Introduction to Linguistics

Marcus Kracht
Department of Linguistics, UCLA
3125 Campbell Hall
450 Hilgard Avenue
Los Angeles, CA 90095–1543
kracht@humnet.ucla.edu

Contents

Lecture 1: Introduction

> Languages are sets of *signs*. Signs combine an exponent (a sequence of letters or sounds) with a meaning. Grammars are ways to generate signs from more basic signs. Signs combine a form and a meaning, and they are identical with neither their exponent nor with their meaning.

Before we start. I have tried to be as explicit as I could in preparing these notes. You will find that some of the technicalities are demanding at first sight. Do not panic! You are not expected to master these technicalities right away. The technical character is basically due to my desire to be as explicit and detailed as possible. For some of you this might actually be helpful. If you are not among them you may want to read some other book on the side (which I encourage you to do anyway). However, linguistics is getting increasingly formal and mathematical, and you are well advised to get used to this style of doing science. So, if you do not understand right away what I am saying, you will simply have to go over it again and again. And keep asking questions! New words and technical terms that are used for the first time are typed in bold-face. If you are supposed to know what they mean, a definition will be given right away. The definition is valid throughout the entire course, but be aware of the fact that other people might define things differently. This applies when you read other books, for example. You should beware of possible discrepancies in terminology. If you are not given a definition elsewhere, be cautious. If you are given a different definition it does not mean that the other books get it wrong. The symbol  in the margin signals some material that is difficult, and optional. Such passages are put in for those who want to get a perfect understanding of the material; but they are not requried knowledge.

(End of note)

Language is a means to communicate, it is a semiotic system. By that we simply mean that it is a *set of signs*. Its A **sign** is a pair consisting—in the words of Ferdinand de Saussure—of a **signifier** and a **signified**. We prefer to call the signifier the **exponent** and the signified the **meaning**. For example, in English the string /dog/ is a signifier, and its signified is, say, doghood, or the set of all dogs. (I use the slashes to enclose concrete signifiers, in this case sequences of letters.) Sign systems are ubiquitous: clocks, road signs, pictograms—they all are parts of

sign systems. Language differs from them only in its complexity. This explains why language signs have much more internal structure than ordinary signs. For notice that language allows to express virtually every thought that we have, and the number of signs that we can produce is literally endless. Although one may find it debatable whether or not language is actually infinite, it is clear that we are able to understand utterances that we have never heard before. Every year, hundreds of thousands of books appear, and clearly each of them is new. If it were the same as a previously published book this would be considered a breach of copyright! However, no native speaker of the language experiences trouble understanding them (apart from technical books).

It might be far fetched, though, to speak of an entire book as a sign. But nothing speaks against that. Linguists mostly study only signs that consist of just one sentence. And this is what we shall do here, too. However, texts are certainly more than a sequence of sentences, and the study of **discourse** (which includes texts and dialogs) is certainly a very vital one. Unfortunately, even sentences are so complicated that it will take all our time to study them. The methods, however, shall be useful for discourse analysis as well.

In linguistics, language signs are constituted of four different levels, not just two: **phonology**, **morphology**, **syntax** and **semantics**. **Semantics** deals with the meanings (what is signified), while the other three are all concerned with the exponent. At the lowest level we find that everything is composed from a small set of sounds, or—when we write—of letters. (Chinese is exceptional in that the alphabet consists of around 50,000 'letters', but each sign stands for a syllable—a sequence of sounds, not just a single one.) With some exceptions (for example tone and intonation) every utterance can be seen as a sequence of sounds. For example, /dog/ consists of three letters (and three sounds): /d/, /o/ and /g/. In order not to confuse sounds (and sound sequences) with letters we denote the sounds by enclosing them in square brackets. So, the sounds that make up [dog] are [d], [o] and [g], in that order. What is important to note here is that sounds by themselves in general have no meaning. The decomposition into sounds has no counterpart in the semantics. Just as every signifier can be decomposed into sounds, it can also be decomposed into words. In written language we can spot the words by looking for minimal parts of texts enclosed by blanks (or punctuation marks). In spoken language the definition of word becomes very tricky. The part of linguistics that deals with how words are put together into sentences is called **syntax**. On the other hand, words are not the smallest meaningful units of

language. For example, /dogs/ is the plural of /dog/ and as such it is formed by a regular process, and if we only know the meaning of /dog/ we also know the meaning of /dogs/. Thus, we can decompose /dogs/ into two parts: /dog/ and /s/. The minimal parts of speech that bear meaning are called **morphemes**. Often, it is tacitly assumed that a morpheme is a part of a word; bigger chunks are called **idioms**. Idioms are /kick the bucket/, /keep taps on someone/, and so on. The reason for this division is that while idioms are intransparent as far as their meaning is concerned (if you die you do not literally kick a bucket), syntactically they often behave as if they are made from words (for example, they inflect: /John kicked the bucket/).

So, a word such as 'dogs' has four manifestations: its meaning, its sound structure, its morphological structure and its syntactic structure. The levels of manifestation are also called **strata**. (Some use the term **level of representation**.) We use the following notation: the sign is given by enclosing the string in brackets: 'dog'. $[\text{dog}]_P$ denotes its phonological structure, $[\text{dog}]_M$ its morphological structure, $[\text{dog}]_L$ its syntactic structure and $[\text{dog}]_S$ its semantical structure. I also use typewriter font for symbols in print. For the most part we analyse language as written language, unless otherwise indicated. With that in mind, we have $[\text{dog}]_P$ = /dog/. The latter is a string composed from three symbols, /d/, /o/ and /g/. So, 'dog' refers to the sign whose exponent is written here /dog/. We shall agree on the following.

Definition 1 *A **sign** is a quadruple $\langle \pi, \mu, \lambda, \sigma \rangle$, where π is its **exponent** (or **phonological structure**), μ its **morphological structure**, λ its **syntactic structure** and σ its **meaning** (or **semantic structure**).*

We write signs vertically, in the following way.

(1) $$\begin{bmatrix} \sigma \\ \lambda \\ \mu \\ \pi \end{bmatrix}$$

This definition should not be taken as saying something deep. It merely fixes the notion of a linguistic sign, saying that it consists of nothing more (and nothing less) than four things: its phonological structure, its morphological structure, its syntactic structure and its semantic structure. Moreover, in the literature there are

numerous *different* definitions of signs. You should not worry too much here: the present definition is valid throughout *this book only*. Other definitions have other merits.

The power of language to generate so many signs comes from the fact that it has rules by which complex signs are made from simpler ones.

(2) `Cars are cheaper this year.`

In (2), we have a sentence composed from 5 words. The meaning of each word is enough to understand the meaning of (2). Exactly how this is possible is one question that linguistics has to answer. (This example requires quite a lot of machinery to be solved explicitly!) We shall illustrate the approach taken in this course. We assume that there is a binary operation •, called **merge**, which takes two signs and forms a new sign. • operates on each of the strata (or levels of manifestation) independently. This means that there are four distinct operations, Ⓟ, Ⓜ, Ⓛ, and Ⓢ, which simultaneously work together as follows.

(3)
$$\begin{bmatrix} \sigma_1 \\ \lambda_1 \\ \mu_1 \\ \pi_1 \end{bmatrix} \bullet \begin{bmatrix} \sigma_2 \\ \lambda_2 \\ \mu_2 \\ \pi_2 \end{bmatrix} = \begin{bmatrix} \sigma_1 \,Ⓢ\, \sigma_2 \\ \lambda_1 \,Ⓛ\, \lambda_2 \\ \mu_1 \,Ⓜ\, \mu_2 \\ \pi_1 \,Ⓟ\, \pi_2 \end{bmatrix}$$

Definition 2 *A **language** is a set of signs. A **grammar** consists of a set of signs (called **lexicon**) together with a finite set of functions that each operate on signs.*

Typically, though not necessarily, the grammars that linguists design for natural languages consist in the lexicon plus a single binary operation • of merge. There may also be additional operations (such as movement), but let's assume for the moment that this is not so. Such a grammar is said to **generate** the following language (= set of signs) L:

① Each member of the lexicon is in L.

② If S and S' are in L, then so is $S \bullet S'$.

③ Nothing else is in L.

(Can you guess what a general definition would look like?) We shall now give a glimpse of how the various representations look like and what these operations are. It will take the entire course (and much more) to understand the precise consequences of Definitions 1 and 2 and the idea that operations are defined on each stratum *independently*. But it is a very useful one in that it forces us to be clear and concise. Everything has to be written into one of the representations in order to have an effect on the way in which signs combine and what the effect of combination is.

For example, Ⓟ is typically concatenation, with a blank added. Let us represent strings by $\vec{x}$, $\vec{y}$ etc., and concatenation by $^\frown$. So,

(4) $\texttt{dac}^\frown\texttt{xy} = \texttt{dacxy}$

(5) $\texttt{adf}^\frown\square^\frown\texttt{xy} = \texttt{adf xy}$

Notice that visually, $\square$ ('blank') is not represented at the end of a word. In computer books one often uses the symbol $\textvisiblespace$ to represent the blank. (Clearly, though the symbol is different from the blank!) Blank is a symbol (on a typewriter you have to press space to get it. So, $\texttt{x}^\frown\square$ is *not* the same as x! Now we have

(6) $\vec{x}\,Ⓟ\,\vec{y} := \vec{x}^\frown\square^\frown\vec{y}$

For example, the sign 'this year' is composed from the signs 'this' and 'year'. And we have

(7) $\texttt{this year} = [\text{this year}]_P = [\text{this}]_P\,Ⓟ\,[\text{year}]_P = \texttt{this}^\frown\square^\frown\texttt{year}$

This, however, is valid only for words and only for written language. The composition of smaller units is different. No blank is inserted. For example, the sign 'car' the plural sign 's' (to give it a name) compose to give the sign with exponent /cars/, not /car s/. Moreover, the plural of /man/ is /men/, so it is not at all formed by adding /s/. We shall see below how this is dealt with.

Morphology does not get to see the individual makeup of its units. In fact, the difference between 'car' and 'cat' is morphologically speaking as great as that between 'car' and 'moon'. Also, both are subject to the same morphological rules and behave in the same way, for example form the plural by adding 's'. That makes them belong to the same noun class. Still, they are counted as different morphemes. This is because they are manifested differently (the sound structure is different). Therefore we distinguish between a morpheme and its **morphological**

structure. The latter is only the portion that is needed on the morphological stratum to get everything right.

Definition 3 *A **morpheme** is an indecomposable sign.*

A morpheme can only be defined relative to a grammar. If we have only $\bullet$, then S is a morpheme of there are no S' and S'' with $S = S' \bullet S''$. (If you suspect that essentially the lexicon may consist in all and only the morphemes, you are right. Though the lexicon may contain more elements, it cannot contain less.) A word is something that is enclosed by blanks and/or punctuation marks. So the punctuation marks show us that a morpheme is a word. To morphology, 'car' is known as a noun that takes an s-plural. We write

(8) $$\begin{bmatrix} \textsc{mor} : \text{n} \\ \textsc{cls} : \text{s-pl} \end{bmatrix}$$

to say that the item is of morphological category 'n' (nominal) and that it has inflectional category 's-pl' (which will take care of the fact that its plural will be formed by adding 's').

To the syntactic stratum the item 'cars' is known only as a plural noun despite the fact that it consists of two morphs. Also, syntax is not interested in knowing how the plural was formed. The syntactic representation therefore is the following.

(9) $$\begin{bmatrix} \textsc{cat} & : & \text{N} \\ \textsc{num} & : & \text{pl} \end{bmatrix}$$

This says that we have an object of category N whose number is plural. We shall return to the details of the notation later during the course. Now, for the merge on the syntactic stratum let us look again at 'this year'. The second part, 'year' is a noun, the first a determiner. The entire complex has the category of a determiner phrase (DP). Both are singular. Hence, we have that in syntax

(10) $$\begin{bmatrix} \textsc{cat} & : & \text{D} \\ \textsc{num} & : & \text{sg} \end{bmatrix} \textcircled{\text{L}} \begin{bmatrix} \textsc{cat} & : & \text{N} \\ \textsc{num} & : & \text{sg} \end{bmatrix} = \begin{bmatrix} \textsc{cat} & : & \text{DP} \\ \textsc{num} & : & \text{sg} \end{bmatrix}$$

This tells us very little about the action of $\textcircled{\text{L}}$. In fact, large parts of syntactic theory are consumed by finding out what merge does in syntax!

Semantical representations are too complex to be explained here (it requires a course in model-theory or logic to understand them). We shall therefore not say much here. Fortunately, most of what we shall have to say here will be clear even without further knowledge of the structures. Suffice it to say, for example, that the meaning of 'car' is the set of all cars (though this is a massive simplification this is good enough for present purposes); it is clearly different from the meaning of 'cat', which is the set of all cats. Further, the meaning of 'cars' is the set of all sets of cars that have at least two members. The operation of forming the plural takes a set A and produces the set of all subsets of A that have at least two members. So:

(11) $[\![\text{s}]\!]_S : \{\spadesuit, \heartsuit, \diamond, \clubsuit\} \mapsto$
$\{\{\spadesuit, \heartsuit\}, \{\spadesuit, \diamond\}, \{\spadesuit, \clubsuit\}, \{\heartsuit, \diamond\}, \{\heartsuit, \clubsuit\}, \{\diamond, \clubsuit\},$
$\{\spadesuit, \heartsuit, \diamond\}, \{\spadesuit, \heartsuit, \clubsuit\}, \{\spadesuit, \diamond, \clubsuit\}, \{\heartsuit, \diamond, \clubsuit\}, \{\heartsuit, \diamond, \clubsuit\},$
$\{\spadesuit, \heartsuit, \diamond, \clubsuit\}\}$

With this defined we can simply say that Ⓢ; is function application.

(12) $$M \text{Ⓢ} N := \begin{cases} M(N) & \text{if defined,} \\ N(M) & \text{otherwise.} \end{cases}$$

The function is $[\![\text{s}]\!]_S$ and the argument is $[\![\text{car}]\!]_S$, which is the set of all cars. By definition, what we get is the set of all sets of cars that have at least two members in it. Our typographical convention is the following. For a given word, say 'cat' the semantics is denoted by sans-serife font plus an added prime: cat'.

Here is a synopsis of the merge of 'this' and 'year'.

(13)
$$\left[\begin{array}{c} \mathsf{this}' \\ \left[\begin{array}{ll} \textsc{cat} & \text{D} \\ \textsc{num} & \text{sg} \end{array}\right] \\ \left[\begin{array}{ll} \textsc{mor} & \text{n} \\ \textsc{cls} & \text{abl} \end{array}\right] \\ \texttt{this} \end{array}\right] \bullet \left[\begin{array}{c} \mathsf{year}' \\ \left[\begin{array}{ll} \textsc{cat} & \text{N} \\ \textsc{num} & \text{sg} \end{array}\right] \\ \left[\begin{array}{ll} \textsc{mor} & \text{n} \\ \textsc{cls} & \text{s-pl} \end{array}\right] \\ \texttt{year} \end{array}\right] = \left[\begin{array}{c} \mathsf{this}'(\mathsf{year}') \\ \left[\begin{array}{ll} \textsc{cat} & \text{DP} \\ \textsc{num} & \text{sg} \end{array}\right] \\ \left[\begin{array}{ll} \textsc{mor} & \text{np} \\ \textsc{cls} & \star \end{array}\right] \\ \texttt{this year} \end{array}\right]$$

(Here, 'abl' stands for 'ablaut'. What it means is that the distinction between singular and plural is signaled only by the vowel. In this case it changes from [ɪ] to [i:]. ⋆ means: no value.) One may ask why it is at all necessary to distinguish morphological from syntactic representation. Some linguists sharply divide

between lexical and syntactical operations. Lexical operations are those that operate on units below the level of words. So, the operation that combines 'car' and plural is a lexical operation. The signs should have no manifestation on the syntactical stratum, and so by definition, then, they should not be called signs. However, this would make the definition unnecessarily complicated. Moreover, linguists are not unanimous in rejecting syntactic representations for morphemes, since it poses more problems than it solves (this will be quite obvious for so-called polysynthetic languages). We shall not attempt to solve the problem here. Opinions are quite diverse and most linguists do accept that there is a separate level of morphology.

A last issue that is of extreme importance in linguistics is that of deep and surface structure. Let us start with phonology. The sound corresponding to the letter /l/ differs from environment to environment (see Page 525 of Fromkin et. al.). The 'l' in the pronunication of /slight/ is different from the 'l' in (the pronunciation of) /listen/. If we pronounce /listen/ using the 'l' sound of /slight/ we get a markedly different result (it sounds a bit like Russian accent). So, one letter has different realizations, and the difference is recognized by the speakers. However, the difference between these sounds is redundant in the language. In fact, in written language they are represented by just one symbol. Thus, one distinguishes a **phone** (= sound) from a **phoneme** (= set of sounds). While phones are language independent, phonemes are not. For example, the letter /p/ has two distinct realizations, an aspirated and an unaspirated one. It is aspirated in /pot/ but unaspirated in /spit/. Hindi recognizes two distinct phonemes here. A similar distinction exists in all other strata, though we shall only use the distinction between **morph** and **morpheme**. A morpheme is a set of morphs. For example, the plural morpheme contains a number of morphs. One of them consists in the letter /s/, another in the letters /en/ (which are appended, as in /ox/:/oxen/), a third is zero (/fish/:/fish/). And some more. The morphs of a morpheme are called allomorphs of each other. If a morpheme has several allomorphs, how do we make sure that the correct kind of morph is applied in combination? For example, why is the plural of /car/ not /caren/ or /car/? The answer lies in the morphological representation. Indeed, we have proposed that morphological representations contain information about word classes. This means that for nouns it contains information about the kind of plural morph that is allowed to attach to it. If one looks carefully at the setup presented above, the distinction between deep and surface stratum is however nonexistent. There is no distinction between morpheme and morph. Thus, either there are no morphs or there are no morphemes.

Both options are theoretically possible.

Some notes. The idea of stratification is implicit in many syntactic theories. There are differences in how the strata look like and how many there are. **Transformational grammar** recognizes all four of the strata (they have been called **Logical Form** (for the semantical stratum) **S-structure** (for syntax) and **Phonetic Form** or **PF** (for phonological stratum). Morphology has sometimes been considered a separate, lexical stratum, although some theories (for example **Distributed Morphology**) try to integrate it into the overall framework. **Lexical Functional Grammar** (**LFG**) distinguishes **c(onstituent)-structure** (= syntax), **a(rgument)-structure**, **f(unctional)-structure** and **m(orphological)-structure**.

There has also been **Stratificational Grammar**, which basically investigated the stratal architecture of language. The difference with the present setup is that Stratificational Grammar assumes independent units at all strata. For example, a morpheme is a citizen of the morphological stratum. The morpheme 'car' is different from the morpheme 'cat', for example. Moreover, the lexeme 'car' is once again different from the morpheme 'car', and so on. This multiplies the linguistic ontology beyond need. Here we have defined a morpheme to be a sign of some sort, and so it has just a manifestation on all strata rather than belonging to any of them. That means that our representation shows no difference on the morphological stratum, only on the semantical and the phonological stratum.

Alternative Reading. I recommend [Fromkin, 2000] for alternative perspective. Also [O'Grady *et al.*, 2005] is worthwhile though less exact.

Lecture 2: Phonetics

> Phonetics is the study of sounds. To understand the mechanics of human languages one has to understand the physiology of the human body. Letters represent sounds in a rather intricate way. This has advantages and disadvantages. To represent sounds by letters in an accurate and uniform way the International Phonetic Alphabet (IPA) was created.

We begin with phonology and phonetics. It is important to understand the difference between phonetics and phonology. Phonetics is the study of actual sounds of human languages, their production and their perception. It is relevant to linguistics for the simple reason that the sounds are the primary physical manifestation of language. Phonology on the other hand is the study of sound *systems*. The difference is roughly speaking this. There are countless different sounds we can make, but only some count as sounds of a language, say English. Moreover, as far as English is concerned, many perceptibly distinct sounds are not considered 'different'. The letter /p/, for example, can be pronounced in many different ways, with more emphasis, with more loudness, with different voice onset time, and so on. From a phonetic point of view, these are all different sounds; from a phonological point of view there is only one (English) sound, or *phoneme*: [p].

The difference is very important though often enough it is not evident whether a phenomenon is phonetic in nature or phonological. English, for example, has a basic sound [t]. While from a phonological point of view there is only one phoneme [t], there are infinitely many actual sounds that realize this phoneme. So, while there are infinitely many different sounds for any given language there are only finitely many phonemes, and the upper limit is around 120. English has 40 (see Table 7). The difference can be illustrated also with music. There is a continuum of pitches, but the piano has only 88 keys, so you can produce only 88 different pitches. The chords of the piano are given, so that the basic sound colour and pitch cannot be altered. But you can still manipulate the loudness, for example. Sheet music reflects this state of affairs in the same way as written language. The musical sounds are described by discrete signs, the keys. Returning now to language: the difference between various different realizations of the letter /t/, for example, are negligeable in English and often enough we cannot even tell the difference between them. Still, if we recorded the sounds and mapped them out in a spectrogram we could actually see the difference. (Spectrograms are one

Table 1: The letter /x/ in various languages

Language	**Value**
Albanian	[dʒ]
Basque	[x]
English	[gz]
French	[gz]
German	[ks]
Portuguese	[ʃ]
Spanish	[ç]
Pinyin of Mandarin	[ɕ]

important instrument in phonetics because they visualize sounds so that you can see what you often even cannot hear.) Other languages cut the sound continuum in a different way. Not all realizations of /t/ in English sound good in French, for example. Basically, French speakers pronounce /t/ without aspiration. This means that if we think of the sounds as forming a 'space' the so-called basic sounds of a language occupy some region of that space. These regions vary from one language to another.

Languages are written in alphabets, and many use the Latin alphabet. It turns out that not only is the Latin alphabet not always suitable for other languages, orthographies are often not a reliable source for pronunciation. English is a case in point. To illustrate the problems, let us look at the following tables (taken from [Coulmas, 2003]). Table 1 concerns the values of the letter /x/ in different languages: As one can see, the correspondence between letters and sounds is not at all uniform. On the other hand, even in one and the same language the correspondence can be nonuniform. Table 2 lists ways to represent [ə] is English by letters. Basically any of the vowel letters can represent [ə]. This mismatch has various reasons, a particular one being language change and dialectal difference. The sounds of a language change slowly over time. If we could hear a tape recording of English spoken, say, one or two hundred years ago in one and the same region, we would surely notice a difference. The orthography however tends to be conservative. The good side about a stable writing system is that we can (in principle) read older texts even if we do not know how to pronounce them. Second, languages with strong dialectal variation often fix writing according to

Table 2: The sound [ə] in English

Letter	Example
a	about
e	believe
i	compatible
o	oblige
u	circus

one of the dialects. Once again this means that documents are understood across dialects, even though they are read out differently.

I should point out here that there is no unique pronunciation of any letter in a language. More often than not it has quite distinct vaues. For example, the letter /p/ sounds quite different in /photo/ as it does in /plus/. In fact, the sound described by /ph/ is the same as the one normally described by /f/ (for example in /flood/). The situation is that we nevertheless ascribe a 'normal' value to a letter (which we use when pronouncing the letter in isolation or in reciting the alphabet). This connection is learned in school and is part of the writing system, by which I mean more than just the rendering of words into sequences of letters. Notice a curious fact here. The letter /b/ is pronounced like /bee/ in English, with a subsequent vowel that is not part of the value of the letter. In Sanskrit, the primitive consonantal letters represent the consonant plus [a], while the recitation of the letter is nowadays done without it. For example, the letter for "b" has value [bə] when used ordinarily, while it is recited [b]. If one does not want a pronunciation with schwa, the letter is augmented by a stroke.

In the sequel I shall often refer to the pronunciation of a letter; by that I mean the standard value assigned to it in reciting the alphabet, however without the added vowel. This recipe is, I hope, reasonably clear, though it has shortcomings (the recitation of /w/ reveals little of the actual sound value).

The disadvantage for the linguist is that the standard orthographies have to be learned (if you study many different languages this can be a big impediment) and second they do not reveal what is nevertheless important: the sound quality. For that reason one has agreed on a special alphabet, the so-called **International Phonetic Alphabet** (**IPA**). In principle this alphabet is designed to give an accurate

written transcription of sounds, one that is uniform for all languages. Since the IPA is an international standard, it is vital that one understands how it works (and can read or write using it). The complete set of symbols is rather complex, but luckily one does not have to know all of it.

The Analysis of Speech Sounds

First of all, the continuum of speech is broken up into a sequence of discrete units, which we referred to as sounds. Thus we are analysing language utterances as sequences of sounds. Right away we mention that there is an exception. **Intonation** and **stress** are an exception to this. The sentences below are distinct only in intonation (falling pitch versus falling and rising pitch).

(14) `You spoke with the manager.`

(15) `You spoke with the manager?`

Also, the word /`protest`/ has two different pronunciations; when it is a noun the stress is on the first syllable, when it is a verb it is on the second. Stress and intonation obviously affect the way in which the sounds are produced (changing loudness and / or pitch), but in terms of decomposition of an utterance into segments intonation and stress have to be taken apart. We shall return to stress later. Suffice it to say that in **IPA** stress is marked not on the vowel but on the syllable (by a [ˈ] before the stressed syllable), since it is though to be a property of the syllable. **Tone** is considered to be a suprasegmental feature, too. It does not play a role in European languages, but for example in languages of South East Asia (including Chinese and Vietnamese), in languages of Africa and Native American languages. We shall not deal with tone.

Sounds are produced in the vocal tract. Air is flowing through the mouth and nose and the characteristics of the sounds are manipulated by several so-called **articulators**. A rough picture is that the mouth aperture is changed by moving the jaw, and that the shape of the cavity can be manipulated by the tongue in many ways. The parts of the body that are involved in shaping the sound, the **articulators**, can be **active** (in which case they move) or **passive**. The articulators are as follows: **oral cavity**, **upper lip**, **lower lip**, **upper teeth**, **alveolar ridge** (the section of the mouth just behind the upper teeth stretching to the 'corner'), **tongue tip**, **tongue blade** (the flexible part of the tongue), **tongue body**, **tongue**

Table 3: IPA consonant column labels

	Articulators involved
bilabial	the two lips, both active and passive
labiodental	active lower lip to passive upper teeth
dental	active tongue tip/blade to passive upper teeth
alveolar	active tongue tip/blade to passive front part of alveolar ridge
postalveolar	active tongue blade to passive behind alveolar
retroflex	active tongue tip raised or curled to passive postalveolar (difference between postalveolar and retroflex: blade vs. tip)
palatal	tongue blade/body to hard palate behind entire alveolar ridge
velar	active body of tongue to passive soft palate (sometimes to back of soft palate)
uvular	active body of tongue to passive (or active) uvula
pharyngeal	active body/root of tongue to passive pharynx
glottal	both vocal chords, both active and passive

root, **epiglottis** (the leaf-like appendage to the tongue in the pharynx), **pharynx** (the back vertical space of the vocal tract, between uvula and larynx), **hard palate** (upper part of the mouth just above the tongue body in normal position), **soft palate** or **velum** (the soft part of the mouth above the tongue, just behind the hard palate), **uvula** (the hanging part of the soft palate), and **larynx** (the part housing the vocal chords). For most articulators it is clear whether they can be active or passive, so this should not need further comment.

It is evident that the **vocal chords** play a major role in sounds (they are responsible for the distinction between **voiced** and **unvoiced**), and the sides of the tongue are also used (in sounds known as **laterals**). Table 3 gives some definitions of phonetic features in terms of articulators for consonants. Column labels here refer to what defines the **place of articulation** as opposed to the **manner of articulation**. The **degree of constriction** is roughly the distance of the active articulator to the passive articulator. The degree of constriction plays less of a role in consonants, though it does vary, say, between full contact [d] and 'close encounter' [z], and

Table 4: Constriction degrees for consonants

stop	active and passive articulators touch an hold-to-seal (permitting no flow of air out of the mouth)
trill	active articulator vibrates as air flows around it
tap/flap	active and passive articulators touch but don't hold (includes quick touch and fast sliding)
fricative	active and passive articulators form a small constriction, creating a narrow gap causing noise as air passes through it
approximant	active and passive articulators form a large constriction, allowing almost free flow of air through the vocal tract

it certainly varies during the articulation (for example in affricates [dz] where the tongue retreats in a slower fashion than with [d]). The **manner of articulation** combines the degree of constriction together with the way it changes in time. Table 4 gives an overview of the main terms used in the IPA and Table 5 identifies the row labels of the IPA chart. Vowels differ from consonants in that there is no constriction of air flow. The notions of active and passive articulator apply. Here we find at least four degrees of constriction (**close**, **close-mid**, **open-mid** and **open**), corresponding to the height of the tongue body (plus degree of mouth aperture). There is a second dimension for the horizontal position of the tongue body. The combination of these two parameters is often given in the form of a two dimensional trapezoid, which shows with more accuracy the position of the tongue. There is a third dimension, which defines the rounding (**round** versus **unrounded**, which is usually not marked). We add a fourth dimension, **nasal** versus **nonnasal**, depending on whether the air flows partly through the nose or only through the mouth.

Naming the Sounds

The way to name a sound is by stringing together its attributes. However, there is a distinction between naming vowels and consonants. First we describe the names of consonants. For example, [p] is described as a voiceless, bilabial stop, [m] is

Table 5: IPA consonant row labels

plosive	a pulmonic-egressive, oral stop
nasal	a pulmonic-egressive stop with a nasal flow; not a plosive, because not oral
fricative	a sound with fricative constriction degree; implies that airflow is central
lateral fricative	a fricative in which the airflow is lateral
approximant	a sound with approximant constriction degree; implies that the airflow is central
lateral approximant	an approximant in which the airflow is lateral

Table 6: IPA vowel row and column labels

close	compared with other vowels, overall height of tongue is greatest; tongue is closest to roof of mouth (also: high)
open	compared with other vowels, overall height of mouth is least; mouth is most open (also: low)
close-mid, open-mid	intermediate positions (also: mid / uppermid / lowermid)
front	compared with other vowels, tongue is overall forward
central	intermediate position
back	compared with other vowels, tongue is overall back (near pharynx)
rounded	lips are constricted inward and protruded forward

called a (voiced) bilabial nasal. The rules are as follows:

(16) **voicing place manner**

Sometimes other features are added. If we want to describe [p^h] we say that it is a voiceless bilabial aspirated stop. The additional specification 'aspirated' is a manner attribute, so it is put after the place description (but before the attribute 'stop', since the latter is a noun). For example, the sequence 'voiced retroflex fricative' refers to [ʐ], as can be seen from the IPA chart.

Vowels on the other hand are always described as 'vowels', and all the other features are attributes. We have for example the description of [y] as 'high front rounded vowel'. This shows that the sequence is

(17) **height place lip-attitude [nasality]** vowel

Nasality is optional. If nothing is said, the vowel is not nasal.

On Strict Transcription

Since IPA tries to symbolize a sound with precision, there is a tension between accuracy and usefulness. As we shall see later, the way a phoneme is realized changes from environment to environment. Some of these changes are so small that one needs a trained ear to even hear them. The question is whether we want the difference to show up in the notation. At first glance the answer seems to be negative. But two problems arise: (a) linguists sometimes *do* want to represent the difference and there should be a way to do that, and (b) a contrast that speakers of one language do not even hear might turn out to be distinctive and relevant in another. (An example is the difference between English [d] (alveolar) and a sound where the tongue is put between the teeth (dental). Some languages in India distinguish these sounds, though I hardly hear a difference.) Thus, on the one hand we need an alphabet that is highly flexible on the other we do not want to use it always in full glory. This motivates using various systems of notation, which differ mainly in accuracy. Table 7 gives you a list of English speech sounds and a phonetic symbol that is exact insofar that knowing the IPA would tell an English speaker exactly what sound is meant by what symbol. (I draw attention however to the sound [a], which according to the IPA is not used in American English; instead, we find [ɑ].) This is called **broad transcription**. The dangers

of broad transcription are that a symbol like [p] does not reveal exact details of which sounds fall under it, it merely tells us that we have a voiceless bilabial stop. Since French broad transcription might use the same symbol [p] for that we might be tempted to conclude that they are the same. But they are not.

Thus in addition to broad transcription there exists strict or narrow transcription, which consists in adding more information (say, whether [p] is pronounced with aspiration or not). Clearly, the precision of the IPA is limited. Moreover, the more primitive symbols it has the harder it is to memorize. Therefore, IPA is based on a set of a hundred or so primitive symbols, and a number of diacritics by which the characteristics of the sound can be narrowed down.

Notes on this section. The book [Rodgers, 2000] gives a fair and illuminating introduction to phonetics. It is useful to have a look at the active sound chart at

`http://hctv.humnet.ucla.edu/departments/linguistics/VowelsandConsonants/course/chapter1/chapter1.html`

You can go there and click at symbols to hear what the corresponding sound is. A very useful source is also the Wikipedia entry.

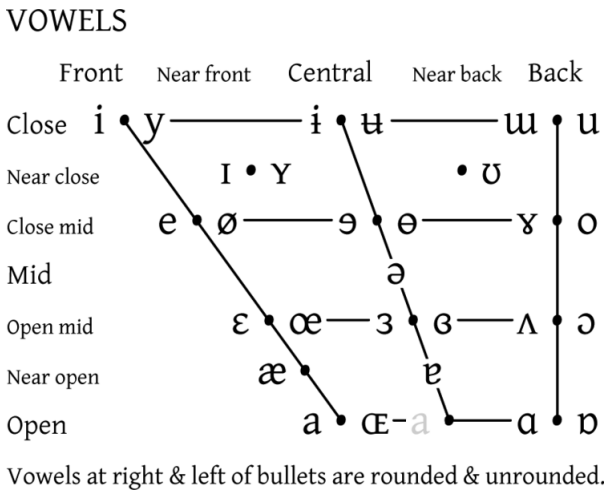
`http://en.wikipedia.org/wiki/International_Phonetic_Alphabet`

Table 7: The Sounds of English

	Phonetic Symbol	**Word illustrating it**
1	p	pope
2	b	barber
3	m	mum
4	f	fife
5	v	vital, live
6	t	taunt
7	d	deed
8	n	nun
9	ɹ	rare
10	θ	thousandth
11	ð	this, breathe
12	s	source, fuss
13	z	zanies
14	ʃ	shush
15	ʒ	measure
16	l	lul
17	tʃ	church
18	dʒ	judge
19	j	yoke
20	k	cook
21	g	gag
22	ŋ	singing
23	w	we
24	h	he
25	i	easy
26	ɪ	imitate
27	e	able
28	ɛ	edge
29	æ	battle, attack
30	a	father
31	ɔ	fought
32	o	road
33	ʊ	book, should
34	u	food
35	ə	aroma
36	ʌ	but
37	ɚ (or ɝ or r̩)	bird
38	aɪ	ride
39	aʊ	house
40	oɪ	boy

Figure 1: IPA Consonant Chart

Figure 2: IPA Vowel Chart



Phonology I: Features and Phonemes

This chapter will introduce the notions of *feature* and *phoneme*. Moreover, we show how the formalism of *attribute value structures* offers a succinct way of describing phonemes and phoneme classes. A *natural class* is one which can be described by a single attribute value structure.

Distinctiveness

There is a continuum of sounds but there is only a very limited set of distinctions that we look out for. It is the same with letters: although you can write them in many different ways most differences do not matter at all. There are hundreds of different fonts for example, but whether you write the letter 'a' like this: *a* or like this: ᴀ, it usually makes no difference. Similarly, some phonetic contrasts are relevant others are not. The question is: what do we mean by relevance? The answer is: if the contrast makes a difference in meaning it is relevant. The easiest test is to find to words that mean different things but differ only in one sound. These are called **minimal pairs**. Table 8 shows some examples of minimal pairs. We see from the first pair that the change from [h] to [k] may induces a change in meaning. Thus the contrast is relevant. In order for this to be meaningful at all we should spell out a few assumptions. The first assumption, established in the last section, is that the sound stream is segmentable into unique and identifiable units. The sound stream of an utterance of /hat/ will thus consist of three sounds, which I write as [h], [æ] and [t]. Similarly, the sound stream of an utterance of /cat/ consists of three sounds, [k], [æ] and [t]. The next assumption is that the two sequences are of equal length, which allows us to align the particular sounds with each other:

(18)

h	æ	t
•	•	•
k	æ	t

And the third assumption is that we can actually 'exchange' particular occurrences of sounds in the stream. (Technically, one can do this nowadays by using software allowing at manipulate any parts of a spectrogram. From an articulatory point of view, exchanging exact sounds one by one is impossible because of adaptations made by the surrounding sounds. The realisation of [æ] will in likelihood be

slightly different whether it is preceded by [h] or by [k].) Given all this, we declare the sound stream to be a minimal pair just in case they have different meaning. Clearly, whether they do or not is part of what the language is; recall that a language is a relation between exponents (here: sound streams) and meanings.

Definition 4 (Minimal Pair) *Two sound streams form a minimal pair, if their segmentations are of the same length, and one can be obtained from the other by exchaning just one sound for another, and that the change results in a change of meaning.*

It is to be stressed that minimal pairs consist of two entire words, not just single sounds (unless of course these sounds *are* words). I should emphasise that by this definition, for two words to be minimal pair they must be of equal length in terms of how many basic sounds constitute them, not in terms of how many alphabetic characters are needed to write them. This is because we want to establish the units of speech, not of writing. The same length is important for a purely formal reason: we want to be sure that we correctly associate the sounds with each other. Also, there is no doubt that the presence of a sound constrasts with its absence, so we do not bother to check whether the presence of a sound makes a difference, rather whether the presence of *this* sound makes a difference other the presence of some *other* sound at a given position.

Likewise, (b) shows that the contrast [p]:[t] is relevant (from which we deduce that the contrast labial:dental is relevant, though for other sounds it need not make a difference). (c) shows that the contrast [æ]:[ʌ] is relevant, and so on. Many of the contrasts between the 40 or so basic sounds of English can be demonstrated to be relevant by just choosing two words that are minimally different in that one has one sound and the other has the other sound. (Although this would require to establish $(40 \times 39)/2 = 780$ minimal pairs, one is usually content with far less.) Let us note also that in English certain sounds just do not exist. For example, retroflex consonants, lateral fricatives are not used at all by English speakers. Thus we may say that English uses only *some* of the available sounds, and other languages use others (there are languages that have retroflex consonants, for example many languages spoken in India). Additionally, the set of English sounds is divided into 40 groups, each corresponding to one letter in Table 7. These groups are called **phonemes** (and correspond to the 40 letters used in the broad transcription). The letter 'l' for example pretty much corresponds to a phoneme of English, which in turn is realized by many distinct sounds. The IPA actually allows to represent the

Table 8: Some Minimal Pairs in English

(a)	`hat`	[hæt]	: `cat`	[kʰæt]
(b)	`cat`	[kʰæt]	: `cap`	[kʰæp]
(c)	`cap`	[kʰæp]	: `cup`	[kʰʌp]
(d)	`flight`	[flaɪt]	: `fright`	[fɹaɪt]
(e)	`flight`	[flaɪt]	: `plight`	[plaɪt]

different sounds to some degree:

(19)

`file` [ˈfaɪɬ]	`slight` [ˈs͡l̥lait]	`wealth` [ˈwɛɬ̪θ]	`listen` [ˈlisən]
`fool` [ˈfuɬ]	`flight` [ˈf͡l̥lait]	`health` [ˈhɛɬ̪θ]	`lose` [ˈluz]
`all` [ˈaɬ]	`plow` [ˈp͡l̥laʊ]	`filthy` [ˈfɪɬ̪θi]	`allow` [əˈlaʊ]

The phoneme therefore contains the 'sounds' [ɬ], [͡l̥l], [l̪] and [l]. (In fact, since the symbols are again only approximations, they are themselves not sounds but sets of sounds. But let's ignore that point of detail here.) The following picture emerges. Utterances are strings of sounds, which the hearer (subconsciously) represents as sequences of phonemes:

(20)

sounds	→	σ_1	σ_2	σ_3	σ_4	...
phonemes	→	p_1	p_2	p_3	p_4	...

The transition from sounds to phonemes is akin to the transition from narrow ((21)) to broad ((22)) transcription:

(21) [ˈd̪ðɪs iz ə fəˈnɛɾɨˀk t͡ʂʰɹɛənˈskɹ̥ɪpʃɨn]

(22) [ðɪs iz ə foʊnɛdɪk tɹænskɹɪpʃən]

(23) `this is a phonetic transcription`

The conversion to phonemic representation means that a lot of information about the actual sound structure is lost, but what is lost is immaterial to the message itself.

We mention right away that the different sounds of a phoneme do not always occur in the same environment. If one sound σ can always be exchanged by σ' of the same phoneme, then σ and σ' are said to be in **free variation**. If however σ and σ' are not in free variation, we say that the realization of the phoneme as either σ or σ' is **conditioned by the context**.

Table 9: Phonemes of English

Consonants

		bila-bial	labio-dental	den-tal	alve-olar	palato-alveolar	pala-tal	velar	glot-tal
stops	vl	/p/			/t/	/tʃ/		/k/	
	vd	/b/			/d/	/dʒ/		/g/	
frica-	vl		/f/	/θ/	/s/	/ʃ/			/h/
tives	vd		/v/	/ð/	/z/	/ʒ/			
nasals		/m/			/n/			/ŋ/	
appro-	lat				/l/				
ximants	cnt	/w/			/ɹ/		/j/		

vl = voiceless, vd = voiced, lat = lateral, cnt = central

Vowels and Diphthongs

	front	central	back		diphthongs
	unrounded	unrounded	unrounded	rounded	
upper high	/i/			/u/	/aɪ/, /aʊ/,
lower high	/ɪ/			/ʊ/	/oɪ/
upper mid	/e/	/ə/		/o/	syllabic
lower mid	/ɛ/		/ʌ/		consonant
low	/æ/		/ɑ/		/ɚ/

Table 9 gives a list of the phonemes of American English. The slanted brackets denote phonemes not sounds, but the sounds are nevertheless given in IPA. On the whole, the classification of phonemes looks very similar to that of sounds. But there are mismatches. There is a series of sounds, called **affricates**, which are written as a combination of a stop followed by a fricative: English has two such phonemes, [tʃ] and [dʒ]. Similarly, **diphthongs**, which are written like sequences of vowels or of vowel and glide, are considered just one phoneme. Notice also that the broad transcription is also hiding some diphthongs, like [e] as in `/able/`. This is a sequence of the vowel [e] and the glide [j]. The reason is that the vowel [e] is *obligatorily followed* by [j] and therefore mentioning of [j] is needless. (However, unless you know English well you need to be told this fact.) The sequence [aɪ] is different in that [a] is not necessarily followed [ɪ], whence writing the sequence is unavoidable.

Some Concerns in Defining a Phoneme

So while a phone is just a sound, a concrete, linearly indecomposable sound (with the exception of certain diphthongs and affricates), a phoneme on the other hand is a set of sounds. Recall that in the book a phoneme is defined to be a basic speech sound. It is claimed, for example, that in Maasai [p], [b] and [β] are in complementary distribution. Nevertheless Maasai is said to have a phoneme /p/, whose feature specification is that of [p]. This means among other that it can only be pronounced as [p]. This view has its justification. However, the theoretical justification is extremely difficult. There is no reason to prefer one of the sounds over the other. By contrast we define the following. Let $\hat{}$ denote concatenation.

Definition 5 (Phoneme) *A **phoneme** is a set of phones (= speech sounds). In a language L, two sounds a and b belong to the same phoneme if and only if for all strings of sounds $\vec{x}$ and $\vec{y}$: if both $\vec{x}\hat{}a\hat{}\vec{y}$ and $\vec{x}\hat{}b\hat{}\vec{y}$ belong to L, they have the same meaning. a and b are **allophones** if and only if they belong to the same phoneme.*

We also say the following. If $\vec{x}\hat{}a\hat{}\vec{y} \in L$ then the pair $\langle \vec{x}, \vec{y} \rangle$, which we write $\vec{x}\_\_\vec{y}$ is an **environment** for a in L. Another word for environment is **context**.

So, if a and b belong to the same phoneme, then either in a given word (or text) containing a one cannot substitute b for a, or one can but the result has the same meaning; and in a text containing b somewhere either one cannot substitute a for b or one can and the result has the same meaning. Take the sounds [t] and [ɾ] in (American) English (see Page 529 of [Fromkin, 2000]). They are in complementary distribution, that is, in a context $\vec{x}\_\_\vec{y}$ at most one of them can appear. So, we have [ˈdeʲ ɾə] but not [ˈdeʲtə] (the context is ˈdeʲ__ə). (The second sounds British.) On the other hand we have [ˈtæn] but not [ˈɾæn] (the context is ˈ__æn). (Notice that to pronounce /`data`/ [ˈdeʲtə] or even [ˈdeʲtʰə] is actually not illegitimate; this is the British pronunciation, and it is understood though not said. The meaning attributed to this string is just the same. The complications arising from the distinction between how something is pronounced correctly and how much variation is tolerated shall not be dealt with here.) On the other hand, if we change the position of the tongue slightly (producing, say [t̻] in place of [t]), the resulting string is judged to be the same. Hence it also means the same. We say that [t] and [t̻] are in **free variation**. So, two allophones can in a given context either be in complementary distribution (or can occur and the other cannot) or in free variation (both can occur). This can vary from context to context, though.

Definition 6 (Phoneme) *If L is a language, and p a specific sound then* $/p/_L$ *denotes the phoneme containing p in L.*

The definition in [Fromkin, 2000] of a phoneme as *one of the basic speech sounds of a language* is different from ours. So it needs comment why we do it differently. First, it needs to be established what a basic speech sound is. For example, in Maasai [p] and [β] are in complementary distribution. By our definition, the sounds instantiating either [p] or [β] all belong to the same Maasai phoneme, which we denote by $/p/_{\text{Maasai}}$. But is $/p/_{\text{Maasai}}$ a basic speech sound? How can we know? It seems that Fromkin et al. do not believe that it is. They take instead the phoneme to be [p], and assume that the context distorts the realization. Now look at English. The sound [p] is sometimes pronounced with aspiration and sometimes not. The two realizations of the letter /p/, [p] and [p^h] do not belong to the same phoneme in Hindi. If this is the case it is difficult to support the idea that $/p/_{\text{English}}$ can be basic. If we look carefully at the definition above, it involves also the notion of meaning. Indeed, if we assume that a word, say /car/, has endlessly many realizations, the only way to tell that we produced something that is not a realization of /car/ is to establish that it does not mean what a realization of /car/ means. Part of the problem derives from the notation [p], which suggests that it is clear what we mean. But it is known that the more distinctions a language makes in some dimension, the narrower defined the basic speech sounds are. English, for example, has only two bilabial stops, which we may write [p] and [b]. Sanskrit (and many languages spoken in India today) had four: [p], [p^h], [b] and [b^h]. There is thus every reason to believe that the class of sounds that pass for a 'p' in English is dissimilar to that in Sanskrit (or Hindi or Thai, which are similar in this respect). Thus, to be perfect we should write $[p]_{\text{English}}$, $[p]_{\text{Sanskrit}}$ and so on. Indeed, the crucial parameter that distinguishes all these sounds, the Voice Onset Time is a *continuous parameter*. (The VOT is the delay of the onset of voicing after the airstream release. The larger it is the more of an aspiration we hear.) The distinction that is binary on the abstract level turns out to be based on a continuum which is sliced up in a somewhat arbitrary way. The discussion also has to do with the problem of narrow versus wide transcription. When we write [p] we mean something different for English than for Hindi, because it would be incorrect to transcribe for a Hindi speaker the sound that realizes /p/ in /pal/ by [p]; we should use [p^h] instead.

Features

By definition, any set of sounds can constitute a phoneme. However, it turns out that phonemes are constituted by classes of sounds that have certain properties in common. These are defined by **features**. Features are phonetic, and supposed to be not subject to cross-language variation. What exactly is a feature? The actual features found in the literature take a (more or less) articulatory standpoint. Take any sound realizing English /b/. It is produced by closing the lips, thereby obstructing the air flow ('bilabial') and then releasing it, and at the same time letting the vocal cords vibrate ('voiced'). If the vocal cords do not vibrate we get the sound corresponding to /p/. We can analyse the sound as a motor program that is executed on demand. Its execution is not totally fixed, so variation is possible (as it occurs with all kinds of movements that we perform). Second, the motor program directs various parts of the vocal tracts, some of which are independent from each other. We may see this as a music score which has various parts for different 'instruments'. The score for the voicing feature is one of them. The value '+' tells us that the cords have to vibrate during the production of the corresponding sound, while '−' tells us that they should not. We have to be a bit cautious, though. It is known, for example, that /b/ is not pronounced with immediate voicing. Rather, the voicing is delayed by a certain onset time. This onset time varies from language to language. Hence, the actual realization of a feature is different across languages, a fact that is rather awkward for the idea that phonemes are defined by recourse to *phonetic* features. The latter should namely be language independent. The problem just mentioned can of course be resolved by making finer distinctions with the features. But the question remains: just how much detail do the phonetic features need to give? The answer is roughly that while phonetically we are dealing with a *continuous* scale (onset time measured in milliseconds), at the phonemic level we are just looking at a binary contrast.

We shall use the following notation. There is a set of so-called **attributes** and a set of so-called **values**. A pair [ATT : val] consisting of an attribute and a value is called a **feature**. We treat +voiced as a notational alternative of [VOICED : +]. An attribute is associated with a **value range**. For phonology, we may assume the following set of attributes:

(24) PLACE, MANNER, VOICED, CONSONANTAL, ASPIRATED, APERTURE, . . .

and we may assume the following set of values:

(25) bilabial, labiodental, plosive, approximant, high, mid, +, −, . . .

The range of PLACE is obviously different from that of MANNER, since 'dental' is a value of the former and not of the latter. A set of features is called an **attribute value structure** (**AVS**). You have seen AVSs already in the first lecture. The notation is as follows. The attributes and values are arranged vertically, the rows just having the attribute paired with its value, separated by a colon:

(26)
$$\begin{bmatrix} \text{PLACE} & : \text{dental} \\ \text{MANNER} & : \text{fricative} \\ \text{VOICE} & : + \end{bmatrix}$$

Notice that the following are also legitimate AVSs:

(27)
$$\begin{bmatrix} \text{PLACE} & : \text{dental} \\ \text{PLACE} & : \text{dental} \\ \text{MANNER} & : \text{fricative} \\ \text{VOICE} & : + \end{bmatrix} \qquad \begin{bmatrix} \text{PLACE} & : \text{dental} \\ \text{PLACE} & : \text{uvular} \\ \text{VOICE} & : + \end{bmatrix}$$

The first is identical to (26) in the sense that it specifies the same object (the features are read conjunctively). The second however does not specify any sound, since the values given to the same feature are incompatible. (Features must have one and only one value.) We say that the second AVS is **inconsistent**. Notice that AVSs are not sounds, they are just representations thereof, and they may specify the sounds only partly. I add here that some combinations may be formally consistent and yet cannot be instantiated. Here is an example:

(28)
$$\begin{bmatrix} \text{CONSONANTAL} & : - \\ \text{VOICE} & : - \end{bmatrix}$$

This is because vowels in English are voiced. There are a few languages, for example Mokilese, which have voiceless vowels. To understand how this is possible think about whispering. Whispering is speaking without the vocal chords vibrating. In effect, whispering is systematically devoicing every sound. That this does not remove the distinction between [p] and [b] shows you that the distinction is not exclusively a voicing contrast! One additional difference is that the lip tension is higher in [p].

The following is however illegitimate because it gives a value to PLACE that is outside of its value range.

(29)
$$\begin{bmatrix} \text{PLACE} & : \text{fricative} \\ \text{VOICE} & : + \end{bmatrix}$$

There is a number of arguments that show that features exist. First and foremost the features encode a certain linguistic reality; the features that we have spoken about so far have phonetic content. They speak about articulatory properties. It so happens that many rules can be motivated from the fact that the vocal tract has certain properties. For example, in German the final consonants of words (to be exact, of syllables) are all voiceless (see the discussion on Page 49). This is so even when there is reason to believe that the consonant in question has been obtained from a voiced consonant. Thus, one proposes a rule of devoicing for German. However, it would be unexpected if this rule would turn [g] into [t]. We would rather expect the rule to turn [g] into [k], [b] into [p] and [d] into [t]. The questions that arise are as follows:

① Why is it that we expect matters to be this way?

② How can we account for the change?

The first question is answered as follows: the underlying rule is not a rule that operates with a lookup table, showing us what consonant is changed into what other consonant. Rather, it is encoded as a rule that says: simply remove the voicing. For this to make sense we need to be able to independently control voicing. This is clearly the case. However, it is one thing to observe that this is technically possible and another to show that this is effectively the rule that speakers use. One way to check that this is effectively the rule is to make Germans speak a different language. The new language will have new sounds, but we shall observe Germans still devoice them at the end of the word. (You can hear them do this in English, for example. The prediction is this that—if they can at all produce these sounds—at the end of a word [ð] will come out as [θ].) Moreover, they will not randomly choose a devoiced consonant but will simply pick the appropriate voiceless counterpart.

Ideally, we wish to write the rule of devoicing in the following way.

(30) $$\begin{bmatrix} \textsc{consonantal} : + \\ \textsc{voice} : + \end{bmatrix} \rightarrow \begin{bmatrix} \textsc{consonantal} : + \\ \textsc{voice} : - \end{bmatrix} \; / \; \_\_\#$$

It will turn out that this can indeed be done (the next chapter provides the details of this). This says that a consonant becomes devoiced at the end of a word. The part before the arrow specifies the situation before the rule applies; the part to the right and before the slash show us how it looks after the application of the rule. The

part after the slash shows in what context the rule may be applied. The underscore shows where the left part of the rule must be situated (and where the right part will be substituted in its place). Here, it says: it must occur right before #, which signals the end of a word. The way this rule operates needs to be explained. The German word /grob/ is pronounced [gʏo:p] (the colon indicates a long vowel; ʏ is a voiced velar fricative, the fricative equivalent of g). The letter /b/ however indicates an underlying [b]. Thus we expect this to be an instance of devoicing. So let's look at [b]:

(31)
$$\begin{bmatrix} \text{CONSONANTAL} & : & - \\ \text{VOICE} & : & + \\ \text{PLACE} & : & \textit{bilabial} \\ \text{MANNER} & : & \textit{stop} \end{bmatrix}$$

As the sound occurs immediately before #, the rule applies. When it applies, it matches the left hand side against the AVS and replaces that part with the right hand side of the rule; whatever is not matched *remains the same*.

(32)
$$\begin{bmatrix} \text{CONSONANTAL} & : & - \\ \text{VOICE} & : & + \\ \text{PLACE} & : & \textit{bilabial} \\ \text{MANNER} & : & \textit{stop} \end{bmatrix} \rightarrow \begin{bmatrix} \text{CONSONANTAL} & : & - \\ \text{VOICE} & : & - \\ \text{PLACE} & : & \textit{bilabial} \\ \text{MANNER} & : & \textit{stop} \end{bmatrix}$$

Thus, the resulting sound is indeed [p]. You may experiment with other AVS to see that the rule really operates as expected. Notice that the rule contains [CONSONANTAL : +] on its left but does not change it. However, you cannot simply eliminate it. The resulting rule would be different:

(33) $\left[\text{VOICE} : + \right] \rightarrow \left[\text{VOICE} : - \right] \; / \; \_\_\#$

This rule would apply to vowels and produce voiceless vowels. Since German does not have such vowels, the rule would clash with the constraints of German phonology. More importantly, it would devoice every word final vowel and thus—wrongly—predict that German has no word final vowels (counterexample: /Oma/ [oma] 'grandmother').

Suppose that you have to say this without features. It is not enough to say that the voiced consonants are transformed into the voiceless ones; we need to know which voiceless consonant will replace which voiced consonant. The tie between [p] and [b], between [k] and [g] and so on needs to be established. Because features have an independent motivation the correspondence is specified uniformly

for all sounds ('voice' refers to the fact whether or not the vocal cords vibrate). As will be noted throughout this course, some rules are not really specific to one language but a whole group of them (final devoicing is a case in point). This seems to be contradictory, because the rules are stated using phonemes, and phonemes are language dependent, as we have seen. However, this need not be what is in fact going on. The fact that language has a contrast between voiced and voiceless is independent of the exact specification of what counts, say, as a voiced bilabial stop as opposed to a voiceless bilabial stop. Important is that the contrast exists and is one of voicing.

For example, Hungarian, Turkish and Finnish both have a rule called vowel harmony. Modulo some difficulties all rules agree that words cannot both contain a back vowel and a front vowel. On the other hand, the front close-mid rounded vowel of Finnish (written /ö/) is pronounced with more lip rounding than the Hungarian one (also written /ö/). Nevertheless, both languages systematically oppose /ö/ with /o/, which differs in the position of the tongue body (close-mid back rounded vowel). The situation is complicated through the fact that Hungarian long and short vowels do not only contrast in length but also in a feature that is called **tension**. Finnish /ö/ is tense even when short, while in Hungarian it is **lax** (which means less rounded and less close). However, even if short and long vowels behave in this way, and even if back and front vowels are different across these languages, there is good reason to believe that the contrast is between 'front' and 'back', no matter *what else is involved*. Thus, among the many parameters that define the actual sounds languages decide to systematically encode only a limited set (which is phonologically relevant and on which the rules operate) even though one still needs to fill in details as for the exact nature of the sounds. Precisely this is the task of **realization rules**. These are the rules that make the transition from phonemes to sounds. They will be discussed in the next lecture.

Natural Classes

Suppose we fix the set of attributes and values for a language. On the basis of this classification we can define the following.

Definition 7 (Natural Class; Provisional) *A **natural class** of sounds is a set of sounds that can be specified by a single AVS.*

This is still not as clear as I would like this to be. First, we need to something about the classification system used above. Let P be our set of phonemes. Recall that this set is in a way abstract. It is not possible to compare phonemes across languages, except by looking at their possible realisations (which are then sounds). We then define (using our theoretical or pretheoretical insights) some features and potential values for them. Next we specify which sounds have which value to which attribute. That is to say, for each attribute A and value v there is a set of phonemes written $[A : v]$ (which is therefore a subset of P). Its members are the phonemes that are said to have the value v to the attribute A. This set must be given for each such legitimate pair. However, not every such system is appropriate. Rather, we require in addition that the following holds.

① For each phoneme p, and each feature A there is a value v such that $p \in [A : v]$, that is to, p has A-value v.

② If $v \neq v'$ then $[A : v] \cap [A : v'] = \varnothing$. In other words: the value of attribute for a given sound is unique.

③ For every two different phonemes p, p' there is a feature A and values v, v' such that $v \neq v'$ and $p \in [A : v]$ and $p' \in [A : v']$.

If these postulates are met we speak of a **classification system** for P. The last condition is especially important. It says that the classification system must be exhaustive. If two phonemes are different we ought to find something that sets it apart from the other phonemes. This means among other that for each p the singleton $\{p\}$ will be a natural class.

First, notice that we require each sound to have a value for a given feature. This is a convenient requirement because it eliminates some fuzziness in the presentation. You will notice, for example, that vowels are classed along totally different lines as consonants. So, is it appropriate to say, for example, that vowels should have some value to MANNER? Suppose we do not really want that. Then a way around this is to add a specific value to the attribute, call it $\star$, and then declare that all vowels have this value. This 'value' is not a value in the intended sense. But to openly declare that vowels have the 'non value' helps us be clear about our assumptions.

Definition 8 (Natural Class) *Let S be a classification system for P. A subset U*

*of P is **natural in** S if and only if it is an intersection of sets of the form $[A : v]$ for some attribute and some legitimate value.*

I shall draw a few conclusions from this.

1. The set P is natural.

2. For every $p \in P$, $\{p\}$ is natural.

3. If P has at least two members, $\emptyset$ is natural.

To show the first, an intersection of no subsets of P is defined to be identical to P, so that is why P is natural. To show the second, let H be the intersection of all sets $[A : v]$ that contain p. I claim that $H = \{p\}$. For let $p' \neq p$. Then there are A, v, and v' such that $v \neq v'$ $p \in [A : v]$ and $p' \in [A : v']$. However, $p' \notin [A : v']$, since the sets are disjoint. So, $p' \notin H$. Finally, for the third, let there be at least two phonemes, p and p'. Then there are A, v and v' such that $p \in [A : v]$, $p' \in [A : v']$ and $v \neq v'$. Then $[A : v] \cap [A : v'] = \emptyset$ is natural.

The Classification System of English Consonants

I shall indincate now how 9 establishes a classification system and how it is written down in attribute value notation. To make matter simple, we concentrate on the consonants. There are then three attributes: PLACE, MANNER, and `voice`. We assume that the features have the following values:

(34)

PLACE*bilabial*, *labiodental*, *dental*, *alveolar*, *palatoalveolar*, *palatal*, *velar*, *glottal*MANNER

The sounds with a given place features are listed in the columns, and can be read off the table. However, I shall give them here for convenience:

(35)
$$\begin{aligned}
[\textsc{place} : \textit{bilabial}] &= \{/\mathrm{p}/, /\mathrm{b}/, /\mathrm{m}/, \mathrm{w}/\} \\
[\textsc{place} : \textit{labiodental}] &= \{/\mathrm{f}/, /\mathrm{v}/\} \\
[\textsc{place} : \textit{dental}] &= \{/\theta/, /\eth/\} \\
[\textsc{place} : \textit{alveolar}] &= \{/\mathrm{t}/, /\mathrm{d}/, /\mathrm{s}/, \mathrm{z}/, /\mathrm{n}/, /\mathrm{l}/, /\text{ɹ}/\} \\
[\textsc{place} : \textit{palatoalveolar}] &= \{/\text{tʃ}/, /\text{dʒ}/, /\text{ʃ}/, /\text{ʒ}/\} \\
[\textsc{place} : \textit{palatal}] &= \{/\mathrm{j}/\} \\
[\textsc{place} : \textit{velar}] &= \{/\mathrm{k}/, /\mathrm{g}/, /\text{ŋ}/\} \\
[\textsc{place} : \textit{glottal}] &= \{/\mathrm{h}/\}
\end{aligned}$$

The manner feature is encoded in the row labels.

(36)
$$\begin{aligned}
[\textsc{manner} : \textit{stop}] &= \{/\mathrm{p}/, /\mathrm{b}/, \mathrm{t}/, /\mathrm{d}/, /\text{tʃ}/, /\text{dʒ}/, /\mathrm{k}/, /\mathrm{g}/\} \\
[\textsc{manner} : \textit{fricative}] &= \{/\mathrm{f}/, /\mathrm{v}/, /\theta/, /\eth/, /\mathrm{s}/, /\mathrm{z}/, /\text{ʃ}/, /\text{ʒ}/\} \\
[\textsc{manner} : \textit{nasal}] &= \{/\mathrm{m}/, /\mathrm{n}/, /\text{ŋ}/\} \\
[\textsc{manner} : \textit{l approx}] &= \{/\mathrm{l}/\} \\
[\textsc{manner} : \textit{c approx}] &= \{/\mathrm{w}/, /\mathrm{j}/, /\text{ɹ}/\}
\end{aligned}$$

(37)
$$\begin{aligned}
[\textsc{voice} : +] =\{&/\mathrm{b}/, /\mathrm{d}/, /\mathrm{g}/, /\text{dʒ}/, /\mathrm{v}/, /\eth/, /\mathrm{z}/, /\text{ʒ}/, /\mathrm{m}/, /\mathrm{n}/, /\text{ŋ}/, \\
&/\mathrm{w}/, /\mathrm{l}/, /\text{ɹ}/, /\mathrm{j}/\} \\
[\textsc{voice} : -] =\{&/\mathrm{p}/, /\mathrm{t}/, /\mathrm{k}/, /\text{tʃ}/, /\mathrm{f}/, /\theta/, /\mathrm{s}/, /\text{ʃ}/\}
\end{aligned}$$

So, one may check, for example, that each sound is uniquely characterized by the values to the attributes; /p/ has value *bilabial* for PLACE, *stop* for MANNER and − for VOICE. So we have

(38)
$$\begin{bmatrix} \textsc{place} & : \textit{bilabial} \\ \textsc{manner} & : \textit{stop} \\ \textsc{voice} & : - \end{bmatrix} = \{/p/\}$$

If we drop any of the three specifications we get a lager class. This is not always so. For example, English has only one palatal phoneme, /j/. Hence we have

(39)
$$\begin{bmatrix} \textsc{place} : \textit{palatal} \end{bmatrix} = \begin{bmatrix} \textsc{place} & : & \textit{palatal} \\ \textsc{manner} & : & \textit{c approximant} \end{bmatrix} = \{/j/\}$$

I note here that although a similar system for vowels can be given, I do not include it here. This has two reasons. One is that it makes the calculations even more difficult. The other is that it turns out that the classification of vowels proceeds along different features. We have, for example, the feature ROUNDED, but do not classifiy the consonants according to feature. If we are strict about the execution of the classification we should then also say which of the consonants are rounded and which ones are not.

Notice also that the system of classification is motivated from the phonetics but not entirely. There are interesting questions that appear. For example, the phoneme /ɹ/ is classified as voiced. However, at closer look it turns out that the phoneme contains both the voiced and the voiceless variant, written [ɹ̥]. (The pronunciation of /bridge/ involves the voiced [ɹ], the pronunciation of /trust/ the voiceless [ɹ̥].) In broad transcription (which is essentially phonemic) one writes [ɹ] regardless. But we need to understand that the term 'voiced' does not have its usual phonetic meaning. The policy on notation is not always consistently adhered to; the symbolism encourages confusing [ɹ] and [ɹ̥], though if one reads the IPA manual it states that [ɹ] signifies only the voiced and not the voiceless approximant. So, technically, the left part of that cell should contain the symbol [ɹ̥].

Binarism

The preceding section must have cautioned you to think that for a given set of phonemes there must be several possible classification systems. Indeed, not only are there several conceivable classification systems, phonologists are divided in the issue of which one to actually use.

There is a never concluded debate on the thesis of **binarism** of features. **Binarism** is the thesis that features have just two values: + and −. In this case, also an alternative notation is used; instead of [att : +] one writes [+att] (for example, [+voiced]) and instead of [att : -] one writes [−att] (for example [−voiced]). I shall use this notation as well.

Although any feature system can be reconstructed using binary valued features, the two systems are not equivalent, since they define different natural classes.

Consider, by way of example, the sounds [p], [t] and [k]. They are distinct

only in the place of articulation (bilabial versus alveolar versus velar). The only natural classes are: the empty one, the singletons or the one containing all three. If we assume a division into binary features, either [p] and [t] or [p] and [k] or [t] and [k] must form a natural class in addition. This is so since binary features can cut a set only into two parts. If your set has three members, you can single out a given member by two cuts and only sometimes by one. So you need two binary features to distinguish the three from each other. But which ones do we take? In the present case we have a choice of [+labial], [+dental] or [+velar]. The first cuts $\{[p], [t], [k]\}$ into $\{[p]\}$ and $\{[t], [k]\}$; the second cuts it into $\{[t]\}$ and $\{[p], [k]\}$ and the third into $\{[k]\}$, and $\{[t], [p]\}$. Any two of these features allow to have the singleton sets as natural classes. If you have only two features then there is a two element subset that is not a natural class (this is an exercise).

The choice between the various feature bases is not easy and hotly disputed. It depends on the way the rules of the language can be simplified which classification is used. But if that is so, the idea becomes problematic as a foundational tool. It is perhaps better not to enforce binarism.

In structuralism, the following distinction has been made: a distinction or **opposition** is **equipollent** or **privative**. To begin with the latter: the distinction between a and b is **privative** if (i) a has something that b does not have or (ii) b has something that a does not have. In case that (i) obtains, we call a **marked** (in opposition to b) and in case that (ii) obtains we call b **marked**. An equipollent distinction is one that is not of this kind. (So, neither a nor b can be said to be marked.) We have suggested above that the distinctions between speech sounds is always equipollent; for example, [p] and [b] are distinct because the one has the feature [−voiced] the other has the feature [+voiced]. Since we have both features, by the rules of attribute value structures, a sound must have one of them exactly if it does not have the other. There is thus a complete symmetry. If we want to turn this into a privative opposition, we have to explicitly mark one of the features against the other. Linguists have instead following another approach. They devised a notational system with just one feature, say, 'voiced'. A sound may either have that feature or not. It is marked precisely when it has the feature, and unmarked otherwise. In such a system, [b] is marked (against [p]), because it has the feature, while [p] is not. Had we chosen instead the feature 'voiceless', [b] would have been unmarked, and [p] marked. In the literature this is sometimes portrayed as features having just one value. This use of language is dangerous and should be avoided. A case of markedness is the pronunciation of [ɹ], where the

default pronunciation is voiced, and the marked one is voiceless. However, this applies to the phonetic level, not the phonemics.

Notes. The book [Lass, 1984] offers a good discussion of the theory and use of features in phonology. Feature systems are subject to big controversy. Roman Jakobson was a great advocate of the idea of binarism, but it seems to often lead to artificial results. [O'Grady *et al.*, 2005] offer a binary system for English. Definition 5 is too strong. Typically, one only has *only if* rather that *if and only if* since there are sound pairs that can be exchanged for each other without necessarily being in a phoneme. However, it is better to use the more stringent version to get an easier feel for this type of definition which is typical for structuralist thinking. Further, what is problematic in this definition is that it does not take into account multiple simultaneous substitution. However, such cases typically are beyond the scope of an introduction.

Phonology II: Realization Rules and Representations

> The central concept of this chapter is that of a *natural class* and of a *rule*. We learn how rules work, and how they can be used to structure linguistic theory.

Determining Natural Classes

Let us start with a simple example to show what is meant by a natural class. Sanskrit had the following obstruents and nasals

(40)

p	pʰ	b	bʰ	m
t	tʰ	d	dʰ	n
ʈ	ʈʰ	ɖ	ɖʰ	ɳ
c	cʰ,	ɟ	ɟʰ	ɲ
k	kʰ	g	gʰ	ŋ

(By the way, if you read the sounds as they appear here, this is exactly the way they are ordered in Sanskrit. The Sanskrit alphabet is much more logically arranged than the Latin alphabet!) To describe these sounds we use the following features and values:

(41)

CONS(ONANTAL) :+
MANNER :*stop*, *fric(ative)*
PLACE :*bilab(ial)*, *dent(al)*, *retro(flex)*, *velar*, *palat(al)*
ASP(IRATED) :+, −
NAS(AL) :+, −
VOICE :+, −

We shall omit the specification 'consonantal' for brevity. Also, we shall omit 'manner' and equate it with 'nasal' (= [NAS : +]).

Here is how the phonemes from the first row are to be represented:

(42)

p	p^h	b
PLACE:bilab	PLACE:bilab	PLACE:bilab
ASP :−	ASP :+	ASP :−
NAS :−	NAS :−	NAS :−
VOICE:−	VOICE:−	VOICE:+

b^h	m
PLACE:bilab	PLACE:bilab
ASP :+	ASP :−
NAS :−	NAS :+
VOICE:+	VOICE:+

Let us establish the natural classes. First, each feature (a single pair of an attribute and its value) defines a natural class:

(43)

[PLACE : bilab]	{p, p^h, b, b^h, m}
[PLACE : dental]	{t, t^h, d, d^h, n}
[PLACE : retroflex]	{ʈ, $ʈ^h$, ɖ, $ɖ^h$, ɳ}
[PLACE : palatal]	{c, c^h, ɟ, $ɟ^h$, ɲ}
[PLACE : velar]	{k, k^h, g, g^h, ŋ}
[ASP : +]	{p^h, b^h, t^h, d^h, $ʈ^h$, $ɖ^h$, c^h, $ɟ^h$, k^h, g^h}
[ASP : −]	{p, b, m, t, d, n, ʈ, ɖ, ɳ, c, ɟ, ɲ, k, g, ŋ}
[NAS : +]	{m, n, ɳ, ɲ, ŋ}
[NAS : −]	{p, p^h, b, b^h, t, t^h, d, d^h, ʈ, $ʈ^h$, ɖ, $ɖ^h$, c, c^h, ɟ, $ɟ^h$, k, k^h, g, g^h}
[VOICE : +]	{b, b^h, m, d, d^h, n, ɖ, $ɖ^h$, ɳ, ɟ, $ɟ^h$, ɲ, g, g^h, ŋ}
[VOICE : −]	{p, p^h, t, t^h, ʈ, $ʈ^h$, c, c^h, k, k^h}

All other classes are intersections of the ones above. For example, the class of phonemes that are both retroflex and voiced can be formed by looking up the class of retroflex phonemes, the class of voiced phonemes and then taking the intersection:

(44) $\{ʈ, ʈ^h, ɖ, ɖ^h, ɳ\} \cap \{b, b^h, m, d, d^h, n, ɖ, ɖ^h, ɳ, g, g^h, ŋ\} = \{ɖ, ɖ^h, ɳ\}$

Basically, there are at most 6×3×3×3 = 162 (!) different natural classes. How did I get that number? For each attribute you can either give a value, or leave the value undecided. That gives 6 choices for place, 3 for nasality, 3 for voice, and three

for aspiratedness. In fact, nasality does not go together with aspiratedness or with being voiceless, so some combinations do not exist. All the phonemes constitute a natural class of their own. This is so since the system is set up this way: each phoneme has a unique characteristic set of features. Obviously, things have to be this way, since the representation has to be able to represent each phoneme by itself. Now, 162 might strike you as a large number. However, as there are 25 phonemes there are $2^{25} = 33,554,432$ different sets of phonemes (if you cannot be bothered about the maths here, just believe me)! So a randomly selected set of phonemes has a chance of about 0.00005, or 0.005 percent of being natural!

How can we decide whether a given set of phonemes is natural? First method: try all possibilities. This might be a little slow, but you will soon find some shortcuts. Second method. You have to find a description that fits all and only the sounds in your set. It has to be of the form 'has this feature, this feature and this feature'—so no disjunction, no negation. You take two sounds and look at the attributes on which they differ. Obviously, these ones you cannot use for the description. After you have established the set of attributes (and values) on which all agree, determine the set that is described by this combination. If it is your set, that set is natural. Otherwise not. Take the set {m, p^h, ɖ}.

(45)

m	p^h	ɖ
PLACE:bilab	PLACE:bilab	PLACE:retro
ASP :−	ASP :+	ASP :−
NAS :+	NAS :−	NAS :−
VOICE:+	VOICE:−	VOICE:+

The first is nasal, but the others are not. So the description cannot involve nasality. The second is voiceless, the others are voicedness. The description cannot involve voicing. Similarly for aspiratedness and place. It means that the smallest natural class that contains this set is—the entire set of them. (Yes, the entire set of sounds is a natural class. Why? Well, no condition is also a condition. Technically, it corresponds to the empty AVS, which is denoted by [ ]. Nothing is in there, so any phoneme fits that description.)

The example was in some sense easy: there was no feature that the phonemes shared. However, the set of all consonants is also of that kind and natural, so that cannot be a criterion. To see another example, look at the set {[p], [p^h], [b]}. Agreeing features are blue, disagreeing features red (I have marked the agreeing

features additionally with ☞):

(46)

	p	p^h	b
☞	PLACE:bilab	PLACE:bilab	PLACE:bilab
	ASP :−	ASP :+	ASP :−
☞	NAS :−	NAS :−	NAS :−
	VOICE:−	VOICE:−	VOICE:+

It seems that we have found a natural class. However, when we extract the two agreeing features and calculate the class we get the class of bilabial stops, which is {[p], [p^h], [b], [b^h]}. This class contains one more phoneme. So the original class is not natural.

Now, why are natural classes important and how do we use them? Let us look at a phenomenon of Sanskrit (and not only Sanskrit) called **sandhi**. Sanskrit words may end in the following of the above: p, m, t, n, ʈ, k, and ŋ. This consonant changes depending on the initial phoneme of the following word. Sometimes the initial phoneme also changes. An example is /tat ɟa ri:ram/, which becomes /tac c^hari:ram/. We shall concentrate here on the more common effect that the last phoneme changes. The books give you the following look-up table:

(47)

	word ends in:						
	k	ʈ	t	p	ŋ	n	ṃ
p, p^h	k	ʈ	t	p	ŋ	n	ṃ
b, b^h	g	ɖ	d	b	ŋ	n	ṃ
t, t^h	k	ʈ	t	p	ŋ	n	ṃ
d, d^h	g	ɖ	d	b	ŋ	n	ṃ
ʈ, $ʈ^h$	k	ʈ	ʈ	p	ŋ	ṃʂ	ṃ
ɖ, $ɖ^h$	g	ɖ	ɖ	b	ŋ	ɳ	ṃ
c, c^h	k	ʈ	c	p	ŋ	ṃɟ	ṃ
ɟ, $ɟ^h$	g	ɖ	ɟ	b	ŋ	ɲ	ṃ
k, k^h	k	ʈ	t	p	ŋ	n	ṃ
g, g^h	g	ɖ	d	b	ŋ	n	ṃ
n, m	ŋ	ɳ	n	m	ŋ	n	ṃ

[ʂ] is a voiceless retroflex fricative, [ɟ] is a voiceless palatal fricative. There is one symbol that needs explanation. ṃ denotes a nasalisation of the preceding vowel (thus it is not a phoneme in the strict sense—see below on a similar issue concerning vowel change in English). Despite its nonsegmental character I take it here at face value and pretend it is a nasal.

We can capture the effect of Sandhi also in terms of *rules*. A rule is a statement of the following form:

(48)
X	→ Y	/	C____D
Input	→ Output		Context

For the understanding of rules is important to stress that they represent a step in a sequence of *actions*. In the rule given above the action is to replace the input (X) by the output (Y) in the given context. If the context is arbitrary, nothing is written. The simplest kind of rule, no context given, is exemplified by this rule:

(49) a → b

This rule replaces a by b wherever it occurs. Thus, suppose the input is

(50) The visitors to Alhambra are from abroad.

then the output is

(51) The visitors to Alhbmbrb bre from bbrobd.

Notice that A, being a different *character* is not affected by the rule. Also, b is not replaced by a, since the rule operates only in one direction, from left to right.

If we want to restrict the action of a rule to occurrences of letters at certain places only, we can use a context condition. It has the form C__D. This says the following: if the specified occurrence is between C (on its left) and D (on its right) then it may be replaced, otherwise it remains the same. Notice that this is just a different way of writing the following rule:

(52) CXD → CYD

I give an example. The rules of spelling require that one uses capital letters after a period (that's simplifying matters a bit since the period must end a sentence). Since the period is followed by a blank—written ␣, in fact, maybe there are several blanks, but let's ignore that too—the context is .␣__. D is omitted since we place no condition on what is on the right of the input. So we can formulate this for the letter a as

(53) a → A/.␣__

This rule says that `a` is changed to `A` if it is preceded by a blank which in turn is preceded by a period. Alternatively we could use

(54) `.␣a` → `.␣A`

Let's return to Sandhi. As we have done in the previous chapter, a word boundary is denoted by #. This is not a printed character, and may in fact come out in different ways (look at the way it comes out before punctuation marks). Also, since we are mostly dealing with spoken language, there is no real meaning in counting blanks, so we leave the precise nature of blank unspecified. Suppose we want to write rules that capture Sandhi. Each entry of the table presents one individual rule. For example, if te previous word ends in /k/ and the following word begins with /b/, then rather than the sequence /k#b/ we will see the sequence /g#b/.

Thus we find that Sandhi is among many others the rule

(55) /k#b/ → /g#b/

We can reformulate this into

(56) k → g/____#b

To be precise, it is perhaps useful to think that Sandhi also erases the word boundary, so we should write really the rule as follows.

(57) k# → g/____b

However, once we understand where I have simplified matters, we can move on to the essential question, namely, how to best represent the Sandhi using abstract rules.

If you do the calculations you will find that this table has 154 cases (and I haven't even given you the whole table). In 60 cases an actualu change occurs. It is true that if there is no change, no rule needs to be written, unless you consider the fact that in all these cases the word boundary is erased. However, in any case this is unsatisfactory. What we want is to represent the regularities directly in our rules.

There is a way to do this. Notice for example the behaviour of /k/, /t/ and /p/. If the consonant of the following word is voiced, they become voiced, too. If the

consonant is voiceless, they remain voiceless. This can be encoded into a single rule. Observe that the last consonant of the preceding word is a stop; and the first consonant of the following word is a stop, too. Using our representations, we can capture the content of all of these rules as follows.

$$(58) \qquad \begin{bmatrix} \text{VOICE}{:}- \\ \text{NAS} \ :- \end{bmatrix} \rightarrow \begin{bmatrix} \text{VOICE}{:}+ \\ \text{NAS} \ :- \end{bmatrix} / \underline{\quad\quad} \# \begin{bmatrix} \text{VOICE}{:}+ \\ \text{NAS} \ :- \end{bmatrix}$$

As we explained in the previous lecture, this is to be read as follows: given a phoneme, there are three cases. (Case 1) The phoneme does not match the left hand side (it is either voiced or a nasal); then no change. (Case 2) The phoneme matches the left hand side but is not in the context required by the rule (does not precede a voiceless stop). Then no change. (Case 3) The phoneme matches the left hand side and is in the required context. In this case, all features that are not mentioned in the rule will be left unchanged. This is the way we achieve generality. I will return below to the issue of [t] shortly. (Notice that every consonant which is not a nasal is automatically a stop in this set. This is not true in Sanskrit, but we are working with a reduced set of sounds here.)

I remark here that the rule above is also written as follows.

$$(59) \qquad \begin{bmatrix} \text{NAS}{:}- \end{bmatrix} \rightarrow \begin{bmatrix} \text{VOICE}{:}+ \\ \text{NAS} \ :- \end{bmatrix} / \underline{\quad\quad} \# \begin{bmatrix} \text{VOICE}{:}+ \\ \text{NAS} \ :- \end{bmatrix}$$

The omission of the voicing specification means that the rule applies to any feature value. Notice that on the right hand side we find the pair [VOICE : +]. This means that whatever voice feature the original sound had, it is *replaced* by [VOICE : +].

Next, if the consonant of the following word is a nasal, the preceding consonant becomes a nasal. The choice of the nasal is completely determined by the place of articulation of the original stop, which is not changed. So, predictably, /p/ is changed to /m/, /t/ to /n/, and so on.

$$(60) \qquad \begin{bmatrix} \text{NAS}{:}- \end{bmatrix} \rightarrow \begin{bmatrix} \text{VOICE}{:}+ \\ \text{ASP} \ :- \\ \text{NAS} \ :+ \end{bmatrix} / \underline{\quad\quad} \# \begin{bmatrix} \text{NAS}{:}+ \end{bmatrix}$$

The reason we specified voicing and aspiratedness in the result is that we want the rule to apply to all obstruents. But if they are voiceless and we change only nasality, we get either a voiceless nasal or an aspirated. Neither exists in Sanskrit.

We are left with the case where the preceding word ends in a nasal. The easy cases are /ŋ/ and /m̥/. They never change, and so no rule needs to be written. This leaves us with two cases: /t/ and /n/. Basically, they adapt to the place of articulation of the following consonant provided it is palatal or retroflex. (These are the next door neighbours, phonetically speaking.) However, if the following consonant is voiceless, the nasal changes to a sibilant, and the nasalisation is thrown onto the preceding vowel.

Let's do them in turn. /t/ becomes voiced when the following sound is voiced; we already have a rule that takes care of it. We need to make sure that it is applied, too. (Thus there needs to be a system of scheduling rule applications, a theme to which we shall return in Lecture 6. If we are exact and state that the rules remove the word boundary, however, the rules cannot be applied sequentially, and we need to formulate a single rule doing everything in one step.)

(61) $$\begin{bmatrix} \text{NAS} & :- \\ \text{PLACE} & :\textit{dent} \end{bmatrix} \rightarrow \begin{bmatrix} \text{NAS} & :- \\ \text{PLACE} & :\textit{retro} \end{bmatrix} / \underline{\quad\quad} \# \begin{bmatrix} \text{NAS} & :- \\ \text{PLACE} & :\textit{retro} \end{bmatrix}$$

A similar rule is written for the palatals. It is a matter of taste (and ingenuity in devising new notation) whether one can further reduce these two rules to, say,

(62) $$\begin{bmatrix} \text{NAS} & : & - \\ \text{PLACE} & :\textit{dent} \end{bmatrix} \rightarrow \begin{bmatrix} \text{NAS} & :- \\ \text{PLACE} & :\alpha \end{bmatrix} / \underline{\quad\quad} \# \begin{bmatrix} \text{NAS} & :- \\ \text{PLACE} & :\alpha \end{bmatrix}$$
$$(\alpha \in \{\textit{retro}, \textit{palat}\})$$

Let us finally turn to /n/. Here we have to distinguish two cases: whether the initial consonant is voiced or unvoiced. In the voiced case /n/ assimilates in place:

(63) $$\begin{bmatrix} \text{NAS} & : & + \\ \text{PLACE} & :\textit{alv} \end{bmatrix} \rightarrow \begin{bmatrix} \text{NAS} & :+ \\ \text{PLACE} & :\alpha \end{bmatrix} / \underline{\quad\quad} \# \begin{bmatrix} \text{NAS} & :- \\ \text{PLACE} & :\alpha \end{bmatrix}$$
$$(\alpha \in \{\textit{retro}, \textit{palat}\})$$

If the next consonant is voiceless, we get the sequence /m̥/ plus a fricative whose place matches that of the following consonant. (This is the only occasion where we are putting in the manner feature, since we have to describe the resulting

phoneme.)

(64) $$\begin{bmatrix} \text{NAS} : + \\ \text{PLACE}: alv \end{bmatrix} \rightarrow \underset{\cdot}{\text{m}} \begin{bmatrix} \text{NAS} : - \\ \text{VOICED} : + \\ \text{MANNER}: fric \\ \text{PLACE} : \alpha \end{bmatrix} / \underline{\quad\quad} \# \begin{bmatrix} \text{NAS} : - \\ \text{VOICE}: - \\ \text{PLACE}: \alpha \end{bmatrix}$$

$(\alpha \in \{retro, palat\})$

Thus, we use the rules (58), (60), (62), (63) and (64). Given that the latter three abbreviate two rules each this leaves us with a total of 8 rules as opposed to 60.

Neutralization of Contrast

There are also cases where the phonological rules actually obliterate a phonological contrast (one such case is stop nasalization in Korean). We discuss here a phenomenon called *final devoicing*. In Russian and German, stops become devoiced at the end of a syllable. It is such rules that cannot be formulated in the same way as above, namely as rules of specialization. This is so since they involve two sounds that are not allophones, for example [p] and [m] in Korean or [k] and [g] in German. We shall illustrate the German phenomenon, which actually is rather widespread. The contrast between voiced and voiceless is phonemic:

(65)

`Kasse`	[ˈkasə]	(*cashier*)	:	`Gasse`	[ˈgasə]	(*narrow street*)
`Daten`	[ˈda:tən]	(*data*)	:	`Taten`	[ˈta:tən]	(*deeds*)
`Peter`	[ˈpe:tɛ^{a}]	(*Peter*)	:	`Beter`	[ˈbe:tɛ^{a}]	(*praying person*)

Now look at the following words: `Rad` (*wheel*) and `Rat` (*advice*). They are pronounced alike: [ˈʁa:t]. This is because at the end of the syllable (and so at the end of the word), voiced stops become unvoiced:

(66) $$\begin{bmatrix} +\text{stop} \end{bmatrix} \rightarrow \begin{bmatrix} +\text{stop} \\ -\text{voiced} \end{bmatrix} / \_\_ \#$$

(Here, # symbolizes the word boundary.) So how do we know that the sound that underlies `Rad` is [d] and not [t]? It is because when we form the genitive the [d] actually reappears:

(67) (`des`) `Rades` [ˈʁa:dəs]

The genitive of `Rat` on the other hand is pronounced with [t]:

(68) (`des`) `Rates` [ˈʁa:təs]

This is because the genitive adds an [s] (plus an often optional epenthetic schwa) and this schwa suddenly makes the [t] and [d] nonfinal, so that the rule of devoicing does not apply.

Phonology: Deep and Surface

The fact that rules change representations has led linguists to posit two distinct sublevels. One is the level of **deep phonological representations** and the other is that of **surface phonological representation**. The deep representation is more abstract and more regular. For German, it contains the information about voicing no matter what environment the consonant is in. The surface representation however contains the phonological description of the sounds that actually appear; so it will contain only voiced stops at the end of a syllable. The two representations are linked by rules that have the power of changing the representation. In terms of IPA–symbols, we may picture the change as follows.

(69)
[ˈʁa:d]
↓ (Final Devoicing)
[ˈʁa:t]

However, what we should rather be thinking of is this:

$$\begin{bmatrix} -\text{vowel} \\ +\text{approximant} \\ +\text{velar} \\ +\text{voiced} \end{bmatrix}\begin{bmatrix} +\text{vowel} \\ +\text{open} \\ +\text{front} \\ +\text{long} \end{bmatrix}\begin{bmatrix} -\text{vowel} \\ +\text{stop} \\ +\text{labiodental} \\ +\text{voiced} \end{bmatrix}\#$$

(70)

$$\downarrow$$

$$\begin{bmatrix} -\text{vowel} \\ +\text{approximant} \\ +\text{velar} \\ +\text{voiced} \end{bmatrix}\begin{bmatrix} +\text{vowel} \\ +\text{open} \\ +\text{front} \\ +\text{long} \end{bmatrix}\begin{bmatrix} -\text{vowel} \\ +\text{stop} \\ +\text{labiodental} \\ -\text{voiced} \end{bmatrix}\#$$

(I mention in passing another option: we can deny that the voicing contrast at the end of a syllable is phonological—a voiced stop like [b] is just realized (= put

into sound) in two different ways, depending on the environment. This means that the burden is on the phonology–to–phonetics mapping. However, the evidence seems to be that the syllable final [b] is pronounced just like [p], and so it simply *is* transformed into [p].)

We may in fact view all rules proposed above as rules that go from deep to surface phonological mapping. Some of the rules just *add* features while some of them *change* features. The view that emerges is that deep phonological structure contains the minimum specification necessary to be able to figure out how the object sounds, while preserving the highest degree of abstraction and regularity. For example, strings are formed at deep phonological level by concatenation, while on the surface this might not be so. We have seen that effect with final devoicing. While on the deep level the plural ending (or a suffix like `chen`) are simply added, the fact that a stop might find itself at the end of the word (or syllable) may make it change to something else. The picture is thus the following: the word `Rad` is stored as a sequence of three phonemes, and no word boundary exists because we might decide to add something. However, when we form the singular nominative, suddenly a word boundary gets added, and this is the moment the rule of final devoicing can take effect.

The setup is not without problems. Look at the English word /`mouse`/. Its plural is /`mice`/ (the root vowel changes). How is this change accounted for? Is there are a phonological rule that says that the root vowel is changed? (We consider the diphthong for simplicity to be a sequence of two vowels of which the second is relevant here.) The answer to the second question is negative. Not because such a rule could not be written, but because it would be incredibly special: it would say that the sequence [maʊss#] (with the second 's' coming from the plural!) is to be changed into [maɪs#]. Moreover, we expect that phonological rules can be grounded in the articulatory and perceptive quality of the sounds. There is nothing that suggests why the proposed change is motivated in terms of difficulty of pronunciation. We could equally well expect the form [maʊsəz], which is phonologically well–formed. It just is no plural of the word [maʊs] though English speakers are free to change that. (Indeed, English did possess more irregular plurals. The plural of [bʊk] was once [be:k], the plural of [ˈtunge] was [ˈtungan], and many more. These have been superseded by regular formations.) So, if it is not phonology that causes the change, something else must do that. One approach is to simply list the singular [maʊs] and the plural [maɪs], and no root form. Then [maɪs] is not analyzable into a root plus a plural ending, it is just is a

simple form. Another solution is to say that the root has *two* forms; in the singular we find [maʊs], in the plural [maɪs]. The plural has actually *no* exponent, like the singular. That plural is signaled by the vowel is just a fact of choosing an alternate root. The last approach has the advantage that it does not handle the change in phonology; for we have just argued that it is not phonological in nature.

There is—finally—a third approach. Here the quality of the second half of the diphthong is indeterminate between ʊ and ɪ. It is specified as 'lower high'. If you consult Table 9 you see that there are exactly *two* vowels that fit this description: ʊ and ɪ. So, it is a natural class. Hence we write the representation as follows.

(71)
$$\begin{array}{cccc} \text{m} & \text{a} & \text{ʊ/ɪ} & \text{s} \\ \left[\begin{array}{c}+\text{stop}\\+\text{nasal}\\+\text{bilab}\end{array}\right] & \left[\begin{array}{c}+\text{vowel}\\+\text{low}\\+\text{back}\\-\text{rounded}\end{array}\right] & \left[\begin{array}{c}+\text{vowel}\\+\text{lower high}\end{array}\right] & \left[\begin{array}{c}+\text{fric}\\+\text{alveol}\\-\text{voiced}\end{array}\right] \end{array}$$

This representation leaves enough freedom to fill in either front or back so as to get the plural or the singular form. Notice however that it does not represent a phoneme. In early phonological theory one called these objects **archiphonemes**.

We shall briefly comment on this solution. First, whether or not one wants to use two stems or employ an underspecified sound is a matter of generality. The latter solution is only useful when it covers a number of different cases; in English we have at least /this/:/these/, /woman/:/women/, /foot/:/feet/. In German this change is actually much more widespread (we shall return to that phenomenon). So there is a basis for arguing that we find an archiphoneme here. Second, we still need to identify the representation of the plural. Obviously, the plural is not a sound in itself, it is something that makes another sound become one or another. For example, we can propose a notation $^{\leftarrow}$[+ front] which says the following: go leftward (= backward) and attach yourself to the first possible sound. So, the plural of /mouse/ becomes represented as follows:

(72)
$$\begin{array}{cccc} \text{m} & \text{a} & \text{ʊ/ɪ} & \text{s} \\ \left[\begin{array}{c}+\text{stop}\\+\text{nasal}\\+\text{bilab}\end{array}\right] & \left[\begin{array}{c}+\text{vowel}\\+\text{low}\\+\text{back}\\-\text{rounded}\end{array}\right] & \left[\begin{array}{c}+\text{vowel}\\+\text{lower high}\end{array}\right] & \left[\begin{array}{c}+\text{fric}\\+\text{alveol}\\-\text{voiced}\end{array}\right]^{\leftarrow}[+\text{ front}] \end{array}$$

By convention, this is

(73)
$$\begin{array}{cccc} \text{m} & \text{a} & \text{ɪ} & \text{s} \\ \left[\begin{array}{c}+\text{stop}\\+\text{nasal}\\+\text{bilab}\end{array}\right] & \left[\begin{array}{c}+\text{vowel}\\+\text{low}\\+\text{back}\\-\text{rounded}\end{array}\right] & \left[\begin{array}{c}+\text{vowel}\\+\text{lower high}\\+\text{front}\end{array}\right] & \left[\begin{array}{c}+\text{fric}\\+\text{alveol}\\-\text{voiced}\end{array}\right] \end{array}$$

Note that the singular has the representation ←[+ back]. You may have wondered about the fact that the so-called **root** of a noun like /cat/ was nondistinct from its singular form. This is actually not so. The root does not have a word boundary while the plural does (nothing can attach itself after the plural suffix). Moreover, there is nothing wrong with signs being empty.

This, however, leaves us with a more complex picture of a phonological representation. It contains not only items that define sounds but also funny symbols that act on sounds somewhere else in the representation.

Notes on this section. You may have wondered about the use of different slashes. The slashes are indicators showing us at which level of abstraction we are working. The rule is that /·/ is used for sequences of phonemes while [·] is used for sequences of actual sounds. Thus, one rule operates on the *phonemic level* while another operates on the *phonetic level*. Furthermore, writers distinguish a surface phonological level from a deep phonological level. (There could also be a deep phonetical and surface phonetical level, though that has to my knowledge not been proposed.) In my own experience textbooks do not consistently distinguish between the levels, and for the most part I think that writers themselves do not mentally maintain a map of the levels and their relationships. This therefore is a popular source of mistakes. Since I have not gone into much detail about the distinction between the two, obviously there is no reason to ask you to be consistent in using the slahes. It is however important to point out what people intend to convey when they use them.

In principle there are two levels: phonetic and phonemic. At the phonemic level we write /p/, and at the phonetic level we write [p]. Despite the transparent symbolism there is no connection between the two. Phonemics does not know how things sound, it only sees some 40 or so sounds. Writing /p/ is therefore just a mnemonic aid; we could have used ♠ instead of /p/ (and no slashes, since now it is clear in which level we are!!). Phonemes may be organised using features, but the phonetic content of these features is unknown to phonemics. To connect

the (abstract) phoneme /p/ with some sound we have realisation rules: /p/ → [p] is a rule that says that whatever the phoneme /p/ is, it comes out as [p]. The trouble is that [p] is ambiguous in English. We have a different realisation of /p/ in /spit/ than in /pit/. There are two ways to account for this. One is to write the realisation rules so as to account for this different behaviour (V stands for vowel), for example by writing

(74) /p/ → [p]/s__*V*

Another is to translate /p/ into the 'broad' vowel [p] and the leave the specification to phonetics. Both views have their attraction, though I prefer the first over the second. I have made no commitment here, though to either of the views, so the use of /·/ versus [·] follows no principle and you should not attach too much significance to it.

I also use the slashes for written language. Here they function in a similar way: they quote a stretch of characters, abstracting from their surface realisation. Again, full consistency is hard to maintain (and probably mostly not so necessary). A side effect of the slashes is that strings become better identifiable in running text (and are therefore omitted in actual examples).

With regards to our sketch in Lecture 1 I wish to maintain (for the purpose of these lectures at least) that phonology and phonetics are just one level; and that phonology simply organises the sounds via features and contains the abstract representations, while phonetics contains the actual sounds. This avoids having to deal with too many levels (and possibilities of abstraction).

Phonology III: Syllable Structure, Stress

> Words consist of syllables. The structure of syllables is determined partly by universal and partly by language specific principles. In particular we shall discuss the role of the sonoricity hierarchy in organising the syllabic structure, and the principle of maximal onset.

Utterances are not mere strings of sounds. They are structured into units larger than sounds. A central unit is the *syllable*. Words consist of one or several syllables. Syllables in English begin typically with some consonants. Then comes a vowel or a diphthong and then some consonants again. The first set of consonants is the **onset**, the group of vowels the **nucleus** and the second group of consonants the **coda**. The combination of nucleus and coda is called **rhyme**. So, syllables have the following structure:

(75) [onset [nucleus coda]]

For example, /strength/ is a word consisting of a single syllable:

(76)

[stɹ	[ɛ	ŋθ]]
Onset	Nucleus	Coda
	Rhyme	
	Syllable	

Thus, the onset consists of three consonants: [s], [t] and [ɹ], the nucleus consists just of [ɛ], and the coda has [ŋ] and [θ]. We shall begin with some fundamental principles. The first concerns the structure of the syllable.

> **Syllable Structure I** (*Universal*)
> Every syllable has a nonempty nucleus. Both coda and onset may however be empty.

A syllable which has an empty coda is called **open**. Examples of open syllables are /a/ [e] (onset is empty), /see/ [si] (onset is nonempty). A syllable that is not open is **closed**. Examples are /in/ [ɪn] (onset empty) and /sit/ [sɪt] (onset nonempty). The second principle identifies the nuclei for English.

> **Vowels are Nuclear** (*English*)
> A nucleus can only contain a vowel or a diphthong.

This principle is not fully without problems and that is the reason that we shall look below at a somewhat more general principle. The main problem is the unclear status of some sounds, for example [ɚ]. They are in between a vowel and consonant, and indeed sometimes end up in onset position (see above) and sometimes in nuclear position, for example in /bird/ [bɚd].

The division into syllables is clearly felt by any speaker, although there sometimes is hesitation as to exactly how to divide a word into syllables. Consider the word /atmosphere/. Is the /s/ part of the second syllable or part of the third? The answer is not straightforward. In particular the strident (that is, the sounds [s], [ʃ]) enjoy a special status. Some claim that they are **extrasyllabic** (not part of any syllable at all), some maintain that they are **ambisyllabic** (they belong to *both* syllables). We shall not go into that here.

The existence of rhymes can be attested by looking at verses (which also explains the terminology): words that rhyme do not need to end in the same syllable, they only need to end in the same rhyme: /fun/ – /run/ – /spun/ – /shun/. Also, the coda is the domain of a rule that affects many languages: For example, in English and Hungarian, within the coda the obstruents must either all be voiced or unvoiced; in German and Russian, all obstruents in coda must be voiceless. (Here is an interesting problem caused among other by nasals. Nasals are standardly voiced. Now try to find out what is happening in this case by pronouncing words with a sequence nasal+voiceless stop in coda, such as /hump/, /stunt/, /Frank/.) Germanic verse in the Middle Ages used a rhyming technique where the onsets of the rhyming words had to be the same. (This is also called alliteration. It allowed to rhyme two words of the same stem; German had a lot of Umlaut and ablaut, that is to say, it had a lot of root vowel change making it impossible to use the same word to rhyme with itself (say /run/ – /ran/).) It is worthwhile to remain with the notion of the **domain** of a rule. Many phonological constraints are seen as conditions that concern two adjacent sounds. When these sounds come into contact, they undergo change to a smaller or greater extent, for some sound combinations are better pronounceable than others. We have discussed sandhi at length in Lecture 4. For example, the Latin word /in/ ‘in’ is a verbal prefix, which changes in writing (and therefore in pronunciation) to /im/ when it precedes a labial (/impedire/). Somewhat more radical is the change from [ml] to [mpl] to avoid the awkward combination [ml] (the word /templum/ derives from /temlom/, with /tem/ being the root, meaning ‘to cut’). There is an influential theory in phonology, **autosegmental phonology**, which assumes that

Table 10: The Sonoricity Hierarchy

dark vowels	> mid vowels	> high vowels
[a], [o]	[æ], [œ]	[i], [y]
> r-sounds	> nasals; laterals	> vd. fricatives
[r]; [ɹ]	[m], [n]; [l]	[z], [ʒ]
> vd. plosives	> vl. fricatives	> vl. plosives
[b], [d]	[s], [ʃ]	[p], [t]

phonological features are organized on different scores (**tiers**) and can spread to adjacent segments independently from each other. Think for example of the feature [±voiced]. The condition on the coda in English is expressed by saying that the feature [±voiced] spreads along the coda. Clearly, we cannot allow the feature to spread indiscriminately, otherwise the total utterance is affected. Rather, the spreading is blocked by certain constituent boundaries; these can be the coda, onset, nucleus, rhyme, syllable, foot or the word. To put that on its head: the fact that features are blocked indicates that we are facing a constituent boundary. So, voicing harmony indicates that English has a coda.

The nucleus is the element that bears the stress. We have said that in English it is a vowel, but this applies only to careful speech. In general this need not be so. Consider the standard pronunciation of /beaten/: [ˈbi:tʰn̩] (with a syllabic [n]). For my ears the division is into two syllables: [bi:] and [tʰn̩]. (In German this is certainly so; the verb /retten/ is pronounced [ˈʁɛtʰn̩]. The [n] must therefore occupy the nucleus of the second syllable.) There are more languages like this. (Slavic languages are full of consonant clusters and syllables that do not contain a vowel. Consider the island /Krk/, for example.) In general, phonologists have posited the following conditions on syllable structure. Sounds are divided as follows. The sounds are aligned into a so-called **sonoricity hierarchy**, which is shown in Table 10 (vd. = voiced, vl. = voiceless). The syllable is organized as follows.

Syllable Structure II
Within a syllable the sonoricity strictly increases and then decreases

again. It is highest in the nucleus.

This means that a syllable must contain at least one sound which is at least as sonorous as all the others in the syllable. It is called the **sonoricity peak** and is found in the nucleus. Thus, in the onset consonants must be organized such that the sonority rises, while in the coda it is the reverse. The conditions say nothing about the nucleus. In fact, some diphthongs are increasing ([ɪə] as in the British English pronunciation of /here/) others are decreasing ([aɪ], [oɪ]). This explains why the phonotactic conditions are opposite at the beginning of the syllable than at the end. You can end a syllable in [ɹt], but you cannot begin it that way. You can start a syllable by [tɹ], but you cannot end it that way (if you to make up words with [tɹ], automatically, [ɹ] or even [tɹ] will be counted as part of the following syllable). Let me briefly note why diphthongs are not considered problematic in English: it is maintained that the second part is actually a glide (not a vowel), and so would have to be part of the coda. Thus, /right/ would have the following structure:

(77)

r	a	j	t
O	N	C	C

The sonoricity of [j] is lower than that of [a], so it is not nuclear.

A moment's reflection now shows why the position of stridents is problematic: the sequence [ts] is the only legitimate onset according to the sonoricity hierarchy, [st] is ruled out. Unfortunately, both are attested in English, with [ts] only occurring in non-native words (eg /tse-tse/, /tsunami/). There are phonologists who even believe that [s] is part of its own little structure here ('extrasyllabic'). In fact, there are languages which prohibit this sequence; Spanish is a case in point. Spanish avoids words that start with [st]. Moreover, Spanish speakers like to add [e] in front of words that do and are therefore difficult to pronounce. Rather than say [stɹeɪnʒ] (/strange/) they will say [estɹeɪnʒ]. The effect of this maneuver is that [s] is now part of the coda of an added syllable, and the onset is reduced (in their speech) to just [tɹ], or, in their pronunciation most likely [tr]. Notice that this means that Spanish speakers apply the phonetic rules of Spanish to English, because if they applied the English rules they would still end up with the onset [stɹ] (see below). French is similar, but French speakers are somehow less prone to add the vowel. (French has gone through the following sequence: from [st] to [est] to [et]. Compare the word /étoile/ 'star', which derives from Latin /stella/ 'star'.)

Representing the Syllabification

The division of words into syllables is called **syllabification**. In written language the syllable boundaries are not marked, so words are not explicitly syllabified. We only see where words begin and end. The question is: does this hold also for the representations that we need to assume, for example, in the mental lexicon of a speaker? There is evidence that almost no information about the syllable boundaries is written into the mental lexicon. One reason is that the status of sounds at the boundary change as soon as new material comes in. Let us look at the plural /cats/ of /cat/. The plural marker is /s/, and it is added at the end. Let's suppose the division into syllables was already given in the lexicon. Then we would have something like this: /†cat†/, where † marks the syllable boundary. Then the plural will be /†cat†s†/, with the plural 's' forming a syllable of its own. This is not the desired result, although the sonoricity hierarchy would predict exactly that. Let us look harder. We have seen that in German coda consonants devoice. The word /Rad/ is pronounced [ˈʁaːt] as if written /Rat/. Suppose we have added the syllable boundary: /†Rad†/. Now add the genitive /es/ (for which we would also have to assume a syllabification, for example /†es†/, or /es†/). Then we get /†Rat†es†/, which by the rules of German would have to be pronounced [ˈʁaːt ʔəs], with an inserted glottal stop, because German (like English) does not have syllables beginning with a vowel (whereas French does) and prevents that by inserting the glottal stop. (Phonemically there is no glottal stop, however!) In actual fact the pronunciation is [ˈʁaːdəs]. There are two indicators why the sound corresponding to /d/ is now at the beginning of the second syllable: (1) there is no glottal stop following it, (2) it is pronounced with voicing, that is, [d] rather than [t].

We notice right away a consequence of this: syllables are *not* the same as morphemes (see Lecture 7 for a definition). And morphemes neither necessarily are syllables or sequences thereof, nor do syllable boundaries constitute morpheme boundaries. Morphemes can be as small as a single phoneme (like the English plural), a phonemic feature (like the plural of /mouse/), they can just be a stress shift (nominalisation of the verb /protest/ ([proˈtɛst] into /protest/ [ˈprotɛst]) or they can even be phonemically zero. For example, in English, you can turn a noun into a verb (/to google/, /to initial/, /to cable/, /to TA/). The representation does not change at all neither in writing nor in speaking. You just verb it...

Of course it may be suggested that the syllabification *is* explicitly given and changed as more things come in. But this position unnecessarily complicates matters. Syllabification is to a large extent predictable. So there is reason to believe it is largely not stored. It is enough to just insert syllable boundary markers in the lexicon where they are absolutely necessary and leave the insertion of the other boundary markers to be determined later. Another reason is actually that the rate of speech determines the pronunciation, which in turn determines the syllable structure.

Language Particulars

Languages differ in what types of syllables they allow. Thus, not only do they use different sounds they also restrict the possible combinations of these sounds in particular ways. Finnish allows (with few exceptions) only one consonant at the onset of a syllable. Moreover, Finnish words preferably end in a vowel, a nasal or 's'. Japanese syllables are preferably CV (consonant plus vowel). This has effects when these languages adopt new words. Finns for example call the East German car /Trabant/ simply [rabant:i] (with a vowel at the end and a long [t]!). The onset [tr] is simplified to just the [r]. There are also plenty of loanwords: /koulu/ [kɔulu] 'school' has lost the 's', /rahti/ [rahti] 'freight' has lost the 'f', and so on. Notice that it is always the last consonant that wins.

English too has constraints on the structure of syllables. Some are more strict than others. We notice right away that English does allow the onset to contain several consonants, similarly with the coda. However, some sequences are banned. Onsets may not contain a nasal except in first place (exception: [sm] and [sn]). There are some loanwords that break the rule: /mnemonic/ and /tmesis/. The sequence sonorant-obstruent is also banned ([mp], [ɹk], [ɹp], [lt] and so on; this is a direct consequence of the sonoricity hierarchy). Stridents are not found other than in first place; exceptions are [ts] and [ks], found however only in nonnative words. The cluster [ps] is reduced to [s]. It also makes a difference whether the syllable constitutes a word or is peripheral to the word. Typically, inside a word syllables have to be simpler than at the boundary.

Syllabification is based on expectations concerning syllable structure that derive from the well-formedness conditions of syllables. However, these leave room. [ˈpleito] (/Plato/) can be syllabified [ˈplei.to] or [ˈpleit.o]. In both cases the sylla-

bles we get are well formed. However, if a choice exists, then preference is given to creating either open syllables or syllables with onset.

All these principles are variations of the same theme: if there are consonants, languages prefer them to be in the onset rather than the coda.

> **Maximise Onset**
> Put as many consonants into the onset as possible.

The way this operates is as follows. First we need to know what the nucleus of a word is. In English this is easy: vowels and only vowels are nuclear. Thus we can identify the sequences coda+onset but we do not yet know how to divide them into a coda and a subsequent onset. Since a syllable must have a nucleus, a word can only begin with an onset. Therefore, say that a sequence of consonants is a **legitimate onset** of English if there is a word such that the largest stretch of consonants it begins with is that sequence.

> **Legitimate Onsets**
> An sequence of sounds is is a legitimate onset if and only if it is the onset of the first syllable of an English word.

For example, [spɹ] is a legitimate onset of English, since the word /`sprout`/ begins with it. Notice that /`sprout`/ does not show that [sp] is a legitimate onset, since the sequence of consonants that it begins with is [spɹ]. To show that it is legitimate we need to give another word, namely /`spit`/. In conjunction with the fact that vowels and only vowels are nuclear, we can always detect which sounds belong to either a coda or an onset without knowing whether it is onset or coda. However, if a sound if before the first nucleus, it must definitely be part of an onset. In principle, there could be onsets that never show up at the beginning of a word, so the above principle of legitimate onsets in conjunction with Maximise Onsets actually says that this can never happen. And this is now how we can find our syllable boundaries. Given a sequence of consonants that consists a sequence coda+onset, we split it at the earliest point so that the second part is a legitimate onset, that is, an onset at the beginning of a word.

From now on we denote the syllable boundary by a dot, which is inserted into the word as ordinarily spelled. So, we shall write /`in.to`/ to signal that the syllables are /`in`/ and /`to`/. Let us now look at the effect of Maximise Onset. For example, take the words /`restless`/ and /`restricted`/. We do not find /`re.stless`/

nor /res.tless/. The reason is that the there is no English word that begins with /stl/ or /tl/. There are plenty of words that begin in [l], for example /lawn/. Hence, the only possibility is /rest.less/. No other choice is possible. Now look at /restricted/. There are onsets of the form [stɹ] (/strive/). There are no onsets of the form [kt], so the principle of maximal onsets mandates that the syllabification is /re.stric.ted/. Without Maximize Onset, /res.tric.ted/ and /rest.ric.ted/ would also be possible since both coda and onset are legitimate. Indeed, maximal onsets work towards making the preceding syllable open, and to have syllables with onset.

In the ideal case all consonants are onset consonants, so syllables are open. Occasionally this strategy breaks down. For example, /suburb/ is syllabified /sub.urb/. This reflects the composition of the word (from Latin /sub/ 'under' and /urbs/ 'city', so it means something like the lower city; cf. English /downtown/ which has a different meaning!). One can speculate about this case. If it truly forms an exception we expect that if a representation in the mental lexicon will contain a syllably boundary: /sʌb†ɚb/.

Stress

Syllables are not the largest phonological unit. They are themselves organised into larger units. A group of two, sometimes three syllables is called a **foot**. A foot contains one syllable that is more prominent than the others in the same foot. Feet are grouped into higher units, where again one is more prominent than the others, and so on. Prominence is marked by **stress**. There are various ways to give prominence to a syllable. Ancient Greek is said to have marked prominence by pitch (the stressed syllable was about a fifth higher (3/2 of the frequency of an unstressed syllable). Other languages (like German) use **loudness**. Other languages use combination of the two (Swedish). Within a given word there is one syllable that is the most prominent. In IPA it is marked by a preceding ['']. We say that it carries **primary stress**. Languages differ with respect to the placement of primary stress. Finnish and Hungarian place the stress on the first syllable, French on the last. Latin put the stress on the last but one (**penultimate**), if it was long (that is to say, had a long vowel or was closed); otherwise, if was a syllable that preceded it (the **antepenultimate**) then that syllable got primary stress. Thus we had pe.re.gri.nus ('foreign') with stress on the penultimate (gri) since the vowel was long, but in.fe.ri.or with stress on the antepenultimate /fe/ since the /i/

in the penultimate was short. (Obviously, monosyllabic words had the stress on the last syllable.) Sanskrit was said to have free stress, that is to say, stress was free to fall anywhere in the word.

Typically, within a foot the syllables like to follow in a specific pattern. If the foot has two syllables, it consists either in an unstressed followed by a stressed syllable (**iambic metre**), or vice versa (**trochaic metre**). Sometimes a foot carries three syllables (a stressed followed by two unstressed ones, a **dactylus**). So, if the word has more than three syllables, there will be a syllable that is more prominent than its neighbours but not carrying main stress. You may try this with the word /antepenultimate/. You will find that the first syllable is more prominent than the second but less than the fourth. We say that it carries **secondary stress**: [ˌantɛpənˈʌltɪmɛt]. Or [əˌsɪməˈleɪʃn]. The so-called **metrical stress** theory tries to account for stress as follows. The syllables are each represented by a cross (×). This is a Layer 0 stress. Then, in a sequence of cycles, syllables get assigned more crosses. The more crosses, the heavier the syllable. The number of crosses is believed to correspond to the absolute weight of a syllable. So, a word that has a syllable of weight 3 (three crosses) is less prominent than one with a syllable of weight 4. Let's take

(78)

Layer 0	×	×	×	×	×
	ə	sɪ	mə	leɪ	ʃn

We have five syllables. Some syllables get extra crosses. The syllable [sɪ] carries primary stress in /assimilate/. Primary stress is always marked in the lexicon, and this mark tells us that the syllable must get a cross. Further, heavy syllables get an additional cross. A syllable counts as **heavy** in English if it has a coda or a diphthong or long vowel. So, [leɪ] gets an extra cross. [ʃn] is not heavy since the [n] is nuclear. So this is now the situation at Layer 1:

(79)

Layer 1		×		×	
Layer 0	×	×	×	×	×
	ə	sɪ	mə	leɪ	ʃn

Next, the nominalisation introduces main stress on the fourth syllable. So this syllable gets main stress and is therefore assigned another cross. The result is this:

(80)

Layer 2				×	
Layer 1		×		×	
Layer 0	×	×	×	×	×
	ə	sɪ	mə	leɪ	ʃn

If larger units are considered, there are more cycles. The word /maintain/ for example has this representation by itself:

(81)

Layer 2		×
Layer 1	×	×
Layer 0	×	×
	meɪn	teɪn

To get this representation, all we have to know is where the primary stress falls. Both syllables are heavy and therefore get an extra cross at Layer 1. Then the main syllable gets a cross at Layer 2. Now, if the two are put together, a decision must be made which of the two words is more prominent. It is the second, and this is therefore what we get:

(82)

Layer 3							×	
Layer 2		×					×	
Layer 1	×	×		×			×	
Layer 0	×	×	×	×	×	×	×	×
	meɪn	teɪn	ə	sɪ	mə	leɪ	ʃn	

Notice that the stress is governed by a number of heterogeneous factors. The first is the **weight** of the syllable; this decides about Layer 1 stress. Then there is the position of the main stress (which in English must to a large extent be learned—equivalently, it must explicitly be given in the representation, unlike syllable structure). Third, it depends on the way in which the word is embedded into larger units (so syntactic criteria play a role here). Also, morphological formation rules can change the location of the main stress! For example, the suffix (a)tion attracts stress ([ˌkɔmˈbaɪn] and [ˌkɔmbɪˈneɪʃn]) so does the suffix ee (as in /employee/), but /ment/ does not (['gavəˠn] and ['gavəˠnment]). The suffix /al/ does move the accent without attracting it (['ænɛkdot] versus [ˌænɛkˈdotal]).

Finally, we mention a problem concerning the representations that keeps coming up. It is said that certain syllables cannot receive stress because they contain a vowel that cannot be stressed (for example, schwa: [ə]). On the other hand, we can also say that a vowel is schwa because it is unstressed. Take, for example, the pair ['ɹiəlaɪz] and [ˌɹiəlɪˈzeɪʃn]. When the stress shifts, the realisation of /i/ changes. So, is it rather the stress that changes and makes the vowel change quality or does the vowel change and make the stress shift? Often, these problems find no satisfactory answer. In this particular example it seems that the stress shift is first, and it induces the vowel change. It is known that unstressed vowels undergo

reduction in time. The reason why French stress in always on the last syllable is because it inherited the stress pattern from Latin, but the syllables following the stressed syllable eventually got lost. Here the stress was given and it drove the development.

Phonology IV: Rules, Constraints and Optimality Theory

Scheduling Rules

It is important to be clear about a few problems that arise in connections with rules. If you have some string, say /teatime/ and a rule, say t → d, how should you go about applying it? Once, twice, as often as you can? And if you can apply it several times, where do you start? Suppose you can apply the rule only once, then you get either /deatime/ or /teadime/, depending on where you apply it. If you apply if as often as you can, then you can another round and apply the rule. This time the result is the same, /deadime/. This is not always so. Consider the rule a → x/__a; with input /aaa/ you have two choices: you can apply it to the first occurrence of /a/ or to the second. The third is not eligible because of the context restriction. If you apply the rule to the first occurrence you get /xaa/. If you apply it to the second you get /axa/. With /xaa/ you can another round, giving /xxa/; but when you did the second, the rule can no longer apply. Often a proper formulation of the rule itself will be enough to ensure only the correct results will be derived. Sometimes, however, an explicit scheduling must be given, such as: *apply the rule going from left to right as long as you can*, or similar statements. In this lecture we shall not go into the details of this.

Instead we shall turn to another problem, namely the interaction between two rules. I give a very simple example. Suppose we have two rules, R_1: a → b and the other R_2: b → c. Then we can schedule them in many ways.

① R_1 or R_2 can applied once;

② R_1 or R_2 can be applied any number of times;

③ R_1 must be applied before R_2;

④ R_2 must be applied before R_1.

All these choices give different results. We shall exemplify this with an interesting problem, the basic form of the past tense marker in English.

A Problem Concerning the Underlying Form

Regular verbs form the past tense by adding one of the following three suffixes: [t], [d] or [əd]. The choice between these forms is determined by the root.

(83)

[t]		[d]		[əd]	
licked	[lɪkt]	bugged	[bʌgd]	mended	[mɛndəd]
squished	[skwɪʃt]	leaned	[lind]	parted	[pɑɹtəd]
kept	[kɛpt]	buzzed	[bʌzd]	feasted	[fistəd]
laughed	[læft]	played	[pleɪd]	batted	[bætəd]

We ask: what is the source of this difference? Surely, it is possible to say that the past has three different forms and that depending on the verb a different form must be chosen. This, however, misses one important point, namely that the choice of the form is determined solely by the phonological form of the verb and can be motivated by phonological constraints of English. The facts can be summarized as follows.

① [d] is found if the last sound of the verb is voiced but unequal to [d].

② [t] is found if the last sound of the verb is voiceless but unequal to [t].

③ [əd] is found if the verb ends in [d] or [t].

We mention here that it is required that the verb is *regular*. Thus, /run/ and /catch/ are of course not covered by this rule. We may think of the latter as entered in the mental lexicon as unanalysed forms. Thus, rather than seeing /caught/ as a sequence of two forms, namely /catch/ plus some past tense marker, I think of the form entered as a whole, though otherwise functioning in the same way. It is, if you will, an idiom. (Compare this with earlier discussions of the plural formation.)

It seems that the choices can be accounted for solely by applying some general principles. First, notice that in a coda, with the exception of sonorants ([l], [m], [n]), all consonants agree in voicing. An **obstruent** is a consonant that is either a stop (or an affricate), or a fricative.

> **Voice Agreement Principle**
> Adjacent obstruent sequences must either be both [VOICE : +] or both [VOICE : −] at the end of a word.

(The principle is less general than possible: the constraint is valid not only at the end of a word but in any coda.) This goes half way in explaining the choice of the suffix form. It tells us why we see [d] after voiced consonants. But it does not tell us why it is that we get [lind] rather than [lint], because either of them is legitimate according to this principle. Furthermore, we do not know why we find the inserted schwa. The latter can be explained as follows: suppose there was no schwa. Then the present and the past forms would sound alike (/mendd/ would be [mɛnd]). Languages try to avoid double consonants (although they never completely manage), and English employs the strategy to insert schwa also in the plural. We find [bʌsəz] /busses/ (or /buses/), not [bʌsː] (/buss/, with a long /s/). (Another popular strategy is **haplology**, the dropping of one of the consonants.)

It is possible to recruit the Voice Agreement Principle if we assume that there is just a single form not three, and that variants arise only as the result of a repair. The repair is performed by applying some rules. Various analyses are possible.

Analysis 1. We assume that the underlying form is [d]. There is a rule that devoices [d] right after a voiceless obstruent. There is a second rule which inserts a schwa right before [d]. For the purpose of the definition of the rules, two consonsants are called **similar** if they differ at most in the voicing feature (for example, [t] is similar to both [t] and [d], but nothing else).

(84a) $$\begin{bmatrix} +\text{voice} \\ +\text{obstruent} \end{bmatrix} \rightarrow \begin{bmatrix} -\text{voice} \\ +\text{obstruent} \end{bmatrix} \quad /[-\text{voice}]\_\_\#$$

(84b) $$\varnothing \rightarrow [\text{ə}] \quad /\text{C}\_\_\text{C}' \quad (\text{C and C}' \text{ similar})$$

The symbol ∅ denotes the empty string. The first rule is actually two rules in our feature system. I reproduce here the reproper formulation:

(85) $$\begin{bmatrix} \text{voice} : + \\ \text{manner:}\textit{stop} \end{bmatrix} \rightarrow \begin{bmatrix} \text{voice} : - \\ \text{manner:}\textit{stop} \end{bmatrix} /[-\text{voice}]\_\_\#$$
$$\begin{bmatrix} \text{voice} : + \\ \text{manner:}\textit{fricative} \end{bmatrix} \rightarrow \begin{bmatrix} \text{voice} : - \\ \text{manner:}\textit{fricative} \end{bmatrix} /[-\text{voice}]\_\_\#$$

The second rule effectively says that it is legal to *insert* schwa anywhere between similar consonants. Since we have two rules, there is a choice as to which one shall be applied first. We shall first schedule (84a) before (84b). This means that

the rule (84a) is applied to the original form F_0, giving us an output form F_1, and then we apply rule (84b) to get F_2. Each rule applies only once, so the output is F_2. This gives the following result:

	root	/bʌgd/	/lɪkd/	/mɛndd/	/staɹtd/
(86)	(84b)	/bʌgd/	/lɪkd/	/mɛndəd/	/staɹtəd/
	(84a)	/bʌgd/	/lɪkt/	/mɛndəd/	/staɹtəd/

Notice that when we say that a rule does not apply this does not mean that no output is generated. It means that the form is left *unchanged* by the rule. Textbooks sometimes indicate this by a hyphen as if to say that there is not output. But there is an output, it just is the same as the input. So I do not follow the practice. You may figure out for yourselves in which cases this has happened! Notice, too, that sometimes rules do apply and do not change anything.

Now, suppose we had instead scheduled (84a) before (84b). Then this would be the outcome:

	root	/bʌgd/	/lɪkd/	/mɛndd/	/staɹtd/
(87)	(84a)	/bʌgd/	/lɪkt/	/mɛndd/	/staɹtt/
	(84b)	/bʌgd/	/lɪkt/	/mɛndəd/	/staɹtət/

If the last consonant is /t/, the rule (84a) would first assimilate the past tense marker, and we get the suffix /ət/, contrary to fact. Thus, the order in which the rules apply is relevant here. There is, however, nothing intrinsic in the system of the rules that tells us in which order they have to apply. This has to be *stipulated*. In the present analysis we see that the second ordering always gives us an output form, but sometimes the wrong one.

Analysis 2. The underlying form is assumed to be [t]. In place of (84a) there now is a rule that voices [t] right after a voiced obstruent or a vowel. There is a second rule which inserts a schwa right before [d] or [t].

(88a) $\begin{bmatrix} -\text{voice} \\ +\text{obstruent} \end{bmatrix} \rightarrow \begin{bmatrix} \text{voice} : + \\ \text{MANNER:stop} \end{bmatrix}$ /[+voice]__#

(88b) $\varnothing \rightarrow$[ə] /C__C′
(C and C′ similar)

	root	/bʌgt/	/lɪkt/	/mɛndt/	/staɹtt/
(89)	(88b)	/bʌgt/	/lɪkt/	/mɛndət/	/staɹtət/
	(88a)	/bʌgd/	/lɪkt/	/mɛndəd/	/staɹtəd/

If we schedule (88a) before (88b), this will be the outcome:

	root	/bʌgt/	/lɪkt/	/mɛndt/	/staɹtt/
(90)	(88a)	/bʌgd/	/lɪkt/	/mɛndd/	/staɹtt/
	(88b)	/bʌgd/	/lɪkt/	/mɛndəd/	/staɹtət/

Once again, we see that schwa insertion must take place first.

Analysis 3. The underlying form is [əd]. There is a rule that devoices [d] right after a voiceless obstruent. There is a second rule which *deletes* schwa in between dissimilar consonants.

(91a) $\begin{bmatrix} +\text{voice} \\ +\text{obstruent} \end{bmatrix} \rightarrow \begin{bmatrix} -\text{voice} \\ +\text{obstruent} \end{bmatrix}$ /[−voice]__#

(91b) [ə] → ∅ /C__C′
(C and C′ dissimilar)

	root	/bʌgəd/	/lɪkəd/	/mɛndəd/	/staɹtəd/
(92)	(91b)	/bʌgd/	/lɪkd/	/mɛndəd/	/staɹtəd/
	(91a)	/bʌgd/	/lɪkt/	/mɛndəd/	/staɹtəd/

If we schedule (91a) before (91b), this would be the outcome:

	root	/bʌgəd/	/lɪkəd/	/mɛndəd/	/staɹtəd/
(93)	(91a)	/bʌgəd/	/lɪkəd/	/mɛndəd/	/staɹtəd/
	(91b)	/bʌgd/	/lɪkt/	/mɛndəd/	/staɹtət/

We conclude that schwa deletion must preceded voice assimilation.

In principle, there are many more analyses. We can assume the underlying form to be anything we like (say, even [ð] or [əʒ]). However, one clearly feels that such a proposal would be much inferior to any of the above. But why?

The principal difference between them is solely the extent to which the rules that transform them can be motivated language internally as well as language externally. And this is also the criterion that will make us choose one analysis over the others. Let's look carefully. First, let us go back to the Voice Agreement Principle. It says only that adjacent obstruents agree in voicing, it does not claim that obstruents must agree with the preceding vowel, since we do actually find forms like [kæt]. Analysis 2 incorporates the wrong version of the Voice Agreement Principle. Rule (88a) repairs some of the forms without need, while Analysis 1 repairs the forms if and only if they do not conform to the Voice Agreement Principle. Now look at Analysis 3: it does not conflict with the Voice Agreement Principle. However, it proposes to eliminate schwa in certain forms such as ['lɪked]. There is however no reason why this form is bad. There are words in English such as /wicked/ that have such a sequence. So, it repairs forms that are actually well-formed. Thus the best analysis is the first, and it proposes that the underyling fom is [d].

Let us summarize: an analysis is preferred over another if it proposes laws of change that are widely attested (schwa insertion is one of them, and final devoicing is another). Also, an analysis is dispreferred if its rules change representations that are actually well-formed. Thus, rules of the kind discussed here are seen as *repair* strategies that explain why a form sometimes does not appear in the way expected. What we are looking at here is, by the way, the mapping from deep phonological form to surface phonological form. The last bit of evidence that makes us go for Analysis 1 is the following principle:

> **Not-Too-Similar Principle**
> No English word contains a sequence of subsequent similar obstruents.

Namely, Rule (84a) devoices an obstruent in coda if the preceding consonant is voiceless. And in that case, the Voice Agreement Principle is violated. After the rule has applied, the offending part is gone. Rule (84b) applies if there are two subsequent similar consonants, precisely when the Not-Too-Similar Principle is violated. After application of the rule the offending part is gone.

Which Repair for Which Problem?

The idea that we are pursuing is that deep-to-surface mappings institutionalize a repair of impossible situations. Thus, every rule is motivated from the fact that the input structure violates a constraint of the language and the output structure removes that offending part. Unfortunately, this is not all of the story. Look at the condition that onsets may not branch. This constraint exists in many languages, for example in Japanese and Finnish. But the two languages apply different repair strategies. While Japanese likes to insert vowels, Finnish likes to cut the onset to the last vowel. Both repair strategies are attested in other languages. However, we could imagine a language that simplifies an onset cluster to its first element: with this strategy /trabant/ would become not /rabantti/ but /tabantti/ in Finnish, Germanic /hrengas/ would become not /rengas/, but /hengas/ instead, and so on. This latter strategy is however *not* attested. So, we find that among the infinite possibilities to avoid forbidden cluster combinations, only some get used at all, while others are completely disfavoured. Among those that are in principle available, certain languages choose one and not the other, but in other languages it is the opposite.

Some general strategies can actually be motivated to a certain degree. Look at schwa insertion as opposed to schwa deletion. While the insertion of a schwa is inevitably going to improve the structure (because languages all agree in that CV is a syllable...) the deletion of a schwa can in principle produce clusters that are illegitimate. Moreover, it is highly unlikely that deleting a schwa will make matters betters. Thus, we would expect that there is a bias towards the repair by insertion of schwa. Yet, all these arguments have to be taken with care. For example, if a rule changes stress, this can be a motivating factor in reducing or eliminating a vowel.

Optimality Theory

Several ways to look at the regularities of language have been proposed:

- ☞ The **generative approach** proposes **representations** and **rules**. The rules shall generate all and only the legitimate representations of the language.
- ☞ The **descriptive** or **model-theoretic** approach proposes only **representa-**

tions and conditions that a representation has to satisfy in order to belong to a given language.

Note that generative linguists do not always propose that rules are real, that is, in the head of a speaker and that derivations take place in time. They would say that the rules are a way to systematize the data. If that is so, however, it is not clear why we should not adopt a purely descriptive account, characterizing all only the legal representations rather than pretending that they have been derived in a particular way. The more so since many arguments drawn in favour of a given analysis comes from data on child development and language learning. To interpret the data coming from learning we need to have a theory of the internal knowledge of language ('language faculty'). This knowledge may either consist in representations and conditions on the representations, or in fact in representations plus a set of rules.

The discussion concerning the problem whether we should have rules or not will probably go on forever. **Optimality Theory** (**OT**) adds a new turn to the issue. OT tries to do away with rules (though we shall see that this is an illusion). Also, rather than saying exactly which representations are legitimate it simply proposes a list a desiderata for an optimal result. If a result is not optimal, still it may be accepted if it is the best possible among its kin. Thus, to begin, OT must posit two levels: **underlying representation** (**UR**) and **surface representation** (**SR**). We start with the UR [bætd] (/batted/). Which now is the SR? OT assumes that we generate all possible competitors and rank them according to how many and how often they violate a given constraint.

Here is how it can work in the present situation. We shall assume that the SR deviates in the least possible way from the UR. To that effect, we assign each segment a slot in a grid:

(94)

b	æ	t	d
•	•	•	•

We may subsequently insert or delete material and we may change the representations of the segments, but we shall track our segments through the derivation. In the present circumstances we shall require, for example, that they do not change order with respect to each other. (However, this sometimes happens; this is called

metathesis.) Here is an example where a segment changes:

(95)
b	æ	t	d
•	•	•	•
↓	↓	↓	↓
•	•	•	•
b	æ	t	t

Here is an example where we one segment is dropped:

(96)
b	æ	t	d
•	•	•	•
↓	↓	↓	
•	•	•	
b	æ	t	

And here is an example where one segment is added:

(97)
b	æ	t		d
•	•	•		•
↓	↓	↓		↓
•	•	•	•	•
b	æ	t	ə	d

Now, we look at pairs $\langle I, O \rangle$ of representations. Consider the following constraints on such pairs.

> **Recover the Morpheme**
> At least one segment of any given morpheme must be preserved in the SR.

This principle is carefully phrased. If the underlying morpheme is empty, there need not be anything in the surface. But if it is underlyingly not empty, then we must see one of its segments.

> **Recover Obstruency**
> If an underlying segment is an obstruent, it must be an obstruent in the SR.

The last principle says that the segment that is underlyingly hosting an obstruent must first of all survive; it cannot be deleted. Second, the phoneme that it hosts must be an obstruent.

> **Recover Voicing**
> If a segment is voiced in UR, it must be voiced in SR.
>
> **Syllable Equivalence**
> The UR must contain the same number of syllables than the SR.
> **Recover Adjacency**
> Segments that are adjacent in the UR must be adjacent in the SR.

These principles are too restrictive in conjunction. The idea is that one does not have to satisfy all of them; but the more the better. An immediate idea is to do something like linear programming: for each constraint there is a certain penalty, which is 'awarded' on violation. For each violation, the corresponding penalty is added. (In school it's basically the same: bad behaviour is punished, depending on your behaviour you heap up more or less punishment.) Here however the violation of a more valuable constraint cannot be made up for. No matter how often someone else violates a lesser valued constraint, if you violate a higher valued constraint, you loose.

To make this more precise, for a pair $\langle I, O\rangle$ of underlying representation (I) and a surface representation (O), we take note of which principles are violated. Each language defines a partial linear order on the constraints, such as the ones given above. It says in effect, given two constraints C and C', whether C is more valuable than C', or C' is more valuable than C, or whether they are equally valuable.

Now suppose the UR I is given:

① Suppose that $\pi = \langle I, O\rangle$ and $\pi' = \langle I, O'\rangle$ are such that for all constraints that π violates there is a more valuable constraint that π' violates. Then O is called **optimal with respect to** O'.

② Suppose that $\langle I, O\rangle$ and $\pi' = \langle I, O\rangle$ are such that C is the most valuable constraint that π violates; and that C is also the most valuable constraint that π' violates. Then if π' violates C more often than π, O is **optimal with respect to** O'.

③ *O* is **optimal** if it is optimal with respect to every other pair with same UR.

The actual output form for *I* is the optimal output form for *I*. (If there are several optimal output forms, all of them are chosen.) So, given *I*, if we want to know which SR corresponds to it, we must find an *O* which is optimal. Notice that this *O* need not be unique. (OT uses the following talk. *O* and *O′* are called **candidates** and they get **ranked**. However, candidacy is always relative to the UR.)

Let's apply this for [bætd]. We rank the constraints NTS (Not-Too-Similar), RObs (Recover Obstruency), RMorph (Recover the Morpheme) and RAdj (Recover Adjacency) as follows:

(98) NTS, RObs, RMorph > RAdj

The first three are ranked equal, but more than the fourth. Other principles are left out of consideration.

(99)

/bætd/	NTS	RObs	RMorph	RAdj
[bætd]	★			
[bænd]		★		
[bæt]			★	
[bætəd]				★

The forms (a), (b), (c) all violate constraints that are higher ranked than RAdj. (d) violates the latter only. Hence it is optimal among the four. (Notice that we have counted a violation of Recover Obstruency for (c), even though one obstruent was dropped. This will concern us below.) Note that the optimal candidate still violates some constraint.

We now turn to the form /sæpd/. Assume that the ranking is (with VAgr = Voice Agreement, RVoice = Recover Voicing)

(100) VAgr, RMorph, RObs > RAdj > RVoice

This yields the following ranking among the candidates:

(101)

/sæpd/	VAgr	RMorph	RObs	RAdj	RVoice
[sæpd]	★				
[sæp]		★			
[sæmp]			★		
[sæpəd]				★	
[sæpt]					★

Some Conceptual Problems with OT

First notice that OT has no rules but its constraints are not well-formedness conditions on representations either. They talk about the relation between an UR and an SR. They tell us in effect that certain repair strategies are better than others, something that well-formedness conditions do not do. This has consequences. Consider the form [bænd]. We could consider it to have been derived by changing [t] to [n]:

(102)

b	æ	t	d
•	•	•	•
↓	↓	↓	↓
•	•	•	•
b	æ	n	t

Or we could think of it as being derived by deleting [t] and inserting [n]:

(103)

b	æ	t		d
•	•	•		•
↓	↓			↓
•	•		•	•
b	æ		n	d

More fancyful derivations can be conceived. Which one is correct if that can be said at all? And how do we count violations of rules? The first one violates the principle that obstruents must be recoverable. The second does not; it does not even violate adjacency! Of course, we may forbid that obstruents be deleted. But note that languages do delete obstruents (Finnish does so to simplify onset clusters). Thus obstruents can be deleted, but we count that also as a violation of the principle of recoverability of obstruents. Then the second pair violates that principle. Now again: what is the punishment associated with that pair? How does it get ranked? Maybe we want to say that of all possibilities takes the most favourable one for the candidate.

The problem is a very subtle one: how do we actually measure violation? The string [cat]—how many constraints must we violate how often to get it from [staɹtd]? For example, we may drop [s], then [d] and change the place of articulation of the first sound to [+velar]. Or may we drop [s] and [d] in one step, and then change [t] to [k]—or we do all in one step. How many times have we violated

Rule C? The idea of violating a constraint 3 times as opposed to once makes little sense unless we assume that we apply certain rules.

Notes on this section. The transition from underlying /d/ to /t/, /ə/ or /d/ has mostly been described as morphophonemics. This is a set of rules that control the spelling of morphs (or morphemes?) in terms of phonemes. Under this view there is one morph, which is denoted by, say, /PAST/$_M$ (where $/\cdot/_M$ says: this is a unit of the morphological level) and there are rules on how to realise this morph on the phonological level (recall a similar discussion on the transition from the phonological to the phonetic level). Under that view, the morph /PAST/$_M$ has three realisations. The problem with this view is that it cannot motivate the relationship between these forms. Thus we have opted for the following view: there is one morph and it is realised as /d/ phonologically. It produces illicit phonological strings and there are therefore rules of combination that project the illicit combinations into licit ones.

Morphology I: Basic Facts

> Morphemes are the smallest parts that have meaning. Words may consist of one or several morphemes in much the same way as they consist of one or more syllables. However, the two concepts, that of a morpheme and that of a syllable, are radically different.

Word Classes

Morphology is the study of the minimal meaningful units of language. It studies the structure of words, however from a semantic viewpoint rather than from the viewpoint of sound. Morphology is intimately related to syntax. For everything that is larger than a word is the domain of syntax. Thus within morphology one considers the structure of words only, and everything else is left to syntax. The first to notice is that words come in different classes. For example, there are verbs (/to imagine/) and there are nouns (/a car/), there are adverbs (/slowly/) and adjectives (/red/). Intuitively, one is inclined to divide them according to their meaning: verbs denote activities, nouns denote things adverbs denote ways of performing an activities and adjectives denote properties. However, language has its own mind. The noun /trip/ denotes an activity, yet it is a noun. Thus, the semantic criterion is misleading. From a morphological point of view, the three are distinct in the following way. Verbs take the endings /s/, /ed/, and /ing/, nouns only take the ending /s/. Adjectives and adverbs on the other hand do not change. (They can be distinguished by other criteria, though.)

(104) We imagine.

(105) He imagines.

(106) We are imagining.

(107) He imagined.

Thus we may propose the following criterion: a word *w* is a verb if and only if we can add [z] (/s/), [d] (/ed/) and [ɪŋ] (/ing/ and nothing else; *w* is a noun if and only if we can add [s] (/s/) and nothing else.

This distinction is made solely on the basis of the possibility of changing the form alone. The criterion is at times not so easy to use. Several problems must be noted. The first is that a given word may belong to several classes; the test

using morphology alone would class anything that is both a noun and a verb, for example /fear/ as a verb, since the plural (/fears/), is identical to the third singular. Changing the wording to replace 'if and only if' to 'if' does not help either. For then any verb would also be classed as a noun. A second problem is that there can be false positives; the word /rise/ [ɹaɪz] cannot be taken as the plural of /rye/ [ɹaɪ]. And third, there some words do not use the same formation rules. There are verbs that form their past tense not in the way discussed earlier, by adding [d]. For example, the verb /run/ has no form */runned/. Still, we classify it as a verb. For example, the English nouns take a subset of the endings that the verb takes. The word /veto/ is both a noun and a verb, but this analysis predicts that it is a verb. Therefore, more criteria must be used. One is that of taking a context and looking which words fit into it.

(108) The governor ____ the bill.

If you fill the gap by a word, it is certainly a verb (more exactly a transitive verb, one that takes a direct object). On the other hand, if it can fill the gap in the next example it is a noun:

(109) The ____ vetoed the bill.

When we say 'fill the gap' we do not mean however that what we get is a meaningful sentence when we put in that word; we only mean that it is grammatically (= syntactically) well-formed. We can fill in /cat/, but that stretches our imagination a bit. When we fill in /democracy/ we have to stretch it even further, and so on. Adjectives can fill the position between the determiner (/the/) and the noun:

(110) The ____ governor vetoed the bill.

Finally, adverbs (/slowly/, /surprisingly/) can fill the slot just before the main verb.

(111) The governor ____ vetoed the bill.

Another test for word classes in the combinability with affixes. (**Affixes** are parts that are not really words by themselves, but get glued onto words in some way. See Lecture 13 for details.) Table 11 shows a few English affixes and lists the word classes to which it can be applied. We see that the list of affixes is heterogeneous, and that affixes do not always attach to all members of a class with equal ease (/anti-house/, for example, is yet to be found in English). Still, the test reveals a lot about the division into different word classes.

Table 11: English Affixes and Word Classes

Affix	Attaches to	Forming	Examples
anti-	nouns	nouns	anti-matter, ant-aircraft
	adjectives	adjectives	anti-democratic
un-	adjective	adjectives	un-happy, un-lucky
	verbs	verbs	un-bridle, un-lock
re-	verbs	verbs	re-establish, re-assure
dis-	verbs	verbs	dis-enfranchise, dis-own
	adjectives	adjectives	dis-ingenious, dis-honest
-ment	verbs	nouns	establish-ment, amaze-ment
-ize	nouns	verbs	burglar-ize
	adjective	verbs	steril-ize, Islamic-ize
-ism	nouns	nouns	Lenin-ism, gangster-ism
	adjectives	nouns	real-ism, American-ism
-ful	nouns	adjectives	care-ful, soul-ful
-ly	adjectives	adverbs	careful-ly, nice-ly
-er	adjectives	adjectives	nic-er, angry-er

Morphological Formation

Words are formed from simpler words, using various processes. This makes it possible to create very large words. Those words or parts thereof that are not composed and must therefore be drawn from the lexicon are called **roots**. Roots are 'main' words, those that carry meaning. (This is a somewhat hazy definition. It becomes clearer only through examples.) Affixes are not roots. Inflectional endings are also not roots. An example of a root is /cat/, which is form identical with the singular. However, the latter also has a word boundary marker at the right and (so it looks more like (/cat#/, but this detail is often generously ignored). In other languages, roots are clearly distinct from every form you get to see on paper. Latin /deus/ 'god' has two parts: the root /de/, and the nominative ending /us/. This can be clearly seen if we add the other forms as well: genitive /dei/, dative /deo/, accusative /deum/, and so on. However, dictionaries avoid using roots. Instead, you find the words by their citation form, which in Latin is the nominative singular. So, you find the root in the dictionary under /deus/ not under /de/. (Just an aside: verbs are cited in their infinitival form; this need not be so. Hungarian dictionaries often list them in their 3rd singular form. This is because the 3rd

singular reveals more about the inflection than the infinitive. This saves memory!)

There are several distinct ways in which words get formed; moreover, languages differ greatly in the extent to which they make use of them. The most important ones are

① **compounding**: two words, neither an affix, become one by juxtaposition. Each of them is otherwise found independently. Examples are /`goalkeeper`/, /`whistleblower`/ (verb + noun compound), /`hotbed`/ (adjective + noun).

② **derivation**: only one of the parts is a word; the other is only found in combination, and it acts by changing the word class of the host. Examples are the affixes which we have discussed above (/`anti`/, /`dis`/, /`ment`/).

③ **inflection**: one part is an independent word, the other is not. It does however not change the category, it adds some detail to the category (inflection of verbs by person, number, tense ...).

Compounding

In English, a compound can be recognised by its stress pattern. For example, the main stress in the combination adjective+noun is on the noun if they still form two words (`black board` [blæk ˈbɔɹd]) while in a compound the stress is on the adjective (/`blackboard` [ˈblækbɔɹd]). Notice that the compound simply is one word, so the adjective has lost its status as adjective through compounding, which explains the new stress pattern.

In English, compounds are disfavoured over multiword constructions. It has to be said, though, that the spelling does not really tell you whether you are dealing with one or two words. For example, although one writes /`rest room`/, the stress pattern sounds as if one is dealing with only one word.

There are languages where the compounds are not just different in terms of

stress pattern from multiword constructions. German is such a language.

(112) Regierung-s-entwurf
government proposal

(113) Schwein-e-stall
pig sty

(114) Räd-er-werk
wheel work = mechanism

The compound /Regierungsentwurf/ not only contains the two words /Regierung/ ('government') and /Entwurf/ ('proposal'), it also contains an 's' (called 'Fugen s' = 'gap s'). To make German worthwhile for students, what gets inserted is not always an 's' but sometimes 'e'; /Schweinestall/ is composed from /Schwein/ and /Stall/. /Räderwerk/ is composed from /Rad/ and /Werk/. /Schweine/ and /Räder/ sound exactly like the plural, while /Regierungs/ is not like any case form of /Regierung/. In none of these cases can the compound be mistaken for a multiword construction.

The meaning of compounds often enough is not determinable in a strightforward way from the meaning of its parts. It is characteristic of compounds that they often juxtapose some words and leave it open as to what the whole means (take /money laundry/ or /coin washer/, which are generally not places where you launder money or wash coins; if you did you would be in serious trouble). This is more true of noun+noun compounds than of verb+noun/noun+verb compounds, though.

Sanskrit has far more compounds than even German. There is an entire classification of compounds in terms of their makeup and meaning.

Derivation

English has a fair amount of derivational affixes. Table 11 shows some of them. We said that derivation changes the category of the word; this is not necessarily so. Thus it might be hard to distinguish derivation from inflection in that case. However, derivation is optional (while inflection is not), and can be iterated (inflection cannot be iterated). And, inflectional affixes are typically outside derivational affixes. To give an example: you can form /republican/ from

/republic/. To this you can add /anti/: /antirepublican/ and finally form the plural: /antirepublicans/. You could have added the 's' to /republic/, but then you could not go on with the derivation. There is no word */republicsan/. Similarly, the word /antirepublics/ has only one derivation: first /anti/ is added and then the plural suffix. Whether or not a word is formed by derivation is not always clear. For example, is /reside/ formed by affixation? Actually, it once was, but nowadays it is not. This is because we do not have a verb */side/ except in some idioms. Thus, derivation may form words that initially are perceived as complex, but later lose their transparent structure. This maybe because they start to sound different or because the base form gets lost. Nobody would guess that the word /nest/ once was a complex word */nizdo/ (here the star means: this form is reconstructed), derived from the words */ni/ ('down') and */sed/ ('sit').

Inflection

To fit a word into a syntactic construction, it may have to undergo some changes. In English, the verb has to get an 's' suffix if the subject is third person singular. The addition of the 's' does not change the category of the verb; it makes it more specific, however. Likewise, the addition of past tense. Adding inflection thus makes the word more specific in category, narrowing down the contexts in which it can occur. Inflection is not optional; you must choose an inflectional ending. In Latin, adjectives agree in gender, number and case with the noun they modify:

(115)	discipul-us	secund-us
	student-nom.sg	second-masc.nom.sg
(116)	discipul-orum	secund-orum
	student-gen.pl	second-masc.gen.pl
(117)	puell-arum	secund-arum
	girl-gen.pl	second-fem.gen.pl
(118)	poet-arum	secund-orum
	poet-gen.pl	second-masc.gen.pl

The last example was chosen on purpose: form identity is not required. It is actually true that the forms of adjectives resemble those of nouns. The word /poeta/ belongs a form class of nouns that are mostly feminine, this is why adjectives show this form class if agreeing with a feminine noun (this has historic reasons).

But the form class contains some masculine nouns, and to agree with them adjectives show a different form class. The latter actually is identical to that which contains more masculine nouns. This also explains why we have not added gender specifications to the nouns; unlike adjectives, nouns cannot be decomposed into gender and a genderless root.

The morphological characteristic of inflection is that it is harder to identify an actual affix (morph).

Syntax I: Categories, Constituents and Trees. Context Free Grammars

> Sentences consist of words. These words are arranged into groups of varying size, called *constituent*. The structure of constituents is a *tree*. We shall show how to define the notion of constituent and constituent occurrnce solely in terms of sentences.

Occurrences

Sentences are not only sequences of words. There is more structure than meets the eye. Look for example at

(119) `This villa costs a fortune.`

The words are ordered by their appearances, or if spoken, by their temporal arrangement. This ordering is linear. It satisfies the following postulates:

① No word precedes itself. (Irreflexivity)

② If w precedes w' and w' precedes w'' then w precedes w''. (Transitivity)

③ For any two distinct words w and w', either w precedes w' or w' precedes w. (Linearity)

There is however one thing about we must be very careful. In the sentence

(120) `the dog sees the cat eat the mouse`

we find the word /the/ three times (we do not distinguish between lower and upper case letters here). However, the definitions above suggest that /the/ precedes /the/ since it talks of words. Thus, we must change that and talk of *occurrences*. The best way to picture occurrences is by underlining them in the string:

(121)
`the dog sees the cat eat the mouse`
`the dog sees the cat eat the mouse`
`the dog sees the cat eat the mouse`

Occurrences of same of different strings can either overlap, or precede each other. They overlap when they share some occurrences of letters. Otherwise one precedes the other. The next two occurrences in (122) and (123) overlap, for example, while the occurrences of /the/ above do not.

(122) the dog sees the cat eat the mouse

(123) the dog sees the cat eat the mouse

When we cannot do that, we need another tool to talk of occurrences. Here is one way of doing that.

Definition 9 *Let $\vec{x}$ and $\vec{z}$ strings. An **occurrence** of $\vec{x}$ in $\vec{z}$ is a pair $\langle \vec{u}, \vec{v} \rangle$ such that $\vec{z} = \vec{u}\vec{x}\vec{v}$. Given an occurrences $C = \langle \vec{u}_1, \vec{v}_1 \rangle$ of $\vec{x}_1$ and and occurrence $D = \langle \vec{u}_2, \vec{v}_2 \rangle$ of $\vec{x}_2$ we say that C **precedes** D if $\vec{u}_1\vec{x}_1$ is a prefix of $\vec{u}_2$; we say that C and D **overlap** if C does not precede D and D does not precede C.*

Thus, the word /the/ has the following occurrences in (120) (with spaces shown when important):

(124)
$\langle \varepsilon, \text{␣dog sees the cat eat the mouse} \rangle$
$\langle \text{the dog sees␣}, \text{␣cat eat the mouse} \rangle$
$\langle \text{the dog sees the cat eat␣}, \text{␣mouse} \rangle$

If you apply the definition carefully you can see that the first occurrence precedes the second: ε concatenated with /the/ gives /the/, which is a prefix of /the dog sees/. Notice the notion of overlap; here are two occurrences, one of /the dog/ and the other of /dog sees/. These, consisting of the first and second and second and third occurrences of words, respectively, must overlap. They share the occurrences of the second word.

(125)
$\langle \varepsilon, \text{␣sees the cat eat the mouse} \rangle$
$\langle \text{the␣}, \text{␣cat eat the mouse} \rangle$

The postulates above simply have to be reformulated in terms of occurrences of words and they become correct. This has to do with the assumption that occurrences of words cannot overlap. Let $\vec{z}$ a given sentences, and U the set of occurrences of words in $\vec{z}$. Then the following holds.

① No member of U precedes itself. (Irreflexivity)

② Let C, C' and C'' be in U. If C precedes C' and C' precedes C'' then C precedes C''. (Transitivity)

③ For any two distinct occurrences C and C' from U, either C precedes C' or C' precedes C. (Linearity)

Talk of occurrences of words is often clumsy, but it is very important to get the distinction straight.

A notationally simpler way is the following. We assign to the occurrences words some symbol, say a number. If we use numbers, we can even take advantage of their intrinsic order. We simply count the occurrences from left to right.

Constituents

We claim that, for example, /`a fortune`/ is a sequence of different character than /`costs a`/. One reason is that it can be replaced by /`much`/ without affecting grammaticality:

(126) `This villa costs much.`

Likewise, instead of /`this villa`/ we can say

(127) `This costs much.`

Notice that exchanging words for groups or vice versa does not need to preserve the meaning; all that is required is that it preserves grammaticality: the result should take English sentences to English sentences, and non-English sentences to non-English sentences. For, example if we replace /`costs much`/ by /`runs`/ we are not preserving meaning, just grammaticality:

(128) `This runs.`

Notice that any of the replacements can also be undone:

(129) `This villa runs.`

(130) `This costs a fortune.`

We call a sequence of words a **constituent** if (among other conditions) it can be replaced by a single word. A second condition is that the sequence can be **coordinated**. For example, we can replace /a fortune/ not only by /a lot/ but also by /a fortune and a lot/. The latter construction is called **coordinated** because it involves the word /and/ (a more precise version will follow). On the assumption that everything works as promised, the constituents of the sentence (119) has the following constituents:

(131)
```
{this villa costs a fortune,
 this villa,    costs a fortune,
this, villa, costs,   a fortune,
                      a, fortune}
```

The visual arrangement is supposed to indicate order. However, this way of indicating structure is not precise enough. Notice first that a word or sequence of words can have several occurrences, and we need to distinguish them, since some occurrences may be constituent occurrences, while others are not.

I shall work out a somewhat more abstract representation. Let us give occurrence of a word a distinct number, like this:

(132)
```
this villa costs a fortune
  1     2     3   4     5
```

Each occurrence gets its own number. A sequence of occurrences can now conveniently be represented as a set of numbers. The constituents can be named as follows:

(133) $\{\{1,2,3,4,5\},\{1,2\},\{1\},\{2\},\{3,4,5\},\{3\},\{4,5\},\{4\},\{5\}\}$

Let $w_1w_2w_3\ldots w_n$ be a sequence of words constituting a sentence of English. Then a constituent of that sentence has the form $w_iw_{i+1}w_{i+2}\ldots w_j$, for example w_2w_3, $w_5w_6w_7$ but not $w_2w_4w_7$. It always involves a *continuous stretch of words*.

Continuity of Constituents
Constituents are continuous parts of the sentence.

Non-Crossing
Given two constituents that share a word, one must be completely inside the other.

Words are Constituents
Every occurrence of a word forms its own constituent.

Here is a useful terminology. A constituent C is an **immediate constituent** of another constituent D if C is properly contained in D, but there is no constituent D' such that C is properly contained in D' and D' is properly contained in D. Our sentence (119) has only two immediate subconstituents: /`this villa`/ and /`costs a fortune`/. The latter has the immediate constituents /`costs`/ and /`a fortune`/. It is not hard to see that it is enough to establish for each constituents its immediate subconstituents. Notice also that we can extend the notion of precedence to constituents. A constituent C **precedes** a constituent D if all words of C precede all words of D. So, /`this villa`/ precedes /`a fortune`/ because /`this`/ precedes both /`a`/ and /`fortune`/ and /`villa`/ precedes both /`a`/ and /`fortune`/.

This coincidence opens the way to a few alternative representations. One is by enclosing constituents in brackets:

(134) [[[`this`] [`villa`]] [[[`costs`] [[`a`] [`fortune`]]]]]

Typically, the brackets around single words are omitted, though. This gives the slightly more legible

(135) [[`this villa`] [[`costs` [`a fortune`]]]]

Another one is to draw a tree, with each constituent represented by a node, and drawing lines as given in Figure 3. Each node is connected by a line to its immediate subconstituents. These lines go down; a line going up is consequently from a constituent to the constituent that it is an immediate part of. So, 8 and 9 are the immediate constituents of 7, 6 and 7 are the immediate subconstituents of 3, and so on. It follows that 6, 7 and 8 are subconstituents of 3, which is to say that 3 consists of /`costs`/, /`a`/ and /`fortune`/.

Definition 10 *A **tree** is a pair $\langle T, < \rangle$ where T is a set and $<$ is a set of pairs of elements from T such that (a) if $x < y$ and $x < z$ then either $y < z$, or $y = z$ or $z < y$ and (b) there is an element r such that $x < r$ for all $x \neq r$ (r is called the **root**).*

The tree in Figure 3 consists of the nodes $1, 2, 3, \ldots, 9$ (we ignore the stuff below them; this concerns wordforms and is of no interest to syntax; we add these things for orientation only). The relation $<$ is as follows: $2 < 1$, $3 < 1$, $4 < 1$, $5 < 1$, $6 < 1$, $7 < 1$, $8 < 1$, $9 < 1$; $4 < 2$, $5 < 2$; $6 < 3$, $7 < 3$, $8 < 3$, $9 < 3$; $8 < 7$, $9 < 7$; no other relations hold.

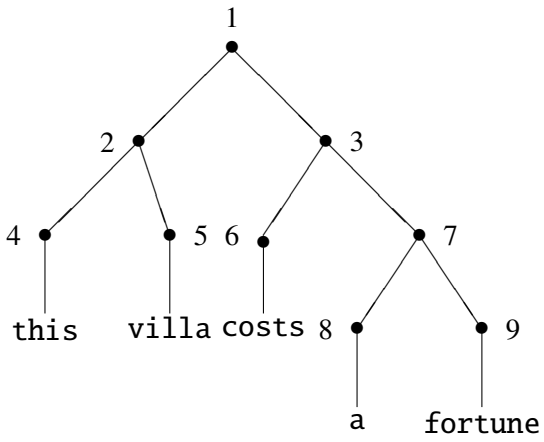


Figure 3: An Unlabelled Syntactic Tree

Notice also that the order in which elements follow each other is also reflected in the tree, simply by placing the earlier constituents to the left of the later ones.

Categories

A **context** is a pair $\langle \vec{x}, \vec{y} \rangle$ of strings. We denote them in a more visual way as follows:

(136) $\vec{x}\_\_\_\_\vec{y}$

An example of a context is

(137) `this ____ a fortune`

(Notice the blanks left and right of the gap!) Given a sentence, every substring is uniquely identified by its context: the part that precedes it and the part the follows it. The missing part in (137) is /`villa costs`/. So, every substring can be denoted by its left context and its right context. Substituting a substring by another is extracting the context and then putting back the replacing string. Replacing /`villa costs`/ by /`misses`/ gives

(138) `this misses a fortune`

Not a good sentence, but grammatical. (Does this show, by the way that /`villa costs`/ is a constituent ...?)

We agree to call arbitrary strings of English words **constituents** just in case they occur as constituents in *some* sentence. The wording here is important: I do not say that they have to occur everywhere as constituent. It may happen that a sequence occurs in one sentence as a constituent, but not in another. Here is an example. The string /`you called`/ is a constituent in (139) but not in (140):

(139) `This is the man you called.`

(140) `This is a song for you called "Sweet Georgia Brown".`

To know why this is to try to substitute /`you called`/ by /`you called and Mary liked`/. So, first of all we need to talk about sequences that occur *as constituents*. If they do, we call the occurrence a *constituent occurrence*.

We are now ready to make matters a bit more precise. A **string** is a sequence of letters (or sounds). A **string language** is a set of strings. The set of sentences of English is an example. Call this set E. We say that a string $\vec{x}$ is a (**grammatical**) **sentence of English** just in case $\vec{x}$ is a member of E. So, /`We stay at home.`/ is a (grammatical) sentence of English while /`Cars a sell if.`/ is not. Based on this set we define constituents and constituent occurrences and all that.

First we give some criteria for constituent occurrences. $\vec{x}$ has a **constituent occurrence** in $\vec{z}$ only if there are $\vec{u}$ and $\vec{v}$ such that:

① $\vec{z} = \vec{u}\frown\vec{x}\frown\vec{v}$;

② $\vec{z} \in E$ (that is, $\vec{z}$ is a grammatical sentence);

③ either $\vec{u}\frown\vec{v} \in E$ or there is a single word $\vec{w}$ such that $\vec{u}\frown\vec{w}\frown\vec{v} \in E$;

④ $\vec{u}\frown\vec{x}\frown\mathsf{and}\frown\vec{x}\frown\vec{v} \in E$.

The reason why in ④ we use $\vec{x}$ twice rather than making up some more imaginative conjunct (as we did above) is that for a formal test it is of use if we do not have to choose an appropriate string. What if we choose the wrong one? Not all of them work in all circumstances, so it is better to use something less imaginative but secure.

In case ① – ④ are satisftied we say that $\langle \vec{u}, \vec{v} \rangle$ is a **constituent occurrence** of $\vec{x}$ (in $\vec{z}$). This list is incomplete. In other words, there are more criteria that need to be satisfied before we say that $\vec{x}$ has a constituent occurrence. But this list shall suffice for us.

Let us apply this to (139) and (140). With $\vec{x}$ = `you called` we have in (139) the occurrence

(141) $\langle$`This␣is␣the␣man␣, ␣.`$\rangle$

① and ② are obviously met. That ③ and ④ are met is witnessed by the following.

(142) `This is the man.`

(143) `This is the man you called and you called.`

We try the same on (140). The following sentence is ungrammatical, so ④ fails.

(144) *`This is a song for you called and you called "Sweet Georgia Brown".`

Definition 11 *A **category** (of a given language) is a set* Δ *of strings such that any constituent occurrence of a member of* Δ *can be replaced by any other member of* Δ *yielding once again a constituent occurrence.*

Thus, given that $\vec{x}$ and $\vec{y}$ are members of the same category, if $\vec{u}\vec{x}\vec{v}$ is a grammatical sentence of English, so is $\vec{u}\vec{y}\vec{v}$ and vice versa.

The definition of a category is actually similar to that of a phoneme; phonemes were defined to be substitution classes up to meaning preservation. Here we define the classes up to grammaticality (more or less, since we have the caveat about 'constituent occurrences' because we are dealing with the substitution of strings for strings, not just of an item for another item). We give an example. The intransitive verbs in the 3rd personal singular form a category:

(145) {`falls`, `runs`, `talks`, ...}

By this definition, however, /`talks`/ and /`talk`/ are not members of the same category, for in (146) we cannot replace /`talk`/ by /`talks`/. The result is simply ungrammatical.

(146) `Mary and Paul talk.`

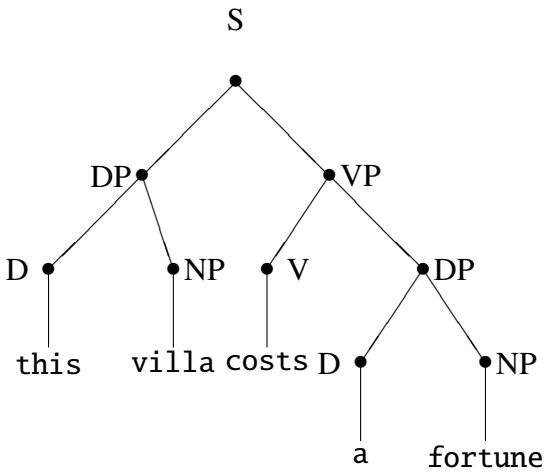


Figure 4: A Labelled Syntactic Tree

It seems that the number of categories of any given language must be enormous if not infinite. It is certainly true that the number of categories is large, but it has a highly regular structure which we shall unravel in part for English.

Now that we have defined the constituents, let us go back to our tree in Figure 3. Each of the nodes in that tree is a constituent, hence belongs to some category. Omitting some detail, the categories are given in Figure 4. (We call S a sentence, D a determiner, DP a determiner phrase, NP a noun phrase, VP a verb-phrase, V verb.)

Context Free Grammars

Now look at Figure 4. If the labelling is accurate, the following should follow: any sequence of a determiner phrase followed by a verb phrase is a sentence. Why is this so? Look at the tree. If DP is the category to which /`this villa`/ belongs, and if VP is the category to which /`costs a fortune`/ belongs, then we are entitled to substitute *any* DP for /`this villa`/, and *any* VP for /`costs a`

fortune/:

(147)
$$\left\{\begin{array}{l}\texttt{this villa}\\ \texttt{a car}\\ \texttt{it}\\ \texttt{tomorrow's sunshine}\end{array}\right\}\left\{\begin{array}{l}\texttt{costs a fortune.}\\ \texttt{walks.}\\ \texttt{catches the bus.}\\ \texttt{sings praise to the Lord.}\end{array}\right\}$$

None of the sentences you get are ungrammatical, so this actually seems to work. We state the fact that a DP followed by a VP forms a sentence in the following way:

(148) S → DP VP

We have seen earlier statements of the form 'something on the left' → 'something on the right', with a slash after which some conditions to the context were added. The condition on context is absent from (148); this is why the rule is called **context free**. Because it can be used no matter what the context is. However, notice one difference, namely that the thing on the left is always a single symbol, denoting a category, while the thing on the right is a sequence of symbols, which may each be either a category or a word. A **grammar** is a set of such rules, together with a special symbol, called **start symbol** (usually S). Consider by way of the example the following grammar. The start symbol is S, the rules are

(149a) S → DP VP
(149b) DP → D NP
(149c) D → a | this
(149d) NP → villa | fortune
(149e) VP → V DP
(149f) V → costs

Here, the vertical stroke '|' is a disjunction. It means 'can be either the one or the other'. For example, the notation D → a | this is a shorthand for two rules: D → a and D → this.

This grammar says that (119) is of category S. How does it do that? It says that /this/ is a determiner (D; Rule (149c)), and /villa/ is a noun phrase (NP; Rule (149d)). By Rule (149b) we know that the two together are a DP. Similarly, it tells us that /a fortune/ is a DP, that /costs a fortune/ is a VP. Finally, using Rule (149a) we get that the whole is an S.

Given a set of rules R, we define a *derivation* as follows. An R-**derivation** is a sequence of strings such that each line is obtained by the previous by doing one replacement according to the rules of R. If the first line of the derivation is the symbol X and the last the string $\vec{x}$ we say that the string $\vec{x}$ has **category** X in R. The following is a derivation according to the set (149a) – (149f).

(150) S
DP VP
D NP VP
D NP V DP
D NP V D NP
`this` NP V D NP
`this villa` V D NP
`this villa costs` D NP
`this villa costs a` NP
`this villa costs a fortune`

Thus, in this rule system, /`this villa costs a fortune`/ is of category S. Likewise, one can show that /`this fortune`/ is of category DP, and so on.

A context free grammar consists of (a) an alphabet A of terminal symbols, (b) an alphabet N of nonterminal symbols, (c) a nonterminal $S \in N$, and (d) a set of context free rules $X \to \vec{x}$, where $X \in N$. