INTRODUCTION TO THEORETICAL LINGUISTICS: LECTURE NOTES
Stanley Dubinsky
University of Iowa

PREFACE

The syllabus and lecture notes presented here are for a course entitled "Introduction to Linguistics." Actually, given the course content one might more appropriately entitle the course "Introduction to Theoretical Linguistics" or "Introduction to General Linguistics." This is so because the course represents an introduction to the core curricula of general linguistics, focusing on phonetics, phonology, morphology and syntax. No attempt is made to present the total range of linguistic sciences: the so-called hyphenated linguistics, such as socio-linguistics and psycho-linguistics. As such, this course is not a lower level survey course in linguistics. Rather, the material presented herein is designed to provide a one semester introduction to general linguistics for upper level undergraduate students, and serve as a prerequisite for courses in the above mentioned core curricula.

This course is designed to introduce students to the basic concepts that play a part in most current theories of language, rather than present to them the ideas of any particular theory. Thus, for example, while constituency and syntactic categories are discussed in detail, syntactic transformations are nowhere mentioned. The lecture notes themselves are divided into two major units, each of which could itself be subdivided into two sections. The first unit, phonetics/phonology, begins with a general introduction and an articulatory description of segmental phonemes. It continues with a treatment of syllabic structure, suprasegmentals, features and markedness. The second unit, syntax/morphology, introduces a model of language having distinct lexical and syntactic components, and presents syntactic categories and phrase structure. The second half of this unit is divided into a discussion of the sentence and of morphology.

The lecture notes themselves were designed to be mimeographed and handed out to students before each lecture. This was done in order to maximize the students' ability to concentrate on understanding the lecture, rather than on trying to keep up with the lecturer in their notebooks. While students were advised to take supplementary notes, every student had an accurate record of the basic content of the lectures, which could be referred to in preparation for examinations. Having lecture notes distributed in class also saved having to present a great deal of material on a chalk board.
While the notes are not designed to take the place of textbook material, they are intended to supplement the texts and to provide a framework within which textbook material from different sources can be understood. In this spirit, references to chapters and sections are made either when the notes are adapted from a specific source, or when reference to that source develops the topic in greater detail. This is done with the idea that students should have access to at least two perspectives on a given topic.
SYLLABUS

Texts: Required

Sloat, Taylor and Hoard (STH) 1978.
Introduction to Phonology. Prentice-Hall.

Lyons, John (LYO) 1968.
Introduction to Theoretical Linguistics. Cambridge.

Recommended

Matthews, P.H. (MAT) 1981.
Syntax. Cambridge.

Ladefoged, Peter (LAD) 1982.
A Course in Phonetics. Harcourt, Brace, Jovanovich.

Unit 1: Phonetics/Phonology

Week     Readings       Topic
          
1         LYO:Ch 1       Introduction to linguistic theory.
          LYO:Ch 2.2     Substance and Form

2         STH:Ch 1       Introduction: Phonetics/phonology
          STH:Ch 2       Vowels: production/characterization

3         STH:Ch 3       Consonants: production/characterization
          LYO:Ch 3       EXAM 1

4         STH:Ch 4       EXAM REVIEW
          LAD:Ch 10      Syllables: definition/description

5         STH:Ch 5       Suprasegmentals: phonetic character
          functions of pitch, length, amplitude

6         STH:Ch 6       Features: definition/motivation
          LAD:Ch 11      classes of features
7     STH:Ch 7        Naturalness and markedness
       syllabic structures/segments
       EXAM 2

Unit 2: Syntax/Morphology

<table>
<thead>
<tr>
<th>Week</th>
<th>Readings</th>
<th>Topic</th>
</tr>
</thead>
</table>
| 8    | LYO:Ch 4.1 | EXAM REVIEW
       |          | Components of language                    |
|      | LYO:Ch 4.2 | Lexicon and grammar                        |
| 9    | LYO:Ch 5.2 | Grammatical acceptability                  |
|      | LYO:Ch 5.3 | Grammatical categories                     |
|      | MAT:Ch 3  | Projection of phrasal categories           |
| 10   | MAT:Ch 4  | Constituency, dependency and word order    |
|      | LYO:Ch 6.1| EXAM REVIEW                                |
|      | LYO:Ch 6.2| Components of language                    |
| 11   | MAT:Ch 2  | Phrase markers                             |
|      |          | Sentences                                  |
| 12   | LYO:Ch 5.2| Sentences (cont.)                          |
|      |          | EXAM 3                                     |
| 13   | LYO:Ch 5.3| EXAM REVIEW                                |
|      | LYO:Ch 5.4| Words and morphemes                        |
| 14   | MAT:Ch 3  | Word and morphemes (cont.)                 |
LECTURE NOTES - UNIT 1

I. Introduction

A. Development of linguistic theory. [Lyons:ch. 1]

1. Classical linguistics.
   Writing of grammars to preserve the purity and correct usage of official/classical languages.
   - Greek grammars (by Greeks and Romans)
   - Latin grammars (by medieval scholars)
   Prescriptive adequacy.

2. Structural linguistics.
   - Indo-European comparative linguistics and reconstruction.
   - Grammars of exotic (non Indo-European languages).
   Descriptive adequacy.

   - Treatment of language as partly inherited.
   - Cross-linguistic typology.
   - Search for the universal component of languages.
   Explanatory adequacy.

B. Descriptive applications.

- Teaching second languages.
- Speech disorders: pathology and therapy.
- Computational linguistics: NLP and speech synthesis.

C. Theoretical goals.

- Identification of UG (versus that which is learned).
- Development of adequate theories of cognition.

- Strong hypothesis: all languages are the same.
- Weak hypothesis: all languages are different.

- Abstractness: the investigational dilemma.
- Language capacity as a 'black box'.
Sources of evidence:
  Aphasia; child language acquisition; speech disorders;
speaker errors; native speaker judgements.

D. Divisions of theoretical linguistics

Phonetics - the study of the substance of speech.
  aspiration after initial [p] but not [sp].

Phonology - the organization of that substance.
  voicing in English: [p] [t] [k] vs [b] [d] [g].

Morphology - the process of word formation.

Syntax - the internal structure of utterances.
  I read this book. vs Kono hon o yonda.
  Vs. This book was read by me.

Semantics - the association of meanings with structure.
  All the students brought one gift.
  One gift was brought by all the students.

II. Substance and form [Lyons:pp. 53-70]

A. Sounds and words

  The COMPOSITION of sounds CORRELATES with words, which are
  themselves COMPOSED to form sentences.

  Sounds are combined.
  Combinations of sounds are ARBITRARILY associated with meanings.
  These units (words) are combined to form utterances.

  The DOUBLE-ARTICULATION (or double composition) of language:
  Independent composition on two levels which are correlated
  arbitrarily.

B. Substance and form.

  Language is FORM not SUBSTANCE. (Saussure)
1. SUBSTANCE is arbitrary (but PHONIC substance is primary).

Sounds [a]:[i] can be replaced by graphic representations A:I; can be replaced by numbers 1:9 (base 10), 1:11 (base 8), 1:1001 (base 2); can be replaced by electrical impulses; or signs.

2. FORM: the systematic relationships between units of the same substance (the material through which language is represented in a given utterance or discourse is generally uniform).

The formal relationship between units rather than their absolute value is what counts:

Stress in shouted speech vs. soft speech.
High/low tones among different speakers.
High/low vowels among different speakers.

3. ISOMORPHISM (having the same form).

One can replace the elements used to express a formal relationship without changing the form or meaning of the expression.

"7 times 4 equals 28" is equivalent to (expresses the same relationship as) "111*100=11100"

The systematic relationship holding between /b/ /d/ /g/ and /p/ /t/ /k/ could be realized by /p/ /t/ /k/ and /p'/ /t'/ /k'/.

\[
\begin{align*}
\text{tab} : \text{tap} & \rightarrow \text{tap} : \text{tap}' \\
\text{made} : \text{mate} & \rightarrow \text{mate} : \text{mat}'e \\
\text{plug} : \text{pluck} & \rightarrow \text{pluk} : \text{pluk}' \\
\end{align*}
\]

The reason that a speaker with a speech impediment is generally understood is due to a partial isomorphism. One sound is uniformly replaced by another: [r] \rightarrow [w]

\[
\begin{align*}
wight & \rightarrow \text{wabbit stwetch} \\
\end{align*}
\]

This breaks down because the isomorphism is only partial:

\[
\begin{align*}
\text{wide(verb, noun)} & \rightarrow \text{wide(adj)} \\
\end{align*}
\]

SEMANTIC isomorphism exists between languages when there is a one-to-one correspondence between two sets of terms and their referents:
Semantic isomorphism between languages occurs within specific semantic fields. No two languages are completely isomorphic semantically.

4. Language is abstract.

A PHONOLOGICAL word is a combination of sounds that represents a GRAMMATICAL word.

A phonological word is a particular combination of sounds.

A grammatical word is a particular meaning.

Two distinct grammatical words can be realized by the same phonetic word: /down/ (prep and noun).

Two distinct phonological words can realize the same grammatical word: /dreamed/ and /dreamt/.

III. Phonetics/Phonology: introduction.

A. Phonetics: the study of the phonic substance of language.
Phonology: the study of its organization (the formal relationship among these sounds).

Sounds can be substantially (or phonetically) distinct, yet formally (phonologically) identical.

English has /p/ and /b/ which are distinct and serve to distinguish words: /lop/ vs /lob/.

The /p/ in /lop/ is slightly aspirated, as distinct from the /p/ in /spoil/. Although the aspirated and unaspirated /p/ are measurably different, they 'count for the same sound /p/ in English. In Chinese, these sounds are both phonetically AND phonemically distinct.

B. Types of phonetic evaluation.

Acoustic phonetics: the description of language sounds from their acoustic qualities. [a] is distinct from [i] in that the F1 of [i] is higher and further apart from F2 than that of [a].
Articulatory (physiological) phonetics: the description of language sounds in terms of how they are produced. [a] is distinct from [i] in that in the production of [a] the tongue is low in the mouth, while in the production of [i] it is high and shifted forward.

