This model implies that the output of one component forms the input to the next component, so the phonological component starts with whatever the morphological component gives it, and applies its own rules (which are then subject to principles of physical interpretation in the phonetic component). The output of the morphological component, which is the input to the phonology, is by definition the underlying form, so we need to know a little bit about what the morphological component does, to understand what is presented to the phonology.

The function of the morphological component is to assemble words, in the sense of stating how roots and affixes combine to form a particular word. Thus the morphological component is responsible for combining a noun root [dag] and a plural affix [z] in English to give the word *dogs* (i.e. /dag-z/), or in Russian the morphology combines a noun root [vagon] with an inflectional ending [a] according to rules of inflection for Russian, to give the genitive word *vagon-a*. Each morpheme is assumed to have a single constant phonetically defined shape coming out of the morphology (there are a few exceptions such as the fact that the third-person-singular form of the verb *be* in English is *[sz]* and the first-person-singular form of that verb is *[æm]*). The phonetic realization of any morpheme is subject to rules of phonology, so while the morphology provides the plural morpheme /z/ (spelled -*a*), the application of phonological rules will make that that morpheme being pronounced as *[s]* as in *cats* or *[z]* as in *bushes*.

It is very important to understand that the grammar does not formally derive one word from another. (Some languages seem to have special morphological processes, which we will not be discussing here, that derive one word from another – clipping such as *Sally* → *Sal* would be an example.) Rather, one word derives from a given abstract root plus whatever affixes are relevant, and a related word derives by adding a different set of affixes to the same abstract root. Accordingly, the plural of a noun in English does not derive from the singular, rather, both the singular and plural forms derive from a common root: no suffix is added to the root in the singular, and the suffix /z/ is added to the root in the plural. The Russian genitive [vagona] also does not derive from the nominative, nor does the nominative derive from the genitive. Rather, both derive from the root /vagon/, where the nominative adds no affix and the genitive adds the affix /a/.

The underlying form of a word is whatever comes out of the morphology and is fed into the phonology, before any phonological rules have applied. The underlying form of the word *[kæts]* is *[kæt-z]*, since that is what results in the morphology by applying the rule that combines a noun root such as *cat* with the plural suffix. The underlying form of the plural word *[kæts]* is not *[kæt]*, because the plural word has to have the plural morpheme. However, *[kæt]* is the underlying form of the singular word *[kæt]*. There is no phonological rule which inserts /z/ or /s/ in order to form a plural. The principles for combining roots and affixes are not part of the phonology, and thus there is no need to include rules such as *insert [z]* in the plural. Be explicit about what you assume about morphology in a language, i.e. that there is a plural suffix /z/ in English or a genitive suffix /-a/ in Russian. As for the mechanics of phonological analysis, you should assume, for example, that the plural suffix is already present in the underlying form, and therefore do not write a rule to insert the plural suffix since that rule is part of morphology. A phonological analysis states the underlying forms of morphemes, and describes changes in the phonological shape of the root or suffix.

We have concluded that the underlying form of the Russian word *[prut]* `pond' is *[prud]*. In arriving at that conclusion, we saw how important it is to distinguish the phonological concept of an underlying form from the morphological concept "basic form," where the singular form, or an uninflected nominative form would be the morphological "basic form." An underlying form is a strictly phonological concept and is not necessarily equivalent to an actually pronounced word (even disregarding the fundamental fact that underlying forms are discrete symbolic representations whereas actually pronounced words are acoustic waveforms). It is a representation that is the foundation for explaining the variety of actual pronunciations found in the morpheme, as determined by phonological context.

The morphologically basic form of the Russian word for pond is the unmarked nominative, *[prut]*, composed of just the root with no inflectional ending. In contrast, the phonological underlying form is *[prud]*, for as we have seen, if we assume the underlying form to be *[prut]*, we cannot predict the genitive *[pruda]*. The word *[prud]*, with a voiced consonant at the end of the word, does not appear as such in the language, and thus the supposition that the underlying form is *[prud]* is an abstraction, given that *[prud]* by itself is never found in the language – it must be inferred, in order to explain the actual data. The basis for that inference is the genitive form *[pruda]*, which actually contains the hypothesized underlying form as a subpart. It is important to understand, however, that the underlying form of a root may not actually be directly attested in this way in any single word, and we will discuss this point in section 4.6.

### 4.3 Finding the underlying form

A similar problem arises in explaining the partitive and nominative forms of nouns in Finnish. The first step in understanding the phonological alternation seen here is to do a standard preliminary morphological analysis of the data, which involves identifying which parts of a word correlate with each aspect of word structure (such as root meaning or grammatical case). The following examples illustrate that the nominative singular suffix is /Ø/ (i.e. there is no overt suffix in the nominative singular) and the partitive singular suffix is /-ae/, which alternates with /-ei/ if there is a back vowel somewhere before it in the word (we will not be concerned with that vowel alternation in the partitive suffix).

