#### INTRODUCTION TO MORPHOLOGY

© 2002 Mark C. Baker and Jonathan David Bobaljik

Rutgers and McGill

Draft – for class use

# 2. DERIVATION

#### 2.1 Introduction

Word-formation is often divided into three types: **derivation**, **inflection** and **compounding**. Compounding is the easiest of these to define: a compound is a word that contains more than one root, for example *workbench* contains the roots *work* and *bench*. Compounding is discussed in the next chapter and we will lay the topic aside until then. This leaves derivation and inflection, which we can think of at this point as different types of affixation. (Some schools of thought have suggested that inflection in particular may involve processes other than affixation, we return to this question in Chapter X—there is more than enough material to discuss if we limit ourselves to roots and affixes for now.) Although the distinction between derivation and inflection is an old one, inherited from traditional grammar, there is no consensus on what criteria should be used to distinguish the two. The traditional notion is that the derivation trades in meaning, while inflection trades in function. Thus, derivational morphology adds new meaning (or changes meaning) and for some, creates a new word (or lexeme). Inflectional morphology signals the function of a word in a grammatical context.

For example, the **conjugation** paradigm of a verb stands as a canonical example of inflectional morphology. The three forms *work, works, worked* are intuitively 'forms of' the verb *work*, and the suffixes –*s* and –*ed* merely mark the verb for an appropriate context (third person singular subject, or past tense, respectively). Derivational morphology by contrast adds or changes what might be called lexical meaning; the canonical examples of derivational morphology change syntactic category as well. Thus *worker* 'one who works' and *workable* 'able to be worked' (as of a metal, or a solution) are not forms of the verb *work* marked simply so as to appear in some context, rather, they are, respectively, a noun and an adjective derived from the verb *work*, but with a more elaborate meaning. Quite typically, when derivational and inflectional morphology occurs in the same word, the derivational morphology occurs closer to the root than the inflectional morphology. For example, in the word *hospitalized* (as in *'We hospitalized the patient'*) the root is *hospital*, a noun, which is followed by the verbalizing suffix –*ize* (category-changing, hence derivational) in turn followed by the past tense.

The intuition underlying the traditional divide separates clear cases of derivation from clear cases of inflection, even though it leaves a substantial gray area at the border. It is not really useful at this point to try to provide a more detailed characterization of the difference. (Indeed, many linguists deny that there is a principled difference between derivation and inflection and therefore they look for a unified theory.) Instead, we will accept the traditional intuitive division of meaning versus grammatical function, and use derivation as a starting point to explore how one might develop a theory of morphology. After we have considered properties of the clear cases of both derivation and inflection (Chapter Infl), we will be in a point to revisit the usefulness of making a terminological distinction between the two.

#### 2.2 Selection and constituent structure

Recall from the introduction that one of the most important aspects of any explanatory theory lies in its ability to predict not only what can (and does) occur, but also to accurately distinguish this from what is impossible. A good starting point, then, is to compare some set of acceptable words with a set of minimally different combinations of morphemes which are systematically unacceptable. By doing so, we can identify the property that distinguishes the two classes. Working with a small set of data first, we state a hypothesis based on our initial observations, and then derive predictions which we can test against more data.

As a very simple starting point, take the division of affixes into prefixes and suffixes which we established in the previous chapter. Recall that the division is not by absolute position, but rather by the side of the root on which a given affix occurs. An affix which follows the root is a suffix, even if it is in turn followed by other affixes (by definition, all of these would also be suffixes). This is a useful start for description—and it marks the beginning of a theory of morphology—in that it allows us to explain why some combinations of morphemes are impossible as words. For example, the words in (1) contain morphemes aligned on the correct side of their respective roots. The non-words in (2) contain the same morphemes, but aligned incorrectly.

(1) a. undoable

- b. hopefulness, hopefully, carelessness, carelessly
- (2) a. \*do-able-un, \*do-un-able, \*able-un-do, \*un-able-do
  - b. \*fulhopeness, \*lyhopeful

3

Whether a given affix is a prefix or suffix is a property that must be learned for each affix (or for some languages, for classes of affixes). Some languages, such as Inuktitut, spoken across the North American Arctic, have only suffixes while others, such as some Athabaskan languages of the Northwest part of North America have mostly prefixes. Many languages, like English, have a mixture of prefixes and suffixes. While there are some discernable patterns, there is no general theory which will predict universally (either from phonology or meaning) which affixes will occur as suffixes and which as prefixes. Anything that occurs as a suffix in one language can be found surfacing as a prefix in another language. However, once it has been learned (and thus stored in the appropriate lexical entry) for any given morpheme, whether it is a prefix or suffix, this information will serve to explain why morpheme combinations which are misaligned, as in (2), are unacceptable.

But this is only the tip of the iceberg. There are vast numbers of combinations of morphemes which respect the prefix or suffix properties of the affixes involved, but nevertheless systematically fail to generate potential words of the language in question. Again, we can illustrate this point with English examples. The forms in (3a) are made up of the same morphemes as the words in (1b), and these are clearly recognized as gibberish just as are the forms in (2).

- (3) a. \*hopenessful, \*hopelyful \*carenessless, \*carelyless
  - b. \*goodnessful, \*goodnessly \*carefulless, \*hopefulless
  - c. \*goodlessly
  - d. \*enjoyful, \*dethroneless

4

Direction of attachment does not provide an explanation for the unacceptability of the forms in (3). In all of the examples, the affixes involved are aligned on the proper side of the roots. For example, we know from the acceptable forms *hopefulness* and *hopefully* that *-ful, -ness,* and *-ly* are all suffixes. In (3a), this property is not violated, the affixes in all the examples follow the root *hope* and *care*, yet nevertheless these are not potential words.

### 2.2.1 Selection

The first thing we may note is that affixes care not only about the side of the root on which they occur, but also about what they occur next to. For example, the English suffix -ful attaches to nouns (especially to nouns designating emotions); it does not attach to other parts of speech, such as verbs or adjectives. Examples are given in (4).

(4) a. from nouns: *joyful, sorrowful, gleeful, hopeful, careful* 

- b. \* from verbs: \*exciteful, \*exhilirateful ...
- c. \* from adjectives: \*happyful, \* sadful, \*friendlyful ...

(Note that some of the words to which -ful attaches exist as nouns and verbs, e.g., *(the) hope* ~ *(to) hope*. By establishing on the basis of the unambiguous cases, such as *joy*, that the suffix can only attach to nouns, we must conclude that the adjective *hopeful* is constructed from the noun *hope* and not from the homophonous verb.)

The technical term to express restrictions on what an affix may legitimately attach to is *selection*. (To *select for* some property is to *attach to or combine with something having that property*, a terminology shared with syntax. Selection is also sometimes called *subcategorization* in both morphology and syntax; with that alternative terminology, selectional restrictions are called *subgategorization frames*. The difference is just terminological, but you will find both used in the primary literature and should recognize them.)

Example (5) provides a revised lexical entry for *-ful* incorporating its selectional restrictions. Note that we have also introduced a notation to condense the information. As in phonological rules, the underline specifies the legitimate position for *-ful* relative to other elements, this line indicates that *-ful* is acceptable when occurring to the immediate right of a noun. This incorporates the fact that *-ful* is a suffix and thus we do not need to write that information elsewhere in the lexical entry (a prefix, like *un-*, would have a selectional restriction with the underline before the bracket:  $[\_A]$ ).

(5) LEXICAL ENTRY for -ful

Cat. of Result:	Adj
Selection:	[ N _ ]
Meaning:	'having X
Phon:	/fəl/
Label:	-ful

All and only the acceptable examples in (4) meet the selectional restrictions on -ful. We can represent this more explicitly by adding category information to the examples in (4), and using brackets to delineate the morphemes more carefully. The first example from each row of (4) illustrates this here:

- (6) a. [joy]<sub>N</sub> <u>ful</u>
  - b. \*[ excite ]<sub>V</sub> <u>ful</u>
  - c. \*[ happy ]<sub>Adj</sub> ful

Armed with this small amount of formalism, we are able to give a precise explanation of why *\*hope-ly-ful* and the other words in (3) are impossible as words of English. They all involve one or more violations of the selectional restrictions of the affixes involved. In (7) we give the selectional restrictions for the other suffixes in (3a-c), along with the syntactic category of the result. You should be able to construct lists analogous to (4) to verify this.

(7) a. ness [A\_] Result: N
b. ly [A\_] Result: Adv
c. less [N\_] Result: Adj

Thus, although *goodness* and *careful* are legitimate combinations of morphemes, the further combinations in (3b) *\*goodnessly, \*carefulless* violate the selectional restrictions of the outermost affixes. *Goodness* is a noun, and *-ly* selects for adjectives.

Likewise, *\*hopelyful* and *\*hopeliness* are impossible because the selectional restrictions of -ly have been violated at the first stage of combination: *hope* is a noun (or verb), and *-ly* thus cannot attach to it.

