Theoretical Architecture, p. 1

This handout outlines the fundamental theoretical questions involved in modeling morphology formally. We will consider what different theoretical approaches have to say about two very simple words in Hebrew: the noun form תקציבים *takcivim* 'budgets' and the adjective form תקציביים *takciviim* 'budgetary (pl)'. These are both formed from the root *takciv*, a noun. (We will be assuming that *takciv* is monomorphemic.) The noun form *takcivim* involves the inflectional processes of affixing the suffix *-im* (marking plural); the adjective form *takciviim* involves the derivational process of affixing the suffix *-i* (marking an adjective) and the inflectional process of affixing the suffix *-im* (marking plural).

Note: In order not to complicate matters more than necessary, we will be ignoring gender.

Morpheme-based Model or Word-based Model

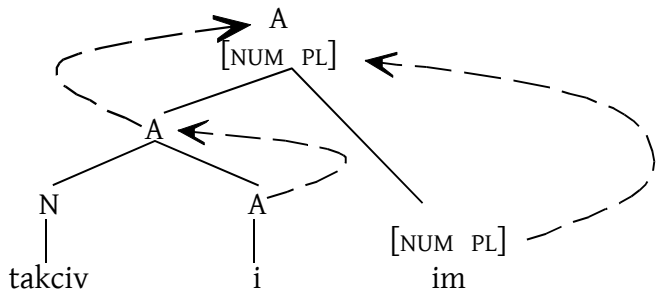
The traditional approach to morphology is based on the idea that the basic unit of morphology is the morpheme. Words are created by re-write rules:

word-form → (inflection) stem (inflection)
stem → (derivation) root (derivation)
stem → stem stem

The exact way in which the elements combine is determined by their idiosyncratic properties, as indicated in the morpheme lexicon. In other words\, each morpheme has a lexical entry: a list of properties in the mental lexicon.

תקציב	-י	-ים
/takciv/	/i/	/im/
N	A_{aff}	inflection
⟨TAKCIV⟩	[N ___]	[N/A___]
'budget'	'of N'	[NUM PL]

The properties of the words are the result of taking the features of the component morphemes and copying them to the higher node. This is called **feature percolation**. Properties of affixes take precedence over properties of the stem; using the analogy of syntax, we can say that the affix is the **head** of the word. Note the following example.



The morpheme-based model is a very attractive approach to morphology. It is economical, as it is based on the notion of a morpheme lexicon. It also maximizes the similarity between morphology and syntax.

Theoretical Architecture, p. 2

An alternative approach considers the basic unit of morphology to be the word (i.e. word-form), not the morpheme. Rules of morphology are not word-structure rules; instead, they relate words to each other. More precisely, they relate word schemas—patterns representing classes of words—and are written as bidirectional relationships between skeletal lexical entries.

$$\begin{bmatrix} /X/ \\ N \\ \text{`x'} \end{bmatrix} \leftrightarrow \begin{bmatrix} /Xi/ \\ A \\ \text{`of x'} \end{bmatrix} \text{ relates } \overset{\text{תקציב}}{\begin{bmatrix} /takciv/ \\ N \\ \langle \text{TAKCIV} \rangle \\ \text{`budget'} \end{bmatrix}} \text{ and } \overset{\text{תקציבי}}{\begin{bmatrix} /takcivi/ \\ A \\ \langle \text{TAKCIVI} \rangle \\ \text{`of the budget'} \end{bmatrix}}$$

$$\begin{bmatrix} /X/ \\ N/A \\ \langle X \rangle \end{bmatrix} \leftrightarrow \begin{bmatrix} /Xim/ \\ N/A \\ \left\langle X, \begin{bmatrix} \text{NUM} & \text{PL} \end{bmatrix} \right\rangle \end{bmatrix} \text{ relates } \overset{\text{תקציב}}{\begin{bmatrix} /takciv/ \\ N \\ \langle \text{TAKCIV} \rangle \\ \text{`budget'} \end{bmatrix}} \text{ and } \overset{\text{תקציבים}}{\begin{bmatrix} /takcivim/ \\ N \\ \left\langle \text{TAKCIV}, \begin{bmatrix} \text{NUM} & \text{PL} \end{bmatrix} \right\rangle \\ \text{`budget'} \end{bmatrix}}$$