Auditory phonetics: the description of sounds in terms of how they are recorded by the human ear and transmitted to the brain. Constrained by the limitation of current research in neurobiology and cognition.

C. Phonological system of a particular language.

1. The INVENTORY (set) of sounds that the language uses (drawn from the larger set of possible human sounds)

   English has /p/, /b/, and /f/; but not /φ/ (except for some speakers in nymph /nimφ/).
   Japanese has /p/, /b/ and /φ/; but not /f/.

2. The COMBINATORIAL POSSIBILITIES (possible co-occurrence of these sounds).

   English has: /s/, /p/, and /b/ which can each occur word initially: sat:pat:bat. However, /p/ but not /b/ can occur between /s/ and a vowel: spot *shot.

3. SOUND PROCESSES: rules that affect sound changes.

   In English (and many other languages), unstressed vowels are often reduced to /ə/ (schwa).

   /sōfa/ --> /səfa/
   /inspiråtion/ --> /inspiråtion/

D. Universals

1. True UNIVERSALS: characteristics shared by all human languages.
   The presence of both vowels and consonants.
   The organization of these sounds into syllables.
   The existence of the syllable shape CV.

2. NEAR universals: characteristics shared by nearly all human languages. The presence of the sound [n].

3. IMPLICATIONAL universal:
   If a language has X, then it also has Y.
If a language has [d], then it has [t].
If a language has [e], then it has [i].

Provides evidence for the relative complexity (MARKEDNESS) of sounds: [d] is the same as [t], with one additional complexity (FEATURE) - it is voiced.

If we think of [d] as being '[t] plus voicing' then we can understand how its existence entails [t].

IV. Vowels

A. Organs of production


2. Articulators: vocal chords (in larynx); tongue; lips; pharynx.

B. Production

1. Vocal chords (in larynx) provide the vibration basic to nearly all vowel sounds.
   Without this we produce voiceless (whispered) vowels.

2. The tongue moves higher/lower and forward/back in the mouth to divide the oral-pharyngeal cavity into two major resonating chambers.

For [a], the tongue is low and back.
For [i], tongue is high and front.

\[
\begin{array}{cccccc}
F2 & 4.25 & 4 & 3.4 & 2.6 & 2.5 \\
& & 2.5 & 1.5 & 1.5 & 1.5 \\
F1 & 1 & 1 & 1 & 1 & 1 \\
\end{array}
\]

(for French; numbers are octaval increments)

In [a], the two resonators are most nearly equal in size; [a] is COMPACT.
In [i], they are most different in size; [i] is DIFFUSE.
3. The velum normally closes off the nasal cavity during speech. When relaxed, sound resonates in this cavity as well, producing nasalized vowels.

   English: 'bone' /bon/ vs French: 'bon' /bɔ̃/

   For this reason, we cannot produce nasal sounds when our nasal passages are blocked.

   I have a cold in my nose. --> I have a cold iby dose.

4. Lip rounding provides an additional resonating area to the oral cavity, resulting in rounded vowel.

   In English, rounding normally accompanies non-low back vowels: /u/ /o/.
   In Japanese, it does not: /w/ /ŋ/.
   In German, rounding is distinctive: /i/ /ʊ/.

5. Pharynx: tightening of the muscles in the vocal tract produces a tense/lax distinction:

   /i/ in Pete: /ɪ/ in pit

C. Phonetic characterization (features):

1. The basic division of the oral cavity is articulated by the tongue:

   front  central  back
   high
   mid
   low

2. Additional features are induced by the other articulators:

   Nasal - by relaxation of the velum.
   Rounded - by the rounding of the lips.
   Tense - by the tightening of the vocal tract.

D. Secondary vowels: vowels which are UNcharacteristically rounded or unrounded.

   Non-low back vowels are usually rounded.
   Low vowels and non-low non-back vowels are normally unrounded.

   Secondary vowels: unrounded - non-low back
   rounded - non-low non-back and low back
   (low front rounded does not exist).
E. Monopthongs/Dipthongs

Monopthong is a vowel articulated in one position of the vowel quadrilateral. Dipthongs begin in one position and finish in another.

F. Other distinctions

1. Length: not restricted to vowels. Many languages which distinguish long and short vowels, also have long and short consonants.

Japanese: /sosai/ 'vegetable'  
/so:sai/ 'counterbalance'  
/dango/ 'dumpling'  
/dango:/ 'confer'  
/kata/ 'style,shape'  
/kat:a/ 'purchased'

2. Voiceless vowels: production of a vowel with no vibration of the vocal chords.

In Japanese, /i/ and /u/ have voiceless alternates following voiceless consonants (especially word finally):

/swi/ 'sushi'
/desu/ 'is'

V. Consonants

A. Organs of production

1. Articulators: lower lip, tongue, pharynx, glottis (velum), (vocal chords)

2. Points of articulation: (front --> back)

bilabial:labio-dental:dental:alveolar:  
palatal:velar:uvular:pharyngeal:glottal

B. Production

1. Obstruents: an articulator (lower lip, tongue, glottis) is brought into contact with or proximity to a point of articulation.
If the constriction is complete, the result is a STOP.  
[p], [b], [k]

If it is nearly complete and creates turbulence, the result is a FRICATIVE.  
[f], [s], [z], [x]

A stop followed closely by a fricative is an AFFRICATE.  
[c], [j]

All these may be voiced or unvoiced.

2. Resonants: All vowels are resonants. In addition, some consonants are resonants.

Resonant consonants:  
Differ from vowels in that they FUNCTION as consonants (and are generally more constricted).  
Differ from obstruents in that they involve the shaping of the vocal tract into a RESONATING shape rather than obstructing it.  
By definition they entail voicing.

Nasals:  
Voiced stops with the velum relaxed, allowing sound to resonate in the nasal cavity.  
Possible points of articulation: labial --> uvular  
NO pharyngeal/glottal nasal

Lateral resonants [l]:  
Less severe constriction than lateral fricatives.  
Air resonates around sides of tongue.  

r-sounds:  
All involve pharyngeal constriction in addition to some articulation by tongue.

Glides:  
[y] and [w] phonetically similar to the vowels [i] and [u].  
Possibly slightly higher (more constricted).  
Crucially, they FUNCTION as consonants.

3. Other features.

To the above we can add other features that serve to distinguish sounds from each other.

In addition to VOICING, mentioned above, there are SECONDARY ARTICULATIONS (accompanying characteristics) such as
Aspiration: release of air  
Labialization: rounding of lips  
etc.

The air stream may be modified. While most sounds are produced with air brought up from the lungs: Egressive  
Sounds can also be produced (rarely) while air is brought INTO the lungs: Ingressive.

Consonants (like vowels) may be long or short:  
Italian: Gianni; stacatto  
Japanese: kippu (ticket); Hokkaido

VI. The Syllable  [LAD Ch. 10]

"'Sar-ah,' he called. 'Sar-ah,' spacing the syllables with an unbearable falsity." Graham Greene [Lowenstamm (1981)]

A. Syllables vs. segments

After segmental sounds, syllables are the next larger unit of phonological analysis.

The psychological reality of the syllable is greater than that of the segment for the speakers of almost any language:

Children are conscious of the syllable but not of the segment until the learn a segmental writing system.

Most writing systems in the world are syllabic:  
Chinese characters use one symbol per syllable.  
Korean - hangul system (most efficient in world).  
Japanese, Cherokee ... use syllabary one symbol for each CV unit plus syllable final consonants such as nasals.

Alphabetic writing systems are all derived from Greek innovation of 2000 b.c.e. - they adapted the Semitic syllabary using one sound per segment.

The number of syllables in a given word can be identified by any speaker, even when that speaker is unsure how many segmental sounds it contains. Most speakers of (the same dialect) of a language will agree on the number of syllables in almost all words.

Differences in the perception of syllables in a particular word by different speakers can be attributed to regular differences in pronunciation, and is usually predictable.
How many syllables are there in:

mysticism, communism; real, meal; hire, fire; mirror, error;
heavier, mediate.

hire:fire   hour:tower

B. Defining the syllable

There is NO agreed upon definition of what is a syllable.
Possible definitions include:

1. Unit containing a single burst of air (chest pulse)
   [definition proposed by psychologist R.H. Stetson].
   However phonetic facts data do not support this theory:
   - city = 2 syllables/1 chest pulse
   - pots = 1 syllable/2 chest pulses

2. Unit containing a vowel (or diphthong).
   - Chinese  [ŋi] 'yes'
   - English  [mɛcɛ] 'mention'

3. Unit containing a resonant.

   What about unvoiced vowels? Recognition of whispered speech?
   - [ræn] 'ran' is 1 syllable containing 3 resonants

   Ladefoged: to define syllable one must -
   1) account for all words where there is agreement about syllable
      structure.
   2) explain disagreement in other words.

4. SONORITY: relative loudness of a sounds with accent, pitch, and
length held constant. In other words, the relative amount of
sound energy inherent in a particular sound.

   Open vowels are highest, followed by close vowels, then liquids,
nasals, voiced fricatives, unvoiced fricatives and finally stops.
Relative sonority for different sounds differs somewhat from one speaker to the next.

One way of defining syllables is by noting that a SYLLABLE NUCLEUS coincides with a SONORITY PEAK:

'visit'    'divide'

Confusion over number of syllables may arise due to differences in sonority peaks for different speakers.

'prism' [z] vs. [m]

However sonority peaks cannot completely account for syllabification:
'spa' technically has two sonority peaks, but only one syllable

Cannot account for why 'frightening' is 2 syllables for some and 3 syllables for others:

[fra t-n n] vs. [fra t-n- n]

5. The problems with SONORITY as a measure of syllabic identity can be partly solved by taking PROMINENCE PEAKS to be the defining characteristic of syllables.