It is important to note that an early violation cannot be saved by satisfying selectional requirements further out. This is shown, for example by *\*goodlessly* in (3). The suffix *–less*,

when it attaches correctly, makes adjectives, and thus can precede the attachment of -ly (*carelessly, hopelessly*). But the selectional restrictions of *-less* are not met in combination with good = \*goodless, and no further operation can salvage any attempt to build on this base. Failure to meet selectional restrictions is fatal, at any point. All the words in (3) are excluded for the same reason.

### 2.2.1.1 Selection: Some extra practice

In Chapter "Basic Concepts" we identified the English suffix written *er* with the meaning "one who / that which does X" and suggested on the basis of the examples in (8a) that it attaches to verbs to make the corresponding agentive noun. Thus, we would posit the selectional restriction  $[V_{-}]$ . Do the examples in (8b) and (8c) challenge this?

(8) a. teacher, worker, driver, lecturer, reader, softener, whitener, blender, modernizer

- b. softer, whiter ...
- c. Londoner, Hamburger, ...

The words in (8b) and (8c) have suffixes with the phonological string [ər], and they are attached to roots that are not verbs (*soft* etc. are adjectives, and *London, Hamburg* are place names). Yet we must bear in mind, as discussed in the previous chapter, that a morpheme is not just a phonological string, but rather a correspondence of sound and meaning. The suffix morphemes in (8b) and (8c) have the same sound as the agentive suffix [ər], but they do not have the meaning part of this suffix. The word *softer* in (8b) cannot mean "something which makes things soft", and a *Londoner* is someone **from** London, not "one who does a London" (whatever that

would mean). Each of these suffixes has its own selectional restrictions. For example, the selectional restriction for the comparative is  $[A_{-}]$  (it has additional phonological restrictions, typically attaching only to adjectives of one or two syllables). Since none of the stems in (8a) is an adjective (although some contain adjectives), none of these words can be used as a comparative adjective (cf., *\*This is softener than that*). Deriving the comparative meaning requires the comparative suffix, but the comparative suffix can only attach directly to adjectives and not to verbs, even if the verb (like *soften*) contains an adjective inside it.

Note also that within (8a), four of the stems (*work, drive, lecture* and *blend*) are all usable as either nouns or verbs. Should this lead us to suggest that the agentive suffix *-er* can attach to nouns or verbs? The question to ask at this point is: does the suffix ever attach to stems which are unambiguously nouns? Since it does attach to stems that are only verbs, such as *teach, soften*, the most restrictive hypothesis that is consistent with the data would be to say that it attaches only to verbs stems. This hypothesis would correctly exclude words such as *\*catter*, or *\*windower*, while allowing *worker*, *driver* etc. Moreover, as we discussed in the previous chapter, this suffix can be added to nonce verbs, but not to nonce nouns. (Note also that English allows quite a bit of latitude in converting nouns to verbs and vice versa. Thus, the word *birder* exists, for example, to mean a dog which is good at catching birds. But, like *work* etc, this is only possible to the extent that *bird* can also be used as a verb: *This dog birds*.)

Consider now the English suffix —*ous*, as in *monstrous*, *glorious*. What are its selectional restrictions? (The best way to approach this is with a reverse dictionary.) Do all words that end in —*ous* involve suffixation? [Note: this is in fact a very tricky question, one which it might not be possible to answer definitively here.]

#### 2.2.2 Lexical Restrictions and Productivity

Selectional restrictions, including the distinction between prefixes and suffixes, do a substantial amount of work in eliminating unacceptable combinations of morphemes from the class of possible words generated by the grammar of any language. In the illustration we have given here, we have focused on only a small set of selectional restrictions for a fragment of the English lexicon. We have hinted in passing at the existence of further restrictions, for example, *-ful* doesn't attach to just any noun, but in particular to nouns designating an emotional state: a chair with holes is not *\*holeful*, nor is a garden with weeds *\*weedful*).

Further restrictions of this sort are quite common in derivational morphology cross-linguistically. In particular, many derivational morphemes impose **lexical restrictions** on what they can combine with (in addition to restrictions on syntactic category, for example). The English suffix *–ity* is a case in point. This suffix attaches to adjectives, and the result is a noun, meaning roughly 'the degree to which something is Adj' or 'something which instantiates the property of Adj'. Thus: *curious – curiosity, adverse – adversity, odd-oddity, able-ability*, etc. But *–ity* does not attach to all adjectives. While the degree to which someone is *curious* is their *curiosity*, one cannot analogously say that the degree to which someone is *young* is their *\*youngity*. Pairs such as *odd-oddity* versus *sad-\*saddity* and *adverse-adversity* versus *terse-\*tersity* suggest that there is no systematic phonological basis for the excluded forms. Rather, *-ity* simply has further lexical restrictions beyond simply selecting adjectives. It is often suggested that the lexical restrictions on *-ity* are that it attaches to the [+latinate] roots of English, that is, those roots which entered the language from Latin (typically via Norman French) as opposed to the Germanic portion of the vocabulary passed down from Old English. Though this classification is not strictly accurate, it

will suffice for the point to be made here, and we will therefore indicate this with a further line in the selectional restrictions in the lexical entry of -ity (in section 2.4 we will suggest a refinement of this treatment).

(9) 
$$-ity$$
 Select: [A\_]

where Adj = [+latinate] (and further lexical restrictions)

Lexical restrictions are especially common in derivational morphology, and there is some common terminology to be learned here. A morpheme is said to be **productive** if it is free from lexical restrictions. A morpheme that has lexical restrictions is said to be **unproductive** (or sometimes **semi-productive**). There is some variation in the use of these terms in the literature, but the basic notion is that a productive affix can be combined with any appropriate stem, including nonce/novel ones (since it has no lexical restrictions), whereas unproductive affixes cannot since novel stems cannot be in list of lexical restrictions (how would they get there?). This is not quite accurate, as we will see immediately below, but it will do for now. (Semi-productive affixes are those that are productive within a limited domain, for example, an affix might attach only to nouns denoting humans, but freely to such nouns).

Let's come back to -ity. In addition to selecting for specific adjectives, -ity also combines with some adjective-forming suffixes. Notable among these is the suffix -able which forms adjectives from (transitive) verbs. Thus: *learnable* "able to be learned"  $\rightarrow$  *learnability* "the degree to which something is able to be learned" (an important term in linguistics). Now, -ity itself is unproductive (the list of things it can attach to is actually a subset of the [+latinate] vocabulary), but, *-able* is fully productive. It can attach to essentially any transitive verb and can even attach to nonce words. Here's an illustration. Remember our friend the cat-owner from Chapter BC? We had concluded that, if her cat *kreels* regularly. The grammar of English (specifically, the fact that the agentive suffix *–er* is productive) licenses the inference that the cat is a *kreeler* (the inference is valid, even though we don't know what that *kreel* means). Suppose we extend this scenario and we tell you that: *My friend's cat kreels mice*. From this (still without knowing what *kreel* means) you are certainly able to conclude that *Mice are kreelable*.

It is, moreover, possible to then talk about the *kreelability* of mice versus, say, dogs. As English speakers, we know that this must mean "the degree to which something is able to be kreeled." In this way, *-ity* can attach to novel stems. However, it does so by virtue of selecting for a productive suffix; *-ity* is not productive, but inherits the productivity of *-able*.

This type of scenario is not at all uncommon cross-linguistically. Here is an example from XX

# EXX FROM OTHER LANG

Some further restrictions on morpheme combinations will be discussed in section XX below. Reverse dictionaries (such as Muthmann XX) are particularly useful in establishing the selectional restrictions of suffixes.

#### 2.2.3 Locality

Selectional restrictions (including lexical restrictions) are **local**. What this means is that the restrictions must be met in a specific configuration. One hypothesis about the constraint on selection is that it holds under the condition of **adjacency**. Consider the non-words *\*carefulless* and *\*carelessful* made up of the root *care* and the suffixes *-ful*, and *-less*. Each of these suffixes

may legitimately attach to the root *care*, yielding *careful* and *careless*, respectively. (It is also worth noting that both of these complex words can serve as the input to further derivation, cf. *carefully, carelessly*.) The requirement that selectional restrictions be met under adjacency correctly excludes the impossible forms. Due to this restriction, the root *care* cannot satisfy the selectional restrictions of both suffixes simultaneously. This is perhaps easier to see in the partially bracketed forms in (10).

- (10) a. \*care ]<sub>N</sub> ful ]<sub>Adj</sub> less
  - b. \*care  $]_N$  less  $]_{Adj}$  ful

Because of the adjacency condition on selection (as implied already in the notation we have used in (5) and (7)) only the innermost of the two suffixes in each of the examples in (10) has its selectional restrictions. The outermost suffix attaches not to a noun, but to an adjective, in violation of its selectional restrictions met. Imposing the adjacency condition thus correctly predicts the unacceptability of the words in (10) and any other words formed in an analogous manner.

