APPENDIX I

What is a morpheme?

Morphemes as structural units of words that are typically but not necessarily meaningful.

Cf. Bloomfield (1933, 161):

"A linguistic form which bears no partial phonetic-semantic resemblances to any other form is ... a morpheme."

OTHER VIEW: Hockett (1958, 123): "Morphemes are the smallest individually meaningful elements in the utterances of a language."

The main problem with Hockett's definition: it makes idioms morphemes, e.g. in he kicked the bucket, meaning 'he died,' there would be three morphemes: he, kick ... the bucket 'die,' and -ed 'past.'

But it seems strange to consider kick ... the bucket a single morpheme (a discontinuous one at that) when it consists of parts that would be morphemes in other contexts, namely kick, the, and bucket.

By Bloomfield's definition -ceive of perceive, receive, deceive, etc. can be considered a morpheme, since it occurs in multiple combinations, even though -ceive has no meaning by itself.

Evidence for considering-ceive a separate morpheme: it has a common irregular allomorph -cept in all its combinations: perception, reception, deception.

The Thematic vowels.
APPENDIX II

HEURISTIC PRINCIPLES IN MORPHOLOGICAL ANALYSIS

Basic principle (Jensen (1990)):

(1) A word must be exhaustively divided into morphemes.

This principle means that, if a word is divided into morphemes at all, each part must be a morpheme, i.e., there must not be any unanalyzed residue, as there would be if otter, for example, were analyzed as ott-er (where -er is identified as 'agentive') since ott- cannot be analyzed as a morpheme.

Nida (1949, Chapter 2) provides six principles for the identification of morphemes.

(2) Nida’s Principle 1:

Forms which have a common semantic distinctiveness and an identical phonetic form in all their occurrences constitute a single morpheme.

This principle allows us to isolate the morpheme -er 'agentive' (doer of an action) in such words as writer, worker, dancer, singer. This form has a single phonetic shape [-er] and the same meaning wherever it occurs. It is semantically distinct from another form of this same phonetic form [-er], the comparative suffix -er of higher, longer, thinner, tastier. These two forms, -er, 'agentive' and -er 'comparative' bear a phonetic similarity (here, in fact, identity) but no semantic similarity and must therefore be counted as separate morphemes.

(3) Nida’s Principle 2

Forms which have a common semantic distinctiveness but which differ in phonemic form (i.e. the phonemes or order of phonemes) may constitute a morpheme provided the distribution of formal differences is phonologically definable.

Italian negative adjectives:

(4) sfortunato 'unlucky'
disabile 'unable'
inelegante 'inelegant'
immangiabile 'inedible'
irragionevole 'unreasonable'
illogico 'illogical'
amorale 'immoral'
anabbaglianti 'not dazzling' (low beam headlights)
Inelegante= 'non elegante'
Immangiabile= 'non mangiabile' ecc.

If we try to isolate the elements that express negativity, we obtain the following analysis:

(5)  
  s  (+fortunato)  
  dis (+abile)  
  in (+ elegante)  
  im (+ mangiabile)  
  ir (+ ragionevole)  
  il (+logico)  
  a (+morale)  
  an (+ abbagliante)

Form/meaning correlations:

(6)  
  a.  s/Negation  
  b.  dis/Negation  
  c.  in/Negation  
  d.  im /Negation  
  e.  ir //Negation  
  f.  il/Negation  
  g.  a /Negation  
  h.  an /Negation

8 morphemes of negation?

Let us first consider the morphemes in (6 c-f), and provide more data:

(7)  
  a.  in + abile  
      in + elegante  
      in + utile  
      in + intellegibile  
      in + operoso  
  b.  in + tollerabile  
      in + distruttibile  
      in + numerevole  
  c.  im + mangiabile  
      in + probabile  
      in + battuto  
  d.  ir + ragionevole  
      ir + razionale  
      ir + reale  
  e.  il + logico  
      il + letterato  
      il + legale
Predictable features of pronunciation can be found both in cases in which there is no alternation, within morphemes, and in cases in which there is a morphemic alternation:

Conclusion: In (7) we are dealing with different allomorphs of the same underlying morpheme /in/.

Let us now consider the forms s and dis. Are they allomorphs of the same morpheme?