PROMINENCE: the SONORITY (or characteristic sound energy of a sound) plus actual LENGTH, STRESS, and PITCH.

It is LENGTH and STRESS of the first [n] sound that distinguishes the words

'lightning' (2 syl.) and 'lightening' (3 syl.)

for some speakers. And distinguishes the sounds

[hid-η-eimz] 'hidden aims' and [hid ne lýzm] 'hid names'

6. Problem with using relative prominence as a way of defining syllables is that there is no formula to compute prominence from the combined values of a word's sonority, length, stress, and pitch. Thus, RELATIVE PROMINENCE is a purely SUBJECTIVE concept.

The alternative to a phonetic definition of syllable is to say
that the syllable is a PHONOLOGICAL (or mental) structure and therefore cannot be defined phonetically.

Evidence for the syllable in mental representations comes from speech errors:

malt whiskey --> walt miskey
flue shots --> shu flats
split brain --> sprit blain
my sock has a hole --> my sole has a hock

Syllable initial consonants always switch with initials and finals with finals.

C. Description of a syllable

A syllable is built around a PEAK (or NUCLEUS), it is delimited by a JUNCTURE (or BOUNDARY), and it has internal STRUCTURE.

1. The most PROMINENT segment of a syllable is that syllable's PEAK (or NUCLEUS). These segments can also be referred to as SYLLABIC SEGMENTS.

In general, VOWELS are syllabic and CONSONANTS are not. However, there are exceptions to this. Thus, NON-syllabic vowels and syllabic consonants are specially marked:

[i] [u] of diphthongs
syllabic [n] [r] and [l]

2. We can often use phonetic evidence to accurately identify syllable JUNCTURE. For example, since we know that [t] is aspirated at the beginning of a word and unaspirated after [s] ([top] vs [stop]), we can determine that the [s] in 'mistake' belongs to the second syllable (the [t] is not aspirated), and that the [s] in 'misstime' belongs to the first syllable (the [t] is aspirated):

'mistake' [mɪ-stēk] 'misstime' [mɪs-tʌˈaim]

3. INTERNAL STRUCTURE: The SYLLABLE is divided into an ONSET and a RHYME (or CORE). The RHYME of the syllable is divided into the PEAK (or NUCLEUS) and CODA.
CV syllables are universal (exist in all human languages).

Therefore, ALL languages contain SOME syllables that do not have a CODA.

A syllable without a coda is called an OPEN SYLLABLE.

OPEN SYLLABLES can be CV or V (or CCV, etc.)

CODA: In SOME languages ALL the syllables have NO coda. Hawaiian, for example, has ONLY open syllables.

Japanese is close to this ALL syllables are either OPEN, or have a nasal coda [n].

English allows for quite complex sequences:

CVCCCC --> [siksθs] 'sixths'

ONSET: Since the CV syllable is universal, ALL languages have SOME syllables WITH onsets.

English allows three consonants clusters (CCC) initially, but only certain very restricted combinations.

The first C of a 3 consonant cluster must be [s].
The second C must be [p] [t] or [k].
The third C must be a liquid [l] or [r].

Serbo-Croatian allows even more complex onset structure. It allows CCC:

First C must be fricative or [g].
Second C must be a stop, [m] or [v].
Third C must be a glide, a liquid, an [n] or a [v].
D. Syllable vs Mora

Where syllables are structural units and MORA are timing units.

A syllable may contain a single mora or more than one. The Japanese writing and accentual systems are based on the mora rather than the syllable.

WEAK syllables = one mora
STRONG syllables = > 1 mora

Strong syllables have:

\[ a) \text{ long } V : VV \\
   \text{ b) short } V + CCC \\
   \text{ c) short } V + \text{long } C \]

VII. Suprasegmentals

A. Phonetic characterization

SUPRASEGMENTALS are sound qualities which cannot (or are often not) identified with particular segments of sound. There are three (or more) sound qualities which can be identified as suprasegmentals.

1. Relative LENGTH: While length is a feature which can distinguish sets of segmental phonemes.

   Arabic: /a/, /i/, /u/ vs /a:/, /i:/, /u:/

   It can also be applied to sequences of phonemes, words, or sentences.

   /ri:li/ 'Really.' (It's so.)
   /ri:li:/ 'Really?!' (I don't believe it.)

2. Relative PITCH: The same sounds or words may be uttered at a higher or lower frequency.

   Low: Look out the window and tell me what you see.
   High: Look out!

3. Relative AMPLITUDE (Loudness): The same sounds may be uttered more or less loudly with respect to each other.

   permit (VERB) vs permit (NOUN)
4. Other qualities: There are other phonetic features which, although normally associated with particular segments, can be spread over sequences of sounds:

Nasalization: /pænɪk/ → /pænɪk/ 

B. Functions of Suprasegmentals

Suprasegmental features are special in that they may assume a variety of phonological (and non-phonological) functions.

1. They may function as PHONEMES:

* As we have seen, sets of consonants and vowels may be distinguished by LENGTH. Such is the case in Japanese and Arabic. While not distinguishing phonemes, LENGTH can differentiate across word boundaries in English:

/t/ [wai taj] 'why tie' /t:/ [wait taj] 'white tie'
[wai t a] 'white eye'

* In certain languages, relative PITCH is phonemic and distinguishes words from each other. Such languages are called TONE languages.

In Chinese, for example, each syllable has segmental phonemes plus a DISTINCTIVE TONE:

ma (55: high level) 'mother'
ma (35: mid rising) 'hemp'
ma (215: low dipping) 'horse'
ma (41: high falling) 'to curse'

Languages which are not TONAL generally have an ACCENTUAL or STRESS system, whereby different syllables in a word are assigned different prominence to each other. Such systems serve, among other things, to distinguish words in an utterance
from each other. In English, for example, a word cannot have more than one primary stress. Thus, the number of words in an utterance must be at least as many as the number of syllables in it bearing primary stress.

'an òre shipper' vs. 'a wòrshipper'

Stress applies to whole syllables rather than to segments. Languages which always have stress on the same syllable are called FIXED STRESS languages.

Czech: accent on first syllable (almost always).

French: accent on last syllable - Bloomingt6n, Indianá

In English and German, accent is not fixed on any particular syllable and are called VARIABLE ACCENT languages.

* STRESS ACCENT languages are those which use these suprasegmental qualities in a combination as a basis for ACCENT. Stress is indicated by an increase in the total amount of acoustic energy (length, amplitude, and pitch) rather than loudness alone.

* Languages which rely primarily on pitch to indicate accent are called PITCH ACCENT languages. Among these are Japanese and Norwegian.

Japanese: Hâna 'flower'
Hanâ 'nose'

3. These sound qualities may also assume a DISCOURSE function. Rather than distinguishing words from each other or identifying the word units of an utterance, they can add meaning to an entire utterance.

* When PITCH is used in this way it is called SENTENCE INTONATION. Rising intonation at the end of a sentence can distinguish a question from an assertion.

John's finished this book. vs. John's finished this book?

* AMPLITUDE is used for CONTRASTIVE STRESS, which provides information about a situation rather than any single element of the sentence.

John is outside. (a simple statement)
John is outside. (John, as opposed to anyone else)
VIII. Features

A. What is a feature?

A phonetic feature is any one of the gestures of which a single sound is composed.

The sound [b] is a "voiced bilabial stop". Rather than a single unit, its production involves several simultaneous gestures (inasmuch as the gestures are simultaneous it is a single segment).

* There is a momentary complete blockage of the airstream: [+interrupted]

* The velum is NOT lowered to allow the passage of air through the nasal cavity: [-nasal]

* The vocal chords are vibrated during the articulation: [+voiced]

* The closure is relatively long from front to back: [+distributed]

* The closure is in the front of the mouth: [+anterior]

* The closure is NOT made with the apex or front of the tongue: [-coronal]

A (partial) feature matrix for [b] would be:

```
+interrupted
-nasal
+voiced
+distributed
+anterior
+coronal
```

B. Reality of Features

If any one (or more) of the gestures is changed, the result is a distinct segmental sound.

<table>
<thead>
<tr>
<th></th>
<th>[b]</th>
<th>[β]</th>
<th>[m]</th>
<th>[v]</th>
<th>[g]</th>
<th>[p]</th>
<th>[d]</th>
</tr>
</thead>
<tbody>
<tr>
<td>+int</td>
<td>-INT</td>
<td>-INT</td>
<td>-INT</td>
<td>+int</td>
<td>+int</td>
<td>+int</td>
<td>+int</td>
</tr>
<tr>
<td>-nas</td>
<td>-nas</td>
<td>+NAS</td>
<td>-nas</td>
<td>-nas</td>
<td>-nas</td>
<td>-nas</td>
<td>-nas</td>
</tr>
<tr>
<td>+vcd</td>
<td>+vcd</td>
<td>+vcd</td>
<td>+vcd</td>
<td>+vcd</td>
<td>-VCD</td>
<td>+vcd</td>
<td>+vcd</td>
</tr>
<tr>
<td>+dst</td>
<td>+dst</td>
<td>+dst</td>
<td>-DST</td>
<td>-DST</td>
<td>+dst</td>
<td>+dst</td>
<td>+dst</td>
</tr>
<tr>
<td>+ant</td>
<td>+ant</td>
<td>+ant</td>
<td>+ant</td>
<td>-ANT</td>
<td>+ant</td>
<td>+ant</td>
<td>+ant</td>
</tr>
<tr>
<td>-cor</td>
<td>-cor</td>
<td>-cor</td>
<td>-cor</td>
<td>-cor</td>
<td>-cor</td>
<td>-cor</td>
<td>+COR</td>
</tr>
</tbody>
</table>
The fact that each of these gestures can independently distinguish sounds supports the notion that segmentals are in fact FEATURE MATRICES/BUNDLES (composites of several simultaneous WELL-DEFINED gestures).

D. Why a feature?

Economy of Representation.

* As we have seen, single features may distinguish between phonemes. Single features, moreover, usually serve to distinguish between more than one pair of phonemes in a phonemic system.