<table>
<thead>
<tr>
<th></th>
<th>Nominative sg</th>
<th>Partitive sg</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td><em>aamua</em></td>
<td><em>hopea</em></td>
</tr>
<tr>
<td>hopea</td>
<td><em>morning</em></td>
<td><em>silver</em></td>
</tr>
</tbody>
</table>
We might assume that the underlying form of the root is the same as the nominative (which has no suffix). The problem which these data pose is that in some nouns, the partitive appears to be simply the nominative plus the suffix -a (for example, muuri ~ muuria), but for other nouns the final vowel alternates, with [i] in the nominative and [e] in the partitive (e.g., yoki ~ yokea). It is obvious that the nature of the following vowel does not explain this alternation, since the same surface-quality suffix appears before both [a] and [u], versus muuria, kirehhtia where [i] appears before these same vowels. Nor can the preceding consonant be called upon to predict what vowel will appear in the partitive, as shown by pairs such as tukki, kulkia versus lehtia, aitiia.

This is an area where there is (potentially) a difference between language-learning pedagogy and a formal linguistic analysis. Faced with the problem of learning the inflectional distinction muuri ~ muuria versus yoki ~ yokea, a second-language class on Finnish might simply have the student memorize a list of words like yoki ~ yokea where the vowel changes in the inflectional paradigm. From the point of view of linguistic analysis this is the wrong way to look at the question, since it implies that this is not a rule-governed property of the language. However, second-language learning is not the same as linguistic analysis: a class in foreign-language instruction has a different goal from a class in analysis, and some students in a language class may receive greater practical benefit from just memorizing a list of words. Thus it is important to distinguish the teaching method where one learns arbitrary lists, and a theoretically based analysis. One simply cannot predict what vowel will appear in the partitive if one only considers the pronunciation of the nominative. This means: nominative forms are not the same as underlying forms (something that we also know given the previous Russian example). The underlying representation must in some way contain that information which determines whether there will be a vowel alternation in a given word.

In looking for the phonological basis for this vowel alternation, it is important to realize that the alternation in stem-final vowels is not chaotic, for we find precisely two possibilities, either i in the nominative paired with i in the partitive, or i in the nominative paired with e in the partitive — never, for example, i paired with o or e paired with o. Moreover, only the vowel i enters into such a vowel alternation in Finnish; so there are no nouns with o in the nominative which is replaced by u in the partitive, nor is there in the nominative ever replaced by a or any other vowel in the partitive. One final fact about the data in (6) suggests exactly how the right underlying representations can explain this alternation: of the eight vowels of Finnish (i, u, e, o, a, u, o, a), all of them appear at the end of the word except the vowel e. Now, since the stem of the word for 'name,' which appears as nimi in the nominative, actually appears on the surface as nime in the partitive, it is not at all unreasonable to assume that the underlying form of the stem is in fact [nime]. It would be a bit bizarre to assume an underlying form such as [nimi], since the vowel [i] never appears in that position in any form of this word: the most natural assumption to make is that the underlying form of a morpheme is actually composed of segments found in some surface manifestation of the morpheme. On the other hand, the stem of the word for 'wall' is pronounced nuuri in both the nominative and the partitive, and therefore there is no reason to assume that it is underlyingly anything other than [nuuri].

We will then assume that the underlying vowel at the end of the stem is actually reflected by the partitive form, and thus we would assume underlying representations such as [yokej], [nimej], [kivej], [lehtej], [jokej] and so on, as well as [muuri], [naapuri], [kaappi], [tukki] and so on. The underlying form of partitive [yokea] would thus be [yokea], that is, no rule at all is required to explain the partitive. Instead, a rule is needed to explain the surface form of the nominative: word-final e is raised to i.

7 Final vowel raising

\[ e \rightarrow i \]
This rule is neutralizing since the distinction between /j/ and /e/ is neutralized by applying this rule: an underlying /e/ becomes phonetic [i].

Apart from illustrating how important correct underlying forms are, these two examples have also shown that it is dangerous, and incorrect in these two cases, to assume that the “most basic” form of a word according to morphological criteria is also the underlying form of the word. To reiterate: the underlying form of a morpheme is a hypothesis set forth by the analyst, a claim that by assuming such-and-such an underlying form, plus some simple set of rules (which need to be discovered by the analyst), the observed variation in the shape of morphemes can be explained.

