of verbs than there is for number inflection on nouns. This makes tense less prototypically inflectional than number. Nevertheless, tense is still clearly inflectional since it does not change the meaning of a verb significantly. It simply **GROUNDS** the meaning of the scene described by the verb in terms of time.

The "big ten" morphological processes

Chapter 1 distinguished three general expression types that languages use to accomplish communicative work. These were lexical expression, syntactic patterns, and morphological processes. In this chapter we are concentrating on morphological processes. First we will describe and exemplify ten morphological processes that will be important in the rest of this book. We will refer to these as the "big ten." These are listed below, with brief explanations and examples:

• Morphological process #1, Prefixation: (English) selfish → unselfish

PREFIXATION involves the addition of a morpheme (a prefix) to the beginning of a root. In English the morpheme *un-* is a prefix. Often languages allow several prefixes to be attached to one root. An example of this in English would be a word like: *antidisestablishment*. This word has at least two prefixes, *anti-* and *dis-*.

• Morphological process #2, Suffixation: (Spanish) hablar → hablaré

SUFFIXATION involves the addition of a morpheme (a suffix) to the end of a root. In English, the past tense is often expressed with a suffix spelled *-ed* as in *called*. As with prefixes, there can be more than one suffix on a word. A word like *establishments* in English has a suffix *-ment* and another suffix *-s*.

• Morphological process #3, Infixation:
(Bontoc) fikas 'strong' → fumikas 'strength'

INFIXATION involves the addition of a morpheme (an **INFIX**) in the middle of a root. Standard Englishes do not employ infixation, but many other languages do. The example above is from Bontoc, an Austronesian language spoken in the Philippines. Various spoken varieties of English do employ infixes, usually for emotive, humorous, or social solidarity reasons. The following naturally occurring examples are from "Rapper" English, an oral variety of African American Vernacular English (Mufwene *et al*. 1998).[2] In this variety of English, the infix *-izz-* is inserted, normally after the first consonant or consonant sequence of a word. It may have originated as a way to increase the number of syllables in a word in order to make it fit the rhythm of a line of rap music. However, it is now clearly used to express a range of emphatic, emotive, and/or humorous effects:

(12) From an internet chat room:

a.	i mean, that movie sizzucked.	'That was a terrible movie.'
b.	i knizzow.	'I wholeheartedly agree.'

In these examples, the infix *-izz-* intensifies the intended effect of the verbs *sucked* and *know*. If the word starts with a vowel, *-izz-* is the first syllable of the word:

(13) a. izzengland 'England, for heaven's sake.'
b. dat's izzall 'That's absolutely all.'
c. cuz' he be who he izzis. 'Because he is none other than who he is.'

This is a completely regular and common process that (as of 2005) is becoming more widespread in spoken varieties of English, chiefly in the USA.

It is important to distinguish infixation from multiple prefixation or suffixation. If we look at a very long English word like *antidisestablishmentarianism*, we will not find infixation – only multiple layers of prefixes and suffixes. This word can be broken up into morphemes as follows:

(14) anti-dis-e-stabl-ish-ment-ari-an-ism

The root of this word is *-stabl-*.[3] There are no morphemes inserted inside of this root, so there are no infixes. There are, however, two or three "layers" of prefixation and five layers of suffixation. These "layers" can be considered to apply in order:

Root:	-stabl-
Prefixation, layer 1:	e-stabl
Suffixation, layer 1:	e-stabl-ish
Prefixation, layer 2:	dis-e-stabl-ish
Suffixation, layer 2:	dis-e-stabl-ish-ment
Prefixation, layer 3:	anti-dis-e-stabl-ish-ment
Suffixation, layer 3:	anti-dis-e-stabl-ish-ment-ary
Suffixation, layer 4:	anti-dis-e-stabl-ish-ment-ari-an
Suffixation, layer 5:	anti-dis-e-stabl-ish-ment-ari-an-ism

As you can see, layers of prefixation and suffixation proceed outwards from the root. There are no affixes that appear inside the root, so there are no infixes.

In Bontoc, on the other hand, the "f" in *fumikas* is not a morpheme distinct from the *-ikas* part. The affix *-um-* can appear after the first consonant of a number of roots in order to express a certain kind of meaning. The first consonant must be considered part of the root, so *-um-* must be an infix.