The feature [voiced] distinguishes SEVEN pairs of consonants in English:

\[
/p/:/b/ \quad /t/:/d/ \quad /k/:/g/ \\
/f/:/v/ \quad /θ/:/ð/ \quad /s/:/z/ \quad /ʃ/:/ʒ/
\]

The feature [tense] distinguishes FIVE pairs of English vowels:

\[
/i/:/i/ \quad /e/:/ε/ \quad /æ/:/æ/ \quad /u/:/u/ \quad /o/:/o/
\]

* Although it is highly uneconomical to replace a single symbol with a set of six to ten features, a relatively limited number of features can distinguish a wide range of sounds. We can distinguish the following TWENTY vowels using only FIVE features.

\[
\text{ae \quad a \quad ã \quad i \quad ñ \quad ë \quad ø \quad Æ \quad æ \quad ø \quad æ \quad ñ \quad ò \quad ë \quad ø \quad Æ \quad æ \quad ñ}
\]

D. Binarism

The five features in the table above could systematically distinguish up to 31 different types of sounds. Unfortunately, the phonetic system is not quite as binary as one would wish it to be:

* There is no class of vowels which are [+high] and [+low]. This rules out seven possibilities. \((31-7=24)\)
High = [+high] [-low]
Mid  = [-high] [-low]
Low  = [-high] [+low]
*    = [+high] [+low]

* [+low] [-back] [+round] vowels are ruled out for physiological reasons. This precludes an additional two vowel sounds.
(24-2=22)

* Otherwise features ARE binary. A sound is either [+nasal] or [-nasal], and NEVER [PARTLY nasal].

E. Redundancy

The fact that there are no [+high] [+low] sounds leads to the notion of REDUNDANCY.

If a sound is specified as [+low], then there is no need to further identify it as [-high]. It is, of course, [-high], but it is redundant to say so.

Similarly, [+interrupted] is by definition [-nasal], and vice versa.

Thus: [+high] --> [-low]  [+int] --> [-nasal]
      [+low] --> [-high]  [+nasal] --> [-int]

F. Classes of Features

1. Major Class Features

- [consonantal]
  - [sonorant]
    - [obstruents]
    + [liquids/nasals]
  + [syllabic]
    - [glides]
    + [vowels]
NOTE: [+consonantals] (obstruents, liquids, nasals) are
redundantly [-syllabic].

[-consonantals] (glides, vowels) are redundantly [+sonorant].

ALTERNATIVE Major Class Features:

<table>
<thead>
<tr>
<th>[-consonantal]</th>
<th>[+consonantal]</th>
</tr>
</thead>
<tbody>
<tr>
<td>[+vocalic]</td>
<td>vowels</td>
</tr>
<tr>
<td></td>
<td>liquids/nasals</td>
</tr>
<tr>
<td>[-vocalic]</td>
<td>glides</td>
</tr>
<tr>
<td></td>
<td>obstruents</td>
</tr>
</tbody>
</table>

2. Tongue Body Features: [low], [high], [back]
( and sometimes [front])

Primarily used to distinguish vowels:
[+back] distinguishes [u] from [i] (both are [+high])

Also used (less crucially) to distinguish consonants:
[k] is distinguished from [q] by [+high].

3. Other Vowel Features

[rounded]: [i] is [-round], [u] is [+round]

[tense]: [e] is [+tense], [ɛ] is [-tense]

4. Consonants: Manner Features

[interrupted]: [+interrupted] stops and affricates
[-interrupted] fricatives

[strident]: [+strident] affricates, some fricatives ([s])
[-strident] stops, some fricatives ([θ])

[nasal]: [+nasal] nasals ([m], [n], etc.)
[-nasal] liquids

[lateral]: [+lateral] l-sounds ([l], [l], etc)
[-lateral] r-sounds

[voiced]: [+voiced] [b], [v], [g], etc.
[-voiced] [p], [f], [k], etc.

[HSP]: [+HSP] aspirated stops and affricates
[-HSP] unaspirated sounds
5. Consonants: Place Features

[coronal]: [+coronal] sounds produced with the apex or front of the tongue - dentals, palatals, retroflex
[-coronal] sounds not produced in this fashion - labials, velars

[anterior]: [+anterior] sounds produced in the alveolar region or in front of it - labials, dentals, alveolars
[-anterior] sounds not produced in this area - velar, uvular, palatal

Secondary Place Features:

Tongue Body

[high]: [+high] velars - [k]
[-high] uvulars - [q]

[back]: [+back] retroflexed alveolars
[-back] unretroflexed alveolars

[low]: [+low] pharyngeals, glottals
[-low] uvulars

[distributed]: [+distributed] sounds for which the front to back closure is LONG - bilabials, velar fricatives
[-distributed] sounds for which the closure is short - labiodentals

6. Prosodic Features:

[long]

[stress]

(for detailed tables of features, see LAD Ch. 11. Figures 11.1, 11.2, 11.4, 11.5, 11.6, 11.7)

IX. Naturalness and Markedness

A. The choice of features reflects a linguist's assumptions about the system being described.

We are more likely to think of a lamp as being [+on] or [-on] than [+off] and [-off].
[-off] = [+on]
[+off] = [-on]

Either feature could be used to accurately describe the lamp's state of being switched on or off. However, off is in some sense the more natural state than on. So, rather than refer to the on state as being [-off], we would call the off state [-on]. This is because we feel that when the lamp is [+on], it possesses some extra characteristic (that of having electricity flowing through its wires).

Applying this analogy to phonetic characteristics, we can distinguish [b] and [p] as [+voiced] and [-voiced]; or as [+voiceless] and [-voiceless]. We feel that [b] has some extra quality that [p] lacks. We therefore call [b] a stop having the feature [+voiced], rather than [p] a stop having the feature [+voiceless]. [b] is, in some sense, more marked or more complex in comparison to [p]. [p] is more natural in comparison to [b].

What about other sounds and other features? Are things always this simple and clearcut?

Of course not!

If we think of vowels with respect to the feature [voiced], we will come to some very different conclusions. Voicing is a natural characteristic of vowels, and we would think of a [-voiced] vowel as being more marked, and a [+voiced] vowel as being more natural.

Thus, we cannot think of features as being marked in an absolute sense, but rather marked relative to other features.

[+voiced] is marked (less natural) for consonants.
[-voiced] is marked for vowels.

For other features, it is not as easy to determine markedness. Is [+coronal] or [-coronal] more natural for consonants? Is [b] more natural or more marked than [d]?

We cannot answer these questions by guessing. Rather, we can examine sounds systems of many languages for evidence.

CROSS-LINGUISTIC evidence: What sounds, combinations of sounds, and syllabic structures can occur in human language? Which are universal? Which are nearly universal?
UNIVERSALS are most natural (least marked).

NEAR UNIVERSALS are highly natural (relatively unmarked).

The rarer a sound or structure is across languages, the more marked it is likely to be.

INTRA-LINGUISTIC evidence: What is the inventory and distribution of sounds and syllabic structures in a particular language? What sounds are universally implied by the existence of what other sounds? What is the range of distribution, or combinatorial possibilities for any particular sound?

Sounds that IMPLY the existence of other sounds are more marked than the sounds that are implied by them.

Sounds and structures that have a wider distribution in a language are more natural (less marked) than sounds and structures whose distribution is highly restricted.

B. Markedness of syllabic structures.

CV is universal. Syllable = CV
CVC or V are more marked. Syllable [+coda] = CVC
VC implies V. Syllable [-onset] = V
Syllable [+coda, -onset] = VC

CCC onset  -->  CC onset
CCC coda  -->  CC coda
etc.

For Onset: C[+son] C[+int]  -->  C[+int] C[+son]
For Codos: C[+int] C[+son]  -->  C[+son] C[+int]

C. Markedness of segments  (More NATURAL type in upper case. More marked type in lower case.)

```
SEMENTS

[+CONSONANTAL]  [-CONSONANTAL]

[-SONORANT]  [+sonorant]  [+SONORANT]  [-sonorant]

obstruents  resonants  vowels  voiceless vowels

[p]  [r]  [a]  [ə]
```
OBSTRUENTS

[+INTERRUPTED]  [-interrupted]

[-STRIDENT]  [+strident]  fricatives

stops  affricate  [+STRIDENT]  [-strident]

[p]  [y]  [s]  [θ]

Affricates imply Fricatives imply Stops

[+CORONAL]

[+ANTERIOR]  [-anterior]

[t]  [retroflex t]

[+BACK]

[+HIGH]  [-high]

[k]  [q]

SONORANTS

VOWELS

[+LOW]  [-low]

[a]  [+HIGH]  [-high]

[-BACK]  [+back]

[u]

[i] < [e] < [o]
What is grammar (syntax)?

In order to answer this question, we must examine the mental organization of language:

The Mental Components of Language

---

**LEXICON**
(contains lexical morphemes)

- semantic representation
- phonological representation

---

**GRAMMAR**
(contains: rules syntactic morphemes)

---

Articulatory mechanisms

---

Distinction between Words and Morphemes (see LYO 5.3 and 5.4):

What is a morpheme? A minimal unit of meaning.

Some words contain only a single morpheme, such as 'bird' and 'construct', and there is a one-to-one correspondence between word and morpheme.

Other words can be analyzed into more than one meaningful unit.
Thus, 'construction' contains the units 'construct' and 'tion'; 'blackbird' contains 'black' and 'bird'; 'swimming' the units 'swim' and 'ing'. Here we see that words contain morphemes, but not vice versa.

The morpheme is a unit of lesser complexity than the word in the same way that the phoneme is of lesser complexity than the syllable. Even though a syllable might contain only one phoneme (like 'bird' contains only one morpheme), the syllable is on a higher order of structure.

Not all morphemes are created equal!

Some morphemes may stand alone as words (just as some phonemes (vowels) may stand alone as syllables. 'black' and 'bird', 'construct', and 'swim' are all words as well as morphemes. They are called FREE MORPHEMES.