$$\begin{bmatrix} /X/ \\ N/A \\ \langle X \rangle \end{bmatrix} \leftrightarrow \begin{bmatrix} /Xim/ \\ N/A \\ \left\langle X, \begin{bmatrix} \text{NUM} & \text{PL} \end{bmatrix} \right\rangle \end{bmatrix} \text{ also relates } \overset{\text{תקציבי}}{\begin{bmatrix} /takcivi/ \\ A \\ \langle \text{TAKCIVI} \rangle \\ \text{`of the budget'} \end{bmatrix}} \text{ and } \overset{\text{תקציביים}}{\begin{bmatrix} /takciviim/ \\ A \\ \left\langle \text{TAKCIVI}, \begin{bmatrix} \text{NUM} & \text{PL} \end{bmatrix} \right\rangle \\ \text{`of the budget'} \end{bmatrix}}$$

This kind of model suggests that morphology is fundamentally different from syntax. However, it has certain strengths relative to the morpheme-based model.

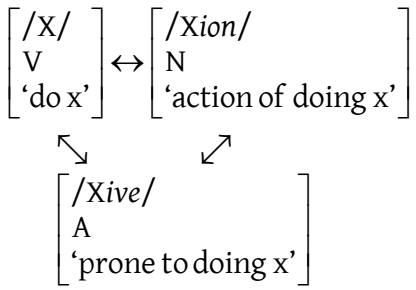
In the first place, it can handle non-concatenative morphology, including those versions that do not plausibly involve morphemes. A morpheme-based model would require a completely different mechanism to create words that involve non-concatenative operations. For example, in English one way to form the past tense of verbs with the vowel [iy] (e.g. *freeze*, *speak*, *steal*) is to change the [iy] to [ow].

$$\begin{bmatrix} /XiyC/ \\ V \\ \langle X \rangle \end{bmatrix} \leftrightarrow \begin{bmatrix} /XowC/ \\ V \\ \left\langle X, \begin{bmatrix} \text{TENSE} & \text{PAST} \end{bmatrix} \right\rangle \end{bmatrix}$$

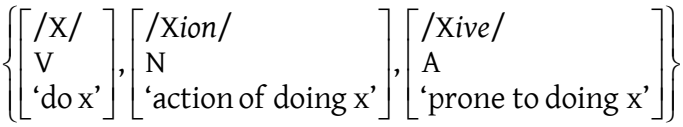
Also, the word-based approach is non-directional. It thus explains the phenomenon of back-formation, a historical process where an affix is removed from a word, such as the verb *edit* which is historically a back-formation from *editor*.

Another important advantage of the word-based model is the ability to express multilateral relations between morphological forms. An example from English, in the textbook.

Theoretical Architecture, p. 3



or, in a more compact format:



This is particularly important when the stem does not exist as an independent word. For example, the words *aggression* and *aggressive* are clearly related to each other morphologically, but the stem **aggress* is not a word.

Finally, based as it is on some form of a word lexicon, it has all the advantages that a word lexicon has over a morpheme lexicon.

Many morphologists today have adopted the word-based model of morphology. One issue it leaves open to which we will be returning later, is the status of the concept "morpheme". It is, at best, a derivative concept, although it is still useful enough for us to continue using it.

Incremental Model or Realizational Model

(This is not discussed in the textbook. The terminology comes from Stump 2001a)

In inflectional morphology, there is a second distinction that can be drawn between theories of morphology: the nature of the relationship between inflectional features and their exponence. Taking *midbarim* as an example, the two models are as follows:

Incremental (or compositional): *Takcivim* is plural because it includes the suffix *-im*. The properties of inflected forms are determined incrementally as morphological processes are applied, and morphological processes are applied in order to achieve the proper combination of features.

Realizational: *Takcivim* includes the suffix *-im* because it is plural. Morphological processes are the realizations of inflectional features. (Features determine exponence.)

This distinction is independent of the morpheme-based/word-based distinction.