(8)  
\begin{align*}
\text{s + Aggettivo} & \quad \text{dis + Aggettivo} \\
\text{a.} & \quad *s+onesto \quad \text{a'.} \quad \text{dis+onesto} \\
& \quad *s+abile \quad \text{dis+abile} \\
& \quad *s+utile \quad \text{dis+utile} \\
& \quad *s + incantato \quad \text{dis + incantato} \\
\text{b.} & \quad \text{s + conveniente} \quad \text{b'.} \quad *\text{dis + conveniente'} \\
& \quad \text{s+ cortese} \quad \text{dis + cortese} \\
& \quad \text{s + fortunato} \quad *\text{dis + fortanato} \\
& \quad \text{s + leale} \quad \text{dis + leale}
\end{align*}

But consider further data;

(9)  
\begin{align*}
\text{s+continuo} & \quad \text{dis+continuo} \\
\text{a.} & \quad *s+convenevole \quad \text{b.} \quad \text{dis+convenevole} \\
& \quad *s+giungibile \quad \text{dis + giungibile}
\end{align*}

\text{dis} and \text{s} can appear in the same phonological context:

(10)  
\begin{align*}
\text{s+conveniente} & \quad \text{dis + continuo}
\end{align*}

Therefore /s/ and /dis/ are different morphemes.

Consider \text{a/an},

(11)  
\begin{align*}
\text{a.} & \quad \text{a + critico} \quad \text{b.} \quad \text{an + abbagliante} \\
& \quad \text{a + partitico} \quad \text{an + alfabeto} \\
& \quad \text{a + simmetrico} \quad \text{an + archia}
\end{align*}

(12)  
\begin{align*}
*\text{ancritico} & \quad *\text{aarchia}
\end{align*}

/\text{a}/ and /\text{an}/ are allomorphes of the same morpheme:

(13)  
\begin{align*}
morphemes & \quad \text{s} & \quad \text{dis} & \quad \text{a} & \quad \text{in} \\
allomorphs & \quad \text{s} & \quad \text{dis} & \quad \text{a} & \quad \text{an} & \quad \text{in} \quad \text{im} \quad \text{il} \quad \text{ir}
\end{align*}
(14) Nida's Principle 3 (Simplified):

Identify forms as a single morpheme even when their distribution is not phonologically determined provided that they have a common meaning and are in complementary distribution.

(15) \( ox \quad oxen \)
    \( box \quad boxes \)

(16) Complementary Distribution

Two or more allomorphs are in complementary distribution if the environment in which one occurs is not an environment where any of the others can occur.

Notice that the irregular plural allomorphs are not entirely in complementary distribution.

(17) singular        irregular        regular
    cherub          cherubim        cherubs
    schema          schemata        schemas
    myth [miθ]      [miðz]          [miθs]
    thesis          theses          theseses

Nida allows such forms to belong to the same morpheme "if the difference in meaning of the allomorphs reflects the distribution of these forms" (1949, 42). In other words, forms like \( cherubim \) and \( cherubs \) have slightly different connotations which parallel the distribution of the two plural allomorphs.

Another case where two forms may be considered as belonging to a single morpheme occurs when the forms are found in free variation, defined in (18).

(18) Free Variation

Two or more forms are in free variation if they have the same meaning, are phonetically distinct, and may occur in exactly the same environments.

(19) either \(([iðə]/[aɪðə])\),
    \( vase \quad ([va:z]/[veyz]/[veys]) \)
    \( economics \quad ([ɪkənəmlks]/[ɛkənəmlks]) \)

Nida’s Principle 5 discusses the conditions under which homophonous forms are identified as the same or different morphemes. We can quote the two parts of the Principle (1949, 56).
(20) Nida’s Principle 5

1. Homophonic forms with distinctly different meanings constitute different morphemes.

2. Homophonic forms with related meanings constitute a single morpheme if the meaning classes are paralleled by distributional differences, but they constitute multiple morphemes if the meaning classes are not paralleled by distributional differences.

Case 1:

(21) *pair, pare, and pear* constitute separate morphemes.

(22) French *verre* 'glass', *vers* 'verse', *vers* 'toward', *vert* 'green', and *ver* worm'.

Case 2:

Examples of homophonic forms with related meanings.

(23) (a) *fish* (to) *fish*,

We can say that there is a single morpheme *fish* and that the verb to fish is derived from this noun by ZERO DERIVATION.

(24) \[[fish]_N \emptyset]_V

(25) Nida’s principle 6:

A morpheme is isolatable if it occurs under the following conditions:

1. in isolation.
2. in multiple combinations in at least one of which the unit with which it is combined occurs in isolation or in other combinations.
3. in a single combination provided the element with which it is combined occurs in isolation or in combinations with nonunique constituents.

Condition 1 of Principle 6 allows us to isolate as morphemes such forms as *girl, cat, play, work, out, he, that, and hey*, which all occur in isolation, provided there is no way of analyzing the forms into smaller morpheme.

Condition 2 allows us to isolate forms such as agentive *-er* as morphemes. This never occurs in isolation but it does occur with forms that occur in isolation, as *dance, play, work* in *dancer, player, worker*. Of course, *-er* can also be isolated as a morpheme by Principle 1.
Condition 2 also allows us to consider the prefix *con-* a morpheme, since it occurs in *conceive, consume, contain, condense*. Dense occurs in isolation and the other stems occur in other combinations, e.g. *perceive, resume, detain*. These roots and stems occur in multiple combinations, as shown in (26).