Morphemes such as 'ing' and 'tion' do not stand alone as words and may only appear as AFFIXES (suffixes, prefixes and infixes). Such morpheme are called BOUND MORPHEMES.

Morphemes which are stored in the lexicon and express meaning are (what else?) LEXICAL MORPHEMES.

Morphemes which express structural relationships between words are SYNTACTIC MORPHEMES.

Thus, the lexical bound morpheme 'ness' means 'the quality of . . .' and when added as a suffix to an adjective such as 'great' yields a distinct noun meaning 'the quality of being great'.

The bound syntactic morpheme 's' when added to a noun such as 'lion' does not result in different word, but only in plurality.

The Structure of the Mental Lexicon

Morphemes stored in the mental lexicon consist of two representations: a PHONOLOGICAL REPRESENTATION and a SEMANTIC REPRESENTATION.

Thus, for the English speaker, the stimulus associated with visible light at certain lower wavelengths is brought together as the semantic representaion of the morpheme RED. The phonological representation of this morpheme consists of the three phonemes /r/ /ɛ/ and /d/ in that order.

The correspondence (or association) between semantic and
phonological representations is ARBITRARY (conventional). The speakers of another language (Spanish, for example) would have in their mental lexicon the same semantic representation corresponding to a different phonological representation /roxo/.

To demonstrate the independence of semantic and phonological representations, we note that:

Distinct semantic representations may have the identical phonological representation. The result is HOMOPHONY:

\[
\begin{align*}
\text{DOWN (prep)} & \quad \rightarrow /\text{daïn}/ \\
\text{DOWN (noun)} & \quad \rightarrow /\text{daïn}/
\end{align*}
\]

A single semantic representation can have distinct phonological representations. The result is SYNONOMY:

\[
\begin{align*}
/\text{pæɡl}/ & \quad \leftrightarrow /\text{bækɪt}/
\end{align*}
\]

Returning to the original question: What is GRAMMAR?

GRAMMAR takes elements (WORDS) out of the LEXICON and provides the structure to necessary to GENERATE (produce) SENTENCES. This is what is going on when you speak or write.

GRAMMAR also takes strings of words and INTERPRETS them. This is what happens when you listen to language or read.

The GRAMMAR of a language is thus essentially a set of rules for ENCODING and DECODING sentential units of language.

GRAMMAR utilizes four major tools for expressing structural relationships in sentences:

1. ORDER (linear precedence): Since language is linear (words are written or uttered sequentially), the ordering of words can be used to express structural relationships.

   In English, a noun is the subject of a verb if it precedes it, and the object of a verb if it follows it.

   'John kissed Mary.'

2. SYNTACTIC MORPHEMES: Certain morphemes are utilized to express the structural relationships between words. The suffix 's' when added to the present tense form of a verb indicates that the subject is
third person and singular:

John reads German.  *John read German.

3. CONSTITUENCY: The ELEMENTS of a sentence (words) may be grouped together in a variety of ways. The organization of smaller units into larger ones encodes the structure of a sentence.

In the sentence 'The boy swam', the elements 'the' and 'boy' form one CONSTITUENT and 'swam' forms the other:

[ [the boy] [swam] ] - RIGHT    [ [the] [boy swam] ] - WRONG

4. CATEGORIAL INFORMATION: The elements that form a sentence belong to CATEGORIES: NOUN, VERB, ADJECTIVE, PREPOSITION. Certain positions in a sentence are reserved for members of a particular category. Thus, one can replace a noun with another noun but not with a verb:

The boy swam. --> The ____ swam. --> The cat swam.
The cat swam. --> The ____ swam. --> *The go swam.

It should be noted that the first two types of information, ORDER and SYNTACTIC MORPHEMES, are used to express the last two, CONSTITUENCY and CATEGORIAL INFORMATION.

The Independence of Grammar and Lexicon (see LYO 4.1)

Although some linguists disagree, many consider the grammar and the lexicon to be autonomous (but related) components. This view is held to be supported by the fact that semantic meaning and syntactic meaning differ independently.

Two sentences can be semantically distinct, yet structurally identical:

Mike kissed Mary.
Mike invited Mary.

Sentences can also be structurally distinct, yet semantically identical:

Mike kissed Mary.
Mary was kissed by Mike.

How does one go about determining what the Grammar of a language is?

The 'black box' metaphor:
Suppose we wanted to discover the circuitry of a microcomputer without opening it. We could:

1) examine the input/output
2) make a hypothesis about its structure
3) test the input/output predictions made by this hypothesis
4) revise our hypothesis in accordance with observed input/output

Similarly, can:

1) look at language
2) make a hypothesis about its structure
3) test the predictions made by this hypothesis with speakers of the language
4) revise the hypothesis, if necessary

What can one/ does one seek to account for?

1. If we examine a CORPUS OF DATA (written or spoken), and we determine the rules that describe that set of sentences, then we will have a grammar of a SUBSET of all the possible sentences of that language.

   STRUCTURALIST linguists go only this far in trying to account for grammar, and do not consider asking speaker intuitions about sentence data a sound method of investigation. They explicitly reject ELICITATION as an investigative procedure.

2. We can take a theory of grammar, based on a corpus of sentence data, and check to see if our grammar is capable of GENERATING (producing) ALL and ONLY the sentences of the language. We can account for the grammar of a particular speaker of the language or we can account for the grammar of a language which is shared by all speakers of that language. This is called the CORE GRAMMAR.

   Linguists who try to account for language in this way are called GENERATIVIST linguists. They rely on speaker intuitions about NEGATIVE data as well as the POSITIVE data gathered from UNELICITED speech and texts.

3. Modern linguistic theory hypothesizes that there also exists something called UNIVERSAL GRAMMAR, which is that which the core grammars of all human languages have in common.

   Thus, although we may first have to deal with the DESCRIPTION of sentence data sets and the judgements of individual speakers, the real goals of modern theoretical linguistics is EXPLAINING the
Grammatical acceptability (see LYO 4.2)

In attempting to determine the rules of a grammar, we rely on the intuitions of native speakers of the language regarding the grammatical acceptability of sentences. What do we mean by grammatical acceptability? What other kinds of acceptability are there?

SEMANTIC unacceptability: sentences which are structurally sound, but either contain words which are not part of the language in question or use words in unacceptable ways (see MAT ch.1: COLLOCATIONS):

Quadruplicity drinks procrastination.

'Twas brillig and the slythy toves did gyre and gimble in the wabe.

SOCIAL unacceptability: sentences which, although they might be well-formed structures and make sense, violate the social conventions of the language:

My, you look wretched today.

CONTEXTUAL unacceptability: sentences which are structurally sound, make sense and are acceptable in some context, but not within the text where they appear:

"Hello," said the dandelion. (in a news article, as opposed to a children's book).

In analyzing data for grammatical description, we are only concerned with grammatical acceptability.

We will use a '*' to indicate grammatically unacceptable sentences, '?' to indicate that a sentence is questionable, and '%' to indicate that it is acceptable to some speakers.

John threw the ball to Mary.
John threw Mary the ball.
?John threw to Mary the ball.
*John threw ball the Mary to.
**The to threw John Mary ball.

The SEMANTIC unacceptability we saw above (wherein English words were used in unacceptable ways) involves SELECTION RESTRICTION violations (selectional restrictions are what Matthews would refer to as COLLOCATIONS). Somewhere in our mental semantic representation of a
verb, besides the semantic representation of some activity, we have knowledge of the acceptable ACTORS and UNDERGOERS for that activity. Thus,

Destruction drinks productivity.

is unacceptable because 'destruction' is not a noun that is capable of 'drinking' and 'productivity' is not a noun which is capable of being 'drunk'. For the SUBJECT of this verb we require something which is [+animate] (recall phonological features?), and the OBJECT of this verb need be some liquid. While a syntactian must be sensitive to the selection characteristics of verbs, and while selectional restrictions are often coexistent with other grammatical phenomena, the unacceptability detailed here is not directly due to the grammar of a language and does not tell us about it.

Categorial Information: Notional vs. Distributional Labels (LYO 4.2.8)

Notional categorial labels correspond to our traditional grammar book ideas (notions) about what kinds of words normally belong to what categories.

Noun: A noun is a person, place or thing.

However, the notional labels can be misleading. First of all, they are not always correct:

'destruction' is not a person, place, or thing, but an activity.

Regardless of what 'destruction' is, it is a NOUN. Moreover, not every verb represents an activity (or state).

What about the VERB 'to be'?

Although we will keep the traditional labels assigned to parts of speech, we will assign words to these categories based on strictly distributional criteria.

Suppose, we were from another Galaxy and were attempting to find out the structure of a language that we did not understand: English. If we took a set of sentences and substituted letters for words, we might get the following sentences:

\[
\begin{align*}
\text{a} & \quad \text{g} & \quad \text{b} & \quad \text{a} & \quad \text{h} & \quad \text{b} & \quad \text{a} & \quad \text{k} \\
\text{b} & \quad \text{g} & \quad \text{a} & \quad \text{b} & \quad \text{h} & \quad \text{a} & \quad \text{b} & \quad \text{j} & \quad \text{b} & \quad \text{k}
\end{align*}
\]

Just from the distribution, we know that A and B can replace each
other in the same context, as can G and H, and J and K.

Let's say, A and B belong to category X, G and H to Y, and J and K to Z

\{a,b\} = X \quad \{g,h\} = Y \quad \{j,k\} = Z

Essentially, then there are two basic sentence types in this data set,

\[X \quad Y \quad X. \text{ and } X \quad Z.\]

If we substitute English words for these letters:

\(a = \text{cats, } b = \text{rats, } g = \text{hate, } h = \text{eat, } j = \text{swim, } k = \text{sink}\)

Then we have:

Cats hate rats. \quad \text{Cats eat rats.} \quad \text{Cats swim.} \quad \text{Cats sink.}

Rats hate cats. \quad \text{Rats eat cats.} \quad \text{Rats swim.} \quad \text{Rats sink.}

\(X = N; \quad Y = V2(\text{trans}); \quad Z = V1(\text{intr})\)

\(\{\text{cats, rats}\} = N; \quad \{\text{hate, eat}\} = V2; \quad \{\text{swim, sink}\} = V1\)

The basic sentence types in this example are, of course:

\[N \quad V2 \quad N. \quad \text{and } N \quad V1.\]

How can we formally represent these grammatical relationships?