• Morphological process #4, Circumfixation

Circumfixation is a rare morphological process in which one morpheme has two parts – one that appears before the root and another after the root. The only non-controversial examples of circumfixation that have been documented to date involve the expression of negation, as in the following examples from Chukchee, a Chukotko-Kamchatkan language spoken in northeastern Siberia, Russia (Skorik 1961, as cited in Marusic 2002):

(15) a. jatjol ‘fox’ b. *a*-jatjol-*ka* ‘without a fox’
c. cakett ‘sister’ d. *a*-cakettə-*ke* ‘without a sister’

In examples 15b and 15d the two parts, *a- . . . -kV* (where V indicates a vowel that changes its form depending on the context) jointly express the negative **INFLECTION**. Neither one occurs independently. Therefore, this must be considered one morpheme with two separate parts.

Often, examples which may appear to be circumfixation can be analyzed as two separate morphemes – a prefix and a suffix – that just happen to appear frequently together to express a particular meaning. For example, Panare, a Cariban language spoken in Venezuela, expresses negation in a manner that is surprisingly similar to Chukchee. However, this inflection in Panare does not represent true circumfixation:

(16) a. Yu-suru'-sa'
3-worry-PPART
‘He/she’s worried.’

b. A-suruku-'ka.
NEU-worry-NEG
‘He/she doesn’t worry.’

c. Wë-runkamɨ-n yu.
1-have.fever-PAST1 1SG
‘I have a fever.’

d. A-runkamɨ-'ka yu.
NEU-have.fever-NEG 1SG
‘I don’t have a fever.’

This kind of negation in Panare is expressed via an inflection that consists of two parts, *a- . . . -'ka*, as illustrated in examples 16b and 16d. However, this is not true circumfixation, since the *a-* part is in fact a distinct prefix that occurs in a number of other inflections in the language:

(17) a. A-suru'-nëpëj këj. ‘He/she’s worrying.’
NEU-worry-IMPERF 3SG

b. A-runkamɨ-nya. ‘while having a fever . . .’
NEU-have.fever-SIM

If the negative inflection in Panare were considered circumfixation, then all the other inflections that occur with the form *a-* would also need to be considered circumfixation, that just happen to share the same initial portion. There is good evidence, however, that this *a-* is in fact a distinct morpheme.[4] This is quite different from the Chukchee situation, in which both parts of the *a- . . . -kV* inflection only occur in the negative. Therefore, the Chukchee examples do illustrate true circumfixation.

• Morphological process #5, Stem modification: (English) sing → sang

STEM MODIFICATION is a change in shape that does not involve the addition of any affix. The difference in form between *sing* and *sang* in English cannot be called infixation because there is no specific form that has been added to the root. Rather, the root vowel has just changed into something else. One might ask how this is different from “weak stem suppletion” described in chapter 1. The difference is that *sing* and *sang* can be related by a rule (“change *-ing* to *-ang* to

form the past tense"). See chapter 1 for a discussion of how to determine whether a particular alternation is predictable by a rule or not.

- Morphological process #6, Autosegmental variation:
  (English) convért → cónvert

AUTOSEGMENTAL VARIATION is a change in shape that does not involve consonants and vowels. Rather, it consists of adjustments in features such as **STRESS, TONE**, and **NASALIZATION**. The best example of autosegmental variation as a morphological process in English is the difference between some nouns and verbs that is signaled by nothing but a change in stress, as in the above example. This difference is not indicated in the regular English spelling system, so I have placed a stress mark in these words to highlight the difference between *convért* (a verb) and *cónvert* (a related noun).

Here is an example of autosegmental variation marking plurality in Dungra Bhil, an Indo-Aryan language spoken in Gujarat state in India (Matthew and Susan 2000). In this case the autosegmental feature is nasalization:

(18) a.	t̪ija?a	'his'	t̪ĩja?a	'their (masc)'
b.	t̪ijʌ?ʌ	'hers'	t̪ĩjʌ?ʌ	'their (fem)'

Note that the only difference between the singular and plural possessive pronouns is that the plurals have a nasalized vowel (indicated by a tilde "∼") in the first syllable.