The adjacency hypothesis will work well for sequences of affixes which are all aligned on the same side of the root, i.e., all suffixes or all prefixes. In some languages, all affixes are aligned on the same side in this manner. In English, though, we allow words to contain mixture of prefixes and suffixes. For this reason, we need to make the hypothesis more precise: are selectional requirements restricted to **linear** or **structural** adjacency?

### 2.2.4 Order versus Structure

Above, we have been concerned with suffixes. Thus, in talking about locality, we have talked about adjacency in a manner that does not distinguish between **linear adjacency**, and a separate notion, **structural adjacency**. The examples in (11) show that it is structural adjacency, and not linear adjacency, that is important in understanding the locality conditions on selectional restrictions. As always, the relevant conclusion is reached by comparing possible words with impossible words.

- (11) a. enjoyment, indecipherable
  - b. \*enjoyful, \*dethroneless, from (3d)

Consider first the words in (11a). The suffix *-ment* attaches only to verbs, yet in (11a) this suffix is linearly adjacent to the noun root *joy*. This combination alone is, of course, impossible: *\*joyment*. Similarly, the prefix *in-* selects for an adjective, yet it is immediately adjacent to the verbal stem *decipher*(which may in turn be internally complex, though this is not directly relevant). Again, *\*indecipher* is not a legitimate word. This is represented via the selectional restrictions in (12). The restrictions on the prefix *en-* in (12c) correctly block this prefix from attaching to *joyful* in (11b), as we will discuss further below.

- (12) a. *ment* [V\_] result: N
  - b. *in* [\_\_\_A] result: Adj
  - c. en [\_N] result: V

There is no big mystery here. Our theory will make the correct distinction if we can correctly express the intuition that *enjoyment* is not derived directly from *joy*, but rather from the verb *enjoy* (likewise, *indecipherable* is from *decipherable*).

The derivation of *enjoyment* must thus proceed in steps:  $joy \rightarrow enjoy \rightarrow enjoyment$ . A stepwise derivation of this sort is formally called **cyclic** (from "cycle") and the hypothesis that derivations proceed in this manner is thus called the **cyclictiy** hypothesis. The cyclic derivation of *enjoyment* is sketched in Derivation A in (13). Derivation B shows what goes wrong if the affixes are attached in the opposite order.

- - $([en [joy]_N]_V ment]_N ** CRASH **$

It is only if the prefix attaches first that the selectional requirements of each of the affixes under investigation are met. When selectional restrictions are not met, attachment is impossible and the derivation is said to "crash". At a crash, the derivation stops, no further derivation is possible.

The hypothesis of cyclicity together with the locality condition on selection requirements forces us to analyze the complex word *enjoyment* as being derived as in (13a). It turns out that this has some important (and useful) consequences in explaining other instances of unacceptable words.

Note first that we have been somewhat more precise in placing our brackets in (13) than we were above. Of particular importance is the placing of the brackets in the second line of (13). The first set of square brackets encloses the root *joy* and the label to its right indicates the category of the root, here a noun. The next set of brackets represents the result of attaching the prefix *en*-. The

result is a single **constituent** consisting of the prefix and the inner constituent (the root) together. The brackets thus contain the whole word *enjoy* and indicate that this whole word is a verb. By convention, category labels are written as subscripts on the right-hand bracket; this has no theoretical significance and could just as easily have been written on the left (note, for example, that the selectional requirements of *en-* are met in the second line, as the "N" refers to the pair of brackets that it occurs on). The selectional requirements for the suffix *-ment* are met in the third line of the derivation.

As in other branches of linguistics, labeled brackets indicate a **constituent structure**. That is, they indicate the groupings and subgroupings of the elements which make up a complex structure. Labeled brackets are convenient in many ways, but the same information may also be expressed via Word Structure Trees (WST). The trees corresponding to the words in (11a) are given in (14).

b.



Each **node** in the tree in (14a) corresponds to a bracketed constituent in (13a). Category labels are written at each node. Note that the root is the only terminal element dominated by a non-branching node. That is, the node dominating the root dominates only the root. Every higher node has more than one daughter. This is the way of expressing in trees what is expressed by the

innermost set of brackets in the first line of (13). Confusion can arise if this device is omitted, and the formulation of the percolation rules introduced below will depend on this way of drawing the trees. (In section 2.2.5.4, we will provide some more formalism regarding constituent structures, stated over trees, in particular, some definitions, conventions and well-formedness conditions.)

These examples show that structural adjacency, or the **sisterhood**, relation is a sufficient condition for the satisfaction of selection requirements. This suggests that something like (15) holds as a principle. We give two versions.

(15) The Sisterhood Condition

WEAK Selectional requirements can be satisfied under sisterhood.

STRONG Selectional requirements must be satisfied under sisterhood.

The difference between the two versions lies in whether sisterhood is taken to be a **sufficient** or a **necessary** condition on the fulfillment of selectional restrictions. This is an important distinction and for all theoretical conditions of this sort, this is a distinction that ultimately needs to be made explicit. In brief, a condition is said to be *sufficient* if satisfying that condition is enough for some expression to count as well-formed. This is a weak statement, because it says the restrictions *can* be satisfied in this manner, but it allows for there to be ways to satisfy selectional requirements other than sisterhood. Where sisterhood does not obtain, there may be some other way to satisfy selection. The stronger claim would be to say that the condition is a *necessary* condition; this would mean that sisterhood is the only way to satisfy selectional requirements, no other configuration is possible.

The strong sisterhood condition is more readily falsifiable, since it makes some specific predictions about impossible combinations of morphemes. To test this, we should consider examples where there is some plausible alternative relationship, but where sisterhood is not met.

Consider, for example, the words in (16).

- (16) a. *joy enjoy*, *joyful* 
  - b. throne dethrone, throneless
  - c. *bug debug*, *buggy*

The root *joy* may combine with the prefix *en*- or with the suffix *-ful*. If we were to reject the strong version of the sisterhood condition, and to assume that linear adjacency was a sufficient condition for the satisfaction of selectional requirements, then we would predict that the root *joy* should be able to combine with both of these affixes at the same time (a word like *enjoyment* shows that having both a prefix and a suffix is not excluded). This prediction (and the parallel one based on (16b)) is falsified:

# (17) *\*enjoyful, \*dethroneless, \*debuggy* (11b)

The strong sisterhood condition explains why these words are impossible. Consider how one might try to represent this in a WST (or the analogous labeled bracketing). Assume that all structures (at least in English) are binary branching. The relevant WSTs are given in (18).



On the one hand, if the derivation were to begin by combining the prefix en- with the root joy, the result would be a verb. The suffix -ful could not then attach to the derived constituent enjoy, because the selectional restrictions of -ful would not be met. On the other hand, if the derivation were to begin by combining the suffix -ful with the root joy, the selectional requirements of -ful would be met, but the result would be an adjective, and the prefix en- would then be unable to attach to the derived constituent, due to en-'s own selectional requirements. If our theory did not take sisterhood to be a necessary condition on the satisfaction of selectional restrictions, then we might make false predictions about words of this sort. Therefore, we tentatively conclude (in the absence of further arguments) that the strong version of the sisterhood condition is a principle of UG.

In the preceding paragraph, we assumed that branching in WSTs is binary. A "flat", i.e., ternary branching structure would not alter the basic point. An attempt to draw a tree as in (19) raises a number of problems.



First, it is not clear that the selectional restrictions are met. The selectional restriction of *en* is  $[\_N]$ , which can be read as entailing that the constituent containing *en* contains only a Noun, to the immediate right of *en*. The environment in which the prefix finds itself in (19) however is formally described as  $[\_N A]$ . Indeed, it appears to be the case that selectional restrictions always specify only one sister constituent: no affix has a restriction like  $[\_N A]$ . If this is indeed universally true, this would follow if all combinations are binary and would thus constitute a strong argument for binarity.

A second problem with (19) is that it is unclear what the resulting constituent would be. Recall that *en-* attaches to nouns and makes verbs, while -ful attaches to nouns and makes adjectives. Thus, (19) in effect will give conflicting instructions to the topmost node.

For reasons such as these, it is probably possible to derive the assumption that WSTs are binary from other aspects of the theory, at least for the basic cases. However, we will not pursue this any further here and will simply assume that trees are binary.

When we looked only at combinations of multiple suffixes, as in the case of *\*hopelyful*, we saw that we needed to state selectional requirements as part of the lexical entry for individual affixes.

(19)

The examples we have considered subsequently have shown that these requirements must be stated over hierarchical structures, and not over string (i.e, linear) adjacency.

Linear adjacency is neither necessary (*enjoyment*) nor sufficient (\**enjoyful*) for stating the selectional requirements of particular affixes. Structural adjacency, the relationship expressed as sisterhood in the tree structures, appears to be both necessary and sufficient. This accords well with much current thinking about selection in syntax, suggesting that this is indeed a deep principle of grammar. We will consider further consequences of this hypothesis in section 2.2.5.2.