The structures we have been talking about are the structure of sentences and phrases. We will represent these relationships using PHRASE STRUCTURE TREES or MARKERS (P-MARKERS), BRACKETING, and PHRASE STRUCTURE RULES (PS RULES).

Linear precedence and constituency can be represented by BRANCHING TREES or BRACKETING:

Taking the sentence: "The cat chased the rat."
We can divide it first into two major constituents (SUBJECT and PREDICATE): "the cat" "chased the rat"

```
[the cat] [chased the rat]
```
The PREDICATE (or VERB PHRASE) "chased the rat" is divisible into its VERB and OBJECT COMPLEMENT: "chased" "the rat"

```
\[
\text{the cat} \quad \text{chased} \quad \text{the rat} \quad \text{[[the cat] [[chased] [the rat]]]}
\]
```

Each of the NOUN PHRASES "the cat" and "the rat" is divisible into a DETERMINER "the" and NOUN: "the" "cat" "the" "rat"

```
\[
\text{the cat} \quad \text{chased} \quad \text{the rat} \quad \text{[[[the] [cat]] [[chased] [[the] [rat]]]]}
\]
```

The bracketing and tree notation are equivalent, and can be used interchangeably, depending upon whether one needs to clearly represent a complex structure (use trees) or needs to save space (use brackets).

Is this formal notation complete as it stands? NO

In addition to constituency and linear order, we must include categorial information in our notation. This is needed because, although we might substitute words of the same category in the sentence and still maintain the same structural relationships between them,

'a' for 'the' (DETERMINER)
'hat' for 'rat' (NOUN)
'ate' for 'chased' (VERB)

would give us: "A cat ate the hat."

We may not substitute words of distinct categories and have the same structure:
'hat' (NOUN) for 'chased' (VERB)

To express this information, we will label the BRANCHING NODES of the tree, or the brackets, with a label:

S = SENTENCE
N = NOUN
A = ADJECTIVE
V = VERB
P = PREPOSITION
DET = DETERMINER
NP = NOUN PHRASE
AP = ADJECTIVAL PHRASE
VP = VERB PHRASE
PP = PREPOSITIONAL PHRASE

[ s [ np [ det the ] [ n cat ] ] [ vp [ v chased ] [ np [ det the ] [ n rat ] ] ] ]

In the notation used here, what is meant by the term NOUN PHRASE?

More generally what is a PHRASE?

We might take a phrase to be any group of ELEMENTS (words) that form a CONSTITUENT. By this definition we might expect to find PHRASES within other PHRASES, since a constituent might itself be composed of other constituents, rather than only of elements.

Thus, if a group of words taken together form a constituent which is the SUBJECT of a sentence, then it is a phrase. In this case it will be a NOUN PHRASE because all the elements in the phrase are subordinate to, and modify, a NOUN which is taken to be the nucleus of the phrase.

In the phrase:

The big brown DOG in the yard . . .

the noun DOG is what the phrase is about, and all the other elements of the phrase serve to modify it.

To be an NP is to FUNCTION as an NP. Thus, there is no requirement on the contents of an NP, other than that it contain a NOUN.
Since "John" can function as the subject of "was eating":

John was eating.

just as well as "The big brown dog in the yard":

The big brown dog in the yard was eating.

We might say that in this case "John" is an NP whose only INTERNAL CONSTITUENT is the NOUN "John".

\[
\begin{array}{c}
\text{NP} \quad \text{(maximal projection)} \\
\mid \\
\text{N} \quad \text{(head)} \\
\mid \\
\text{John}
\end{array}
\]

The NOUN (N) which is the nucleus of a NOUN PHRASE (NP) is called the HEAD of the PHRASE and we say that the NOUN PHRASE is a MAXIMAL PROJECTION of the NOUN. It is a PROJECTION in that it is a structure PROJECTED (or thrown up) by the NOUN. It is MAXIMAL in that it is the largest grouping of elements directly associated with that NOUN.

The relationship between N and NP can be extended to all categories, such that we can generalize the relationship from any category X to its phrasal marker XP.

Does the phrase have any other internal structure besides the presence of the phrasal head?

Some of the material in a phrase might be thought to be more closely associated with the head than other material:

The VP: discovered America in 1492
consists of: the V 'discover'
the NP 'America'
and the PP 'in 1492'

\[
[\text{vp discovered } [\text{np America } ] [\text{pp in 1492 } ] ]
\]

The NP 'America'is required by the V 'discover' and as such is an ARGUMENT of the verb. We cannot say:

*Columbus discovered in 1492.

The PP 'in 1492' is optional and is called an ADJUNCT.
Columbus discovered America.

Word order facts also indicate that the ARGUMENT 'America' is more closely bound to the verb than the ADJUNCT 'in 1492'. We cannot insert the optional material between the verb and its argument:

*Columbus discovered in 1492 America.

We might thus say that there is an intermediate level of structure between V and VP. We will call this V' (V-prime/also V-bar):

\[
\begin{array}{c}
\text{VP} \\
\text{V'} \\
\text{V} \quad \text{NP} \\
\uparrow \quad \uparrow \\
\text{discovered} \quad \text{America} \\
\text{PP} \\
\text{P} \quad \text{NP} \\
\uparrow \quad \uparrow \\
\text{in} \quad \text{1492}
\end{array}
\]

NPs also exhibit an intermediate level of structure. We can see this if we compare NPs and sentences:

\[
\begin{array}{c}
\text{S} \\
\text{NP} \\
\text{V} \\
\uparrow \\
\text{Caesar destroyed Rome.} \\
\text{NP} \\
\text{N'} \quad \text{PP} \\
\uparrow \quad \uparrow \\
\text{Caesar's destruction of Rome}
\end{array}
\]

The PP 'Of Rome' is the complement of the noun 'destruction' in the same way that the NP 'Rome' is the complement of the verb 'destroy'. Just as VP forms an intermediate level of structure with which the SUBJECT combines to form an S, so the noun and its complement form an intermediate level of structure with which the SPECIFIER 'Caesar' combines to form an NP.

We can now generalize over XPs and say that phrases may have the following types of structure:
How do grammars GENERATE the structures that we have represented here with P-Markers and bracketing?

Phrase Structure Rules (PS-Rules)

These rules are also referred to as rewrite rules.

We can express the fact that a unit of a particular type is composed of certain other units in a certain order, by means of the following type of statement:

\[ S \rightarrow NP + VP \]

This rule states that a sentence (S) rewrites as (breaks down into) a noun phrase (NP) and a verb phrase (VP).

In English, the following PS-rule expresses the relationship between a transitive verb and its object:

\[ VP \rightarrow V + NP \]

The verb phrase REWRITES as a verb plus a noun phrase.

eats the cake \rightarrow eats + the cake

English also has the following PS-rules:

\[ VP \rightarrow V \]
\[ VP \rightarrow V + NP + PP \]
\[ NP \rightarrow Det + N \]

Notice that a PS tree (or marker) is produced by a number of PS-rules:

The cat chased the rat.

\[ S \rightarrow NP + VP \]
\[ VP \rightarrow V + NP \]
\[ NP \rightarrow Det + N \]

The phrase markers of all the sentences in English contain at least the following rule:
S \rightarrow VP + NP

EXERCISE

Draw P-markers for each of the sentences, and a set of PS-rules which will generate (at least) all of them.

Are there any ambiguous sentences? Do they have alternate P-markers?

1. The large brown canary are my calico cat.

2. Germaine put the diamonds in the safe.

3. John drowned in a vat of jello.

4. Judy has punched the fellow in the back row.

5. The slythy toves gyred and gimbled in the wabe.

Sentences [see MAT Ch.2; LYO 5.2]

In order to talk about a SENTENCE GRAMMAR, we must determine what we mean by the term SENTENCE.

Bloomfield (famous linguist) defines a sentence as "the largest unit of grammatical description". By this, he meant to say that INTRAsentential relationships (relationships within a sentence between its constituents) are the domain of GRAMMAR. INTERsentential relationships (relationships between sentences) are the domain of rules of DISCOURSE, and not part of the grammar.

SENTENCES are required to be internally GRAMMATICAL.
DISCOURSE is required to be COHERENT.

This still does not tell us how to identify the sentence in its natural habitat (speech/text).

The traditional NOTIONAL definition of sentence, like the notional definition of noun, does not help much:

"A sentence expresses a complete thought"

This definition simply takes the burden of providing a formal definition of SENTENCE away from the linguist, and hands the problem over to the psychologist. What is a COMPLETE THOUGHT, anyway? Do you break up your complete thoughts with little pauses, as you might
break up your sentences?

The problem of providing a formal definition of a "complete thought" is probably ALOT more difficult than defining a complete sentence, so we'll stick with the latter.

Sentence and Utterance

If we want to speak about SENTENCE as a FORMAL object, then we need a neutral term for stretches of speech (phonological continuums) that are separated by pauses.

Zelig Harris (another famous linguist) defined UTTERANCE as "any stretch of talk, by one person, before and after which there is silence on the part of that person."

Thus defined, an utterance could consist of a word, a sentence fragment, a complete sentence, or (possibly) more than one sentence.

The term SENTENCE refers to a structural unit which must be formally defined.

If the utterances: "Go away." and "John read the book." are single, complete sentences, why are the utterances: "Goes away." and "John put the book." incomplete?

If the utterance: "Go away before I get mad." is one sentence, why is: "Go away. I'm angry." two?

Formally defining SENTENCE

If we didn't know exactly what a sentence is. We could determine which utterances contained precisely ONE sentence if we could:

1) identify utterances which contain MORE than 1 sentence, and
2) identify utterances which contain LESS than 1 sentence.

