LING201   Winter '23

Louis Meunier

Contents

1 Introductions/Overview

1.1 Descriptivism vs Prescriptivism

Linguists aim to study how language(s) are *actually* used, rather than how they *should* be used. The difference between these two "forms" of study are as follows:

- **Descriptivism**: describing how language is actually used

- **Prescriptivism**: describing how language *should* be used (think of grammar rules, standardizing spelling, etc)

In studying language this way, linguists do not need to fluently speak the language they are studying, and instead can rely purely on the patterns and regularity (ie, implicit grammar) of said language to make their observations.

1.2 Subfields

- **Phonetics**: the study of speech sounds

- **Phonology**: the study of sound systems

- **Morphology**: the study of internal word structure

- **Historical**: the study of how languages change over time

- **Typology**: the study of how languages vary/differ

- **Syntax**: the study of word order and sentence structure

- **Semantics**: the study of linguistic meaning

- **Pragmatics**: the study of language use

2 Phonetics

Definition 1 (*Phonetics*)

The study of speech sounds

2.1 Preliminaries

Before studying phonetics, its important to understand the **speech chain**, or the process by which speech is (at a subconscious level) produced, and perceived.

From here, we can subdivide speech sounds into three processes:

- **articulatory process:** air stream movements, phonation, and articulation

- **acoustic process:** vibration of air

- **perceptual process:** auditory processes and mental categorization

Speech can be considered to be a *continuous* stream of sound, but in order to effectively study it, linguistcs subdivide it into discrete speech "units", or "segments". These segments are the smallest units of speech that can be distinguished by the human ear. The standard "collection" of all these segments is called the **International Phonetic Alphabet** (IPA), which charts all possible (or, rahter, considered possible) distinguishable sounds that humans can produce. The following sections will essentially be a brief overview of the IPA, how it is organized, and the information it contains.

Before getting into the IPA, it is important to understand the general **articulatory phonetics** by which humans produce speech. These can be broken down as follows:

1. **airstream mechanism:** sound is either produced *egressively* (exhaling), or *ingressively* (inhaling; much rarer)

2. **phonation:** this step is all about the *larynx*, a complex structure containing the *vocal folds* and the *glottis* (the gap through which air can pass). As air passes through the larynx, the vocal folds vibrate, creating a buzzing sound that eventually becomes speech.

   At this stage, the *glottal state*, or the configuration of the glottis, is the primary modifier of the air, and as a result, the sound:

   - **voiceless:**

     – vocal folds *apart*, glottis wide

     – no vibration of vocal folds

     – denoted with a ç̥

   - **modal:**

     – vocal folds *somewhat close*, such that as air comes through, the glottis opens and closes *quickly*

     – the vibration of the folds changes the frequency of the resulting speech

   - **creaky:**

     – vocal folds *compressed* and slack, vibrating slowly and irregularly

     – this results in a lower pitch

     – *in English, this voice does not distinguish speech, but does have social implications*

     – denoted with ç̰

   - **breathy:**

     – aka "murmur"

     – vocal folds *together*, but further than in modal

– *in English, this voice does not distinguish speech, but does have social implications*

– denoted with c̦

3. **filtering sound:** the upper vocal tract *filters* sound from the vocal folds by modifying the shape of different *articulators*. This is where the difference between vowels and consonants becomes clear. In general:

  - **consonants** have a **completely/partially closed** v.t., **no/turbulent** airflow, and **less energy**

  - **vowels** have an **open** v.t., **laminar** airflow, and **more energy**

2.2 Describing Consonants

Definition 2 (*Consonant*)

A speech sound created via a closure (partial/complete) of the vocal tract.

2.2.1 Articulators

Consonants are made and differentiated based off of the movement of the **upper v.t.** (containing the oral and nasal cavities). The relevant parts, called **articulators**, are as follows:

- **active:** in movement

  - tongue

  - bottom lip

- **passive:** destination of active articulators

  - top lip

  - teeth

  - alveolar ridge

  - palate

  - velum

  - uvula

  - back of pharyngeal wall

2.2.2 Parameters

- **voicing:**

  - *voiced:* vocal folds together → vibration

  - *voiceless:* vocal folds apart → no vibration

- **place of articulation:** where the sound "comes" from, generally described by where the tongue is/is moving to. (*bilabial, labiodental, interdental alveolar, alveopalatal, palatal, velar, uvular, pharyngeal, glottal*)

- **manner of articulation:** "how closed is closed"; subdivided based on how much constriction is present:

  - **stop** or *plosive*: a complete constriction. Stages: approach → hold → release

    In English, stops are often *aspriated*, meaning that the stop is followed by a burst of air. This is denoted with a superscript "h", ie [p] becomes [p^h].

  - **fricatives:** narrow constriction and high air volume.

    Within this, there are also **affricates**: a stop + fricative combination, denoted with a bar: [t ʃ] becomes [t͡ʃ].

  - **nasals:** produced by directing air through the nasal cavity, by lowering the velum.

  - **approximants, laterals:** "blur" between consants and vowles, created by drastically changing the the oral cavity resonator

    includes **glides:** "semi-vowels", such as [j], [w]; very similar to vowels, but with slightly more constriction

2.3 Describing Vowels

Definition 3 (*Vowel*)

A speech sound created via an open vocal tract.

Articulatorily:

- resonant

- very open v.t.

- (mostly) all voiced

- articulated with the tongue body

Acoustically:

- louder than consonants

- resonant frequencies

2.3.1 Source Filter Theory

While consonants differ *articulatorily*, vowels differ *accoustically*; ie, by different resonant frequencies. These frequences are determined by the shape of the v.t., the **filter**, which modify the air from the **source**, the glottal state.

2.3.2 Differentiating Vowels

The primary ways that vowels are described on the IPA are as follows:

- **height:** how high the tongue is in the mouth. The height of a vowel also determines the "closeness" of the vowel: a "high" vowel is "close", while a "low" vowel is "open".

- **backness:** how far back the tongue is in the mouth

- **roundedness:** how rounded the lips are

In some languages, vowels are further differentiated by whether they are **oral or nasal** (such as in French; nasal vowels are denoted with a tilde, ie [e] becomes [ẽ]), and their **length** (longer vowels are denoted with a colon, ie [a] becomes [a:]).

2.3.3 Monophthongs vs Diphthongs

Definition 4 (*Monophthong*)

A single vowel sound, with the position of articulators remaining constant.

Definition 5 (*Diphthong*)

A sound made of a sequence of vowel-glide/glide-vowel, with a noticable change in the articulators. Note that despite being denoted with two symbols, diphthongs are still a single sound segment. These include sounds like the Canadian English front mid-high vowel, [ej].

2.4 Suprasegmentals

Definition 6 (*Suprasegmental*)

Suprasegmentals are sounds that are not segments, but are still phonetically relevant. They can distinguish meanings in certain languages as well.

- **stress:** making a particular syllable(s) more prominent in a word, ie, stressed. In English, this can affect meaning, lexically; consider "insult" (verb) vs "insult" (noun), where there is a clear difference in stress.