Kerewe. To better understand the reasoning that leads to correct underlying forms, we investigate other examples. Consider the following data from Kerewe.

<table>
<thead>
<tr>
<th>Infinitive</th>
<th>1sg habitual</th>
<th>3sg habitual</th>
<th>Imperative</th>
</tr>
</thead>
<tbody>
<tr>
<td>kupaamba</td>
<td>mpaamba</td>
<td>apaamba</td>
<td>piamba</td>
</tr>
<tr>
<td>kupaanga</td>
<td>mpaanga</td>
<td>apaanga</td>
<td>paanga</td>
</tr>
<tr>
<td>kupiima</td>
<td>mpima</td>
<td>apima</td>
<td>pima</td>
</tr>
<tr>
<td>kupupa</td>
<td>mpupa</td>
<td>apupa</td>
<td>puupa</td>
</tr>
<tr>
<td>kupiuka</td>
<td>mpeuka</td>
<td>apeuka</td>
<td>peuka</td>
</tr>
<tr>
<td>kupiinda</td>
<td>mpiinda</td>
<td>apienda</td>
<td>piinda</td>
</tr>
<tr>
<td>kuhiga</td>
<td>mpiiga</td>
<td>ahiga</td>
<td>hiiga</td>
</tr>
<tr>
<td>kuleeka</td>
<td>mpeeka</td>
<td>aheeka</td>
<td>heeka</td>
</tr>
<tr>
<td>kuhanga</td>
<td>mpaanga</td>
<td>apaanga</td>
<td>haanga</td>
</tr>
<tr>
<td>kueebeba</td>
<td>mpeeba</td>
<td>aheeba</td>
<td>heeba</td>
</tr>
<tr>
<td>kukiima</td>
<td>mpiima</td>
<td>ahima</td>
<td>hiima</td>
</tr>
<tr>
<td>kuhuuaa</td>
<td>mpuuaa</td>
<td>ahuuaa</td>
<td>huuaa</td>
</tr>
</tbody>
</table>

We notice that every infinitive begins with ku-, which we surmise is the prefix for the infinitive; the third-singular habitual form has the prefix a-, and the first-singular habitual has the prefix m-; the imperative involves no prefix. In addition to segmental prefixes, there is a change in the first consonant of the stem in some verbs, in some contexts. The initial consonant of the verb meaning ‘guide’ alternates between [h] and [p], with [p] appearing in the first-singular habitual after [m] and [h] appearing elsewhere. Since this stem appears in two surface variants, [heeba] and [peeba], two plausible hypotheses are immediately possible: the stem is underlyingly [peeba], or the stem is underlyingly [heeba]. If we assume that the stem is underlyingly [heeba], we require a rule to explain the divergence between the predicted form of the first-singular habitual form - we would expect [mheeba], [mhima], etc. - and the actual form of the verb, [mpeeba], [mpima] and so on. Since in fact we do not see the sequence /mh/ anywhere in the data, we might assume the following neutralizing rule.

\[
\text{H} \rightarrow \text{p} / \text{nasal}
\]

If, on the other hand, we assume that the root is underlyingly [peeba], we would need a rule which changes /p/ into [h] when not preceded by a nasal - in other words, when preceded by a vowel or by nothing. There is no single property which groups together word-initial positions and vowels. Thus, the supposed rule changing /p/ to [h] would have to be a disjunction of two separate environments.

\[
(10) \ p \rightarrow h \left\{ \begin{array}{c} # \\ v \end{array} \right. \]

This suggests that rule (10) is wrong.

More important than the greater complexity of the rule entailed by assuming that the word for ‘guide’ is underlyingly [peeba], it is empirically wrong: rule (10) implicitly claims that /p/ should always become [h] word initially or after a vowel, but this is falsified by forms such as kupuamba, apuamba, puamba ‘adorn’ and kupaanga, apaanga, paanga ‘line up.’ If we assume the stems uniformly begin with /p/, then we cannot predict whether the imperative or infinitive has [h] (kuhaanga) or [p] (kupaanga).

On the other hand, if we assume an underlying contrast between initial /h/ and initial /p/ - i.e. haanga ‘create’, paanga ‘arrange’ - then we can correctly distinguish those stems which begin with /h/ from those which begin with /p/ when no nasal precedes, as well as correctly neutralizing that distinction just in case the stem is preceded by a nasal (mpeanga ‘I create’; ‘I arrange’).

English plurals. A further illustration of how to determine the correct underlying representation comes from English. As the following examples illustrate, the surface form of the plural suffix varies between [s] and [z] (as well as [es], to be discussed later).