Thus, if we had a set of utterances:

\{a, b, c, d, e, f, g, h, i, j, k\}

and we could show that \{b, c, h, k\} contained more than 1 sentence, and that \{d, f\} contained less than 1 sentence,

then would could determine that utterances \{a, e, g, i, j\} were sentences.
We find three formal principles which we can use to determine whether an utterance contains more than one sentence.

1. **INCLUSION** - The principle here is that if a unit contains more than one sentence, then it cannot be included within another sentence.

   Words, phrases, sentence fragments, and sentences can all easily be included within larger sentences:

   - goes ---\> Tom GOES there every day.
   - in the kitchen ---\> He's standing IN THE KITCHEN.
   - watching TV ---\> Bill spent the day WATCHING TV.
   - he's leaving town ---\> I wonder whether HE'S LEAVING TOWN.

   Utterances containing more than one sentence cannot:

   - go away I'm busy ---\> *John wants to GO AWAY I'M BUSY.

2. **RULE BASED/RESTRICTED RELATIONSHIPS** - If an utterance contains more than one sentence then there will not (in general) be regular rules that restrict the distribution of one part on the basis of the other.

   In other words, if an utterance contains two sentences A and B, then we can in principle replace sentence B with ANY other sentence C, D, etc. However, if A and B are parts of the SAME sentence (or fragment) then we expect that A will impose restrictions on what material can replace B.

   In the utterance: "I'm not going to the park, because it's raining." the second part of the utterance is restricted to be a REASON for the first part. We could thus replace "because it's raining" with "because I'm tired" but NOT with "because are you going". The RESTRICTED RELATIONSHIP holding between the first and
second parts suggest that they are parts of a single sentence.

In the utterance: "I'm not going to the park. It's raining." there are almost NO restrictions on the replacement of either the first or the second parts of the utterance.

"I'm not going to the park. Are you going?" would be perfectly acceptable.

The only necessary relationship that must always hold between units an utterance that are independent sentences is that of COHERENCE.

"I'm not going to the park. Chile has an oppressive government."

This utterance consists of two GRAMMATICAL sentences, but seems INCOHERENT. However, it COULD be coherent in the appropriate CONTEXT. Suppose for instance that the Chilean ambassador was making a speech in the park, the second sentence could explain my reasons for not going. COHERENCE, unlike GRAMMATICALITY, is dependent upon CONTEXT.

3. INTONATION - This principle is that sentences generally carry a characteristic intonational pattern. Using this evidence we can determine if an utterance consists of more than one sentence.

Declarative sentences in English usually have a fall in pitch to indicate the end of a sentence. The utterance "Go away. I'm busy." contains TWO sentence final intonations, whereas the utterance "Go away because I'm busy." only has ONE sentence final fall in pitch.

Obviously, this test is only applicable to spoken language, and is of no help in determining the status of written data.

Now that we have determined which utterances contain more than one sentence, how can we determine which utterances are SENTENCE FRAGMENTS?

There are many utterances which seem to be structurally incomplete, yet we would want to count them as complete sentences. On the other side of the same coin, there are many utterances which are complete in a given CONTEXT which we will call INCOMPLETE SENTENCES.

Before approaching the central issue, we note that Matthews distinguishes between INCOMPLETE SENTENCES and INCOMPLETE UTTERANCES.

According to Matthews, an incomplete utterance is a sentence that, for some reason or other has been interrupted.

You're looking for your umbrella and start to ask,
"Has anyone seen my . . . "

and then you spot it in a corner and break off in mid-sentence.

Incomplete utterances can be easily identified, because they lack SENTENCE FINAL INTONATION.

A SENTENCE is a MENTAL CONSTRUCT which is the OUTPUT of the GRAMMAR, while an UTTERANCE is the PHYSICAL ARTICULATION of that CONSTRUCT. Thus, in the case of an incomplete utterance, the grammar may have generated a complete sentence which never got realized:

You had mentally constructed the complete sentence:

"Has anyone seen my folding umbrella?"

You just didn't get to finish saying it.

Because of this we must exclude incomplete UTTERANCES from our discussion of incomplete sentences.

In order to discuss incomplete sentences we must first understand the notion of ANAPHORA.

ANAPHORA is a relationship of IDENTITY holding between an element and something that precedes it. The ANAPHR is the name given to the thing that precedes.

There are two kinds of elements that can have anaphors: GAPS (or ELISIONS) and VARIABLES.

A GAP is simply a place from which material is missing. Such as the SUBJECT in:

"__ Did it yesterday."

or the VERB PHRASE in:

"John went to Hawaii for Thanksgiving, but I didn't ___"

A VARIABLE is, just as in algebra, something that stands for something else and can be set to different values. In mathematics, we use Xs and Ys. In language, we have PRONOUNS.

"I hope HE'll be here on time."

A VARIABLE generally MUST have an ANAPHOR. Most but not all GAPS
require an ANAPHOR.

In the sentence: John is tired, isn't he? 'John' is the anaphor for the pronoun 'he'.

Here, the anaphor for the VARIABLE 'he' is in the same sentence. This is called INTRA-sentential anaphorra. Note that 'John' is an anaphor because it stands in an ANAPHORIC relation (is identified with) 'he'. It is not INHERENTLY an anaphor. This is similar to the situation in "John loves Mary." 'John' is Mary's LOVER (stands in the LOVE relation to Mary), but 'John' is not inherently a 'lover'.

In the exchange (DISCOURSE): 
"Where's John?"
"He's in the kitchen."

the anaphor of the variable 'he' occurs in the previous sentence. This is an example of INTER-sentential (or DISCOURSE) anaphora.

A variable MUST have an anaphor; either in the same sentence or in discourse. You can't just walk into a room where noone knows you and announce,

"He's on his way here."

unless you want people to assume that you're a preacher.

GAPs too can have anaphors within a sentence.

I [vp wanted to go home early ], but Julie didn't [vp ].

The anaphor for the gap in this sentence is the entire VP [vp wanted to go home early ]

They can also be identified with material in previous discourse:

"Where'd John go?"
"_____ Into the kitchen."

In this exchange, the missing material 'John went' is supplied by the DISCOURSE CONTEXT.

However, not all gaps have anaphors:

"_____ Got a match?"

In this sentence the missing information is NOT dependent on CONTEXT.
I cannot ask this question, meaning "has HE got a match?" It can only mean "Have YOU got a match?" We could therefore say that the sentence with the gap can be related BY RULE to the complete version.

We can now discuss the difference between CONTEXTUAL COMPLETENESS and GRAMMATICAL COMPLETENESS.

If an utterance is understandable in a given context, then it is CONTEXTUALLY COMPLETE. The same utterance could vary dependent on the discourse context:

In the context: [ "Where's John?"

"In the kitchen." is contextually complete.

In the context: [ "What's new?"

it is NOT.

The same would apply for the utterance:

"He's in the kitchen."

Utterances with variables or gaps MAY be CONTEXTUALLY COMPLETE, and this completeness is dependent on discourse.

Utterances with variables such as "He's in the kitchen." are STRUCTURALLY (or GRAMMATICALLY) COMPLETE, independently of CONTEXT. Likewise "In the kitchen." is structurally INCOMPLETE, regardless of whether or not it is contextually complete.

So utterances with variables are GRAMMATICALLY COMPLETE and may or may not be CONTEXTUALLY COMPLETE.

Are all utterances with GAPS grammatically INCOMPLETE? Well . . no.

If a gap can be filled by grammatical RULE, independently of CONTEXT, then we will say that it is GRAMMATICALLY COMPLETE. There is a REGULAR (RULE BASED) correspondence between:

"Got a match?"  <---->  "Have you got a match?"

Therefore, "Got a match?" (and "Go away!", etc) is GRAMMATICALLY COMPLETE.
On the other hand, "In the kitchen." cannot be related BY RULE to any particular structurally complete sentence. Dependent on CONTEXT, it could mean:

"JOHN'S in the kitchen."
"I LEFT MY NOTEBOOK in the kitchen."
"HE PARKED HIS MOTORBIKE in the kitchen." or anything else.

It is GRAMMATICALLY INCOMPLETE, even if it is CONTEXTUALLY COMPLETE.

We can therefore say that a GRAMMATICALLY INCOMPLETE SENTENCE is one which contains a GAP that CANNOT be filled by RULE.

<table>
<thead>
<tr>
<th>CONTEXTUALLY COMPLETE</th>
<th>CONTEXTUALLY INCOMPLETE</th>
</tr>
</thead>
<tbody>
<tr>
<td>GRAMMATICALLY COMPLETE</td>
<td>&quot;Got a match?&quot;</td>
</tr>
<tr>
<td></td>
<td>&quot;It was fun.&quot;</td>
</tr>
<tr>
<td></td>
<td>&quot;Did you see Bill?&quot;</td>
</tr>
<tr>
<td></td>
<td>&quot;What's new?&quot;</td>
</tr>
<tr>
<td></td>
<td>&quot;He's in the garage.&quot;</td>
</tr>
<tr>
<td></td>
<td>&quot;He's in the garage.&quot;</td>
</tr>
<tr>
<td>GRAMMATICALLY INCOMPLETE</td>
<td>&quot;Where'd Sue go?&quot;</td>
</tr>
<tr>
<td></td>
<td>&quot;What's new?&quot;</td>
</tr>
<tr>
<td></td>
<td>&quot;To the library.&quot;</td>
</tr>
<tr>
<td></td>
<td>&quot;To the library.&quot;</td>
</tr>
</tbody>
</table>

Morphemes and words [see MAT Ch.3; LYO 5.3,5.4]

Having discussed the identity of the SENTENCE, which is the OUTPUT of a GENERATIVE SENTENCE GRAMMAR, we can now turn to the INPUT to this grammar:

WORDS

It is obvious that WORDS, which are the constituent ELEMENTS of the SENTENCE, have INTERNAL STRUCTURE, much as the sentence itself has internal structure. Since we have identified the MORPHEME as the MINIMAL UNIT of MEANING, and since WORDS are composed of MORPHEMES, we will speak of the MORPHEMIC structure of WORDS.