- **tone:** differing pitch to distinguish meaning. At the simplest level, this involves making a low, mid, or high pitch, but can be more complex (such as in Mandarin, which has combinations of these simpler tones).

   There also exists **intonation**; when a pitch is associated with a particular stress syllable, this is called **pitch accent**; when a pitch is associated with the edge of an utterance, this is called **boundary tone**.

- **rhotacization:** an "r" coloring of a vowel, denoted [e˞] (such as in the last sound of "after")

3 Phonology

Definition 7 (*Phonology*)

The study of the mentally represented categories of speech sounds; an abstract representation of sounds.

3.1 Distribution

Definition 8 (*Contrastive Distribution*)

Two sounds are contrastive if replacing one with the other changes the meaning of the word. For instance, [t^h] and [d] are (in English) both found at the beginning of words, and which we "pick" affects the meaning.

Definition 9 (*Complementary Distribution*)

Two sounds are complementary if they are not found in the same phonological environment. In other words, the environment that a particular sound is used in determines which sound is used. For instance, in English, [t] is only found "within" words, so the sound at the beginning of words is [t^h]; these two sounds are in complementary distribution.

When we say that words are either contrastive or complementary, we can equivalently say that they are either **phonemes** or **allophones**. Specifically, a **phoneme** is a sound that is needed to distinguish words, and are transcribed with /.../. Phonemes can be considered as described how a sound is "thought of". An **allophone** is a sound that is contextually predictable, transcribed with [...]. Allophones are considered as how a sound is "actually produced". More specifically, we can describe two ways of transcription; **phonemic**:

- transcribe only info containing meaning

- ex, /paj/

and **phonetic**:

- transcribe all info (more specific)

- can be broad (ex [paj]) or more narrow (ex [pʰaj:])

- in narrow transcription, we use **diacritics** to denote the exact pronunciation of a sound

3.2 Evidence for Phonemes

Definition 10 (*Minimal Pair*)

A pair of words that differ by only one phoneme, and are contrastive.

3.3 Phonological Rules

We can state a rule to derive the expected variant of an allophone in a particular phrase/word. These require:

1. **Underlying Representation (UR)**: the "memorized" or lexically stored form of the word, *phonemically* transcribed, ie /l/ is the UR of [l], [ɫ], and [l̥].

2. **Surface Form (SF)**: the phonetic representation of a sound

3. **Context**: the context in which a sound is pronounced a particular way.

How do we get from the phonemic representation to the surface form ("actual" pronunciation)? We have to write a rule, specifying how a phoneme is pronounced in a given context.

For instance; if /l/ is at the end of a syllable, it becomes [ɫ], and it becomes [l̥] after a voiceless consonant; we don't necessarily have to write a rule for every possible context, and the default assumed context is simply [l]

Formally, linguists use the following notation:

$$/A/ \rightarrow [B]/X\_Y$$

Where A is the input (UR), B is the output (SF), and X_Y is the context, or "conditioning rule". For instance, with the English /l/ from above, we could say:

$$/l/ \rightarrow [ɫ] / \text{"voiceless consonant"}$$

Note that, when we write phonological rules like this, it is important to include a **useful name**, **description**, and then the **formal rule**. For instance:

- **Name:** North American Flapping

- **Description:** /t,d/ become flaps when between stressed and unstressed vowels

- **Rule:**
$$/t,d/ \rightarrow [ɾ]/\text{'stressed vowel'}\_\text{'unstressed vowel'}$$

- **Name:** Canadian Raising

- **Description:** /aj/ and /aw/ are raised (become [ʌj] and [ʌw]) when after voiceless consonants

- **Rule:**
$$/aj/,/aw/ \rightarrow [ʌj], [ʌw]/\_\text{"voiceless consonants"}$$

3.4 Syllables

Definition 11 (*Syllable*)

Phonological unit that groups consonants and vowels into larger structures, forming the basis for metrical structure (ie, rhythm). They typically consist of a vowel preceded and/or followed by a number of consonants, and can be denoted by CVC, CCVC, etc.

Many phonological processes are sensitive to syllable structure; for instance, stops are aspirated (in english) when they start syllables which are word-initial or stressed.

Some languages have more inclusive/exclusive possible syllable structures than others, which can affect learning the language or how loanwords are modified. For instance:

- Hawaiian has a strict (C)V pattern

- Korean has a (C)V(C) pattern

- French has a (C)(C)V(C)(C) pattern

3.4.1 Internal Structure of Syllables

Syllables are **not just** flat sequences of segments.

- **Onset**: all pre-nuclear consonants; they are optional in many languages. In English, they can get rather complex.

- **Rhyme**, or **rime**: important for stress-assignment, and create the rhythm of a word.

  - **Nucleus**: the "core" of the syllable, and is usually a vowel; in English, /n,l, ɹ(r)/ can also be used, and are denoted with a diacritic .

  - **Coda**: all post-nuclear consonants; they are optional in many languages. In English, they can get rather complex.

3.5 Syllabification

Definition 12 (*Syllabification*)

The process of dividing a word into syllables.

σ (syllable)

Onset   Rhyme
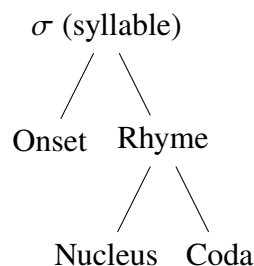
Nucleus   Coda

1. **Nucleus formation**: find the nuclei in the syllable. For the most part, is is simply the vowel in the syllable.

2. **Onset formation**: simply the consonants before the nucleus

3. **Coda formation**: (if it exists), simply the consonants after the nucleus

4. **Rhyme formation**: group the nucleus and coda together

In some multi-syllabic word, it may be difficult to distinguish an onset vs coda. In these cases, we use the following principles:

- **Maximum Onset Principle**: assign as many consonants to onset as possible

- **Sonority Contour Principle**: *sonority* (constriction degree in oral tract) rises before the nucleus and declines

11

after the nucleus. The typical hierarchy is as follows:

$$\text{vowels} > \text{glides} > \text{liquids} > \text{nasals} > \text{obstruents}$$

- **Binarity Principle**: for most languages, codas and onsets contain at **most** 2 segments.

3.6 Natural Classes

When we look across languages, we find patterns in which sounds occur together in stating rules; for instance, we expect to find rules applying to /p,t,k/, but not to /p,g,n,i/. This observation is formalized as **natural classes**.

Definition 13 (*Natural Class*)

Group of sounds with similar phonological properties, within a language; *groups will differ across languages.*

Phonological process often involve natural classes, allowing us to make more general rules about a larger group of sounds rather than individual sounds.

3.7 Derivation Tables

Derivation tables allow use to show the effect of rules in representative contexts. They contain particular words/sounds, a set of rules, and their output based off the effect of the rules. For instance:

UR	'mice' (/majs/)	'time' (/tajm/)
Canadian Raising	[mʌjs]	-
Output	[mʌjs]	/tajm/

We can use derivation tables with more rules, by simply adding more rows to the table.

