Morphotactics, affix ordering, the mirror principles, and the relevance principle

David R. Mortensen

June 28, 2023

Morphotactics and Affix Ordering

The ordering of morphs/morphemes in a word is not free:

1. Changing the order of inflectional affixes results in an ILL-FORMED word
2. Changing the order of derivational affixes changes the meaning of a word and may result in ill-formedness
3. Changing the order of words in a compound changes the meaning and may result in ill-formedness

The principles that govern how morphemes are complex, but can be motivated by a few basic principles.

Subcategorization and Scope

Scope (how broad a swath an operations applies to) is central to morphotactics, as is SUBCATEGORIZATION (selectivity on the part of operations about what they will apply to). If thing A has scope over thing B in grammar, the morphological equivalent of B will occur closer to the root than A. From this following the prediction that inflection should almost always occur farther from the root than derivation. This is because inflection tells one about the whole word, whereas derivation applies to a narrower scope. Consider the following example:

(1) establish -ment -ary -ian -ism

The suffix *-ment* must attach to a verb and has scope only over *establish*. It makes a noun that means, roughly, the abstract concept corresponding to the patient of verb. The suffix *-ary*, on the other hand, attaches to nouns, so it must attach to *establish-ment* as a whole and have scope over it. It yields adjectives. *-ian* takes an adjective and yields a nouns (signifying a person related to the property associated with the adjective). These are sometimes understood in terms of a tree-structure of items which the grammar can manipulate in various ways (as with the inflectional structures that motivated the Mirror Principle as seen in the next section). However, to capture a related insight, let us suppose that each word is a sign (a tuple of signifier and signified) and each affix is a function (from word to word). [1] Parts of speech are represented as types. In our example, we start with a root **establish** of type Verb. We then have the following five functions (with type signatures):

The Mirror Principle, in its original formulation, is couched in Generative Grammar, a theory of language that was developed by Noam Chomsky and his collaborators starting in the 1950s. In Generative Grammar, syntax (the grammar of phrases and sentences) is central and this is understood, in large part, in terms of tree structures and operations on them.

[1] This is an oversimplification because applying an affix to a base may not result in a complete word (without the application of one or more other affixes).

(2) a. MENT: Verb → Noun
b. ARY: Noun → Adj
c. IAN: Adj → Noun
d. ISM: Noun → Noun

We can then imagine the word *establishmentarianism* as the composition of these functions and their subsequent application to **establish**.

(3) a. (ISM ∘ IAN ∘ ARY ∘ MENT)(**establish**)
b. ISM(IAN(ARY(MENT(**establish**))))

Note that the types of the functions constrain how they can be composed and that the outputs of the functions at the right are the inputs to the functions at the left (the inner functions output is the outer functions input). This means that the order of the affixes is constrained by their type signatures as well as their effect on the meaning of a word. This is know in linguistics as SUBCATEGORIZATION.

The Mirror Principle

One way of modeling the order of morphemes in a word, with tools from theoretical linguistics, is the Mirror Principle. The Mirror Principle holds that morpheme order should be based on the hierarchies of syntactic projections. What is a syntactic projection? You can think of it as a scope that can be represented in the grammar. Consider the difference between the following sentences:

(4) Micah liked Tamar.

(5) a. Tamar was liked by Micah. (passive)
b. Rahel caused Micah to like Tamar. (causative)

(6) a. Micah was caused to like Tamar by Rahel. (passive of causative)
b. Rahel caused Tamar to be liked by Micah. (causative of passive)

Now, imagine a language (like many Bantu languages) where PASSIVE and CAUSATIVE are represented by morphemes. The Mirror Principle predicts that the causative morpheme will occur closer to the root than the passive morpheme in sentences translating (6a) and that the opposite ordering will occur for sentences translating (6b).