# 2.2.5 Some consequences of constituennt structure

# 2.2.5.1 Structural Ambiguity

Constituent structures also allow us to provide a precise characterization of certain kinds of ambiguity. The words in (20) are ambiguous, as illustrated by the sentences following them.

# (20) a. undoable, unbuttonable, unfoldable, untieable

meaning 1: This exam is so hard, I think it's *undoable*. meaning 2: Don't worry about word processing mistakes, they're all *undoable*.

b. non-fictional, anti-symmetric

meaning 1: concerning the property of anti-symmetry	= [ [ anti- [ symmetry ] ] ic ]	
meaning 2: the opposite of symmetric.	= [ anti- [ [ symmetry ] ic ] ]	

The two meanings of *undoable, unbuttonable* etc. are precisely characterized by the internal constituency, that is, the order in which the prefix and suffix have attached. This is illustrated in (21). If the suffix is combined first as in (21a), the result is an adjective, with a meaning roughly paraphrasable as 'able to be done'. To this, the prefix *un*- may attach, giving the negative adjective, meaning [not [able to be done]]. This is meaning 1 in (20). The WST in (21b) represents the other order of attachment: *un*- attaches to the verb, giving a verb of reversal  $do \rightarrow undo$ . As this is still a verb, it can be the input to *-able* suffixation, with the meaning 'able to be undone', which is meaning 2 in (20).



Note that in some cases, the two structures correspond to meanings that overlap quite substantially, as is the case with *anti-symmetric*.

The structural ambiguity is thus the same as other structural ambiguities in language, such as *English linguistics professor*. This can refer to a professor who teaches English Linguistics (even if the professor herself is not English), or it can refer to a Linguistics Professor who is English, even if that professor doesn't actually teach English Linguistics. The structural ambiguity is reflected in the two trees in (22). In many departments, the two meanings clearly refer to different individuals.

(22) Structural Ambiguity



A teacher of Linguistics who is English A teacher of English Linguistics It is worthwhile noting that English, though often said to be a morphologically poor language, plays an important role here in establishing that selectional requirements must be met under sisterhood, rather than linear adjacency. In languages with only suffixes (or only prefixes) the argument cannot be constructed, this is possible only in languages, like English, which have a mix of prefixes and suffixes.

#### 2.2.5.2 No Lookahead

Above, we introduced the Sisterhood Condition (15) and saw that this was a useful theoretical principle that distinguishes between possible and impossible words. In this section, we will consider another consequence of the (strong) sisterhood condition.

One assumption that we have maintained throughout is that the derivation of a complex word proceeds *cyclically*, that is step by step in the manner discussed above. Combining this assumption with the strong sisterhood condition yields another prediction, which can be called the *No Look-Ahead Corollary* which can be stated as in (23).

(23) The No Look-ahead Corollary

No morpheme can impose selectional restrictions that refer to more peripheral affixes.

or, equivalently:

At every node, all selectional requirements must be met

The intuitive content of the no look-ahead corollary is the following. At each point that an affix is added, the derivation checks to see if the selectional restrictions of that affix are met. The No Look-ahead Corollary states that the derivation fails (and the affix cannot successfully combine)—the technical term, from syntax, is "crashes"—if the selectional restrictions are not met at that point. The derivation cannot "wait" to see if some subsequent affix might be added which will then satisfy selectional requirements.

Like the sisterhood condition, no look-ahead is about the **locality** of relationships among morphemes.

No Look-ahead (23) excludes any selectional restriction of the form:  $\_$  ] X ]<sub>X</sub> (where X is some specific morpheme or morphological property).

So what work does this do in real terms? Well, for one thing it entails that no morpheme can be specified as attaching to a verb only if a subsequent nominalizing suffix will be added later in the derivation. Another type of scenario that is disallowed by No Look-ahead is a pattern that would be the reverse of the -ity and -able combination discussed above. In English, -ity inherits the productivity of the inner suffix -able. No look-ahead entails that it is impossible for an inner suffix to inherit the productivity of an outer suffix. There can be no suffix -ITY' which has severe lexical restrictions, but which restrictions are lifted when it is followed by a more

productive suffix. The excluded situation is not illogical, but it seems not to arise. The No Lookahead Corollary explains its absence.

INCLUDE SOME NO LOOKAHEAD EXAMPLES FROM OTHER LANGUAGES, e.g., Simpson & Withgott; others...

Note that No Look-Ahead is a corollary, that is, it is not a principle which need be stated explicitly in the theory. It follows automatically from the strong interpretation of the Sisterhood Condition in (15), together with the cyclicity hypothesis. If the derivation is cyclic, and all selectional requirements must be met under sisterhood, then (23) follows and need not be stated as an independent condition.

For the same reason, if No Look-Ahead is shown to be wrong (as some linguists have contended), then the strong sisterhood condition and the cyclicity hypothesis cannot both be right. One or the other would have to be weakened or abandoned. In this way, you should begin to see how the pieces of the theory hang together. By making the assumptions and hypotheses relatively precise, a delicate web is woven, and a change to one part has implications for the overall structure. This allows for fairly precise predictions to be stated and tested, and allows for a clear understanding in some cases as to what part of the theory is in need of revision when a prediction fails.

In practice, then, when proposing changes to one part of a theory, linguists must be attentive to the consequences of those changes for other parts of the theory.

#### 2.2.5.3 Some exercises

Provide labeled WSTs for the words in (A).

For the unacceptable combinations of morphemes in (B), identify what condition or conditions are violated.

Using a reverse English dictionary or the intuitions of a native speaker, write lexical entries including selectional restrictions for the English affixes: *–er* (meaning: 'one who does X'), *-ize*, *-dom*, and *re-*.

# 2.2.5.4 Some formal aspects of constituent structures (optional)

Our purpose here is to provide only as much formalism as needed for the discussion at any point. This section therefore has a dual purpose, but is somewhat of an aside to the main text. The first purpose is as a refresher on (or introduction to) tree-drawing, though it should also be useful in making explicit the conventions we will use throughout the book. The second purpose is to show how the formalism allows one to derive empirically testable predictions. We will illustrate a few here in the form of conjectures representing commonly held positions, though we will not explore them in any detail.

#### DEFINITIONS

THIS SECTION IS INCOMPLETE. Refer to Borsley, Robert *Syntactic Theory* pp. 21-23 (inserted here in the course pack), paying special attention to: (Dominance, Immediate Dominance, Constituent, Sister, and the triangle notation for abbreviated trees on p.21).

Using trees is useful in developing a formalism because it constrains what we can propose, and how we have to formulate it. Graph theory (a branch of mathematics) can prove theorems about the trees and this tool can thus be used to explore the formal properties of morphological theories, and in principle to discover unnoticed problems and state and test hypotheses as the theory becomes more complicated.

#### Conventions

There is an unfortunate piece of terminological overlap here. In graph theory, especially as used in syntax, the "root node" is defined as the node which dominates all other nodes, i.e., the topmost node. This is potentially confusing, since in morphology, the "root" is generally the most deeply embedded node. In order to minimize confusion, we will use "root" in this textbook only in the morphological sense, and refer to the topmost node in a WST as the "topmost" node.

Root: A terminal immediately dominated by a non-branching node.

- Morpheme: A terminal. Recall that using the phonological (or orthographic) form of terminals is a shortcut. The terminal is really the lexical entry.
- (More advanced: the terminal can be seen as a *token* of a lexical entry / morpheme, not the entry itself, this is necessary to allow multiple occurrences of a single morpheme, for example. The important point is that the terminal includes all the features, not just the phonology).
- (for later: compound node: a node immediately dominating two non-terminals; compound: a node dominating two or more roots)

These definitions allow us to formulate hypotheses succinctly, and in a manner that we can communicate to other linguists, so that we can be sure that we are talking about the same thing.

For example, one widely held assumption is that there are no words consisting solely of affixes. This can be stated in either of two ways.

(24) The rootedness hypothesis

a. WEAK Every **topmost** node dominates a root morpheme.

b. STRONG Every node dominates a root.

(NB. dominance  $\neq$  immediate dominance).

Both the weak and the strong versions of this hypothesis exclude words that consist only of affixes. [Show this.] Nevertheless, they are not equivalent and differ on another point. Can you see what the difference is? Of the two, the strong (but not the weak) rootedness hypothesis also excludes complex affixes, as in (25).

(25) a. Excluded by (24a-b) b. Admitted by (24a), excluded by (24b)



This is an empirical issue on which there is as yet no firm consensus. However, the formalism we have adopted allows us to frame the issue clearly, to ask the relevant questions, and to make explicit what the consequences of the empirical investigation will be.

# 2.2.6 EXERCISES

#### 2.2.6.1 Generativity

Can you show that the impossibility of the words in (2)-(3) is not a result of memorization? For example, what argument would you give to someone who says that these words are not words of English simply because they are not listed in the dictionary?