The incremental approach is in keeping with notions of compositionality elsewhere in language (including in syntax). Note that it is assumed in the diagram of feature percolation above in the discussion of morpheme-based approaches. These two approaches often go together, although they

Theoretical Architecture, p. 4

need not. (The theory of Distributed Morphology (Halle and Marantz) is an example of a theory which is morpheme-based and realizational.) However, there are two types of phenomena that suggest that a realizational approach is preferable.

• The first is the phenomenon of extended exponence. Consider the English past tense verb forms *sold* and *told*, the past tenses of *sell* and *tell* respectively. Two morphological processes have applied in the derivation of these forms: the stem vowel has changed to *o* [ow] and the suffix *-d* has been added. Either one of these processes by itself involves the past tense feature: the alveolar suffix fairly transparently, and the vowel change as in verbs like *speak-spoke*, *write-wrote* and *break-broke*. Whichever one of these applies first should be enough to establish the verb form as past tense. Since under an incremental approach it is the application of a morphological process that results in the word having the feature [TENSE PAST], there is a rather odd redundancy in doing both.

For another example of extended exponence, note the following data from Swahili:

tu-li-taka 'we wanted'
ha-tu-ku-taka 'we didn't want'

The *tu* is the agreement marker for 'we'. *-li-* marks affirmative past tense, and *-ku-* marks negative past tense. *Ha-* marks negation. An incremental theory has no way to account for the necessity of *ha-*; the word already has the negation feature by virtue of the affix *-ku-*.

A realizational theory has no problem with extended exponence. There is nothing to prevent the realization of some (set of) feature(s) from involving more than one morphological process. All the grammar of Swahili has to say is that the realization of the feature combination [TENSE PAST, POLARITY NEG] involves two affixation processes. The English example is a bit more complicated, since it involves an irregular class of verbs, but it would work essentially the same way.

•Another argument for a realizational theory and against an incremental theory comes from the inflectional properties of unaffixed forms. If *takcivim* is plural by virtue of including the suffix *-im*, why is *takciv* singular?

① We could say that the unaffixed form is given default (or unmarked) features. For example, it is generally acknowledged that [NUM SING] is unmarked and [NUM PL] is marked. The morphology reflects this by not having a singular affix: nouns are singular unless there is some morphological indication that they are not. This approach looks reasonable, but further consideration shows it cannot be correct. Consider the English verb forms *act* ⟨ACT, [TENSE PRES, PERS 1|2, NUM SG]|[TENSE PRES, NUM PL]⟩ and *acts* ⟨ACT, [TENSE PRES, PERS 3, NUM SG]⟩. The default number is singular, and the default person is 3rd. We certainly wouldn't want to say that the default agreement for a verb is 1st person singular, 2nd person singular, and plural! There is no independent notion of default or markedness that would give us the right result. We are thus left with no account of why *act* has the inflectional properties that it has.

② Another approach would be to say that *takciv* does have an inflectional suffix, but that this suffix has no phonological properties. That is to say, *takciv* is really *takciv+Ø*. Such an analysis is artificial, and would require us to postulate a multitude of Ø affixes, one for each inflectional property which is not expressed by an overt affix.

Theoretical Architecture, p. 5

In a realizational theory, we can say that the process of affixing *-im* is the realization of [NUM PL]. An unaffixed form like *takciv* has no inflectional features. However, Hebrew has no expression of the feature [NUM SG]. If one wants to express ⟨TAKCIV, [NUM PL]⟩, one applies the plural rule (affixation of *-im*) to the stem *midbar*. If one wants to express ⟨TAKCIV, [NUM SG]⟩, the best one can do is to use the form *midbar*, which expresses a subset of what one wants, but at least doesn't contradict the desired features. This will also work for the English example of *act* and *acts*: the morphology of English provides a realization of [TENSE PRES, PERS 3, NUM SG], but not for any other persons or numbers in the present tense.

Final Conclusion

A theory of morphology should be word-based and realizational. This is essentially the position taken in the textbook, and what we will be assuming for the rest of the course. A more formalized version of this (though not the only possible one) is Paradigm Function Morphology (Stump 2001a).