<table>
<thead>
<tr>
<th>Infinitive</th>
<th>1sg habitual</th>
<th>3sg habitual</th>
<th>Imperative</th>
</tr>
</thead>
<tbody>
<tr>
<td>kaaps</td>
<td>caps</td>
<td>kebzs</td>
<td>cabs</td>
</tr>
<tr>
<td>kaets</td>
<td>caps</td>
<td>kedzs</td>
<td>cads</td>
</tr>
<tr>
<td>kaks</td>
<td>cockzs</td>
<td>kags</td>
<td>karzs</td>
</tr>
<tr>
<td>pruufs</td>
<td>proofs</td>
<td>hovzs</td>
<td>hooves</td>
</tr>
<tr>
<td>hlys</td>
<td>fleas</td>
<td>plyez</td>
<td>plows</td>
</tr>
<tr>
<td>tlayz</td>
<td>giveaways</td>
<td>pyrez</td>
<td>pureez</td>
</tr>
</tbody>
</table>

The generalization regarding distribution is straightforward: [s] appears after a voiceless segment, and [z] appears after a voiced one (be it an obstruent, a liquid, nasal or a vowel).

This same alternation can be found in the suffix marking the third singular present-tense form of verbs.
If we suppose that the underlying form of the affixes for noun plural and third singular present verbs are /z/, then we would assume the following rule to derive the phonetic variant [s].

(13) obstruent → voiceless / voiceless _

On the other hand, if we were to assume that these suffixes are underlyingly /s/, we would assume the following rule.

(14) obstruent → voiced / voiced _

In terms of the simplicity and generality of these two rules, the analyses are comparable. Both formulations require the same number of phonetic specifications to state the rule, and both formulations apply to general and phonetically natural classes. However, the two analyses differ quite significantly in terms of their overall predictions for English. The implicit prediction of the first rule (13) is that there should be no voiceless obstruents after voiceless segments in English, since that rule would devoice all such obstruents. This generalization seems to be correct: there are no words like [yakd], [ptfz], [sdap]. The implicit prediction of the second rule (14) is different: that rule implies that there should be no voiceless segments after any voiced segments. This is manifestly incorrect, as shown by the existence of words such as [hus] hiss, [p enf] path, [d ens] dance, [f als] false. We prefer a hypothesis which makes the correct prediction about the phonetic structure of the language as a whole, and thus we select the underlying form /z/ and a rule devoicing obstruents after voiceless segments. Looking for such asymmetries plays an important role in determining which of two hypotheses is the correct one.

The alternation z – s is not limited to the two affixes -z ‘plural’ and -z ‘3sg present tense.’ The rule of devoicing can also be seen applying to the possessive suffix -z.

(15) Noun Noun + poss.
ket kets cat
slug slags slug
klam klamz clam
snow snowz snow

Moreover, certain auxiliary verbs such as has [haz] and is [iz] undergo a reduction in casual speech, so that they appear simply as [z] or [z], the choice between these two being determined by the devoicing rule which we have motivated.

(16) Noun + has Reduced Noun + is Reduced
Jack haz iy?η Jack is iy?η Jack haz iy?η Jack
part haz iy?η part is iy?η part is iy?η Pat

The devoicing rule (13) automatically explains the alternation in the surface shape of the consonant here as well.

Jita tone. It is important to look for correlations which may lead to causal explanations, in analyzing data. Consider the following data from Jita, concentrating on the tones of morphemes (H or high tone is marked with acute accent, L or low-toned syllables are unmarked).

(17) a. okujumana 'to hit' okusija 'to block'
okujumira 'to hit for' okusijira 'to block for'
okujumana 'to hit e.o.' okusijana 'to block e.o.'
okojumirana 'to hit for e.o.' okusijirana 'to block for e.o.'
b. okušuma 'to bite' okušija 'to fold'
okušumira 'to bite for' okušijira 'to fold for'
okušumana 'to bite e.o.' okušijana 'to fold e.o.'
okuširana 'to bite for e.o.' okušijirana 'to fold for e.o.'

We can conclude that there is a prefix oku- perhaps marking the infinitive, a suffix -a appearing at the end of every verb, and two suffixes -ir- ‘for’ and -on- ‘each other.’ There are also root morphemes: šum- ‘hit,’ sif- ‘block,’ as well as šum- ‘bite’ and –šif- ‘fold.’ We decide that ‘bite’ and ‘fold’ under­wingly have H tones in part based on the fact that there actually is an H tone on the vowels of these roots in the simplest verb forms.

In addition, we observe that the suffixes -ir- and -on- have H tone when they come immediately after these verb roots. The suffixes do not have H tone after the first set of roots: appearance of H on the suffix is correlated with the preceding root morpheme, it must therefore be an aspect of the underlying form of the preceding morpheme.