Hint: consider the use of *nonce* words to make arguments about the generative nature of grammar in the previous chapter. Recall our friend from page xx. If we were to tell you more specifically that she earns her living by *kreeling* mice, you would not only be able to conclude that she is a *kreeler*, but also that mice are *kreelable*. Moreover, if *kreeling* is a process that can only be done to mice, then everything else must be *unkreelable*. Have you ever seen these words before? Do you note a difference between these words and the unacceptable stings: *\*ablekreelun*, *\*unablekreel*, *\*kreelunable*, etc.?

# 2.2.6.2 Bracketing = WST practise

#### **PRACTICE EXERCISES:**

- 1. Draw Word Structure Trees and Labeled Bracketings for each of the words in XX. Label each node in the tree and each right bracket with the appropriate syntactic category.
- 2. Using only the data given, for each of the affixes in XX, state its selectional restrictions using the notation indicated.
- 3. Repeat 1-2 with data from an 'exotic' language.

### 2.3 headedness and percolation

#### 2.3.1 Nodes, constituents and words

In section 2.2 we introduced Labeled Brackets and Word Structure Trees as devices for expressing constituent structure in words. Each node in the tree defines a *constituent*, and each constituent bears a label indicating its syntactic category.

In English, it happens that most well-formed constituents are also acceptable as words. But it is important to note that this is not a general requirement of UG. In many languages certain classes of words must meet various requirements above and beyond the selectional restrictions of individual derivational morphemes. For example, in inflected languages (see chapter XX, below), all verbs must bear some marking of their role in the sentence, such as having affixes expressing tense and/or agreement. This is true, for example, of Inuktitut—noun root and stems may be used in their bare form, but the root or stem of a verb may not be. This is shown in (26).

- (26) a. pisuk- 'to walk' (bound, not a word)
  pisuktuŋa 'I am walking' (can be used as a word)
  pisukpit 'Are you walking?' (can be used as a word)
  pisuktualuuqqauyami 'because he walked a lot'
  pisuyit 'walk!' (the velar is voiced by regular phonology here).
  \*pisuk
  - tuktu N. caribou. can be used as a word

(For the purposes of the following discussion, we will make two simplifications. First, we will treat the string *-tuya* as if it were a single morpheme, an inflectional suffix meaning [first person singular subject, present tense]. It actually has internal complexity, but this is not relevant here. Second, we will treat the entire inflected word as having the syntactic category Verb. Note that it can be used as a full sentence, since Inuktitut has the common property that pronouns can be freely omitted.)

The requirement that verbs bear inflectional suffixes is true of verb roots and equally true of derived verb stems. Consider the word in (27a). This is formed with the suffix -siuq which has the partial lexical entry in (27c).

(27) a.	tuktusiuqtuŋa	'I am hunting caribou'	(V)
b.	*tuktusiuq	not a word	
	*siuqtuna	noot a word	
c.	Phonology:	/siuq/	
	Meaning:	'to hunt X'	
	Selection:	[ N ]	
	Cat. Result:	Verb	

The word in (27a) is represented by the WST in (28).



In this tree, the lowest node corresponds to the noun root *tuktu* which is a free root, and hence a simple word. The topmost node is a verb, the complex word in (27a). The intermediate node constitutes a constituent, specifically, the verb stem to which the inflectional suffix has attached, but the intermediate node does not itself constitute a word of Inuktitut, see (27b).

Although it is neither a word nor a morpheme, [*tuktusiuq*] it is a constituent and more specifically, a verb. We know this because this constituent can serve as the input to further derivation and can combine with suffixes that select for verbs.

pisuy-vik 'place for dancing'  $[k \rightarrow \gamma = \text{phonology}]$ 

tuktu-siuR-vik 'place for hunting caribou'  $[q \rightarrow R = phonology]$ 

### \*tuktu-vik violates selection

The fact that it has syntactic category Verb also explains why the intermediate constituent is not a word. As noted above, all verbs in Inuktitut, in order to be used as words, require inflectional affixes, and thus *tuktusiuq* cannot be used as a word until further inflectional morphology is added.

In conclusion, we see that WSTs indicate constituency. Each node constitutes a constituent and each constituent has a label expressing information about that constituent. Although most constituents constitute words in English, this is not a requirement of UG, but rather a reflection of the general property of (Modern) English that the language does not have much inflectional morphology. In other languages, intermediate nodes frequently do not correspond to words.

Aside: Why does this example not constitute a problem for the theory we have been developing, specifically for No Look-ahead? In Inuktitut, like English, nouns do not always need to bear inflectional affixes (except in certain cases), and so the noun root corresponds to a simple word. When the suffix *-siuq* is added, the category of the complex consitutent is a verb. The selectional restrictions of -siuq are merely [ N \_\_ ] and so those properties are satisfied at the intermediate node. But the general requirement on verbs bearing inflection is not met. Looking at the picture as a whole, we see then that the No Look-ahead Corollary is very precise in what kinds of relations it does, and does not, admit of. It is possible for the addition of an affix to necessarily trigger the addition of other affixes, but only for general reasons, and not for reasons specific to the selectional properties of the inner affix. In Inuktitut there is a consequence that is somewhat like look-ahead—adding a verbalizing affix is possible only if the proper inflection is added subsequently—but what we see is that the theory admits this, because this is a general property of verbs, and not a specific property of the suffix -siuq. Taken as a whole the theory makes quite a specific prediction: No Look-ahead will hold only of selectional restrictions, i.e., restrictions that must be stated as part of a lexical entry. Various refinements of this position have been

proposed in the literature, but at the level presented here, this consequence of the theory is essentially correct.

#### 2.3.2 Node Labels

In each tree, every node has a label, indicating the syntactic category of the constituent it represents. Thus far, we have simply put the appropriate labels on the nodes from our knowledge of English words. This works well in English, where intermediate nodes typically correspond to words, but our theory should be able to characterize the knowledge that we have that allows us to perform this task. This is particularly important in languages where intermediate nodes do not correspond to well-formed surface words, as discussed in section 2.3.1., since there we cannot simply consult intuitions directly to know the category of some constituent that is not a word. In this section, we will explore in more detail how it is that the nodes are labeled, formally.

Let us start with bi-morphemic words, taken from previous sections. Consider the WSTs in (30).



Start with the lower node, the one immediately dominating the root. The category of this node is always that of the root, as given in the root's lexical entry. We indicate this by writing the category information from the root's lexical entry under the terminal node, and drawing an arrow to the label on the immediately dominating node. The purpose of the arrow is thus to indicate the source of the information displayed at that node.



Now let's look at the topmost node. How do we know what category it has? Does the following set of examples help us to answer this question?

(32) a.	English	b.	Russian
	true Adj		čit- V 'read' (bound root)
	truth N		čit-at <sup>j</sup> V 'to read' (infinitive)
	truism N		čt-enije N '(a) reading'
	truthful Adj		čit-ajuščij Adj 'reading'
	thruthfulness N		

The root, it should be clear, does not determine the syntactic category of the entire word. That category is determined by the (outermost) affix. Similarly, in most cases (though see below) words ending with the same affix will have the same syntactic category. It is worth noting here

that even words with fairly idiosyncratic selectional properties nevertheless retain the basic category of the affix. For example:

(33) marriageable --attaches to N, not V, but still an A < able

hatter --attaches to N, not V, but still an N  $\leq er$ 

thankful --attaches to V(?) not N, but still an A  $\leq ful$ 

Indeed, we have recognized the dominance of the affix in complex words implicitly above by including "category of result" in the lexical entry of each affix. The category of result for an affix plays exactly the same role in the system as the syntactic category does in the lexical entry of a root. Specifically, it determines the category of the node that immediately dominates it. This information can thus also be entered into the tree by means of an annotation at the bottom, and an arrow:



If we keep to these examples, we see that the category label of a node is always determined by the morpheme that that node immediately dominates. This is always unambiguous—no node immediately dominates more than one lexical entry. We formalize this as the Node Labeling Convention (henceforth, NLC) as in (35). For the sake of transparency, we have also indicated in the trees that it is the NLC that licenses the labeling we have provided, by annotating each of the arrows.

#### (35) The Node Labeling Convention (NLC)

The features of a node are those of the lexical entry it immediately dominates.

The NLC is a hypothesis regarding the flow of information in the derivation of a complex word. Such conventions (we will see more below) arise from the observation that constituents in language (not just in morphology) are typically asymmetric. They consist of one prominent member that determines properties of the whole constituent, and one less-prominent member. There is special term for the prominent member of a constituent: the HEAD. For morphology, then, we define the head as:

# (36) Head

The head of a constituent is the member of the constituent that determines the properties of the whole.

When we put this together with the NLC, it turns out that in all the cases discussed so far, an affix is always the head of its immediate constituent. This will turn out to be largely true, with certain principled exceptions to be discussed in the next sections and chapters.

Conventions such as the NLC that determine headedness within morphological constituents are often referred to as **percolation conventions**. They are very similar to **projection** in syntax (the head of the V[erb] P[hrase] is the Verb, since it is the verb which determines the syntactic category of the larger constituent). Indeed, one school of thought holds that there is no difference

between morphology and syntax. This theory being developed here is consistent with this approach, but we will not dwell on the issue at this point. We will take this point up again in Chapter C.

### 2.3.2.1 Affixes as lexical items

At this point, our theory is ready for a quick simplification, a bit of notational housecleaning as it were. Above, we have seen that the "category of result" information in the lexical entry of an affix plays formally the same role as the "category" information of a root. We will therefore simplify our lexical entries by simply writing "category" in all cases (i.e., for roots and affixes).

	Category:	Noun	Adjective
	Selection:		[ N _ ], where N is a noun of emotion
	Meaning:	$\odot$	'having X'
	Phon:	/dʒo <sup>j</sup> /	/ful/
	Label:	joy	ful
(37)		"Root"	"Affix"

In a sense, this means that we are treating the affixes as having categories themselves. The leading idea in doing this is that affixes are essentially the same as other lexical items, i.e., 'signs in the Saussurean sense, distinguished only by their morphological requirement of needing to attach to something.

This line of reasoning, taken to its logical conclusion, suggests that whether something is a root or an affix is a language-particular property, i.e. one that must be learned. That is, we expect that a single meaning can be expressed as an affix in one language but as a separate word in another. In a wide range of cases, this is true.

(38) Inuktitut: adjectives like *big*, *fake*, verbs like *try*, modals, are affixes; *siuq* above

(It is not expected that simple concepts like *dog* will be expressed as affixes, though. The reason is that affixes always select for their complement, and their meaning reflects this. Thus modifiers and verbs taking complements make good affixes, but it would be hard to see what an affix meaning *dog* could reasonably be expected to select for.)

There is another important implication of making this simplification, especially when we consider the sisterhood condition on selection.

By treating affixes as just like roots in terms of their category, we are implicitly assuming first that the category of the complex word is uniquely determined by the affix, and second, that the range of categories of simple words should be the same as that of complex words, derived by affixation.

#### Is this correct?

It appears to be. For the purposes of syntax, an adjective is an adjective, and it does not matter what (if any) the derivational history of that adjective is. Thus, any of the adjectives in (39b) may occur in the frame in (39b)—a frame which admits only adjectives.

(39) a. That movie was very \_\_\_\_\_.

- b. [sad]<sub>Adj</sub>
  - $[[joy]_N ous]_{Adj}$

[[laugh]<sub>V</sub> able]<sub>Adj</sub>

 $[ [ en [ joy ]_N ]_V able ]_{Adj}$ 

This is a very interesting aspect of the theory that emerges as a consequence of the manner in which we have chosen to formalize it. In particular, the sisterhood condition and the assignment of affixes to syntactic categories, when taken together, yield another corollary, which has been called the *Atom Condition* (Williams). This corollary is the prediction that derivational history/internal complexity will not be relevant for selectional or distributional restrictions that appeal to syntactic category. Here, concretely, are some of the logically possible things that are predicted never to occur.

- (40) a. No affix, in any language, will be specified as attaching only to Adjectival stems that contain a Noun (somewhere) inside them.
  - b. No determiner (or other distributional frame) will select only nouns that are internally complex.
  - c. Etc.
- 2.3.2.2 Some exercises with the NLC

Provide the missing labels in the following trees.



# 2.3.3 Backup Percolation

Consider now English words with the prefix non-:

(41) a.	non-invasive	non-trivial	non-essential
b.	non-smoker	non-fiction	non-proliferation

This prefix is interesting because it appears to select for both adjectives (41a) and nouns (41b). Other prefixes with a similar behaviour include *anti-*, *semi-*, etc. Now, what syntactic category might we assign to such a prefix?

To determine this, we must look at the syntactic category of the complex words that include the prefix *non*-. What we see is that the combination of *non*- plus an adjective yields... an adjective and the combination of *non*- plus a noun yields... a noun. Representative WSTs are given in (42).

(42) a.





If the NLC were the only percolation convention we had, then we would have to posit either two homophonous prefixes *non*-, or some kind of split lexical entry as (43).

(43) *non-*:

phonological shape:	/nc	ın/
meaning:	not	X
selection	[N_]	[A_]
category	Noun	Adjective

Positing homophony in this way misses an important generalization. Specifically, *non*- adds meaning, but does not change category. What we would like the theory to formalize is that unlike the affixes above, *non*- allows the node dominating it to inherit the feature/category of the lower node (in apparent violation of the NLC).

To implement this, we will first suggest that *non*-, unlike the affixes considered above, has no syntactic category. Following standard practise in linguistics, we write the zero sign " $\emptyset$ " to indicate this, in the syntactic category part of the lexical entry. It is important to note that  $\emptyset$  is not a category feature, rather, it is a notational device to indicate the absence of a feature. Thus percolation conventions will not copy zeros.

(44) *non*:

phonological shape:	/ nan /
meaning:	not X
selection:	[ {N,A} _ ]
category:	Ø

If we try at this point to draw a tree for a word containing the prefix *non*-, we will find that the NLC can not apply. Therefore, we must introduce a second convention which takes over when the NLC does not apply.

(45) BACKUP PERCOLATION CONVENTION (BPC)

If only one daughter of a given node  $\alpha$  has feature specifications, then the node  $\alpha$  inherits the features of its specified daughter.

The application of the two conventions together is illustrated in (46):

(46) Applying the Conventions:



The BPC allows us to incorporate into our theory an adequate characterization of the behaviour of affixes that add meaning but do not change category. Drawing again the parallel to syntax for those who find it useful, affixes like *non*- have a formal status quite similar to adjoined modifiers (like adverbs). Building on this parallel, we will henceforth refer to constituents in which the NLC applies as **head-complement structures** (where we have defined head above), and constituents in which the BPC applies as **modification structures**.

# 2.3.4 Gender and the percolation conventions

Thus far, we have introduced two labeling conventions, which closely parallel the syntactic notions of head-complement and adjunction-modification in their effects. These labeling conventions have been used to manipulate the flow of information in the WSTs. We have to this point only been concerned with one type of feature, however, namely, the feature of syntactic category.

We will now see that the theory works the same when we consider other types of grammatical features, for example, grammatical gender.

#### 2.3.4.1 Grammatical gender I

In many languages, nouns are divided into two or more morphological classes. In some languages such as many Indo-European languages, the classification is at least partly aligned to natural gender, and hence the classes are referred to as **genders**. Noun classes (including genders) play little (if any) syntactic or semantic role, but they are quite prevalent in morphology, determining agreement and often determining allomorphy on inflectional affixes on nouns (see chapter Infl).

Let's look at an example. Singular nouns (including place names) in Russian come in three genders: masculine, feminine and neuter. Consider the following data (the following conventions are observed: c = IPA / ts/; transcription broadly phonemic, in particular stress-related vowel alternations are not indicated and palatalization not consistently indicated).

(47) leningrad vjetnam 'Leningrad' [Masc]
 wietnam 'Masc]
 amerika 'America' [Fem]

italija 'Italy' [Fem] marokko 'Morocco' [Neut] meksiko 'Mexico' [Neut]

Russian also has suffixes that can be added to place names to create words meaning a person from that place (compare English: America  $\rightarrow$  America-n, London  $\rightarrow$  London-er, etc.). There are distinct suffixes for males and females from a given place. These suffixes are illustrated in (48).

(48)	leningrad <sup>j</sup> ec	'Leningrader' [Masc]	leningradka	'Leningrader' [Fem]
	vjetnam <sup>j</sup> ec	'Vietnamese'[Masc]	vjetnamka	'Vietnamese' [Fem]
	amerikan <sup>j</sup> ec	'American' [Masc]	amerikanka	'American' [Fem]
	italjan <sup>j</sup> ec	'Italian' [Masc]	italjanka	'Italian' [Fem]
	meksikan <sup>j</sup> ec	'Mexican' [Masc]	meksikanka	'Mexican' [Fem]
	marokkan <sup>j</sup> ec	'Moroccan' [Masc]	marokkanka	'Moroccan' [Fem]

Our main interest in the data in (48) is in the distribution of gender features. However, there are a few minor points we need to take care of before we can get to gender.