(26)  
<table>
<thead>
<tr>
<th>English</th>
<th>compel</th>
<th>impel</th>
<th>demit</th>
<th>permit</th>
</tr>
</thead>
<tbody>
<tr>
<td>repel</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>remit</td>
<td>commit</td>
<td></td>
<td>demit</td>
<td>permit</td>
</tr>
<tr>
<td>refer</td>
<td>confer</td>
<td>infer</td>
<td>defer</td>
<td></td>
</tr>
<tr>
<td>resume</td>
<td>consume</td>
<td></td>
<td>deceive</td>
<td>perceive</td>
</tr>
<tr>
<td>receive</td>
<td>conceive</td>
<td>induce</td>
<td>deduce</td>
<td></td>
</tr>
<tr>
<td>reduce</td>
<td>conduce</td>
<td>induce</td>
<td>deduce</td>
<td></td>
</tr>
</tbody>
</table>

Condition 3 permits us to isolate the morpheme *cran-* in *cranberry*. Cran occurs just in this single combination, but the element with which it is combined, *berry*, occurs in isolation and in combinations with other constituents, such as *strawberry, blueberry, raspberry*, etc. Likewise, *boysen* as in *boysenberry* can be isolated as a morpheme, for the same reasons. Notice that we cannot assign any meaning to either *cran-* or *boysen-*. They seem simply to convey the meaning of 'a kind of berry, different from other named berries.'

Nida 1949, 60:  
Not all recurring phonetic elements can be isolated as morphemes.

Principle 6 specifically excludes the element *-er* from morphemic status in *hammer, ladder, otter, badger, under, linger, and bitter* (examples from; other examples include *udder* and *rudder*). In these examples, the elements with which this *-er* occurs, namely *hamm-, ladd-, ott-* etc. occur only in combination with *-er*.

A second type of recurring element that is specifically excluded from morphemic status by Principle 6 goes by the name phonetic symbolism. Nida (1949, 61) gives this example.

"In the English series *slide, slush, slurp, slip, slop, slime, slobber,* and *slick* we can recognize a common phonemic element *sl* with a common meaningful relationship which may be defined as 'smoothly wet.' A series such as *flash, flare, flame, flicker,* flimmer have an analogous relationship involving the meaning 'moving light.' Despite these partial phonetic-semantic resemblances, however, we do not isolate either *sl-* or *fl-* as morphemes, since they do not occur with free forms or with forms which occur in other combinations."

7
APPENDIX III

Portmanteaux and Feature Cumulation

(27) Mezquital Otomi (Jensen 1990).

1. di⁄mpeni⁄ 'I launder'
2. dánñhi 'I hurried'
3. dampení 'he will launder'
4. dínñhi 'I hurry'
5. bimpení 'he laundered'
6. binñhi 'he hurried'
7. giññhi 'you will hurry'
8. dámpení 'I laundered'
9. gimpení 'you launder'
10. bíññhi 'he hurries'
11. gîmpeni⁄ 'you launders'
12. gánñhi 'you hurried'
13. bîmpení 'he launders'
14. gnñhi 'I will hurry'
15. gampení 'I will launder'
16. danñhi 'he will hurry'
17. gámpení 'you laundered'
18. gîññhi 'you hurry'

Two stems: mpeni'laundry' and ntihi 'hurry.'

We cannot, however, isolate affixes meaning 'I,' 'you,' 'he,' 'past,' 'present,' and 'future.' The prefixes combine person and tense, as shown in (28).

(28) Sayula Popoluca (Jensen 1990))

1. tače?mp 'I am seeking you'
2. tanče?mp 'I am seeking him or it'
3. ?inče?mp 'you are seeking him or it'
4. ?iče?mp 'he is seeking him or it'
5. ?išče?mp 'he is seeking you or you are seeking me'
6. tase?mp 'he is seeking me'
7. ta?e?p 'I am looking at you'
8. tan?e?p 'I am looking at him or it'
9. ?in?e?p 'you are looking at him or it'
10. ?i?e?p 'he is looking at him or it'
11. ?iš?e?p 'he is looking at you or you are looking at me'
12. tae?e?p 'he is looking at me'
13. tanhúyp 'I am buying it'
Here we can isolate three stems: če?m 'seek,' ?e? 'see,' and hůy 'buy.'

