3. PRINCIPLES OF MORPHOLOGICAL ANALYSIS: BASICS

Some approaches to morphology

Traditional grammar: Traditionally, “grammar” essentially meant what we call “morphology”. The main concern of traditional grammarians like Panini (the Urgrammarian) and Priscian (mentioned by Haspelmath) was to classify morphemes and morphological construction types. These traditional grammarians created much of the morphological terminology that we use today.

American linguistics: there have been two main approaches to morphology

structuralists (for example, Bloomfield, Nida): methods for separating out and categorizing the elements that make up complex words (*discovery procedures*)

**STRUCTURALIST TASK:** Apply rigorous procedures to isolate the form-meaning distinctions in datasets like *inexact*, *indecent*, *impolite*, *inclement*, *irrelevant*, *illegal* and account for variation in form where necessary.

generativists: Until the mid-1970’s there was no “morphology”! Words were built by rules of *syntax*, and rules of *phonology* provided the pronunciation. Here is an example from Robert B. Lees, *The Grammar of English Nominalizations*, Mouton, 1963.

*He is selling the car.* (by syntactic transformation) ⇒ *(He is the) seller (of the car.)*

*The door weighs X.* (by syntactic transformation) ⇒ *(The) weight (of the door is X.)*

By the mid-1970’s, linguists like Mark Aronoff and Stephen Anderson saw morphology as requiring principles in addition to those needed for syntax and phonology, and the generativists’ goal is now to characterize the knowledge that speakers have of the structure of complex words

**GENERATIVIST TASK:** Propose a model of a speaker’s grammar that allows for the production and comprehension of items such as those in the examples above.

We start with the *structuralist* approach as a way to introduce terminology and analytical techniques that we will use throughout the course. American structuralists spoke of “discovery procedures” that would allow linguists to parse linguistic structures in ways that would avoid guesswork and could be replicated by others such that they would (hopefully) come up with the same analyses for the same data. For the parsing of complex words, Nida (1949) proposed six principles for identifying the component parts of such words and for formulating the nature of those parts. We will organize our approach to morphological data around Nida’s principles.

**Nida’s Principle 1:** “Forms which have a common semantic distinctiveness and an identical phonemic form in all their occurrences constitute a single morpheme.”

---

A recipe for applying Principle 1:

(1) Semantic distinctiveness: Propose the likely candidates, in terms of meaning, that might have morphological expression.

ADVICE: To select “candidates”, we are dependent on translation equivalents, which are often just approximations, so we often have to use our imagination or common sense to isolate an “element of common semantic distinctiveness”. For example, vaccinate “inject with a vaccine”, hyphenate “insert hyphens”, orchestrate “arrange music for an orchestra” don’t share a meaning component that is statable in one word, but clearly the words all end with –ate and each is a verb that connects a noun to the object of the verb. Translations in one language may inject distinctions that are irrelevant for the morphology of the data in the other language, for example, for French je sais ‘I know’ vs. je vais ‘I am going’, the English translations use two tenses (simple present “know” vs. present progressive “am going”), whereas French has only one morphological tense, the “present”. On the other hand, translations in one language may fail to make morphological distinctions found in another, for example, for French rou rond ‘round hole’ vs. pierre ronde ‘round stone’, the adjective “round” in English is invariable, but in French, it varies morphologically depending on gender of the noun.

(2) Phonemic form: Pick one candidate from (1) and look at several items from the available data that have that meaning element, looking for something that is common to the pronunciation of all those items.

(3) Systematic lists: List the items identified in (2) in (tentative) categories according to meaning, function, and/or position in the word. For example, you may make a list of what appear to be roots, another list of prefixes that appear to express “subject of a verb”, etc.

(4) Check for exhaustivity: Once you have the lists in (3), go back through the data and make sure all parts of all words are accounted for. If there are parts that are NOT accounted for, these may be separate morphological elements that don’t come out in the translations (see “ADVICE” under (1) above). If this is the case, see whether there seems to be something about these elements that suggests a “common semantic distinctiveness”.

Practice with Principle 1 and the recipe for applying it

Swahili (data supplied by Chacha Mwita—see the practice problems at the end of this section for some information about Swahili orthography and phonology): Do a morphological analysis by applying Principle 1.