Morpheme and Morph

We have previously made distinctions between semantic and phonological representations in the lexicon. Two distinct semantic representations might have the same phonological representation: resulting in HOMOPHONY.

DOWN [noun] $\leftrightarrow$ /dəʊn/ 
DOWN [adj] $\leftrightarrow$ /dəʊn/
One semantic representation might have two distinct phonological representations: resulting in SYNONYMY

'CYLINDRICAL WOODEN WATER-CARRIER WITH HANDLE' $\leftrightarrow$ /pəˈɡɔːl/ $\leftrightarrow$ /ˈbækɪt/

So far we have observed a one-to-one correspondence (or mapping) between MORPHEMES as units of meaning and their phonological realizations. We have observed words that are SEGMENTABLE into their component morphemes. Such as:

deconstruction $\rightarrow$ de + construct + tion

However, this is not always the case. While

bigger

can be divided into the adjective 'big' and the comparative morpheme 'er', we cannot do the same for

worse

This is problematic. Although we do not want to propose segmentizing the word 'worse', we DO want to maintain that 'worse' stands in the same relation to 'bad' as 'bigger' does to 'big':

\[
\begin{array}{ccc}
[+\text{comparative}] & [+]\text{superlative}] \\
\text{big} & \text{bigger} & \text{biggest} \\
\text{bad} & \text{worse} & \text{worst}
\end{array}
\]

We might approach this problem by distinguishing between the mental representation of a morpheme and its phonological realization.

We will continue to call this MENTAL REPRESENTATION a MORPHEME. We will refer to the PHONOLOGICAL REPRESENTATION by the term MORPH.

Thus the [comparative] morpheme is generally, but not always realized as the morph: 'er'

The word 'bigger' is segmentable into two morphs 'big' and 'er', each of which represent a morpheme.
BIG [comparative]
||
|'big' |'er' |=> |'bigger' |

On the other hand, the word 'worse' consists of a single morph which represents two morphemes:

BAD [comparative]
\+/
|\'worse' |

Another distinction we need to make is between OPEN and CLOSED classes of morphemes:

CLOSED CLASSES are those which can have only a limited number of members.
OPEN CLASSES have theoretically no upper bound on the number of morphemes they might contain.

An example of a closed class would be the morphemes that belong to the category CONJUNCTION:

{ and, but, or, if...then, ... }

This category is closed insofar as there is a finite set of POSSIBLE LOGICAL RELATIONSHIPS that may hold between conjoined clauses.

The category NOUN, on the other hand, is open. There is no limit on the number of distinctions that we might make in referring to the physical (or meta-physical) world.

The Inuit ('Eskimo') peoples of the Arctic have dozens of words for what we generally refer to in English as 'snow'. Furthermore, when someone invents or discovers something new, they generally make up a new name (noun) for it.

Some categories (CONJUNCTION, DETERMINER, and PREPOSITION) are closed, and some (NOUN, VERB, and ADJECTIVE) are open. However, the quality of being an OPEN or CLOSED class is not restricted to categories.

A subset of the category NOUN, the PRONOUN, is a CLOSED class. Languages generally use two or three features in their pronominal system and there are a limited number of possible slots for pronouns in any system.

Most languages utilize features of [person], [number] and [gender] to
form pronouns. While other features such as [animate] or [abstract] might also be used, we would not expect to find a pronominal system that incorporated features such as [metallic] or [gaseous].

Given a limited set of features, say: [person], [number], [gender] and a limited range of possible values for these features:

- [person]: { [1st], [2nd], [3rd] }
- [gender]: { [FEM], [MAS], [NEU] }
- [number]: { [SNG], [PLU] }

there will necessarily be an upper bound on the possible number of pronouns. In this case:

\[
\{[1st],[2nd],[3rd]\} \times \{[FEM],[MAS],[NEU]\} \times \{[SNG],[PLU]\}
\]

\[
3 \times 3 \times 2 = 18
\]

Although it is uncommon for a language to have a pronoun for every possibility in the system, the existence of this limit makes it a closed class.

English, for example, has seven out of these eighteen possible pronouns:

- [SNG]
  - [1st] I ([MAS] and [FEM]) we ([MAS] and [FEM])
  - [2nd] you ([SNG] and [PLU]/[MAS] and [FEM])
  - [3rd] he [MAS] they ([MAS],[FEM],[NEU])
    - she [FEM]
    - it [NEU]

We have considered the mental representation of morphemes versus their phonological realizations. We have distinguished open and closed classes of morphemes. We will now consider the MORPHEMIC STRUCTURE of WORDS.

FREE morphemes are those that can occur independently in a language as words. For instance, in English:

- boot, toy, swim, make, under

Free morphemes do not NECESSARILY occur as isolated words, but they CAN.
'swimming' contains the FREE morpheme 'swim' and the BOUND morpheme 'ing'.

The status of being a free morpheme varies from language to language. A semantic representation that is a free morpheme in one language might be a bound morpheme in another:

In English, the verb 'make' forms causative sentences:

'John made me swim.'

In Japanese, the inflection -sase forms causatives:

Zyon wa watasi o oyogASEta.
John I swim-MAKE-past
'John made me swim.'

Bound morphemes, insofar as they must be attached to other morphemes, are AFFIXES (SUFFIXES, PREFIXES, or sometimes INFIXES).

When two (or more) free morphemes combine to form a word, the process is called: COMPOUNDING

'black' + 'bird' ===> 'blackbird'

What about AFFIXES? Are ALL bound morphemes created equal?

Are the 'un' of 'undo'
the 'ing' of 'doing' and
the 'able' of 'doable' the same?

NO.

We will distinguish between DERIVATIONAL and INFLECTIONAL MORPHOLOGY.

Essentially, a DERIVATIONAL morpheme when added to a morpheme forms a composite that is different in meaning or category from what it is added to. It does NOT mark the structure of the phrase or sentence to which it is added.

INFLECTIONAL morphemes are used to signal the structure of a phrase or sentence.

One type of DERIVATIONAL morpheme is that which changes the category of the unit to which it is attached. '-tion', for example, attaches to verbs and changes them into nouns. '-able' also attaches to verbs, but changes them into adjectives.
Another type of derivational morpheme does NOT change the category of the base form, but rather changes its meaning.

\[
\text{re + construct} \rightarrow \text{reconstruct}
\]

Here the prefix 're-' added to a verb 'V' changes the meaning of the verb from 'V' to 'do V again'.

Notice that, at least for English, the derivational affixes that alter category are SUFFIXES and those that don't are PREFIXES.

An INFLECTIONAL morpheme neither changes the category membership of the base form, nor does it alter the meaning of the word. Rather, it signals structure of the sentence into which the word is put.

The 'third-person singular present' inflection '-S' for example does not change the meaning of the verb 'eat' in:

'He eats lunch in the park every day.'

Nor does it change the category of 'eat' from VERB to anything else. It simply marks the fact that the subject of the verb is 3rd person singular and that the sentence is in the present tense.

The plural suffix '-S' does not change the meaning or the category of a noun that it attaches to either. It simply indicates that there is more than one of such a thing being referred to.

'boy' + '-s' = 'boys'

Derivational morphemes and inflectional morphemes differ in some significant ways:

Derivational morphemes can CHANGE CATEGORY membership, while inflectional morphemes cannot.

Derivational morphemes are generally more restricted in what they can attach to than are inflectional morphemes.

Although '-tion' attaches to verbs to form nouns, it cannot attach to ALL verbs:

'destruction', 'permission' but not *'attachtion', *'governtion'
The inflectional suffix '-ing' on the other hand attaches to almost all verbs (are more PRODUCTIVE):

'destroying', 'permitting', 'attaching', 'governing'

Inflectional morphemes, since they express syntactic structure, often CO-VARY (vary together with) other elements in a sentence.

HE eats lunch here at noon. ==> THEY eat lunch here at noon.
TWO boys sat down. ==> ONE boy sat down.

The inflection on the verb 'eat' varies together with the subject pronoun; the plural inflection on the noun 'boy' varies in accordance with the number. This phenomenon of CO-VARIANCE is referred to as AGREEMENT.

Derivational morphemes never depend on the agreement of another element.

Inflectional morphology always occurs 'OUTSIDE' of derivational morphology. If one thinks of a DERIVED and INFLECTED word as something that is built up one morpheme at a time, one cannot add an inflectional morpheme until all the derivational morphemes are affixed.

'-s' [PLU] is inflectional; '-ism' [N] is derivational

We can add '-s' to 'ideal' ===> 'ideals' [[ideal] s]
We can add '-ism' to 'ideal' ===> 'idealism' [[ideal] ism]
and THEN add '-s' to 'idealism' ===> 'idealisms' [[[ideal] ism] s]

BUT we cannot add '-ism' to 'ideals' ===> *

In this example the inflectional morpheme is INSIDE the derivational one.

Based on this, we might say that the OUTPUT of derivational morphology is INPUT to inflectional morphology.

How can we represent the internal structure of derived words?

If, by analogy to sentences which take words as their elements, we say that derived words take morphemes as their elements, then we can propose WORD-STRUCTURE MARKERS similar to the PHRASE MARKERS we used
earlier.

In the case of category changing morphemes, we could say that the affix bears the CATEGORIAL FEATURE, and that the derived word is a PROJECTION of the affix.

Thus, '-tion' is a NOUN that attaches to VERBS to form NOUNS:

[n [v construct] tion]

'-ism' is a NOUN that attaches to NOUNS and ADJECTIVES to form NOUNS:

[n [n ideal] ism]
[n [a formal] ism]

We can say that affixes which do not change category do not have a category feature and take their category from the base form:

[in [v capacitate]] ==> [v incapacitate]
[in [a formal]] ==> [a informal]