UR	'writer' (/'ɹaj.tɚ/)	'rider' (/'ɹaj.dɚ/)
Canadian Raising	['ɹʌj.tɚ]	-
Flapping	['ɹʌj.ɾɚ]	[ɹaj.ɾɚ]
Output	['ɹʌj.ɾɚ]	[ɹaj.ɾɚ]

Note that order of the rules does matter! For instance, consider the table above, with the rules switched:

UR	'writer', /'ɹaj.tɚ/	'rider', /'ɹaj.dɚ/
Flapping	[ɹaj.ɾɚ]	[ɹaj.ɾɚ]
Canadian Raising	-	-
Output	[ɹaj.ɾɚ]	[ɹaj.ɾɚ]

In this case, it is predicted (*incorrectly*) that 'rider' and 'writer' are homophones, which is not the case (in Canadian English). The order of the rules depends on the context in which the rules are used

3.8 Free Variation

Definition 14 (*Free Variation*)

Variation that is not predicable from context; allophones of the same phoneme can occur in the same segmental environment.

4 Morphology

Definition 15 (*Lexicon*)

The "mental dictionary" of a language, with each lexical item associated with three pieces of information: phonological (form), semantic (meaning), and morpho-syntactic (how it combines with other lexical items).

Words and morphemes have both forms (sounds) and meanings (the concepts these forms express); yet, the relationship between the form and the meaning is **completely arbitrary**. This is clear from the fact that a word maintains meaning but changes form in different languages.

Its important to distinguish the following:

- **Words**: the smallest free-standing units of language

- **Morphemes**: the smallest unit that carries meaning/function

Further specifying the definition of a word becomes difficult across languages; while in some languages such as English we can intuitively say that words are "separated by spaces", this is not the case in other languages. Instead, it is more workable to use with the **morphemes** that make up words.

4.1 The Mental Lexicon

Meaning can be stored in different ways, in idiosyncratic, arbitrary form-meaning pairs;

- **compositional meaning**: can determine meaning from parts

- **non-compositional**: cannot determine meaning directly from parts; includes idioms

Morphology, like phonology, is rule-governed; arbitrary sound-meaning pairs + rules = creativity.

4.2 Allomorphs

Definition 16 (*Allopmorphs*)

Variant pronunciations of a morpheme based on the phonological context;

Allomorphy is not the same as allophony; allomorphy is morpheme-specific, while allophony happens anywhere given a particular lexical environment. **AlloMORPHy is to MORPHemes as alloPHONES are to PHONEmes.**

4.3 Lexicons to Surface Realizations

We can follow a set of rules to find the "actual" pronunciation of a particular word(s).

- **Lexical Entries**

- **Morphological Rules**

- **Allomorphy Rules**

- **Allophonic Rules**

- **Free Variation**

4.4   Morphemes

Bound vs free: a **free** morpheme can stand as an independent word, while a **bound** morpheme cannot (-s, -ing, -ed, etc.).

We distinguish between languages by their frequency of free morphemes; those with mostly free morphemes are **analytic**, while those with mostly bound morphemes are **synthetic**.

Root vs affix: a *root* morpheme is the major component of meaning in a word (the "read" in "reading", for instance), while the morphemes surrounding the root morpheme are *affixes*. In English, roots are typically also free; "bound roots" are rare, such as in "deflate".

We can further break down affixes as *suffixes* (after the root) and *prefixes* (before the root); some languages also have *infixes*, which are inserted within another morpheme ("abso-fucking-lutely").

4.5   What's in the Lexicon; Word Building

Do we store only words? Or words, and productive affixes? Or something else entirely? Regardless of which theory we take to be fact, we still need to be able to explain the internal structure of words. Word meanings are typically **compositional** (transparent from their parts), though not all logical combinations of morphemes are possible.

We formalize the "theory" of word building with the ideas of **compounds** and **affixation**.

4.5.1   Headedness

The **head** morpheme of a word is the one that determines the lexical category of the entire compound; in English, the right-most root is the head. For example, consider "green-house"; "house" is a noun, and thus "green-house" is a noun. More particularly;

- X + noun = noun

- X + verb = verb

- X + adj = adj

- etc.

We call English a "right headed language".

We can represent headedness with tree diagrams or with a bracketed notation. For instance, for the word "blackboard":

$$[[\text{black}]_{\text{Adj}}[\text{board}]_N]_N$$

```
          N
         / \
       Adj   N
        |    |
      black board
```
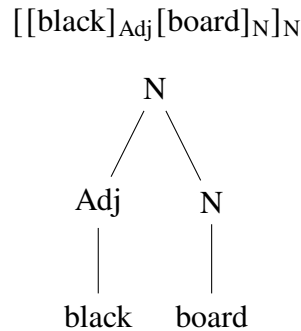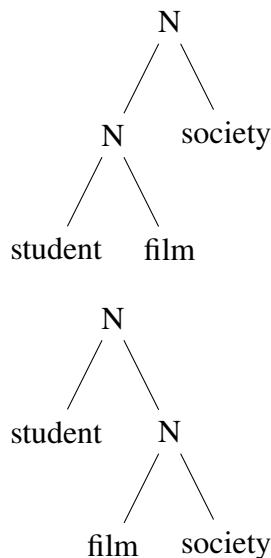
4.5.2 Distinguishing Compound Words and Phrases

Sometimes, a compound word has similar content to a syntactic phrase. We can distinguish these situations:

1. **Stress Pattern:** the stress pattern of a compound word is different from that of a phrase; for instance, "blackboard" has a stress on the first syllable, while "black board" has a stress on the second syllable.

2. **Integrity:**

3. **Semantic drift:** the meaning of a compound may drift over time to be less transparent; a "blackboard" is not necessarily a board that is black.

There may also exist some ambiguity in more complex compounds which consist of more than 2 components; consider, for instance, "student film society"; is this a society *for* student film, or a film society *for* students? These differences can be represented as the difference between $[[[\text{student}]_{\text{Adj}}[\text{film}]_N]_N[\text{society}]_N]_N$ and $[[\text{student}]_{\text{Adj}}[[\text{film}]_N[\text{society}]_N]_N]_N$

```
            N
           / \
          N   society
         / \
    student film
```

```
            N
           / \
     student  N
             / \
          film  society
```

4.5.3 Affixation

We can distinguish between **inflectional** and **derivational** affixes, which represent, respectively, changes in the lexical category of a word, and changes in the meaning of a word.

4.6 Discontinuous Morphology

English morphology is quite linear: suffixes appear the right, prefixes to the left. However, many languages (ie Arabic) have roots consisting solely made up of consonants, and deriving other free-forms requires inserting **vowels** between these consonants in order to create new words; the root and the affix are *interleaved*.

4.7 Inflection vs Derivation

Derivational morphology modifies a lexical item's form and derives **new meaning**, possible changing its **lexical** class.

Inflectional morphology modifies a lexical item's form in order to indicate the **grammatical subclass** to which it belongs.

