1. What is a Morpheme?

As a first approximation, morphemes can be defined as minimal phonetic sequences that recur with a constant meaning. Individual sound segments (such as [k] of *cat*) do not qualify as morphemes because they cannot properly be said to have any meaning. Clauses, phrases, and many words (for example *houses*) fail to qualify as morphemes because they can be divided into smaller meaningful units.

Problem 1. Luiseno.
Isolate the morphemes in the following sentences and state their meanings.

- Noo wukalaq: I am walking
- Noo paa?iq: I am drinking
- Temet caami paa?ivicunin: The sun will make us want to drink
- Noo poy wukalavicuniq: I am making him want to walk
- Noo paa?in: I will drink
- Noo paa?ivicuq: I want to drink
- Temet poy wukalavicuniq: The sun is making him want to walk

Problem 2. English
Isolate the morphemes in the following words and discuss their semantic value.

<table>
<thead>
<tr>
<th>Word</th>
<th>morpheme 1</th>
<th>morpheme 2</th>
<th>morpheme 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Receive</td>
<td>respect</td>
<td>perceive</td>
<td></td>
</tr>
<tr>
<td>Concur</td>
<td>deceive</td>
<td>inspect</td>
<td></td>
</tr>
<tr>
<td>Expect</td>
<td>report</td>
<td>deport</td>
<td></td>
</tr>
<tr>
<td>Transport</td>
<td>conceive</td>
<td>incur</td>
<td></td>
</tr>
<tr>
<td>Recur</td>
<td>export</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

These words involve *bound morphemes*. These are units that are involved in forming words, but which never occur independently as words. There are two reasons for assuming that elements such as *re, ceive, con, port* and others are morphemes. First, they recur with a high frequency, particularly in combination with one another. Second the unit *ceive* has grammatical significance, since it has the special form *cept* in adjectives derived from the verbs in which it appears: *receptive, deceptive, conceptual, perceptive, perceptual*. The fact that *ceive* is a morpheme entails that *re, de, con, and per* are morphemes in *receive, deceive, conceive and perceive*. 
2. Distributional Criteria

Since not all morphemes can be associated with a constant component of meaning, structural linguists were concerned with finding a method that would break phonetic sequences into morphemes based on distributional criteria. The idea is to identify morphemes by looking at their distribution in a collection of sentences or phrases. The goal of a distributional analysis is to try to isolate recurring patterns and try to correlate these recurring patterns with some unit of meaning.

Problem 3. Serbo-Croatian
Isolate morphemes in the following example (without having a translation).

Yacitam
Yapiyem
Ticitas
Vicitate
Oncita
Onipiyu
Tipiyes
Onpiye
Mipiyemo
Onicitayu
Tipušis

We can use the following strategies:

1. Find sequences that repeat in different contexts. For example, ya repeats in the following two contexts: _______citam, and _____piyem. We infer that such recurring sequences are morphemes.

2. Find contexts that repeat with different things in them, that is slots into which a number of different sequences can be substituted. For example, the context oni_______yu repeats twice: onipiyu, onicitayu. We infer that what goes into a repeating context is a morpheme. For example, puši occurs in the repeating context ti___s, therefore it is a morpheme.

3. When several repeating sequences repeat in the same contexts, each of these sequences is a morpheme. For example, the two sequences piye and cita, repeat in the three contexts ya____m, ti_____s, and on_____.

Some definitions:

The environment of X is the sequence in which X occurs, minus X itself. For example, in the sequence yacitam, the environment of cita is ya____m.

A sequence X is independent in environment E if X can be replaced by some other sequence Y, without changing anything in E.

The distribution of a sequence X is the set of all the environments in which X occurs.
For example, the distribution of a sequence \textit{cita} is the following set of four environments:

\begin{align*}
\text{Ya____m} \\
\text{Ti____s} \\
\text{Oni____yu} \\
\text{Vy____te}
\end{align*}

The \textbf{distributional class} of a sequence $X$ is the class of all other sequences which share the same distribution as $X$.

For example, \textit{cita} and \textit{piye} belong to the same distributional class, because they occur in the same set of environments.

\textbf{Morphemes} are independent sequences which share their distribution with other independent sequences.

3. \textbf{Harris’s conditions}.