- Morphological process #7, Reduplication:
  (Ilokano) pingan 'dish' → pingpingan 'dishes'

REDUPLICATION involves the repetition of part or all of a root. Plurality in Ilokano (another Austronesian language spoken in the Philippines) is expressed by reduplicating the first syllable of the root, as in the above example. Here are some further examples:

(19) a.	ulo	'head'	ululo	'heads'
b.	talon	'field'	taltalon	'field'
c.	biag	'life'	bibiag	'lives'
d.	mula	'plant'	mulmula	'plants'

This is called **PARTIAL REDUPLICATION**, because only part of the root is repeated. In the case of Ilokano, the only part that is reduplicated is the first syllable of the word. Some languages, like Indonesian, repeat the whole root. So in Indonesian "child" is *anak* and "children" is *anakanak*. This is called **COMPLETE REDUPLICATION**. Plural is not the only conceptual category that is expressed by reduplication, but it is quite common.

- Morphological process #8, Non-concatenative morphology:
  (Hebrew) sefer 'book' → sfarim 'books'

NON-CONCATENATIVE MORPHOLOGY is common in Semitic languages, such as Hebrew and Arabic, but rare elsewhere. It involves superimposing a

pattern of vowels, and possibly other morphological pieces, on a root that consists only of consonants. For example, 20 illustrates a few of the verb forms for the root *ktb* in Biblical Hebrew. This root can never be pronounced on its own, but must appear in an inflected form (examples courtesy of David Andersen, as cited in van der Merwe, Naudé, and Kroeze 1999):

(20)	a.	ktb	root	(no meaning on its own)
	b.	kətob	imperative	'write!'
	c.	kɑtob	infinitive	'to write'
	d.	kotɛb	present participle	'writing'
	e.	kɑtub	past participle	'written'
	f.	kɑtab	perfective	'wrote'

- Morphological process #9, Subtractive morphology:
  (Murle) nyoon 'lamb' → nyoo 'lambs,' wawoc 'white heron' → wawo 'white herons'

SUBTRACTIVE MORPHOLOGY is another quite rare process, whereby one or more segments are omitted from a word in order to express a particular conceptual category. Murle (along with several other Nilo-Saharan languages of East Africa) is one of the few languages of the world that illustrate true subtractive morphology. In each of the Murle examples above, the stem-final consonant is omitted in order to form the plural.

One has to be careful to distinguish subtractive morphology from simple zero realization of certain categories, especially when those categories have overt marking in another language the linguist is familiar with. For example, in Arbore (a Cushitic language of Ethiopia) for some nouns the singular ends in *-in* while the plural is unmarked (Hayward 1984:159–83, cited in Corbett 2000:17):

		Singular		Plural	
(21)	a.	tiisin	'a maize cob'	tiise	'maize cobs'
	b.	nebelin	'a cock ostrich'	nebel	'ostriches'

One may be tempted to analyze this as an instance of subtractive morphology, because the plural nouns are formally simpler than the singular. However, this would be a mistake. In many languages, including English, the plural is marked and the singular is unmarked. However, in this class of Arbore nouns (items that normally occur in groups), the plural is the unmarked number, while the suffix *-in* is a marker of singular. Sometimes this is called a **SINGULATIVE**. Notice that this is quite different from Murle. In Murle, there is no one suffix that indicates singular. Rather, the last consonant of the stem, no matter what it is, is eliminated in order to form the plural. Here are some more nouns that illustrate this feature of Murle (Arensen 1982:40–41):

		Singular		Plural	
(22)	a.	onyiit	'rib'	onyii	'ribs'
	b.	rottin	'warrior'	rotti	'warriors'

Notice that the last consonant of any stem, whether it is *-t*, *-n*, *-c*, or any number of others, is simply left out to mark the plural. These cannot all be different forms of one "singulative" morpheme, so they must be considered part of the stem, and the morphological process that expresses plurality involves *removing* that last consonant.

- Morphological process #10, Compounding:
  (English) black+bird → bláckbird

COMPOUNDING involves combining roots to form new stems. In the English example above, it is impossible to identify one part as the root and the other as an affix. *Black* and *bird* are both roots that clump together morphologically to form a stem. The new stem, *blackbird*, expresses an idea that is more than simply the combination of the meanings of the two roots – this word does not refer to any bird that happens to be black, but rather to a specific species of bird. Even though this word is formed out of two roots, it functions just like other noun stems in the language.