1. nilianguka I fell
2. ulianguka you (sg) fell
3. tulianguka we fell
4. mlianguka you (pl) fell
5. ninaanguka I am falling
6. unaanguka you (sg.) are falling
7. tunaanguka we are falling
8. mnaanguka you (pl) are falling
9. nitaanguka I will fall
10. utaanguka you (sg) will fall
11. tutaanguka we will fall
12. mtaanguka you (pl) will fall
13. uliniona you (sg) saw me
14. unaniona you (sg) see me
15. utaniona you (sg) will see me
16. nilikuona I saw you (sg)
17. ninakuona I see you (sg)
18. nitakuona I will see you (sg)
19. mlituona you (pl) saw us
20. mnatuona you (pl) see us
21. mmtatuona you (pl) will see us
22. tuliwaona we saw you (pl)
23. tunawaona we see you (pl)
24. tutawaona we will see you (pl)
English: Identify the morphemes (affixes, roots, bases) for the two data sets, give approximate meanings, and (referring to Principle 1) give reasons for identifying or not identifying items as separate morphemes. **Question:** Do the boxed items at the top of each DATA SET share a morpheme?

<table>
<thead>
<tr>
<th>DATA SET 1</th>
<th>DATA SET 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>farmer</td>
<td>inbreed</td>
</tr>
<tr>
<td>villager</td>
<td>incorrect</td>
</tr>
<tr>
<td>taller</td>
<td>inane</td>
</tr>
<tr>
<td>banner</td>
<td>incense (make angry)</td>
</tr>
<tr>
<td>archer</td>
<td>incense (perfumed smoke)</td>
</tr>
<tr>
<td>better</td>
<td>incite</td>
</tr>
<tr>
<td>chapter</td>
<td>indecent</td>
</tr>
<tr>
<td>cleaner</td>
<td>indeed</td>
</tr>
<tr>
<td>eraser</td>
<td>indent</td>
</tr>
<tr>
<td>follower</td>
<td>index</td>
</tr>
<tr>
<td>foreigner</td>
<td>induct</td>
</tr>
<tr>
<td>further</td>
<td>infer</td>
</tr>
<tr>
<td>hamster</td>
<td>infinite</td>
</tr>
<tr>
<td>Londoner</td>
<td>infuse</td>
</tr>
<tr>
<td>Rhode Islander</td>
<td>inlay</td>
</tr>
<tr>
<td>stranger (more strange)</td>
<td>insane</td>
</tr>
<tr>
<td>stranger (unfamiliar person)</td>
<td>insect</td>
</tr>
<tr>
<td>sucker</td>
<td>insist</td>
</tr>
<tr>
<td>teacher</td>
<td>instant</td>
</tr>
<tr>
<td>weather</td>
<td>intrepid</td>
</tr>
</tbody>
</table>

**Bontoc** (Gleason 1955:29): Identify the morphemes and their apparent meanings.

1. **fikas** strong  **fumikas** he is becoming strong
2. **kilad** red     **kumilad** he is becoming red
3. **bato** stone    **bumato** he is becoming stone
4. **fusul** enemy   **fumusul** he is becoming an enemy

**Tamazhaq** (personal field notes): Are there any **circumfixes** in this data?

<table>
<thead>
<tr>
<th>øgmada</th>
<th>‘I went out’</th>
<th>nøgmad</th>
<th>‘we went out’</th>
</tr>
</thead>
<tbody>
<tr>
<td>tagmada</td>
<td>‘you (sg.) went out’</td>
<td>tøgmadalafil</td>
<td>‘you (several males) went out’</td>
</tr>
<tr>
<td>yøgmad</td>
<td>‘he went out’</td>
<td>øgmadan</td>
<td>‘they (several males) went out’</td>
</tr>
<tr>
<td>tøgmad</td>
<td>‘she went out’</td>
<td>øgmadnat</td>
<td>‘they (several females) went out’</td>
</tr>
</tbody>
</table>

---

Practice Problems for Nida’s Principle 1

Ngizim (a Chadic language of northern Nigeria)

Transcription is IPA. Grave accent = low tone (à), circumflex accent = falling tone (â), no accent = high tone.

<table>
<thead>
<tr>
<th></th>
<th>‘shoe’</th>
<th>‘cow’</th>
<th>‘wood’</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘my’</td>
<td>sònògàː</td>
<td>tàgàː</td>
<td>dàm gàː</td>
</tr>
<tr>
<td>‘your (1)’</td>
<td>sònọtʃì</td>
<td>tàtʃì</td>
<td>dàmtʃì</td>
</tr>
<tr>
<td>‘your (2)’</td>
<td>sònọkòm</td>
<td>tàkòm</td>
<td>dàmkòm</td>
</tr>
<tr>
<td>‘your (3)’</td>
<td>sònọkùn</td>
<td>tàkùn</td>
<td>dàmkùn</td>
</tr>
<tr>
<td>‘his’</td>
<td>sònògɔrì</td>
<td>tàgɔrì</td>
<td>dàmǥɔrì</td>
</tr>
<tr>
<td>‘her’</td>
<td>sònògɔrə</td>
<td>tàgərə</td>
<td>dàmǥərə</td>
</tr>
<tr>
<td>‘our (1)’</td>
<td>sònọwə̀</td>
<td>tàwə̀</td>
<td>dàmwə̀</td>
</tr>
<tr>
<td>‘our (2)’</td>
<td>sònọdʒà</td>
<td>tàdʒà</td>
<td>dàmdʒà</td>
</tr>
<tr>
<td>‘their’</td>
<td>sònọkʃì</td>
<td>tàkʃì</td>
<td>dàmkʃì</td>
</tr>
</tbody>
</table>

(i) Possessive pronouns: Apply steps (1-3) of our recipe to isolate the morphemes that express the various possessives meaning ‘my’, ‘your (1)’, etc. We return to step (4) in (iii) below.

(ii) Nouns: Apply steps (1-3) to isolate the parts of these words that seem to express ‘shoe’, ‘cow’, and ‘wood’. There is some variation in tones and vowel lengths. To keep things simple right now, just leave out tone and vowel length in isolating the noun roots. Again, step (4) doesn’t seem to be a concern.

(iii) Back to step (1) and step (4)—there seems to be something left over about meanings:

(a) In using English to elicit these forms, we notice that the word you elicits three forms. However, when looking at various tokens (or maybe just asking our speakers!), we find that ‘your (1)’ is used only when addressing one male, ‘your (2)’ is used only when addressing one female, and ‘your (3)’ is used when addressing a group. Propose labels that tell us more about the meanings than calling them ‘you (1-3)’.

(b) In using English to elicit these forms, we notice that the word we elicits two forms. However, when looking at at various tokens, we find that ‘we (1)’ is used when the person being spoken to is included in the group indicated by ‘we’ but that ‘we (2)’ is used when the person spoken to is not included. Propose labels that tell us more about the meanings than calling them ‘we (1-2)’.

(iv) Accounting for variations in the nouns: As noted in (ii), there is some variation in tones and vowel lengths of nouns. Nida’s Principles (2) & (3) address this, but see whether you can figure out what the variations correlate with.
NOTES ON ORTHOGRAPHY AND PRONUNCIATION: Standard Swahili orthography comes close to being a phonetic representation, about the only exception being that $sh = \left[\mathcal{f}\right]$. Swahili is NOT a tone language, an unusual fact for an African language. Stress is always on the penultimate syllable of a word, which gives a good diagnostic for what counts as a “word” in Swahili. Every vowel represents a separate syllable nucleus, e.g. $u.na.a.n.gu.ka$ ‘you are falling’ or $n.i.ta.ku.o.na$ ‘I will see you’ consist of five syllables each (periods mark syllable breaks). The sequences $[\alpha\text{nasal}][\alpha\text{voiced stop}]$, where “$\alpha$” means “same place of articulation”, and $[\text{nasal}][\text{glide}]$ form the onset of a syllable, but otherwise, a nasal before another consonant is syllabic. Note the syllabification of $m.li.a.n.gu.ka$ ‘you (pl) fell’, $mwa.na.m.ke$ ‘woman’.

1a. mtoto
1b. mtoto mmoja
1c. mtoto mkubwa
1d. mtoto alianguka
1e. nilimwona yule\textsuperscript{4} mtoto; je, ulimwona?

2a. watoto
2b. watoto wawili
2c. watoto wakubwa
2d. watoto walianguka
2e. niliwaona wale watoto; je, uliwaona?

3a. mti
3b. mti mmoja
3c. mti mkubwa
3d. mti ulianguka
3e. niliuona ule mti; je, uliuona?

4a. miti
4b. miti miwili
4c. miti mkubwa
4d. miti ilianguka
4e. niliiona ile miti; je, uliiona?

5a. tunda
5b. tunda moja
5c. tunda kubwa
5d. tunda lilianguka
5e. nililiona ile tunda; je, uliliona?

6a. matunda
6b. matunda mawili
6c. matunda makubwa
6d. matunda yalianguka
6e. niliyaona yale matunda; je, uliyaona?

\textsuperscript{3} Thanks go to Chacha Mwita for providing and discussing the Swahili data.

\textsuperscript{4} In the (e) items, the object noun is preceded by $Xle$. This functions like a definite article.
7a. kiti
7b. kiti kimoja
7c. kiti kikubwa
7d. kiti kilianguka
7e. nilikiona kile kiti; je, ulikiona?
8a. viti
8b. viti viwili
8c. viti vikubwa
8d. viti vilianguka
8e. niliviona vile viti; je, uliviona?
9a. kamba
9b. kamba moja
9c. kamba kubwa
9d. kamba ilianguka
9e. niliiona ile kamba; je, uliiona?
10a. kamba
10b. kamba mbili
10c. kamba kubwa
10d. kamba zilianguka
10e. niliziona zile kamba; je, uliziona?
11a. simba
11b. simba mmoja
11c. simba mkubwa
11d. simba alianguka
11e. nilimwona yule simba; je, ulimwona?
12a. simba
12b. simba wawili
12c. simba wakubwa
12d. simba walianguka
12e. niisiwaona wale simba; je, uliwaona?

5 In coastal dialects of Swahili, items like those in 33-36 that begin in voiceless stops would begin in aspirated stops, e.g. k’amba ‘rope’, k’ubwa ‘big’ (referring to a rope). This would be in contrast to items like those in 20, which would not have aspiration. Comparative Bantu evidence indicates that the aspiration is the effect of there at one time having been a nasal prefix in those items with aspiration. In the dialect represented here, even the aspiration is gone, meaning that all traces of a prefix have been lost.
Aztec of Veracruz (Nida 1949:8-10): Isolate all the morphemes (roots, affixes) and identify their meaning/function. For affixes, try to decide what the stems to which they are attached are. TRANSCRIPTION: The transcription is American structuralist—y = IPA [j], č = IPA [tf]. Plain c in #23 is probably č with the diacritic inadvertently omitted.

1. **ničoka** I cry 13. **titehkawiʔ** you (sg.) climbed
2. **ničokaʔ** I cried 14. **nitehkawiya** I was climbing (and may still be)
3. **nimayana** I am hungry 15. **nitehkawis** I will climb
4. **nimayanaʔ** I was hungry 16. **nikwake** we ate
5. **nimayanaya** I was hungry (and may still be) 17. **nimayanati** I go to be hungry* 
6. **timayana** you (sg.) are hungry 18. **nimayanato** I went to be hungry
7. **nimayananas** I will be hungry 19. **nimayanaki** I come to be hungry
8. **tičoka** you (sg.) cry 20. **nimayanako** I came to be hungry
9. **ničokaya** I was crying (and may still be) 21. **nikmayanati** I cause him to be hungry
10. **ničokas** I will cry 22. **nikmayanatiʔ** I caused him to be hungry
11. **ankwake** you (pl.) ate 23. **nimicmayanatis** I will cause you (sg.) to be hungry
12. **nitehkawi** I climb

*Forms translated with “go to” and “come to” indicate direction away from or toward the speaker.

**English:** HOW MANY MORPHEMES? The sentences in #1 and #2 below contain several instances of underlined affixes with the same spelling and pronunciation. For each sentence, apply Nida’s Principle 1 to decide how many morphemes are involved.

- List each morpheme with a brief characterization of its function or meaning. This could be either a statement of the grammatical function (changes a noun into a verb) or the meaning that the morpheme adds (do an action again, occurring at periodic intervals)
- For each morpheme, pick one example from the sentence and show the base + affix
- Give one additional word from your own knowledge or a dictionary that contains the same morpheme.

1. His arrival at the memorial service was a surprise to everyone because of his repeated denial that the death was accidental.

2. We expected the Vatican to *excommunicate* the *ex*-priest and to *exclude* him from the examination of his *ex*-parish.