The first thing to notice about the data in (48) is that there is some allomorphy here—after a vowel-final stem the affixes begin with an n, whereas that n is not present after consonant-final stems. As this is not fully predictable from Russian phonology, we must include this in the lexical entries of the affixes. Notice also the label we have given to each morpheme. Recall from chapter XX, that the label has no particular theoretical relevance, but is a convenient way to refer to a morpheme that has more than one phonological form. In this case, we have used a standard notation in the label, enclosing the n in parentheses. This indicates that the n is sometimes present and sometimes not. The allomorphy section of the lexical entry gives the distribution of the n.

exical Entries	:			
"label"	-(n) <sup>j</sup> ec	-(n)ka		
phon. /	[n <sup>j</sup> ec] / V	[nka] / V		
allom.				
meaning	someone from N	s.o. from N	-	
position	suffix	suffix	-	
attachment	[N_]	[ N _ ]		
category	Noun [Masc]	Noun [Fem]		

There is some further allomorphy in the data in (48), in particular, the final vowel of the stem of the neuter nouns undergoes a change from *o* to *a*. Phonological alternations in the stem (called **STEM CHANGES**) will be dealt with in a later chapter, so for now, we will simply note them in the interests of accuracy and move on. In other words, in addition to the lexical items in (49), a complete account of the data in (48) must include the information in (50).

### (50) Stem Changes

(49)

 $/o/ \rightarrow /a/$  in the stem, before these suffixes

Now that we have identified the allomorphy in the data, we are ready to look at how information about gender is distributed among the words and constituents in the Russian examples. Of specific interest is the gender of the nouns designating the people from the various place names. What gender do they have? Regardless of the gender of the place name, the noun for a man from X is always masculine and the noun for a woman from X is always feminine. In other words, gender is always determined by the affix, not the stem, just like syntactic category was determined by the affix and not the stem. Again, we see that a feature-bearing affix is the head of the constituent in which it finds itself.

This is, of course, exactly what our theory—more specifically the NLC—predicts. Trees are sketched in (51) illustrating cases where the gender of the affix is different from the gender of the stem, and it is clearly the gender of the affix that determines the gender of the whole word .

(51) Gender and the NLC



The trees work exactly the same in words like *italjanka* 'Italian person (Fem)' when the feminine suffix attaches to a feminine root, see (52):

(52)



How do we know that the feminine feature of the whole word in (52) comes from the suffix and not from the root? This follows from the definitions of NLC and BPC and the lexical entry for the suffix -(n)ka. This suffix must have the feature feminine as a part of its lexical entry, in

order to explain the gender of *leningradka*, *meksikanka* in (48) and other words built from masculine or neuter roots. Since the feminine feature is a part of the lexical entry of -(n)ka, the NLC must apply in (52)—the BPC only applies when the affix lacks features. Or, to put it the other way around, in order for the BPC to have applied in (52), the suffix -(n)ka would have had to lack a gender feature. But if it lacked a gender feature, then we could not explain why the suffix changes gender in the other examples in (48).

# 2.3.4.2 More grammatical gender

It may seem that the Russian example is somewhat obvious, since in the words for people, grammatical gender corresponds with natural gender. While this is certainly a trend, it is by no means always true in Russian that natural and grammatical gender coincide. The words in (53) usually trigger masculine agreement (for example, within a modifying adjective) whether they refer to a man or woman.

(53) agent 'agent'

vrač 'doctor'

Note, though, that we can also replicate the point made in the previous section with inanimate and abstract nouns, for which natural gender is irrelevant. Again, we will see that the gender of the affix always determines the gender of the word, exactly as the NLC predicts.

To make the point, we will turn to German, which, like Russian, has three genders (Masculine, Feminine and Neuter) and in which natural gender and grammatical gender don't always coincide.

buxgalter 'bookkeeper'

Some simple nouns in German are given in (54) along with their corresponding **diminutives**. (The diminutive form is used to refer to small things or things perceived as cute, for example, as when talking to children. English has a few diminutives, such as  $dog \sim doggy$ , which although not very productive, have a similar flair to the much more productive German diminutives. Transcription notes.)

(54)	Simple Noun	Diminutive	Meaning
	esel (Masc)	eselçen (Neut)	'donkey'
	frau (Fem)	frauçen (Neut)	'woman'
	kint (Neut)	kintçen (Neut)	'child'

In (54), it is straightforward to isolate a suffix *-çen* which carries the diminutive meaning. Any word with this suffix has neuter gender, regardless of the gender of the stem. Neuter gender must be a part of the lexical entry of the diminutive suffix, and by the NLC, will always determine the gender of the stem it attaches to.

Another common suffix in German makes abstract nouns from adjectives. This suffix has two allomorphs (*-hait*  $\sim$  *-kait*), and is illustrated in (55). Note that adjectives in German do not have inherent gender. That is, they do not have gender as a part of their lexical entry, though in certain syntactic configurations they may **agree** with the noun that they modify. We will learn more about this in chapter XX. This suffix always makes feminine nouns. There is no a priori reason for this (not all abstract nouns are feminine in German, let alone cross-linguistically), it is simply a fact about this particular suffix that it has the feature feminine and contributes it (by the NLC) to the constituent containing it.

# (55) Adjective Noun

gesund 'healthy' gesundhait (Fem) 'health (or more literally healthiness)' tsertliç 'tender tsertliçkait (Fem) 'tenderness' laŋvailik 'boring' laŋvailikkait (Fem) 'boredom'

The words in (56) are particularly interesting in this regard.

(56) a.	eis	farbe	fluss
	'ice'	'colour'	'flow, river'
	Noun, Neut	Noun, Fem	Noun, Masc
b.	eisik	farbik	flussik
	'icy'	'colourful'	'fluid'
	Adj	Adj	Adj
c.	eisikkait	farbikkait	flussikkait
	'iciness'	'colourfulness'	'fluidity'
	Noun, Fem	Noun, Fem	Noun, Fem

As expected, all of the words in (56c) have feminine gender. They are formed from the adjectives in (56b) by adding the suffix we have been discussing, namely *-kait*. However, the words in (56b) are in turn formed from nouns, which may have any of the three genders of German. The gender of the root, though, does not bear on the gender of the *-kait* nouns. Thus, even though the root *eis* is neuter, the word meaning 'iciness' is feminine, because of the suffix.

As an exercise, we suggest you stop reading at this point and try to draw the trees for the words in (56c), including category and gender information in your labels.

Have you drawn your trees? Your tree for the word *eisigkeit* in (56c) should like this:

(57)



The theory tells us how which features are assigned to which constituents. Since each node corresponds to a real word in German, it happens that we can check the results of the theory, and we find that they are correct (see (56)). There is one point of interest here, namely the intermediate node. Why does the adjective *eisig* not have a gender feature?

It is a property of German adjectives, already noted, that they do not have inherent gender, but why does BPC not apply to the gender feature from the root, bringing it up the tree at least as far as the node marked 'Adj'?

One way to think about this brings us back to the notion of headedness, discussed above. The sufffix -ik has a categorial feature, and thus is the head of the constituent [eis ik]. Because this is a headed constituent, all and only the features from the head percolate. A constituent cannot have more than one head. The BPC is relevant only when the affix is not the head of the complex constituent, and thus contributes no features. Following this line of reasoning to its conclusion leads us to expect that there will be no mixing of features, that is, no constituent can take some

features form one of its daughter nodes and other features from another daughter. Whether or not this is ultimately correct will depend on the treatment of phenomena not discussed here, and we will simply assume that this is correct for the remainder of this chapter.

# 2.3.4.3 Gender and the BPC

The discussion above leaves open a quite specific role for the BPC with respect to gender. Specifically, it should be possible to find suffixes which may combine with nouns, and which do not have category features of their own. These are predicted to be transparent for gender.

To see how this works, let's turn to data from Romance languages such as Italian, beginning with the adjectives in (58). Italian has two genders, feminine and masculine. The gender (and number) is indicated by a **theme vowel**, a special type of inflectional suffix on the adjective.

(58) sg.	bella bello	'good-looking' (fem) 'good-looking' (masc)	stupida stupido	<pre>'stupid' (fem) 'stupid' (masc)</pre>
pl.	belle	'good-looking' (fem)	stupide	'stupid' (fem)
	belli	'good-looking' (masc)	stupidi	'stupid' (masc)

Our basic process of segmentation allows us to isolate the roots and theme vowels clearly:

(59) a. roots: bell- 'good-looking' [Adj] stupid- 'stupid' [Adj]

The adjective roots have no inherent gender of their own. The theme vowel marks **agreement** with the noun that the adjective modifies. Agreement will be discussed in Chapter Infl. For the present, note that our segmentation in (59) is reinforced by the possibility of having other

suffixes between the root and the inflectional suffix. For example, the modifying suffix *–issim* 'very' may occur in this position:

(60) sg. bellissima 'very good-looking' (fem) stupidissima 'very stupid' (fem) bellissimo 'very good-looking' (masc) stupidissimo 'very stupid' (masc)
Nouns in Italian do typically have inherent gender, which like German and Russian may or may not correlate with natural gender. Most nouns in Italian, like adjectives, also consist of a bound root plus a vowel reflecting gender and number. Some noun roots, typically those referring to people, may be of either gender, and combine with the same set of vowels as the adjectives do:

(61) sg.	ragazza 'girl' (fem)	zia	'aunt' (fem)
	ragazzo 'boy' (masc)	zio	'uncle' (masc)
pl.	ragazze 'girls' (fem)	zie	'aunt' (fem)
	ragazzi 'boys' (masc)	zii	'uncle' (masc)

Our principles of segmentation thus lead us to conclude that these nouns are bi-morphemic, and that there are bound roots, *ragazz*-'young person' and *zi*- 'sibling of parent'. It is important to note that the fact that we do not have a single word for the meaning of the Italian root *zi*- does not prevent us from positing this root, nor from assigning a meaning to it.