Having introduced ten major morphological processes, we now turn to a discussion of various methods linguists use to "model" or represent morphological processes and the conceptual categories they express.

Methods for representing morphological processes

Prose

As mentioned in chapter 1, many of the grammatical patterns in language may be expressed in ordinary prose. When doing morphosyntactic analysis, it is very important to be explicit, and sometimes grammatical patterns are so complex that explicit prose statements become difficult to follow. In these cases linguists have found it useful to employ various notational systems. In the following sections we will discuss two mathematically explicit methods for representing patterns of linguistic behavior. These are particularly useful for representing morphological processes, though they could, in principle, be used to express syntactic patterns as well. It should be kept in mind, however, that prose statements are often the most communicative way of expressing facts about grammatical structure.

Position-class diagrams

The second method of representing linguistic knowledge we will discuss is called **POSITION-CLASS DIAGRAMMING**. This method is a variation on a general approach to morphological structure that is called the **ITEM AND ARRANGEMENT** model (Hockett 1958). It can be very useful for describing languages that tend to have lots of morphemes per word, especially if the morphemes tend to fall into well-defined sets, or **PARADIGMS**. While there

are several inadequacies to position-class diagramming, some of which will be discussed at the end of this section, every field linguist needs to be familiar with this method, at least as a beginning point for a full morphosyntactic analysis of a language.

In languages that express many of their conceptual categories morphologically, there are typically several "layers" of prefixes and suffixes, as described above. For example, here are some data from Sierra Nahuatl (from Merrifield *et al.*, 1987). Four of the free translations have been omitted from these data, just to make it a little more interesting:

(23)	a.	nimicita	'I see you.'
	b.	nikita	'I see him.'
	c.	tikmaka	'You give it to him.'
	d.	tinečita	'You see me.'
	e.	nannečmaka	'You (pl) give it to me.'
	f.	tikonmaka	'You give it to him, sir.'
	g.	tikonitatihcinoh	'You see him, most honored sir.'
	h.	tikonmakatihcinohtikah	'You give it to him, most very honored sir.'
	i.	tinečonita	
	j.	tinečonmakatihcinoh	'You give it to me, most honored sir.'
	k.	nannečonmakatikah	'You (pl) give it to me, honored sirs.'
	l.	nannečonitatihcinohtikah	
	m.	tinečonitatikah	
	n.	nannečonmakatihcinoh	

The process of constructing a position-class diagram for data such as these will be given below in step-by-step fashion, though the steps are more a descriptive tool than a "program" for analyzing morphology. As you will discover if you ever have the privilege of doing fieldwork on a real language, morphosyntactic analysis is an art that is trying to be a science. As linguists, we want to be as rigorous and scientific as possible in our research and presentation of findings. However, since our subject matter involves human behavior, there will always be indeterminate cases, educated guesses, and subjective interpretations. So there is no absolute "procedure" that will lead to one correct analysis of a range of linguistic facts. There are only better and worse analyses, and better and worse arguments for them.

Step 1: Isolate the roots. As we look over the free translations of the Sierra Nahuatl data, we see that there seem to be two basic verbs involved: one meaning 'see' and another meaning 'give.' Since we expect similarity of meaning to correlate with similarity in form (see chapter 1, p. 3), we look down the left column to see what elements of form correlate with the meanings 'see' and 'give.' What do you notice? You should see that in the Nahuatl sentences the form *ita* consistently matches 'see' in the translation, while *maka* consistently matches 'give.' Therefore we hypothesize that *ita* and *maka* are the roots meaning 'see'

and 'give' respectively. We list and gloss these roots in a position-class diagram as follows:

ROOT
ita 'see'
maka 'give'

I must emphasize that this is a hypothesis. Your initial impressions need to be held lightly, until they are confirmed by further data. I suggest that everyone use pencil when approaching these kinds of problems.