The first position after the stem permits an optional suffix -ha, which makes the object indirect. The second position after the stem contains a tense morpheme -p 'present' or -w 'past.' The prefix position is for a portmanteau morpheme incorporating the subject and the object:

(30)

<table>
<thead>
<tr>
<th>Subject</th>
<th>Object</th>
</tr>
</thead>
<tbody>
<tr>
<td>[+speaker]</td>
<td>[-speaker, +partic.] [-speaker, -partic.]</td>
</tr>
<tr>
<td>[-speaker, +partip.]</td>
<td>iš -</td>
</tr>
<tr>
<td>[-speaker, -particip.]</td>
<td>taš -</td>
</tr>
</tbody>
</table>

Gaps in (30) accounted for by Disjoint Reference.
APPENDIX IV

MARKEDNESS

The dual is a highly marked number, since it occurs in relatively few languages, it is less frequent than singular or plural in languages that have it, and its presence implies the presence of singular and plural.

Greenberg 1966, Chapter 3 notes the following characteristics of morphological markedness:

1. Frequently, an unmarked term is ambiguous between the generic category and the specific opposite of the marked term, as *man* means either 'a human being' or a 'male human being.'

2. An unmarked term frequently has zero expression, as *author* as opposed to *authoress.*

3. Syncretisms are more frequent in marked categories; for example, the definite article in German distinguishes gender in the singular but not in the plural. In some languages the overt expression of a marked category may be omitted. For example, in Korean, a word without *-tul* may be interpreted as either singular or plural, while a word with *-tul* is only plural.

4. If the opposition between marked and unmarked is suspended, the unmarked term appears. For example, in Hungarian, the singular form of nouns appears after cardinal numbers.

5. Marked forms often exhibit a lesser degree of morphological irregularity.

6. Marked categories often lack certain distinctions present in the corresponding unmarked categories, as Latin has only four tense-aspect distinctions in the subjunctive (marked) compared to six in the indicative (unmarked).

7. A set can be named by the plural (or dual) of the unmarked category of the set, as Spanish *padres* (literally 'fathers') is used for 'parents.'

8. Agreement often requires the unmarked category in a modifier of a phrase that includes both: in Spanish, an adjective modifying a noun phrase that includes both masculine and feminine nouns is masculine.

9. Finally, the unmarked term appears more frequently in running texts.

But English: a suffix marker is found only for the unmarked third person singular
MARKEDNESS THEORY AND CASE SYSTEMS

Implicational relationships between the different Cases (from Blake (1994)):

(31) \( \text{NOM} < \text{ACC} < \text{GEN} < \text{DAT} < \text{LOC} < \text{ABL} < \text{OTHERS} \)

\( \text{INST} \)

(32) Case systems: (The lowest Case in the system usually has the functions of the other Cases appearing in the hierarchy. When this occurs we say that it is \text{Obl} (=\text{Oblique}):)

a. Two Cases:
   \( \text{NOM} - \text{ACC(Obl)} \)
   (e.g. in Chemehuevi, Kabardian)

b. Three Cases:
   \( \text{NOM} - \text{ACC} - \text{GEN} \)
   (e.g. in Classical Arabic, Modern Greek)

c. Four Cases:
   \( \text{NOM} - \text{ACC} - \text{GEN} - \text{DAT(obl)} \)
   (e.g. in Ancient Greek, Nuer)

d. Five Cases:
   \( \text{NOM} - \text{ACC} - \text{GEN} - \text{DAT} - \text{ABL(Obl)} \)
   (e.g. in Latin)
   \( \text{NOM} - \text{ACC} - \text{GEN} - \text{DAT} - \text{INST(Obl)} \)
   (e.g. in O. H. German)

e. Six Cases:
   \( \text{NOM} - \text{ACC} - \text{GEN} - \text{DAT} - \text{LOC} - \text{ABL} \)
   (e.g. in Turkish)
   \( \text{NOM} - \text{ACC} - \text{GEN} - \text{DAT} - \text{LOC} - \text{INST} \)
   (e.g. in Slavic lgs)

f. Seven Cases:
   \( \text{NOM} - \text{ACC} - \text{GEN} - \text{DAT} - \text{LOC} - \text{ABL} - \text{INST} \)
   (e.g. in C. Armenian)

g. Systems with more Cases may include differentiations in the local Cases (allative/ perlatve, etc.), the comitative, the purposive, the comparative, and some other special Cases:

   e.g. Tamil:
   \( \text{NOM} - \text{ACC} - \text{GEN} - \text{DAT} - \text{LOC} - \text{ABL} - \text{INST} - \text{COM} \)

   Toda:
   \( \text{NOM} - \text{ACC} - \text{GEN} - \text{DAT} - \text{LOC} - \text{ABL} - \text{INST} - \text{COM} - \text{PURP} \)

The realization of certain Cases as morphological affixes is disliked. Languages seem to prefer affixal Case-marking for grammatical relations such as subject and object but not for grammatical relations expressing location or instrument. In terms of the theory of markedness we can say that the affixal realization of Cases such as the ablative and the instrumental is marked.