The Mirror Principle does not just apply to valency changing morphology. Now consider inflection: it does not change part of speech and does not change meaning. It just realizes morphosyntactic properties. What determines the ordering of such affixes since part of speech (type) and semantic conditions do not do so? The Mirror Principle provides one answer. In the version of Generative Grammar out of which the Mirror Principle developed,

different morphosyntactic properties are associated with different "projections." For example, tense features reside on the Tense node of the Tense Phrase. These projections are nested so that some are more closely associated

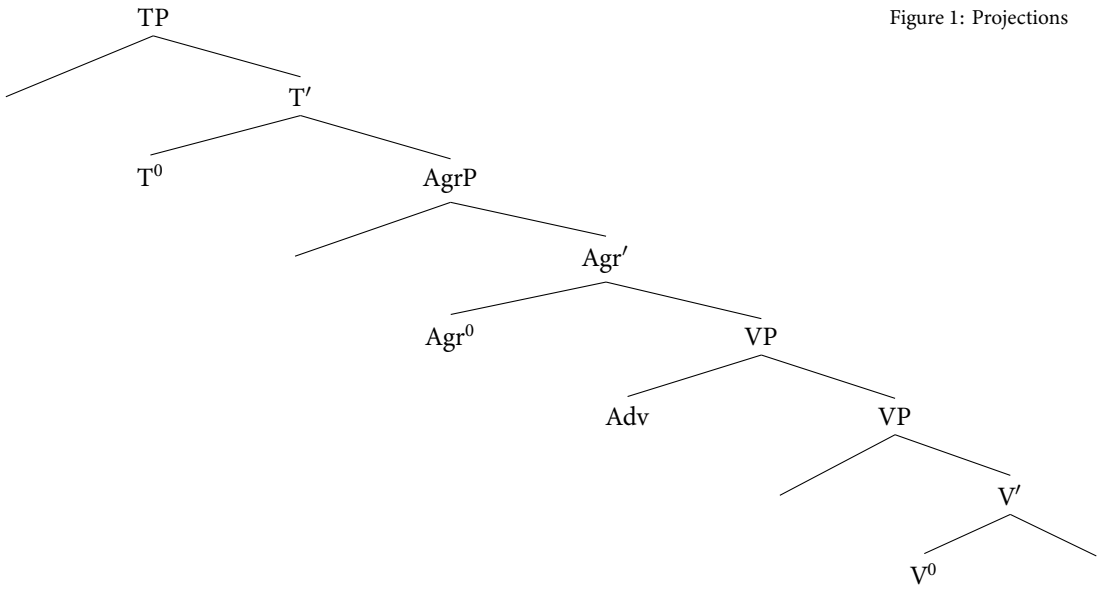

Figure 1: Projections

with the root and some are more distant in the tree. In some sense, these behave like the functions in our example: the features closer to the verb, in a clause, are realized closer to the root.

Baker's Mirror Principle[2] seeks to derive morpheme order from scope in this kind of structure. Think back to the following examples from Bemba:

[2] Mark Baker. The mirror principle and morphosyntactic explanation. *Linguistic Inquiry*, 16(3):373–415, 1985

(7) a. Na- a- mon -an -y -a Mwape na Mutumba
1SG- PAST- see -REC -CAUS -FV 1.Mwape and 1.Mutumba
'I made Mwape and Mutumba see each other.'

b. Mwape na Chilufya ba- a- mon -eshy -an -a
2SG.MASC- PAST- see -CAUS -REC -FV
'Mwape and Chilufya made each other see Mutumba.'

Note that there are too causative affixes in these data, *-y* and *-eshy*.

Baker would say that there are three nested phrases at work in these examples:

1. A verb phrase (VP)
2. A causative phrase (CausP)
3. A reciprocal phrase (RecP)

In both sentences, VP is inside of both CausP and RecP, but in (7a) the reciprocal phrase is nested inside the causative phrase whereas in (7b), the causative phrase is nested inside of the reciprocal phrase.

