kɛ̃sɪ	kɛ̃sɛ	kɛ̃sraa	surpass
kɛtɪ	kɛtɛ	kɛtraa	open
tɛŋgɪ	tɛŋgɛ	tɛŋgraa	accompany
yisi	yise	yisrʌʌ	suffice
yɪsɪ	yɪsɛ	yɪsraa	draw water
birgi	birge	birgrʌʌ	blacken
pasgɪ	pasgɛ	pasgraa	split
mɛntɪ	mɛntɛ	mɛntraa	assemble
gondu	gonde	gondrʌʌ	depart
hɔ̃ŋgʊ	hɔ̃ŋgɛ	hɔ̃ŋgraa	point
sʊrgʊ	sʊrgɛ	sʊrgraa	drop
hɔ̃kʊ	hɔ̃kɛ	hɔ̃kraa	scratch
zullu	zulle	zullrʌʌ	bow
sɪbʊ	sɪbɛ	sɪbraa	die
zambʊ	zambɛ	zambraa	deceive
wufu	wufe	wufrʌʌ	borrow
zɪgamsʊ	zɪgamsɛ	zɪgamsraa	be dirty
hɛ̃msʊ	hɛ̃msɛ	hɛ̃msraa	meet
leli	lele	lellʌʌ	sing
pɪlɪ	pɪlɛ	pɪllaa	trample flat
tarɪ	tarɛ	tataa	plaster
fɛrɪ	fɛrɛ	fɛtaa	cultivate
tʊrʊ	tʊrɛ	tʊtaa	introduce

Further reading

Kenstowicz & Kisseberth 1979, Zwicky 1973, 1974, 1975, Pullum 1976