\textit{Harris condition 1.}

If, in total environment ___$X$, the combination $AB$ occurs, the combination $CD$ occurs and at least one of the combinations $AD$ or $CB$ occurs (Where $A$, $B$, $C$ and $D$ are each phonetically identifiable portions of speech), then it is possible to recognize $A$, $B$, $C$, and $D$ as being each of them discrete morphemic segments in the environment ___$X$.

Example 1.

Let us apply this first condition for word segmentation to an example.

Consider the following expressions:

\begin{align*}
\text{The junkie stole the car stereo} \\
\text{The junkie trashed my car stereo}
\end{align*}

The “total environment” in this case is

\begin{align*}
\text{The junkie _______ car stereo}
\end{align*}

The potential units we are investigating are:
A= stole
B= the
C= trashed
D= my

We can now apply the test. If any of the following is an acceptable utterance, stole, the, trashed and my fulfill the first condition to be considered distinct morphemes.

The junkie stole my car stereo.
The junkie trashed the car stereo

Example 2.
Apply Harris’s first condition to demarcate not just words but morphemes.

The reasoning we saw above can apply to the following utterances:

\[
\begin{array}{c}
\text{The smurf hid my } \\
\text{apple } s \text{ yesterday} \\
\hline
A \quad B \quad X
\end{array}
\]

\[
\begin{array}{c}
\text{The smurf hid my } \\
\text{wallet } \emptyset \text{ yesterday} \\
\hline
C \quad D \quad X
\end{array}
\]

The smurf hid my ___ yesterday

A = apple
B = s
C = wallet
D = \emptyset

Given that the following are valid utterances, we conclude that apple, s, and wallet can be considered distinct morphemes:

The smurf hid my wallets yesterday
The smurf hid my apple yesterday.

Example 3 and problem:
Couldn’t we follow the same reasoning for bug and split it in two morphemes: bu + g?
The smurf washed my bu g yesterday.

X A B X

The smurf washed my ba ck yesterday.

X C D X

The smurf washed my _______ yesterday.

A = bu
B = g
C = ba
D = ck

Intuitively, we don’t want to say that –g is a distinct morpheme (with its own meaning!). Thus, we need to refine this distributional approach.

Harris’ refinement of the distributional method: *Harris condition II.*

Accord morpheme status to sequences A, B, C, if, for example, A, B. and C occur sometimes after morphemes D, E, or F, but never after G or H, where D, E, and F … constitute a distributional class against G, H.

Idea: The potential units isolated by the first condition must be classifiable into a grammatical category/class in order for them to be considered morphemes. This means that the potential units would have to participate in the standard distributional environments typical of a grammatical category/class.

Applying new rule to example 2: apple + s.
We isolated –s as a potential unit by Harris’ condition 1. Now we have to make a set α with the acceptable preceding units and another set β with unacceptable preceding units.

The smurf hid my apples yesterday
The smurf hid my wallets yesterday
The smurf hid my crayons yesterday
The smurf hid my pencils yesterday
The smurf hid my squirrels yesterday

Set α = {apple, wallet, crayon, pencil, squirrel, …}

*The smurf hid my very + s yesterday.
*The smurf hid my late+s yesterday
*The smurf hid my for+s yesterday
*The smurf hid my think+s yesterday

Set $\beta = \{\text{very, late, for, think, \ldots}\}$

Behavior of elements of $\alpha$ in the following environments:

- The good _____ fell
- The _____ hit the ground
- I’ll take the _____ now
- *The _____+s is good.
- *The _____ are good
- The _____+s are good

Applying new rule to example 3: bu + g.

I love my bug
I love my rug
I love my mug
I love my jug
I love my bag
I love my rag
I love my nag

Set $\alpha = \{\text{bu, ru, mu, ju, ba, ra, na, \ldots}\}$

We cannot find a good number of environments in which all the elements in $\alpha$ would behave the same way. This means that the set $\alpha$ does not constitute a grammatical category/class. By Harris’ condition II, this means that –g is not a distinct morpheme.

Notice, however, the hedge in Harris’s procedure. We have to consider a lot of distributions before we can make secure judgments about morpheme demarcation.

Consider bound morphemes $\text{con-}, \text{re-}, \text{ceive}, \text{-cur}$ in the following words: perceive, deceive, conceive; deduct, conduct, perjure, conjure; persist, desist, consist, resist, assist. Does Harris’ method identify the above as morphemes?

4. Special Cases

The units that the method identifies don’t have to be contiguous, that is they don’t have to form a continuous string of elements.