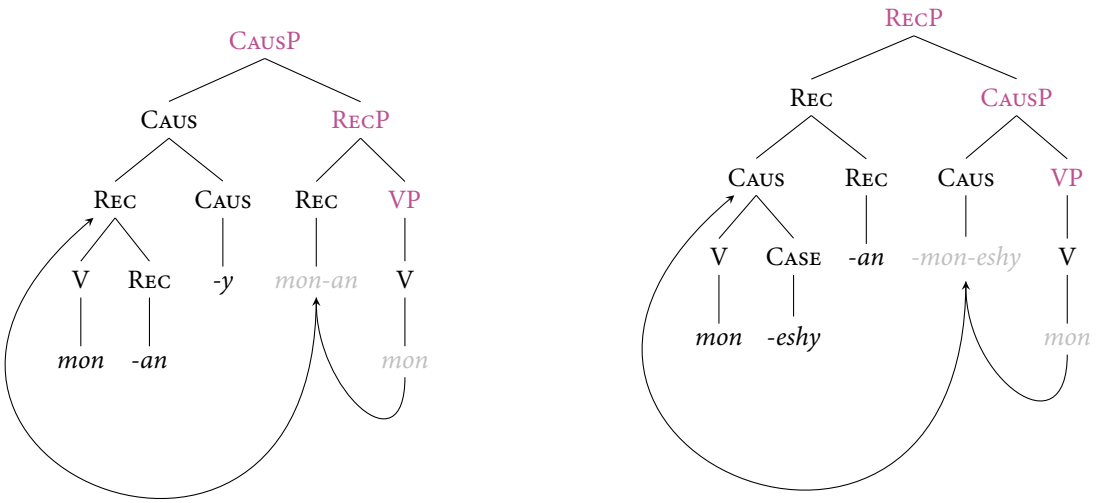

Figure 2: Analysis of the ordering of Bantu causative and reciprocal according to the Mirror Principle

Baker developed a seemingly convoluted way of deriving aspect order from scope, building on a preexisting idea called HEAD MOVEMENT. According to this idea, HEADS (like verbs) move up the three step by step, accumulating material as they go. Subtrees also move[3] This is illustrated in Figure 2. In the first tree, the verb root starts out in the VP, which starts inside the RecP (which is inside the CausP). The verb root raises to the Rec position (the head of the phrase immediately above it) and picks up the *-an* reciprocal suffix. Rec is now the head of RecP. Rec therefore moves to merge with the head of the next phrase up, CausP. The causative suffix *-y* is the head of this phrase. The result is a subtree with the terminal nodes *mon*, *-an*, and *-y* (in that order). This gives us the right affix order based on the scope. The second tree illustrates the same thing, only with CausP nested inside of RecP.

As should be apparent, Baker is operating within a very specific version of a very specific theory of morphology and syntax. The majority of linguists in the world would almost certainly disagree with these assumptions. One is free to agree or disagree with the formal mechanism that Baker used to derive affix order from scope. However, it is clear that—often—affix order *does* follow from scope and Baker's analysis provide insight into this phenomenon.

[3] The full rationale for this is beyond the scope of this course.

The Relevance Principle

[4]

1. The more relevant a category is to the verb, the more likely it is to occur in a synthetic or bound construction with the verb.

[4] Joan Bybee. Diagrammatic iconicity in stem-inflection relations. In *Iconicity in syntax*, pages 11–48. John Benjamins, Amsterdam, 1985

2. The more relevant a morphological category is to the verb, the closer its marker will occupy to the verb stem.

3. The more relevant a morphological category is to the verb, the greater will be the morpho-phonological fusion of that category with the verb stem.

For verbal inflection, this assumes a set of relationships like those in Figure 3. Modality has scope over the whole clause (sentence). It reflects the

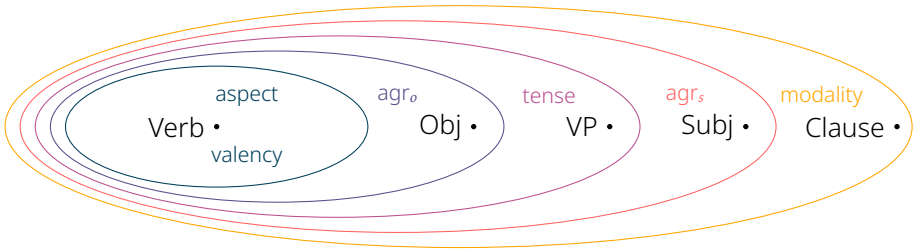

Figure 3: The notion of relevance to the verb illustrated for inflection. Items closer to the center are more relevant to the verb and those farther aware are less relevant.