4.8 Inflectional Morphology

There are two (main) kinds of inflection;

- **Verbal Inflection**: encodes tense and aspect (temporal information), or agreement with argument

- **Nominal Inflection**: encodes *case*, *number*, *grammatical gender/noun class*

4.9 Nominal Inflection

4.9.1 Case

Provides information about the **role** that a noun plays in a sentence. (ie nominative, accusative, dative, genitive, locative, ablative, etc.)

In English, the only case that is "productively" marked on *nouns* is *genitive* (possessive); ie "the man's house".

However, in English, *pronouns* are marked for case; for instance the **nominative case**, "I, he, she, they, it", compared to the **accusative case**, "Me, him, her, them, it".

Generally, languages use a mix of word order and case marking to indicate the roles of nouns.

4.9.2 Noun Class (Grammatical Gender)

Typically called *gender* in Indo-European languages (French, eg); it is generally arbitrary.

More broadly, grammatical gender is a way of classifying nouns in the lexicon into groups; how many classes exist

vary by language (German: three (masc, fem, neuter); French: two (masc, fem); English: one). Some have many more, such as the Bantu languages.

4.9.3 Number

More common system is **singular-plural** (simply one vs more than one). In some languages, both the noun and corresponding pronoun are marked, but in others just the noun is. Another common system is **singular-dual-plural**, distinguishing between one, two, and more than two (Gaelic, eg).

4.10 Verbal Inflection

Many (Indo-European) languages are *temporally marked*, in two main ways;

1. **Tense** indicates the point in time when an event took place (*temporally "when"*).

2. **Aspect** expresses the duration, the telicity, and time of completion (*temporally "how"*); eg perfective, imperfective

Often, tense and aspect along with another category, *mood*, are often expressed together in language (*TAM*).

4.11 Simultaneous Inflectional and Derivational Morphology

What happens if both processes are happening "together"? Inflectional affixes *always* follow derivational affixes.

Inflectional affixes are typically more productive, adn attach to (typically) all instances of the category. Derivational affixes, however, often attach to only a *restricted* set of bases.

Caution: some morphology is *homophonous*; two morphemes sound the same, but have distinct meanings/uses.

5 Historical Linguistics

5.1 Root Internal Changes

Changes inside the root can mark grammatical change. In English, we have **ablaut**, where the vowel changes in the root, but the consonants remain the same.

There is also the idea **suppletion** is a complete change in the morpheme of a word when it changes tense, for example, often due to historical changes; eg, "is" vs "was".

Reduplication copies some part of the root, and adding it to the root.

All languages change, over generations. The changes my be subtle over a single generation, but are quite drastic over time. All aspects of language change;

• **phonetics, phonology**; Old English (OE) had contrastive length, additional consonants phonemes

- **morphology**: OE had was more synthetic/inflectional

- **syntactic**: OE had Sbj-Obj-Verb or VOS, while modern English is SVO

- **lexical/semantic**: words (dis)appear and change meaning

Changes is language are usually regular, and systematic. for instance, a sound change affects all words with that particular sound.

5.2 Sound Change and Variation

Sound change has its root in **synchronic variation**, and is notated $[A] \sim [B]/[C]$.

- **Phonetically-conditioned** sound change (A $\rightarrow$ B/C_D)

- **Phonological change**: add, eliminated, or rearranged the phonemes

5.2.1 Phonological Changes

- **Assimilation**: sounds become more articulatorily similar along some *dimension*, ie voicing assimilation, place assimilation, nasalization, or palatalization (sounds become "more palatal" before front vowels/glides). Palatalization is transcribed either $[C] > [C^j]$ or [k] > [t͡ʃ]

- **Dissimilation**: similar sounds become less similar; less common than assimilation. This comes about due to difficulty of articulation or perception of similar sounds. For instance "surprise" somewhat loses the "r" sound, and becomes "suh-prise". This commonly occurs after liquids.

- **Epenthesis**: consonant sequences which are "hard to pronounce" or not permitted by the language's syllable structure get broken up. For instance, "schola" in Latin became "escuela" in spanish, inserting an "e" at the start. Many loanwords are transformed in this way. *Consonant* epenthesis can serve as an "articulatory bridge" between adjacent consonants; articulation is not instant, and we need time to move our articulators from one configuration to another.

- **Metathesis**: change in relative linear position of segment.

- **Lenition** (weakening): *vowel reduction* often occurs in prosodically weak (unstressed) positions, leading to complete deletion of the segment. In the case of *consonants*, "strength" is typically based on the general scale

    voiceless stops > voiceless fricatives, voiced stops > voiced fricatives > nasals > liquids > glides

    Extreme case: *deletion*.

- **Mergers:** Occur when two or more phonemes collapse into a single phoneme. Eg, th-front, where th become indistinguishable from f. Other mergers include

    - PIN-PEN

18

- – COT-CAUGHT

- – MARY-MARRY-MERRY

- **Splitting:** Splits occur when one phoneme with allophonic variation turns into two district phonemes. For instance, English n and velar n were originally allophones before [g]; when this [g] was lost, the two phonemes split (i.e., in the word "sing").

- **Tonogenesis:** A special case of splitting, where the environment triggers phonetic pitch differences are lost but the pitch differences remain.

- **Shifts:** In vowel systems, a typical kind of change is the *chain shift* where a change in one phoneme leads to further changes. For instance, the *Great English Vowel Shift* involving long vowels in ~1400-1700. Once a shift starts, all others must "make way"; languages like to use a whole quadrangle subsection of the vowel chart. However, it is less clear how these shifts begin.

5.3 Modeling Language Change

Language change is often modeled with either a tree model, or a wave model . The tree model

- captures the historical change from one stage of a langue to another

The wave model

- captures the synchronic variation in a language

- the gradual nature of individual changes

Regardless of what model we use, any kind of reconstruction supposed that changes can be undone; this assumes that sound change is systematic and regular. A strict view of this idea is called the **Neogrammarian hypothesis**, which states

1. all sound change **simultaneously** affects all words where the conditions environment is met;

2. sound change is **exceptionless**.

5.4 Reconstruction Example

- Starting from **cognates**: which segments in each language are shared, or are most common? (An innovation in a single language is more likely than multiple innovations in multiple languages)

- When there are many differences, one must consider the most plausible manner that the sound change occurred; making a sound easier to pronounce (lenition) is more likely.

- What is more likely; many languages undergoing epenthesis, or one language undergoing vowel reduction? At a weak position, vowel reduction is very plausible.

- **Inner subgroupings** can also form when reconstructing languages, where a language had to "split" multiple

times before reaching its "new" form.

5.5   Typology and Variation

The big view: language families consist of *"genetically related"* languages.

5.5.1   Some major language families

- **Indo-European**, including Germanic, Celtic, Italic, Hellenic, Albanian, Armenian, Baltic, Slavic, and Indo-Iranian.

- **Family of Indigenous Canadian Languages**

There are many languages in the world, which can be roughly broken down to **stable**, **endangered**, and **institutional** languages ($\sim$40, $\sim$ 45, $\sim$ 7).

5.6   Typology

Classification of languages based on shared features, such as phonology, morphology, etc; it asks the question, how do languages vary?

5.6.1   Phonological Typology

Some sounds/patterns are much more common than others, and we can make *typological generalizations* about them using *implicational universals*, of the form "if language has X then Y but not v.v.".

5.6.2   Stating Universals

- **Absolute**: all human languages do X; very rare, eg all languages have more than one vowel.

- **Tendencies**: most languages do X, or X is more common than Y; eg most languages have 3+ vowels.

We can categorize these categories with the idea of **markedness**; marked/ "less basic" vs unmarked/ "more basic". We can note this as $a > b$, where a is more marked than b.

6   Syntax

6.1   Constituency Tests: Noun Phrases

Constituents are interpretation-based subunits within sentences. Although they may seem intuitive, certain tests for them are often necessary.

6.1.1 Substitution Test

One characteristic of NP (noun phrase) units is that they can, in context, be replaced with a personal pronoun while maintaining meaning. For instance,

- **The children** are annoying (*Target sentence*)

- **They** are annoying (*Experimental sentence*)

The phrase **the children** here would be a noun phrase in this context. In general, if the phrase at hand can be substituted while maintaining meaning and grammatical structure, it is a noun phrase. As a counter example, *children* in the sentence *"the children are annoying"* is not a noun phrase, as replacing it with *"they"* forms the ungrammatical sentence *"*the they are annoying"*; here, *"the children"* would be a proper noun phrase.

Some symbols: "*" indicates a ungrammatical sentence and "#" indicates a contextually ambiguous sentence.

We can denote noun phrases:

$$\text{She gets } [_{NP}[_{Det}\text{her}][_N\text{wine}]][_{PP}\text{from Spain}]$$

6.1.2 Movement Test

A test by changing the order of word(s) in a sentence to see if the sentence maintains meaning (and grammar). For instance, "The cat plays that funky music" can be rearranged "that funky music is played by the cat", indicating that "that funky music" is a NP.

6.1.3 Coordination Test

If you can combine a string of words with a conjunction, then each original string is a NP ([X and X], [X or X], etc). For instance, *"The cat plays in the yard"* can be written *"The cat and the dog play in the yard"*, thus *"the cat"* is a noun phrase ($[_{NP}\text{NP and NP}]$).

Consider the target sentence *"This big dog will leave"* and the experimental sentence *"This big dog and small cat will leave"*; *"big dog"* and *"small cat"* are both nouns that coordinate to form a noun, meaning there must be a syntactic unit bigger than N and smaller than NP, which is often denoted N' (read as "N-bar").

6.1.4 One-Substitution Test

Used to identify the "middle" level; for instance *"This big cat and that big cat"*, where the second *"big cat"* can be replaced with *"one"*. "One" can only replace the N' level, *not* the lowest (in a syntax tree).

$$NP$$

```
          NP
         /  \
       Det   N'
             |
             N'
             |
             N
```
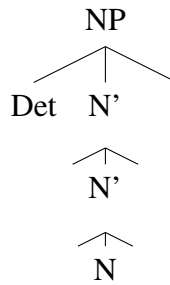
In summary: a "phrase" can be represented by a N-bar if it can be substituted out in a sentence while maintaining sense.

6.2 Constituency Tests: Verb Phrases

6.2.1 Movement Test

Example: "They say they will finish the test".

We can write "they say that they will finish the test AND", which must then be completed by "finish the test they will". Thus, "finish the test" is a VP.

$$\text{They say that they will } [_{VP} \text{ finish the test}]$$

6.2.2 Ellipsis/Deletion Test

In this test, we remove part of a sentence (or replace by ellipsis). For instance, "the parents have arrived at the school". Appending another phrase, "the children have arrived at the school too", we can delete "arrived at the school" and this sentence will still make sense ("the parents have arrived at the school and the children have ~~arrived at the~~ school too"). Thus, "arrived at the school" is a VP.

6.2.3 Do So-Substitution

In this test, we replace a potential VP with "*do so*". For instance, "They will read the book tomorrow"; replied with "Yes; they will *do so* tomorrow", and thus "read the book" is a V'.

We can further test V' by replacing "read the book tomorrow", ie "yes; they will *do so*". Thus, "read the book tomorrow" is a V'.
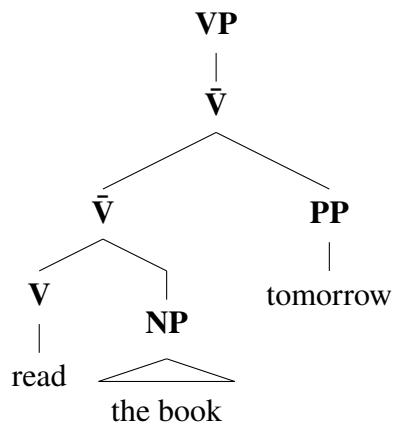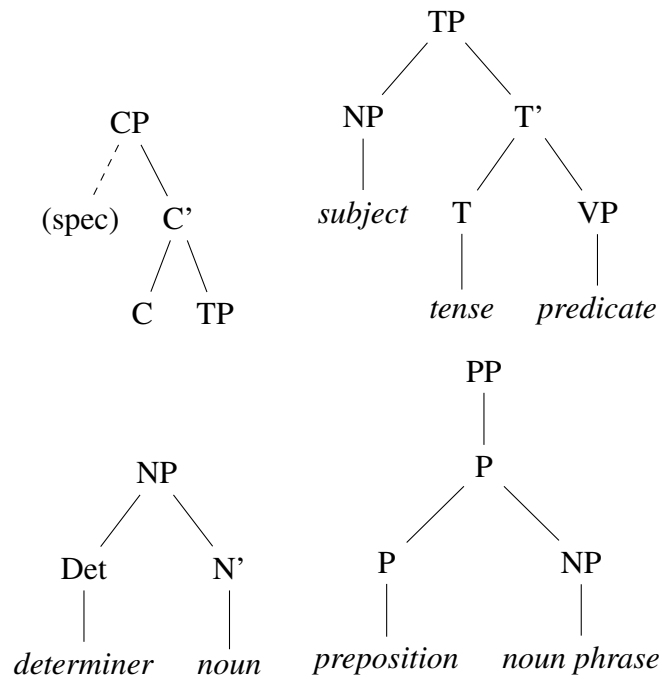
Note that we cannot just replace "read"; this would result in the nonsensical phrase "Yes; they will *do so* the book tomorrow", and thus "read" alone is not a V'.

6.3   Generalization: The X-Bar Schema
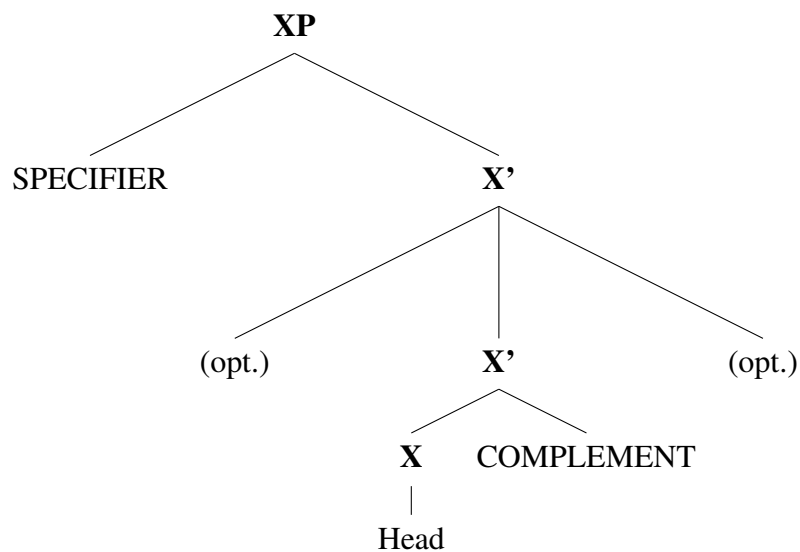
6.3.1   Typical Constituent Structures

Several constituent phrases (*in English*) have a fairly standardized structure, regardless of the context they appear in:

TP
CP
NP   T'
(spec)   C'
subject   T   VP
C   TP
tense   predicate

PP

NP   P

Det   N'   P   NP

determiner   noun   preposition   noun phrase

6.4 Head of a Sentence

Given the sentence "the children have left", we have the following tree:

?

NP
    ?
"Aux"   VP
the children
have   left
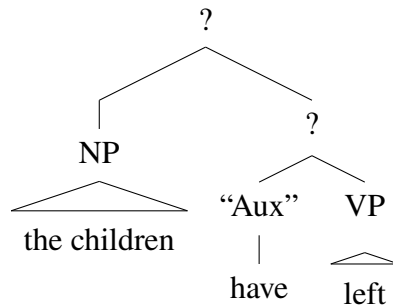
In the place of the "?"'s, we use the new category "T", the "Tense" group, for elements in auxiliary positions;

TP

NP   T'

T   VP
the children
have   left
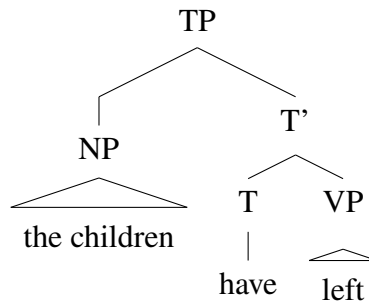
6.4.1 Auxiliary Positions

6.4.2 Empty T

The Tense (T) position of a sentence/phrase need not be occupied; for instance, consider the sentence "the children left", which has the following tree:
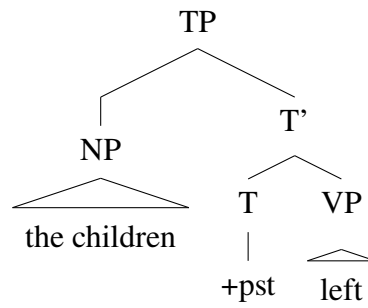
In this case, we fill the T with +pst, for "past tense".

6.5   Complements & Modifiers

Definition 17 (*Specifier*)

A phrase that combines with X' into XP. See tree 6.5.

Definition 18 (*Complement*)

A phrase that combines with X into X'. See tree 6.5.

Definition 19 (*Modifier*)

A phrase that combines with X' into X'. See tree 6.5.

With X-bar:
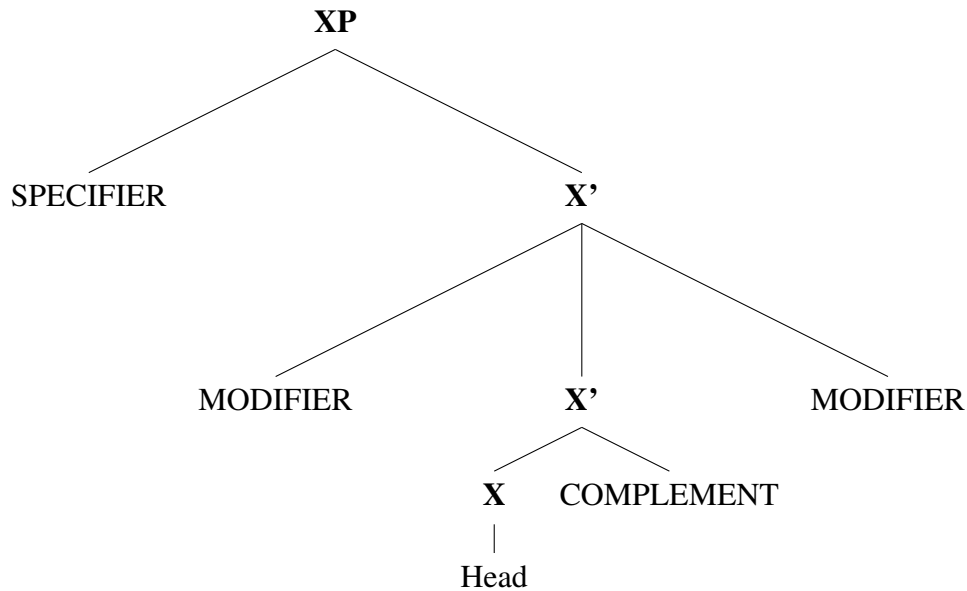
A *modifier* of XP is a phrase that combines with X' into X'

6.5.1   In Noun Phrases

For instance, take the two sentences

- $[_{NP}[_{Det}$the$][_N$student$][_{pp}$**from France**$]]$

- $[_{NP}[_{Det}$the$][_N$student$][_{pp}$**of physics**$]]$

25

With one-substitution test, we can say "the student from ... AND this *one* from France", and "the student of ... AND this *one* of physics". The second of these substitutions is questionable grammatically, but the first is not; thus, "student" is a N' and "from France" is a PP, and a *modifier* of this N'. On the other hand, assuming the second substituted sentence is ungrammatical, "student" is just a N, and "of physics" is a PP, and a *complement* of this N instead.

6.5.2 In Verb Phrases

Take the following phrases, and corresponding do-so substitutions:

1. "he will work *at* school" → "..., he will ~~work~~ **do so** at school"

2. "he will walk *to* school" → "..., he will ~~work~~ **do so** to school"

This second sentence is ungrammatical; while "work" is a V' in the first sentence, it is not in the second. Thus, similarly to with NP's, "at school" is a *modifier* of "work" in the first sentence, and a *complement* of "work" in the second sentence.

In general,

- **modifiers** are sisters of $\bar{X}$'s

- **complements** are sisters of X (NOT $\bar{X}$)

6.6 Complementizer Phrase

In order to describe more "complex" questions (specifically, questions, and "question-like" phrases), we introduce another constituent to the X-bar schema of CP, for complementizer phrase. In English, CP takes the following general form:

```
        CP
       ⁄  \
      X    C'
          ⁄ \
         C   TP
```
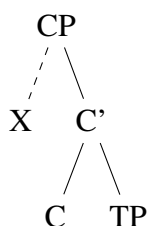
Technically, *any* sentence can be seen as a complementizer phrase at its base (with the TP being the "main" part of the sentence), though it is usually omitted for simplicity. Note too the dashed line to the specifier position of CP. In English, this position is only occupied in the case of a *wh*-phrase, such as "who", "what", "where". Often, it is simply there to be blank, such to allow for the movement of the subject of the sentence to the front of the sentence.

The C position is occupied by the complementizer; in English, there are few, such as "that", "whether", etc. C can also be empty.

When working with CP's and their accompanying phrases, things can often get complicated; in general, however, it is helpful to break phrases down into their constituents, and picture a larger CP by rearranging it such as to answer

26

the question it is posing. For instance, consider the question *For which child is it likely that they will buy a bike?*. This question can be broken down into "subphrases":

1. "it is likely that they will buy a bike for which child"

2. "it is likely that for which child they will buy a bike"

3. "is it likely that they will buy a bike"

This helps (questionably) break down how the structure of the sentence comes together, particularly what movement is going on.

CP

PP — C'

P' — C — TP

P — NP — is — NP — T'

for — Det — N' — N' — T — VP

which — child — N — *t* — V'

it — V — AP

t — A'

A

likely — CP

PP — C'

t — C — TP

that — NP — T'

N' — T — VP

N — will — V' — PP

they — V — NP — *t*

buy — Det — N'

a — N

bike

6.7 Language Variation

Not all languages have the same "rules" concerning complements, modifiers, etc; they often have their own distinct X-bar schemas. analyzing data (and drawing trees, of course) allows us to determine the X-bar schema of a particular language. To do so, recall the definitions of complements, modifiers, and specifiers, and look out for patterns in how they appear; recall that many types of constituents and constituent phrases can have modifiers, etc..

In English, we have the following:

| phrase	head	complement	specifier
VP	V	∅, NP, PP, AP, CP, VP	∅, adv
AP	A	∅, PP, CP	∅, adv
PP	P	∅, PP, NP, CP	∅, adv
NP	N	∅, PP, CP	∅, det
TP	T	VP	NP (subject)
CP	C	TP	∅, wh-phrase

6.7.1 Word Order

Not all languages follow the same word order as English. For instance, Madagasy is a VOS language ("read the book the child"). Nevertheless, the X-bar schema can still be used to analyze the structure; simply with some modification of order of where modifiers and complements appear.

6.7.2 Transformations: T-to-C Movement

Subject-Aux Inversion; movement of parts of a sentence in certain situations. ("Will Yoko sing?" → "Yoko will sing")

Complementizer (CP): a "new" projection that allows for the movement described above
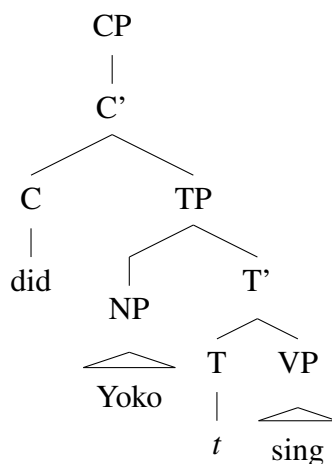
Similarly, we have the idea of **do-support**; "did yoko sing?" → "yoko *did* sing". This is the same support that is needed in answering a yes-no question, where the [+pst] takes the place of the "did" in the question; "yoko [+pst] sang".

6.7.3 Examples: French

Both French and English are SVO (ie "Satoshi helps Pikachu" vs "Satoshi aide Pikachu"). However, in some scenarios, things change;

- *adverb placement*: "she always works at home", *("elle tojours travaille à...")

- *yes-no questions*: *("works she at home?"), "travaille-t-elle à la maison?"

- *negation*: ...

French: If the head can move, it'll move If TENSE position already occupied, can't move!

7   Semantics and Pragmatics

Definition 20 (*Semantics*)

> *Studies the aspects of meaning that are independent of "context", i.e., the language users and the situations in which the language is actually used.*

Definition 21 (*Pragmatics*)

> *Studies the aspects of meaning that depend on or interact with context.*

7.1   Truth Conditions

What kind of "thing" is the meaning of a sentence? We call this the **truth condition** of a sentence. Specifically, to know the meaning of a sentence is to know under what conditions it would be true or false; this is part of your *unconscious semantic knowledge.*

This is a separate concept than the sentence's **truth value**: literally, if the sentence is true (T) or false (F). Knowing one of these does not imply knowing the other; however, knowing the truth value without knowing the truth conditions is less common.

Definition 22 (*Semantic Entailment*)

> *A sentence semantically entails another if its truth condition implies the truth condition of the other, ie "A, therefore, B". Phrase A is the premise, and phrase B is the conclusion.*

Definition 23 (*Pragmatic Entailment*)

> *When A, in conjunction with certain background knowledge, semantically entails B, we say that A pragmatically entails B.*

We can also say that:

- A **semantically contradicts** B if A's truth condition implies the negation of B's truth condition

- A is **semantically equivalent** to B if A and B have the same truth value in all "possible worlds"

7.2 Meaning Composition

Definition 24 (*Principle of Compositionality*)

The meaning of a sentence is determined by the meanings of the words it contains and the way they are syntactically combined.

Connectives:

- *Negation*: ¬ (not)

- *Conjunction*: ∧ (and)

- *Disjunction*: ∨ (or)

- *Conditional*: → (if, then)

Truth tables allow us to show the meanings of connectives.

Negation:

A	$\neg A$
T	F
F	T

Conjunction

A	B	$A \wedge B$
T	T	T
T	F	F
F	T	F
F	F	F

Disjunction:

A	B	$A \vee B$
T	T	T
T	F	T
F	T	T
F	F	F

Conditional:

A	B	$A \rightarrow B$
T	T	T
T	F	F
F	T	T
F	F	T

We can create truth trees for statements as follows:
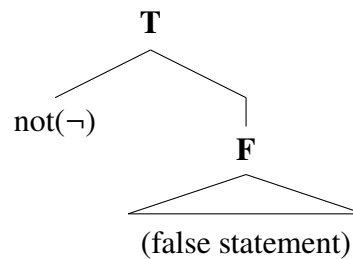
Where the root of the truth represents the *total* truth value of the statement.

7.2.1 Logic vs English

Negation (¬) and conjunction (∧) behave very closely to how "not" and "and" work in English. Conditionals, on the other hand, don't always quite align with the intuitive English logic.

For instance, consider the phrase "I won or she won". In English, this phrase would hold true if either I OR she won; but not both. This is is unlike the "rules" in logic. To describe this case, we introduce the **exclusive or** ($\oplus$) operator, which is the same as "or" except that it is false if both statements are true. Precisely:

A	B	$A \oplus B$
T	T	F
T	F	T
F	T	T
F	F	F

7.3 Presupposition

Some presupposition triggers include "again", "the", "know", "quit", for instance:

- "He went to jail *again*", ie, he's been there before

- "We met in *the* bar in Leacock", ie, there is a bar (and only one) in Leacock

- "He *knows* the answer", ie, the answer exists

- "I *quit* smoking", ie, I was smoking before

Note: for a sentence to be either true or false, the presupposition must be true as well; think of a "tricky" legal question, "would you cheat on the test again?".

For A to presuppose B:

A		B
T	$\Rightarrow$	T
F	$\Rightarrow$	T
T or F	$\Leftarrow$	T
#	$\Leftarrow$	F

In other words, for A to presuppose B, both A and the negation of A must imply that B is true; if B is false, there is no presupposition present. Conversely, if A implies B but the negation of A does not imply B, then A instead entails B.

7.3.1 Presupposition Projection
Definition 25 (*Presupposition Projection*)

The idea that a presupposition can be projected from a sentence to a larger unit, such as a clause or a sentence.

7.4 Implicature
Definition 26 (*Implicature*)

A speaker's meaning that is not part of the sentence's literal meaning, but is implied by the sentence's literal meaning. Total sentence meaning = what is said + what is implied.

Implications can be **cancelled**; ie, the implied "extra" meaning can be explicitly stated as being false. In contrast, presuppositions cannot. For instance, "I took my cat to the vet" implies (implicates?) that my cat is sick; I can clarify, and state that "she isn't sick, she just needs vaccines", *cancelling* the implication. However, I cannot cancel the presupposition (that I have a cat) by saying "I don't have a cat"; it is simply nonsensical. TL;DR: presuppositions *have* to be true, implicatures are *probably* true.

In short: to differentiate if the relationship between statements *A*, *B* is an entailment, presupposition, or implicature:

- *B* is cancellable, ie, *B* being false ⇏ *A* being false; **implicature**.

- *B* being false ⇒ *A* is # (deviant, ungrammatical); **presupposition**.

- *A*: true ⇒ *B*: true *and A*: false ⇒ *B*: true; **presupposition**

- *B*: false ⇒ *A*: false (*not* #); **entailment**

7.5   Gricean Maxims

Definition 27 (*Gricean Maxims*)

> *"Rules" of conversation that are not explicitly stated, but are assumed to be followed by speakers and listeners. They are formed under the* Principle of Cooperation: *people cooperate to effectively converse.*

- **Quantity**: Provide as much information as is required to convey the intended meaning.

- **Quality**: Be truthful.

- **Relation**: Be relevant.

- **Manner**: Be orderly, avoid obscurity and ambiguity.

By assuming that these maxims are being followed, intended/implied sentence meaning can be derived. They can be used to describe implicatures, particularly scalar ones:

7.5.1   Scalar Implicatures

A phrase expressing a scalar quality implies that the described quality is the highest-statable quality. For instance, "The essay is passable" implies that the essay is not "good".

Generally, if phrases A and B are on the same scale and B semantically entails B (but *not* vice versa), then B is implicated to be false, or at least not confidently true.

8   Glossary

Definition 1 (*Phonetics*)

The study of speech sounds

Definition 2 (*Consonant*)

A speech sound created via a closure (partial/complete) of the vocal tract.

Definition 3 (*Vowel*)

A speech sound created via an open vocal tract.

Definition 4 (*Monophthong*)

A single vowel sound, with the position of articulators remaining constant.

Definition 5 (*Diphthong*)

A sound made of a sequence of vowel-glide/glide-vowel, with a noticable change in the articulators. Note that despite being denoted with two symbols, diphthongs are still a single sound segment. These include sounds like the Canadian English front mid-high vowel, [ej].

Definition 6 (*Suprasegmental*)

Suprasegmentals are sounds that are not segments, but are still phonetically relevant. They can distinguish meanings in certain languages as well.

Definition 7 (*Phonology*)

The study of the mentally represented categories of speech sounds; an abstract representation of sounds.

Definition 8 (*Contrastive Distribution*)

Two sounds are contrastive if replacing one with the other changes the meaning of the word. For instance, $[t^h]$ and $[d]$ are (in English) both found at the beginning of words, and which we "pick" affects the meaning.

Definition 9 (*Complementary Distribution*)

Two sounds are complementary if they are not found in the same phonological environment. In other words, the environment that a particular sound is used in determines which sound is used. For instance, in English, $[t]$ is only found "within" words, so the sound at the beginning of words is $[t^h]$; these two sounds are in complementary distribution.

Definition 10 (*Minimal Pair*)

A pair of words that differ by only one phoneme, and are contrastive.

Definition 11 (*Syllable*)

Phonological unit that groups consonants and vowels into larger structures, forming the basis for metrical structure (ie, rhythm). They typically consist of a vowel preceded and/or followed by a number of consonants, and can be denoted by CVC, CCVC, etc.

Definition 12 (*Syllabification*)

The process of dividing a word into syllables.

Definition 13 (*Natural Class*)

Group of sounds with similar phonological properties, within a language; groups will differ across languages.

Definition 14 (*Free Variation*)

Variation that is not predicable from context; allophones of the same phoneme can occur in the same segmental environment.

Definition 15 (*Lexicon*)

The "mental dictionary" of a language, with each lexical item associated with three pieces of information: phonological (form), semantic (meaning), and morpho-syntactic (how it combines with other lexical items).

Definition 16 (*Allopmorphs*)

Variant pronunciations of a morpheme based on the phonological context;

Definition 17 (*Specifier*)

A phrase that combines with X' into XP. See tree 6.5.

Definition 18 (*Complement*)

A phrase that combines with X into X'. See tree 6.5.

Definition 19 (*Modifier*)

A phrase that combines with X' into X'. See tree 6.5.

Definition 20 (*Semantics*)

Studies the aspects of meaning that are independent of "context", i.e., the language users and the situations in which the language is actually used.

Definition 21 (*Pragmatics*)

Studies the aspects of meaning that depend on or interact with context.

Definition 22 (*Semantic Entailment*)

A sentence semantically entails another if its truth condition implies the truth condition of the other, ie "A, therefore, B". Phrase A is the premise, and phrase B is the conclusion.

Definition 23 (*Pragmatic Entailment*)

When A, in conjunction with certain background knowledge, semantically entails B, we say that A pragmatically entails B.

Definition 24 (*Principle of Compositionality*)

The meaning of a sentence is determined by the meanings of the words it contains and the way they are syntactically combined.

Definition 25 (*Presupposition Projection*)

The idea that a presupposition can be projected from a sentence to a larger unit, such as a clause or a sentence.

Definition 26 (*Implicature*)

A speaker's meaning that is not part of the sentence's literal meaning, but is implied by the sentence's literal meaning. Total sentence meaning = what is said + what is implied.

Definition 27 (*Gricean Maxims*)

"Rules" of conversation that are not explicitly stated, but are assumed to be followed by speakers and listeners. They are formed under the Principle of Cooperation: *people cooperate to effectively converse.*