Step 2: Estimate the affix positions. Since in the Nahuatl examples there is material to the left and the right of these roots, we suspect there are prefixes and suffixes. So we want to leave room for these elements:

P3	P2	P1	ROOT	S1	S2	S3
			ita 'see' maka 'give'			

Notice that the affix positions are numbered outward from the root. This diagram leaves room for three prefix positions (P1, P2, P3) and three suffix positions (S1, S2, S3). Again, this is only an estimate. At this stage you want to allow for what you consider to be the maximum number of affixes, within reason. In this case we have guessed the same number of prefixes as suffixes, but this need not be the case. Furthermore, in step 1 you may have noticed that the roots come at the beginning or the end of the structure you are analyzing. You would need no positions for affixes that obviously don't occur.

Step 3: Begin to analyze prefixes. Since in this problem the first examples seem to contain just prefixes, we will start by trying to analyze the prefixes. Remember the basic principle that similarity in form usually expresses similarity in meaning. In examples 23a and b, we see that the formal variation between *nimic-* and *nik-* correlates with a meaning variation between 'I > you' ('I' acting on 'you') and 'I > him' ('I' acting on 'him'). Within these two possibilities, we see a common element, *ni-*, and variation between *mic-* and *k-*. Since the common element of meaning is 'I,' and the variation is between 'you (object)' and 'him,' we suspect that there are two prefix positions, with *ni-* coming earlier, and *mic-* and *k-* following. Since these kinds of prefixes are likely to be grammatical morphemes, we will gloss them according to reasonable guesses as to their conceptual categories:

P3	P2	P1	ROOT	S1	S2	S3
	ni- 1SG.SUBJ	mic- 2SG.OBJ k- 3SG.OBJ	ita 'see' maka 'give'			

The next three examples help us fill out this chart a little more, as follows:

P3	P2	P1	ROOT	S1	S2	S3
	ni- 1SG.SUBJ ti- 2SG.SUBJ nan- 2PL.SUBJ	mic- 2SG.OBJ k- 3SG.OBJ neč- 1SG.OBJ	ita 'see' maka 'give'			

When we get to example 23f, however, we encounter a difficulty. We notice that there is a form *on-* which seems to come in between the root and the prefixes we have posited. We notice that this *on-* correlates with the meaning 'sir.' Therefore, we need to move our two prefix positions to the left, and add *on-* in the P1 position (this is why I suggest you use pencil). We guess that a good grammatical gloss for this *on-* might be 'the speaker expressing respect for the hearer.' Therefore we gloss it with the abbreviation 'RESP':

P3	P2	P1	ROOT	S1	S2	S3
ni- 1SG.SUBJ ti- 2SG.SUBJ nan- 2PL.SUBJ	mic- 2SG.OBJ k- 3SG.OBJ neč- 1SG.OBJ	on- RESP	ita 'see' maka 'give'			

This analysis seems to take care of all of the prefix combinations in the data set. At this point, we can calculate the probable free translation for example 23i. I will let you figure out what that should be.

Step 4: Analyze the suffixes. This step will be very similar to the previous one. Comparing 23g and h, we see that the difference in the forms of the suffixes is related to the difference between 'most honored sir,' and 'most very honored sir.' Since there is some commonality in form (*tihcinoh*) to both of these examples, we suspect that there are two suffix positions. The two suffixes in question may be particularly difficult to gloss, but we can make a stab at it, as follows:

P3	P2	P1	ROOT	S1	S2	S3
ni- 1SG.SUBJ ti- 2SG.SUBJ nan- 2PL.SUBJ	mic- 2SG.OBJ k- 3SG.OBJ neč- 1SG.OBJ	on- RESP	ita 'see' maka 'give'	-tihcinoh HONOR1	-tikah HONOR2	

The glosses 'HONOR1' and 'HONOR2' are impressionistic guesses as to the meanings of these suffixes. The idea is that *-tihcinoh* probably expresses a first degree of honorific status, while *-tikah* expresses the second, higher, degree. At this point these are just educated guesses, so the actual glosses chosen are not terribly significant. The point is to notice what the affixes are, and how they are related to one another in the verb word.

Inspection of the rest of the data reveals that the hypothesis that there are two suffix positions, instantiated by *-tihcinoh* in position S1 and *-tikah* in position S2,

holds true. At this point, you can fill in probable free translations for examples 23l through n.

Step 5: Label the columns. Positions in complex morphological structures tend to be associated with particular sets of conceptual categories. For example, verbs in a highly morphological language might have one position for tense, another position for **ASPECT**, another for **PERSON** and **NUMBER** of the **SUBJECT**, etc. As we inspect the diagram given above, we notice that all the forms in the P3 position express the person and number of the subject. All the forms in the P2 position express the person of the object. In this example, there is only one form in each of the other affix positions, so it is difficult to infer a meaning category for the whole column. Nevertheless, it does not hurt, at this point, to hazard a guess. Here is one possible way of completing our sample position class diagram:

Subject	Object	Respect	ROOT	Honor1	Honor2
ni- 1SG ti- 2SG nan- 2PL	mic- 2SG k- 3SG neč- 1SG	on-	ita 'see' maka 'give'	-tihcinoh	-tikah

Further data may cause us to revise this hypothesis, but just based on the data given in the problem set this seems to be a reasonable analysis. There are many aspects of the morphology of a language that this kind of diagram simply does not capture. For instance, even in this small data set, it seems to be the case that the honorific suffixes always co-occur with the 'respect' prefix. More data would be needed to see if this is always the case, or just an incidental property of the examples chosen. There often are these kinds of long-range "dependencies" between elements (the presence of one element "depends" on the presence of another one somewhere else in the structure). There are ways of annotating position-class diagrams to show this, but it can get rather messy.

Other problems for position-class diagrams are situations where morphemes can occur in more than one place in a structure, or, as we will see in the following chapter, morphemes that are pronounced differently depending on the context. Position-class diagrams are not very useful for describing non-concatenative morphology, autosegmental phonology, stem modification, reduplication, or compounding. In spite of these problems, a position-class diagram is a good start on the road to building a solid and insightful morphosyntactic description of a language, especially if the language uses lots of prefixes and suffixes, which are, after all, the most common kinds of morphological processes found in the world's languages.

Process rules

While position-class diagrams are the mainstay of basic morphological analysis, their shortcomings have led many linguists to devise a number of alternative schemes for representing morphological structures. Most of these

additional schemes can be described as **PROCESS RULES**. A process rule is a representation that describes relationships among the various shapes of words as though they were *changes* that the words undergo. This general approach to morphological structure has been described as the **ITEM AND PROCESS MODEL** (Hockett 1958). Process rules were hinted at in chapter 1. For example, the structure of regular nouns in English can easily be represented in a simple position-class diagram as follows:

Root	Number
cat	ø SG
dog	-S PL
mat	
tree	
. . .	

In a process rule, the formation of the plural (and perhaps the singular as well) would be treated as a "process" that changes a root into the appropriate **INFLECTED** form. For example, one could say:

(24) Singular noun + -s = Plural noun

The insight behind this particular type of rule is that, since some forms are "simpler" than others, it makes sense to think of the more complex forms as being based on the simpler ones. Forms "start out" simple, and "end up" complex.

Though there are many different ways to formulate process rules, the ways we will discuss in this text all involve three parts: the conceptual category that is expressed, a structural description of the form *before* the process occurs, and a description of what the form changes into *after* the process occurs. These can be abbreviated as:

CC = **C**onceptual **C**ategory
SD = **STRUCTURAL DESCRIPTION** (starting form)
SC = **STRUCTURAL CHANGE** (ending form)

The rules themselves will always have the following pattern:

(25) CC: SD → SC

For example, the regular pattern for plural formation of English nouns may be expressed as:

(26) Plural: N → N + *-s*

One way this formula may be read would be: "To express plurality, start with a noun, and end with that same noun plus an *-s* suffix."

How would one employ process rules to describe complex data such as the Sierra Nahuatl verbs given above? Well, each conceptual category would require its own rule. Here is a possible subset of process rules for the Sierra Nahuatl data:

(27)

	CC:	SD:	SC:
a.	Respect:	Verb	→ on- + Verb
b.	2sg object marking:	Verb	→ mic- + Verb
c.	3sg object marking:	Verb	→ ke- + Verb
d.	1sg object marking:	Verb	→ neč- + Verb
e.	1sg subject marking:	Verb	→ ni- + Verb
	etc.		

There are many ways of representing morphological process rules. For example, Haspelmath (2002:47–51) provides a very nice system which he terms the "word-based model." The formulas in Haspelmath's system contain all the elements of process rules and can be very helpful in describing and understanding the morphological patterns of a language that has only a few morphemes per word. However, when analyzing complex data, like Sierra Nahuatl above, that involve affix "positions" in word structure, some form of a position-class diagram (or "morpheme-based model") may be more useful.

Process rules, of the sort illustrated above, can be very useful for describing non-concatenative morphology. For example, here are some data from Arabic (these data have been slightly regularized but are true to the general facts of most varieties of Arabic):

(28)

	Root:	slm		Root:	ktb
a.	muslim	'person of peace'	g.	muktib	'literate person/scribe'
b.	salima	'he was safe'	h.	katiba	'he was reading'
c.	ʔislaamun	'Islam'	i.	ʔiktaabun	'literature'
d.	salaamun	'peace'	j.	kataabun	'book'
e.	saalimun	'safe'	k.	kaatibun	'writing'
f.	salama	'he was calm'	l.	kataba	'he wrote'

The first step in analyzing these data is to determine the conceptual category, as discussed earlier in this chapter. What conceptual category do you think is expressed by examples 28a and g? Can you come up with a good description of the meaning expressed by the particular morphological pattern displayed by these two examples? Looking at the English translations, you may guess that the Arabic words are probably nouns, and you know that they both describe people. If the root *slm* means something like 'peace,' and *ktb* means something like 'read,' then perhaps the pattern can be thought of as expressing the idea of 'someone who is/does the concept described by the root.' What is a good name for such a function? As with the Yup'ik examples discussed earlier, we may want to call this "nominalization," and since this nominalization refers to people, we may want to call it "person nominalization." This is only one possible label for this particular conceptual category.

These data illustrate six conceptual categories as expressed in two distinct roots. Since Arabic is a typical Semitic language, it exhibits non-concatenative morphology. The problem we will now address is how to represent such interesting morphological patterns in terms of a process rule. So far we have a name for the

first conceptual category. Now we need a Structural Description and a Structural Change:

(29)

CC	SD		SC
Person nominalization:	?	→	?

If we think of this process as "starting out" with the roots *slm* and *ktb*, and ending up with the inflected forms, we can describe the data in two separate statements as follows:

(30)

	CC	SD		SC
a.	Person nominalization:	slm	→	muslim
b.	Person nominalization:	ktb	→	muktib

These provisional statements are the first step in formulating rules; however, we cannot really call them "rules" yet. They simply restate the data in terms of a process, but they don't capture the fact that there is really only one process involved – one pattern that applies consistently to both roots (and many more that are not in this data set). Since part of what an Arabic speaker must unconsciously "know" is that this pattern affects different roots in the same way, we want to generalize our representation of this knowledge so that one statement covers all relevant forms. If we can do this, our linguistic description will be consistent with the unconscious knowledge of Arabic speakers.

We can see that both 28a and g begin with a prefix *mu-*, followed by the first consonant of the root, then the second consonant of the root, then an *i* vowel, and finally the last consonant of the root. We can use the symbol C to represent any consonant, and number them C_1, C_2, and C_3 in order to represent the order in which they occur in the root. Then we can simply substitute these symbols for the consonants that enter into the pattern:

(31)

CC	SD		SC
Person nominalization:	$C_1C_2C_3$	→	$muC_1C_2iC_3$

Since C represents any consonant, this rule applies to both *slm* and *ktb*. The subscripted numbers link the consonants in the output (the SC) with the consonants in the input (the SD), no matter what the actual consonant is. If C_1 is an *s* in the SD, C_1 will be an *s* in the SC, etc. If there were no subscripts, there would be no way to distinguish the consonants from one another (since C is an abstract symbol, and not a particular sound).