We thus explain the H tone on these suffix morphemes by positing that [oku-šum-] derives from underlying [oku-šum-an-a], by applying a rule of tone shift which shifts a H tone rightward to the following syllable, as long as the syllable is not word-final. Because of the restriction that H does not shift to a final syllable, the underlying H surfaces unchanged in [oku-šuma].

Now consider the following data.

(18) okumusija 'to block' okuminšia 'to block for'
okumusijira 'to block for' okusijira 'to block for'
okušijana 'to block e.o.' okusijana 'to block e.o.'
okuširana 'to block for e.o.' okuširana 'to block for e.o.'
When the L-toned roots of (17a) stand after the object prefixes -mu- 'him/her' and -ci- 'it', they have an H tone at the beginning of the root. Again, since the presence of the H is correlated unpredictably with the prefixes -mu- and -ci-, we hypothesize that the tones are part of the underlying representation of the prefixes - the prefixes are /mu/ and /ci/, and the H tone shifts to the right by the tone shift rule which we have already posited.

4.4 Practice at problem solving

You should now be able to apply this reasoning to data which pose analogous problems; a series of examples are given in this section for practice.

Chamorro vowel alternations. There are alternations in the quality of vowels in initial syllables in some contexts seen in the following data from Chamorro.

(19) gWihan 'fish' i gWihan 'the fish'
guma? 'house' i guma? 'the house'
katta 'letter' i katta 'a letter (object)'
t‘upa 'cigarettes' i t‘upa 'the cigarettes'
fiño? 'talk' i fiño? 'lots of talk'
tuó? 'to know' i tuó? 'you know'
t‘ügo? 'juice' i t‘ügo? 'lots of juice'
sónsug 'village' i sónsug 'the village'
hulu? 'up' i hulu? 'upward'
pétu 'chest' i pétu 'the chest'
tóm 'knee' i tóm 'the knee'
obtut 'ant' i obtut 'lots of ants'
oKu? 'hill' i oKu? 'at the hill'
dágkulu 'big one' i dágkulu 'the big one'
láhi 'male' i láhi 'the male'
lágu 'north' i lágu 'toward north'
pulónnum 'trigger fish' i pulónnum 'the trigger fish'
mundónggu 'cow's stomach' i mundónggu 'the cow's stomach'
putamóndé 'wallet' i putamóndé 'the wallet'

What underlying representations, and what rule or rules, are required to account for these data? When you answer this question, you should consider two hypotheses which differ in particular about what form is taken to be underlying - what are the two most obvious ways of treating these alternations? One of these hypotheses is clearly wrong; the other is the correct hypothesis.

Korean. Now consider the following data from Korean. The first column, the imperative, seems to involve a vowel suffix. One reason to think that there is an imperative suffix is that every imperative ends either in the vowel a or in a (the choice between a versus a is based on the vowel which precedes that suffix, /a/ or /a/ versus other vowels, and can be ignored here). A second reason comes from comparing the imperative and the plain present forms. Comparing ana and anninta, or kama and kmninta, we can see that for each verb, the portions common to both the imperative and the plain present are respectively an- and km-. From this we deduce that there must be a suffix, either -a or -a, which marks the imperative, and another suffix -ninta which marks the plain present.

(20) Imperative Plain present
ana anninta 'hug'
kama kmninta 'wind'
sina sinhinta 'wear shoes'
t’atimta t’atimninta 'trim'
noma nmninta 'overflow'
nama namninta 'remain'
éama éamninta 'endure'
ipa inninta 'put on'
kupa knninta 'bend'
čapo čamninta 'fold'
tata tanninta 'close'
pútn a punninta 'adhere'
čoča čonninta 'follow'
mača mačninta 'eat'
sak’a saqnninta 'mix'
tak’a tánntinta 'polish'
čuka čunnninta 'die'
iko ignninta 'ripen'

What is the underlying form of these verb stems, and what phonological rule or rules are required to account for the variations that are seen in the surface shape of the various stems?

Koasati. What is the underlying form of the first-singular possessive prefix in Koasati, and what phonological rule applies in these examples?

(21) Noun My N
apahca amapahca 'shadow'
asikči amasikči 'muscle'
ilkan6 amilk6 'right side'
if4 aminf4 'dog'
ap6 amap6 'grandmother'
iski amiski 'mother'
pačokk6:ka ampačokk6:ka 'chair'
towá antowá 'onion'
kastó aşkastó 'flea'
bayána ambayána 'stomach'
táta antá:ta 'father'