1. Consonantal roots in Semitic:

- Kataba ‘He wrote’
- Katabtu ‘I wrote’
- Ka:taba ‘He corresponded’ (‘:’ indicates vowel length)
- Ka:tabtu ‘I corresponded’
The above gives the following units:

\[
\begin{align*}
\text{K} & \quad \text{t} & \quad \text{b} & \quad \text{‘write’} \\
\_ & \quad \text{a} & \quad \text{a} & \quad \text{perfective (completed in the past)} \\
\_ & \quad \text{a} & \quad \text{‘he’} \\
\text{vowel length: reciprocal}
\end{align*}
\]

2. Repeated elements are identified as units

- \text{Filius bonus} - ‘good son’
- \text{Filia bona} - ‘good daughter’

3. Other noncontiguous agreement phenomena

- \text{Le bon-∅ fils}
- \text{La bonne fille}

4. Phoneme replacement:

- \text{Take – took}
- \text{Give – gave}
- \text{Sing – sang}
- \text{Hang - hung}

**Extra exercises.**

Exercise 1. Turkish

Turkish is a language with vowel harmony, meaning that all the vowels in a word must be similar in a certain way. There are eight vowels in Turkish, which can be distinguished in terms of three phonological features: high versus non-high, front versus back, and rounded versus unrounded. A vowel can be characterized in terms of its value for each of these three features: \([i’],\) for instance, is a high front rounded vowel, and \([a]\) is a non-high back unrounded vowel.

<table>
<thead>
<tr>
<th></th>
<th>Rounded</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>i</td>
</tr>
<tr>
<td></td>
<td>ü</td>
</tr>
<tr>
<td></td>
<td>u</td>
</tr>
<tr>
<td></td>
<td>ı</td>
</tr>
<tr>
<td>Non-high</td>
<td>e</td>
</tr>
<tr>
<td></td>
<td>ĕ</td>
</tr>
<tr>
<td></td>
<td>o</td>
</tr>
<tr>
<td></td>
<td>a</td>
</tr>
<tr>
<td>front</td>
<td>back</td>
</tr>
</tbody>
</table>
1. Examine the following words and state the rule of vowel harmony:

<table>
<thead>
<tr>
<th>Word</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evde</td>
<td>‘in the house’</td>
</tr>
<tr>
<td>aldım</td>
<td>‘I took’</td>
</tr>
<tr>
<td>Ankara</td>
<td>‘Ankara’</td>
</tr>
<tr>
<td>odun</td>
<td>‘wood’</td>
</tr>
<tr>
<td>Verdim</td>
<td>‘I gave’</td>
</tr>
<tr>
<td>odası</td>
<td>‘his room’</td>
</tr>
<tr>
<td>köprü</td>
<td>‘bridge’</td>
</tr>
<tr>
<td>göstermek</td>
<td>‘to show’</td>
</tr>
</tbody>
</table>

2. Divide the following sentences into words consistent with the requirement of the vowel harmony rule. Assume that no word begins or ends in a cluster of two consonants, and that a consonant between two vowels at a word boundary belongs to the second word.

Cayığıctik  ‘We drank the tea’
odyabircojukgörđüm “I saw a child in the room”

3. Using the results of parts 1 and 2, isolate the words in the following sentences and state their meanings.

kızlarigörđüm  “I saw the girls”
biradamgörđüm “I saw a man”

Exercise 2. Palaulan
Divide expressions into words.

a?adelmeňitakl  ‘the man that sings’
amalkelmerros  ‘the chicken that crows’
akemæňetelrael  ‘the road that is long’
amerroselmalk  ‘the road that is long’
amenjetakle]?ad  ‘the man that sings’
araelelkema?et  ‘the road that is long’
augile]ad  ‘the man that is good’
aklowelmerros  ‘the chicken that is big’
augilelrael  ‘the road that is good’
amenjetakleldil  ‘the woman that sings’

Exercise 3. Apply Harris conditions I and II to show whether the word undo consists of one single morpheme or it consists of two, namely un + do. The purpose of this exercise is NOT to find out what the true answer is, but to see how a purely distributional analysis would do at deriving the correct result. In particular:

- specify your environment _____
- say what A, B, C, and D are (provide a suitable C and D yourself)
- give the ‘crossed’ examples that prove condition I,
- spell out several environments where you test the categorical behavior required in condition II,
- enunciate the conclusion derived by the distributional method