How would the other patterns illustrated in 28 be formulated using this system? I'm sure you can figure that out by now! Here is one more, just for fun. Examples 28b and h express a conceptual category that may be interpreted as "past continuous." Why would it not be sufficient to call this conceptual category simply "past tense"? The reason is that there is another category, represented in 28f and l, that also expresses events in the past. So we know there are at least two past "tenses." In order to distinguish these two, the term "past continuous" for the category illustrated in 28b and h, and perhaps "simple past" for the category

represented in 28f and l, seem like reasonable labels. So if we use the term "past continuous" for the conceptual category illustrated in 28b and h, and apply the same reasoning as we did above to describe the morphological manifestation of this category, we would come up with the following process rule:

(32)

CC	SD		SC
Past continuous:	$C_1C_2C_3$	→	$C_1aC_2iC_3a$

This rule describes both of the following actual changes:

(33)

s l m	→	s a l i m a
k t b	→	k a t i b a

In addition to non-concatenative morphology, such as found in Semitic languages, process rules of this sort are particularly useful for expressing other kinds of morphological patterns, in particular, morphological patterns that involve reordering, or **METATHESIS**, of sounds in a word, and reduplication (see above).

As far as we know, no language uses metathesis alone to express conceptual categories. This is why it is not one of the "big ten." However, metathesis does occasionally accompany one of the other processes, such as prefixation or suffixation. For example, consider the following data from Yagua:

(34)

a.	rakyáraay	'I go astray.'
b.	hikyáraay	'You go astray.'
c.	sakáray	'He/she goes astray.'
d.	naakyáraay	'We go astray.'
e.	naadakáraay	'They 2 go astray.'
f.	rikyáraay	'They go astray.'

As you inspect these data, you notice that in examples 34a, b, d, and f, there is a *y* sound that appears *after* the first consonant of the root. The fact that this *y* does not appear in all the **INFLECTIONS** suggests that it is not part of the root, but part of the prefix. In fact, this is completely regular in the language, and it is quite easy to show that the prefixes are the following:

(35)

a. ray-	1SG	d. naay-	1PL
b. hiy-	2SG	e. naada-	2DL
c. sa-	3SG	f. riy-	3PL

So in order for Yagua speakers to inflect their verbs, they must "know," subconsciously of course, that if the prefix ends in a *y*, you have to switch the order of this *y* with the first consonant of the root. It turns out that this is a perfectly regular pattern in the language, so we don't need to restrict it to any particular conceptual category (see chapter 3 on the difference between phonological rules and morphophonemic rules). We can simply state it as a rule with no restrictions:

(36) yC → Cy

This rule simply says that whenever a *y* occurs before any consonant, the *y* and the consonant switch places. In this case, we do not need to use a subscripted number after the C, since there is no other unspecified consonant (C) in the rule that may cause confusion.

Finally, process rules can be very useful for representing stem modification and reduplication. For example, consider the following data from Sierra Nahuatl (Elson and Pickett 1988:51):

(37)

a.	se	'one'	f. sehse	'ones/one by one'
b.	ome	'two'	g. ohome	'twos/two by two'
c.	eyi	'three'	h. eheyi	'threes/three by three'
d.	makwil	'five'	i. mahmakwil	'fives/five by five'
e.	čikasen	'six'	j. _____	'sixes/six by six'

For the conceptual category (call it "multiplicative") represented by the words in the right-hand column, the initial **SYLLABLE** (consonant plus vowel, or just a vowel) of the root is reduplicated and followed by an *h* sound. This can be nicely represented with the following process rule:

(38)

CC	SD		SC
Multiplicative:	$(C_1)V_1X$	→	$(C_1)V_1h(C_1)V_1X$

In this formula, the Consonant is placed in parentheses to indicate that it may or may not be present. This notation allows this rule to apply both to consonant-initial stems (examples 37a, d, and e), and vowel-initial stems (examples 37b and c). The V is a cover symbol that stands for any Vowel. The X is a cover symbol that stands for anything, including silence at the end of a word. This notation allows the rule to apply to stems of any length – it essentially says that whatever comes after the first syllable of the root (including silence, as in 37a) does not affect the rule at all.

While this formula may appear quite complex, it is absolutely explicit, and accurately captures something that Sierra Nahuatl speakers must unconsciously "know" about their language. Using this rule, I'm sure you can infer the form of the word meaning 'sixes/six by six' (37j). If you can, then you have assimilated and applied part of the knowledge that every Sierra Nahuatl child has concerning the expression of numbers in this language.

Conceptual outline of chapter 2

I. Conceptual categories:

- A conceptual category exists when there is an expectation of patterned behavior – a consistent relationship between variation in form and variation in function.
- Conceptual category labels (or "glosses") are interpretations designed to help readers of a grammatical description understand