In fact, almost all nouns in Italian (except for borrowings) have this bi-morphemic structure, even if though a great many have a specific gender as part of their lexical entry, and thus combine only with two of the four vowels. (Note that Italian speakers need not be consciously aware of this structure in their nouns, but the fact that they use their grammar in a way that makes direct reference to this structure shows that they have unconscious knowledge of the structure.)

- (62) a. masculine libr-o libr-i 'book(s)'
  - b. feminine scatol-a scatol-e 'box(es)'

Since roots of this sort occur with a fixed gender, it must be that the information about gender is stored in the lexical entry of the root, even though the property is expressed on a separate morpheme—the theme vowel. This distributed realization of information is in fact not at all uncommon. In German, discussed above, or French (related to Italian), every noun has gender, but the gender feature is not expressed on the noun itself, rather it can be detected only by looking at other words in the syntax which agree with the noun, such as the determiner, or adjectives.

Since agreement is not our interest at the moment, let us make a simplifying assumption about how it works. For Italian, we will say that the theme vowel must agree in gender features with the features of its sister node.

(63) a. TREES

Now consider Italian diminutives. Diminutives in Italian differ from German diminutives in that they do not contribute a gender feature. Thus:

### (64) Italian nouns and diminutives:

scatola	'box' (fem)	scatoletta	'little box' (fem)
camicia	'shirt' (fem)	camicietta	'little shirt' (fem)
libro 'book	x' (masc)	libretto	'little book' (masc) <sup>1</sup>
libri 'book	(masc)	libretti	'little books' (masc)
vaso 'vase	'(masc)	vasetto	'little vase' (masc)
vasi 'vase	s' (masc)	vasetti	'little vases' (masc)

<sup>&</sup>lt;sup>1</sup> Also means text to an opera.

Unlike German, the gender of Italian diminutives is (usually) determined by the gender of the stem. The diminutive suffix must therefore not have any gender feature of its own, and must therefore have the lexical entry in (65).

(65)

```
/ - ett - / -
diminutive
suffix
attaches to
noun root
features : Ø
```

The trees below show how the correct predictions about the gender of diminutives are made by assuming that the Italian diminutive is a modifier. It has no features, and thus the BPC is responsible for determining the gender of the node which is a sister to the vowel.

(66)



# 2.3.4.4 Italian diminutive adjectives

The Italian diminutive can also attach to adjectives. What predictions does our theory make for the properties of a constituent consisting of an adjective plus the diminutive suffix? We know that the diminutive suffix has no features of its own—it is a modifier. This is why, unlike German, the gender of the diminutive in Italian is inherited from the noun root. If the diminutive suffix is a modifier, it should also have no effect on category. When the diminutive attaches to a noun, the resulting compound will remain a noun, and when the diminutive suffix attaches to an adjective, the result will remain an adjective. In both cases, this is the result of the BPC. This is illustrated in (67), and in the tree in (68).

(68)



### 2.3.5 Summary

In this section, we have developed a theory of node labeling. The theory involves two conventions, the NLC and the BPC, corresponding to two types of relation between an affix and a stem. If the affix is a head, taking the stem as its complement, then the NLC applies and the features of the affix are those of the whole constituent. If the affix has no features of its own, then it is a modifier. As such, the BPC will ensure that the features of the stem constitute those of

the whole constituent. German and Italian diminutives provide a nice near minimal pair illustrating the application of both conventions for gender.

# 2.4 The separation Hypothesis (EXTRA)

Warning: This section introduces a somewhat advanced concept. The concept will not be crucial until Chapter Inf=Def and we will not make much use of it in the remainder of this chapter. Given that it remains somewhat controversial as of writing (2002) this section should be seen as a "plug in".

In section 0 we discussed the fact that syntactic category alone was not a sufficient predictor for the distribution of individual affixes. In particular, many affixes are not fully productive and do not attach to all of the words that they might have been expected to on the basis of their syntactic category.

We used the device of lexical restrictions to implement this formally. Lexical restrictions constitute arbitrary information, stored in a lexical entry, which must simply be learned by the child. Some aspects of the distribution of affixes are indeed arbitrary. But not all. To the extent that arbitrary devices are used to model systematic behaviour, the theory is in need of improvement. We will therefore sketch one direction that has been taken by some morphologists in order to improve the theory. The hypothesis to be discussed in this section will turn out to be quite important in understanding inflectional systems, though its correctness for derivational morphology has yet to be completely established.

Above, we used the distribution of the affix -ity to present the lexical restrictions. Consider now the table in (69). In this table, we have collected a sample of English adjectives and their corresponding degree or quality nouns.

- (69) Adjective Corresponding (Degree / Quality) Noun
  - a. odd odd-ity rare rar-ity able abil-ity X-able X-abil-ity
  - b. intelligent intelligen-ce (\*intellignetity)

prudent pruden-ce

c. strong streng-th

wide wid-th

d. sad sad-ness \*sadd-ity, \*sad-th, \*sad-ce ...
 happy happi-ness
 weird weird-ness (\*weirdity)
 wetwet-ness (\*wetity)

There are two things to note about this table. First, it is almost without exception the case that any (scalar) adjective in English will have a corresponding degree noun. Second, with the exception of *–ness* each of the affixes that makes a degree noun out of an adjective is restricted to some lexical set; only *–ness* is productive (recall that new words can be formed with *–ity* but only by inheriting the productiveness of *–able*).

Why should these properties hold of this collection of affixes?

One suggestion that has been put forward is that the relationship in which these affixes stand to one another is the same as that among the plural affixes in (70).

(70) a. sheep sheep-Ø

fish fish-Ø

b. child child-ren

ox-en

c. focus foc-i

OX

- locus loc-i
- d. jeep jeep-s fox fox-es crocus crocuses

This collection of word pairs shares many properties with the degree nominalizations in (69). For example, almost without exception, any English (count) noun will have a corresponding plural (perhaps marked by zero). Likewise, with the exception of -(e)s (phonologically /z/) each of the other plural markers is restricted to some lexical set; only -es is productive. In addition, in both sets there is a certain amount of variation among English speakers, with the productive affix being to some degree acceptable in place of one of the restricted affixes. Latin plurals such as those in (70) are gradually being replaced by the "regular" plural marking (%focuses, %locuses); compare %oddness, %rareness.

In Chapter Inf, we will see that the proper analysis of the plural formatives in (70) treats these as allomorphs of a single, abstract plural morpheme, in a manner analogous to the Hungarian abstract 2sg inflection, discussed in Chapter BC.

(71) label: PLURAL  
meaning: PLURAL  
phonology: 
$$\begin{cases} \emptyset / N_{,}, N \in \{\text{sheep, fish ...}\} \\ -(r) en / N_{,}, N \in \{\text{child, ox ...}\} \\ -i / N_{,}, N \in \{\text{foc(us), loc(us) ...}\} \end{cases}$$
  
 $z// N_{,} < elsewhere>$   
selection:  $J_{N}_{,}$   
categ res: N

Parity of reasoning would therefore suggest that the pattern in (69) is also derived by allomorphy. That is, rather than having separate lexical entries for each of the nominalizing affixes in (69), with all but one having lexical restrictions, a better account would have a single abstract morpheme, with (at least) four allomorphs. The lexical restrictions are then not a part of the theory of selection proper, but rather constitute conditions on allomorphy.

(72) label: DEGNOM

meaning: the degree to which something is A

$$\begin{array}{ll} \mbox{phonology:} & \mbox{``ity'' / Adj_, } & \mbox{Adj} \in \{\mbox{odd, rare, -able ...}\} \\ & \mbox{``ce'' / Adj_, } & \mbox{Adj} \in \{\mbox{intelligent, dominant ...}\} \\ & \mbox{``th'' / Adj_, } & \mbox{Adj} \in \{\mbox{strong, wide ...}\} \\ & \mbox{``ness'' / Adj_ } & \mbox{} \end{array} \right\}$$

selection: [A\_] categ res: N

The representations in (71)-(72) instantiate the SEPARATION HYPOTHESIS. It is widely recognized that this hypothesis is crucial to the understanding of inflectional morphology, although it is still somewhat controversial within the domain of derivational morphology. We will therefore postpone further discussion of it until later chapters. We will also not make explicit use off the Separation Hypothesis in the remainder of this chapter, however, you should be aware that the analyses in this chapter in which we use the phonological (or orthographic) form to identify affixes is, on this hypothesis, a convenient shortcut. An apparent suffix like -ity is not a true suffix, but rather an allomorph of the more abstract suffix DEGNOM.