attitude of the speaker to the whole proposition (whether it is true, whether it should be true, what the source of the information is, etc.). Aspect has scope over the verb alone. Valency changing morphology (including voice) is an interesting case. It is relevant to the object and the subject (and sometimes other arguments of the verb), so one might place it at the clause level (the outer layer of the relevance onion) but it really makes reference only to what is called the ARGUMENT STRUCTURE of the verb. It is changing the syntax of the verb by changing (i) the semantics of the verb—what arguments it takes and with what roles—and (ii) how the semantic roles of the entities that participate in the event signified by the verb map onto noun phrases in the clause. Alternations in argument structure are often indicated by changing the verb entirely (see, for example, *kill* which is the causative of *die*). That is to say, valency changing morphology is more relevant to the verb than object agreement (agr$_o$), tense, subject agreement (agr$_s$) or modality.

Note that this hierarchy makes many predictions (all of which should be understood statistically, not deterministically):

- Agreement should appear further from the root than aspect or voice/valency morphology.
- If a language marks both object and subject agreement, object agreement should appear closer to the root than subject agreement. This is borne out by most Bantu languages, as well as Nahuatl.
- If a language marks object and subject agreement as well as tense, tense should appear between object and subject agreement. This is also borne out by Bantu languages, Nahuatl, and others.
- Modality marking should not occur closer to the root than any other kind of marking.

Bybee found, in a typologically balanced survey of 50 languages, that the predictions of this model were largely borne out. However, there are some problematic cases. For example, aspect, tense, and modality are often realized by a single morpheme. While this does not invalidate the predictions of the principle, it does complicate how we view it.

Templatic Ordering

Sometimes, however, affix ordering does not follow from general principles. Consider again Bantu languages: in some of them, the order of the valency changing suffixes is ordered according to scope (following the predictions of the mirror principle and also the relevance principle). However, in others, the order always follows the sequence CARP:

1. Causative
2. Applicative[5]
3. Reciprocal[6]
4. Passive

For example, if a word has both a causative and a passive suffix, the ordering will always be ROOT-CAUS-PASS, regardless of whether the meaning is 'cause to be Ved' or 'be caused to V'.

This is an example of TEMPLATIC ORDERING. It is not uncommon for words of a particular part of speech (e.g., nouns or verbs) the have a set of "slots" in a fixed sequence that may be filled by a morpheme or be empty. The sets of morphemes that can fill a slot are called POSITION CLASSES. In many cases, the ordering of position classes follows from the Mirror Principle or the Relevance Principle, but sometimes—as in the case of Bantu languages that follow the CARP template—the order of positions is arbitrary. In such cases, there is often a historical reason why the morphemes are "stuck" in an order that does not follow from grammatical or functional principles.

Implications

While the Mirror Principle and the Relevance Principle come at the problem from very different angles, but suggest that words with multiple affixes can be seen as composed as nested layers, each of which has scope over the layers inside of it. The Mirror Principle relates these layers to operations in the syntax. The relevance principle relates them to communicative function. But the Relevance Principle, in particular, leads us to believe that morphemes close to the root will be more collocated with it (that is, have higher mutual information). It also suggests that for a word like *re-implement-at-ion-s*, it should be easier to learn a meaningful embedding for *implementation* than for *ations*, even though both consist of the concatenation of three morphemes.

[5] Applicatives add an OBLIQUE, like a beneficiary or an instruct, as an additional argument.

[6] Change a transitive verb into an intransitive verb with an "each other" meaning.

References

Mark Baker. The mirror principle and morphosyntactic explanation. *Linguistic Inquiry*, 16(3):373–415, 1985.

Joan Bybee. Diagrammatic iconicity in stem-inflection relations. In *Iconicity in syntax*, pages 11–48. John Benjamins, Amsterdam, 1985.