Lexical Analysis
Ling 106
September 10, 2003

Lexicon: contains lexical units with phonological form and meaning (words) and their
grammatical categories.

LEXICON (fragment of English):

<table>
<thead>
<tr>
<th>Grammatical category</th>
<th>Lexical Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proper nouns</td>
<td>Joan, Bill, Philadelphia</td>
</tr>
<tr>
<td>Pronouns</td>
<td>He, she, it, they</td>
</tr>
<tr>
<td>Nouns</td>
<td>Dog, cat, candy</td>
</tr>
<tr>
<td>Intransitive verbs</td>
<td>Walk, sleep, snore</td>
</tr>
<tr>
<td>Transitive verbs</td>
<td>See, find, hug</td>
</tr>
<tr>
<td>Auxiliary verbs</td>
<td>Will, would, could, must</td>
</tr>
<tr>
<td>determiners</td>
<td>The, a, some, every</td>
</tr>
<tr>
<td>prepositions</td>
<td>With, in, on, to, before</td>
</tr>
<tr>
<td>adjectives</td>
<td>Tall, short, green</td>
</tr>
<tr>
<td>adverbs</td>
<td>Quickly, carefully, very</td>
</tr>
<tr>
<td>complementizers</td>
<td>That, if, whether</td>
</tr>
<tr>
<td>conjunctions</td>
<td>And, or, but</td>
</tr>
</tbody>
</table>

Consider the following cases:

(1) An infant, knowing no other language, must determine the words in its
caretaker’s speech and assign them to grammatical categories.

(2) A linguist is dropped into the field and must organize a grammar of the language
spoken locally. Consider two cases:
a. The linguist shares at least one language with some cooperative local.
b. The linguist has no language in common with any of the locals and must
   induce words, their meanings and their grammatical behavior from his
   observations in the field.

Problem 1. Spanish
Isolate the words in the following sentences and state their meanings, so that every
segment of every sentence is assigned to some word:

Unombreestaaki ‘A man is here’
Elgatoestaenfermo ‘The cat is ill’
Elombreestaenfermo ‘The man is ill’
Ungatoestaenfermo ‘A cat is ill’
Elgatoestaaki ‘The cat is here’
Ungatoestaaki ‘A cat is here’
Unombreestaenfermo  “A man is ill’
Elombreestaaki      “The man is here’

Strategy: Semantic recurrence:
Find recurrent phonological sequences that always correlate with a constant component of meaning. The units posited as words should be the minimal sequences meeting this condition without leaving meaningless phonological residues. To apply this strategy, we need several sentences which differ minimally from one another.

Problem 2. Luiseno (Uzto-Aztecan language, Southern California)
Isolate the words in the following sentences and state their meanings

nawitmalqaywukalaqpokiik     “The girl is not walking home’
ya?aspoloov                  “The man is good’
ya?aswukalaqpokiik           “The man is walking home’
nawitmalwukalaqpokiik        “The girl is walking home’
ya?asqaywukalaq              “The man is not walking’
nawitmalqaypoloov            “The girl is not good’
ya?aswukalaqpokiik           “A man is walking home’

Problem: There is no word-to-word correspondence between all languages. The meaning component that is expressed in a separate word in English may not correspond to an independent word in another language and vice-versa.

Phonological Criteria

Some phonological evidence can help us determine what parts out of a phonological string are words. Some examples:

- Consonant clusters
  Some consonantal clusters are prohibited word-internally in many languages. If we encounter one such cluster, there must be a word break in between. A linguist might speculate on grounds such as these that the sentence nawitmalqaywukalaqpokiik in Luiseno should be segmented into words as either [Nawitmal qay wukalaq pokii] or [Nawit malqay wukalaq pokii]. Until he uncovered more data, however, he could not determine which segmentation was correct.

- Stress position
  In French, lexical stress falls on the last vowel of a word. In Papago, stress falls on the first vowel of a word.
  Problem 3. Papago
  a) Locate the word boundaries in the following Papago sentences. Assume that no word end in a consonant cluster (that is all words end in a vowel or in a single consonant), and that no word is longer than three syllables.
máagina?omíd     “The car is running”
wákial?ocíkpan   “The cowboy is working”
wísilo?ohúhu?idgmiistol “The calf is chasing the cat”

b) Using the results of part (a), isolate the words in the following sentences and list them with their meaning.

míd?ogmáagina     “The car is running”
cíposid?ogwákialgwísilo “The cowboy is branding the calf”
míistol?ohúhu?idgwákial “The cat is chasing the cowboy”
cíposid?og?áaligmíistol “The child is branding the cat”
?áali?ocíkpan     “The child is working”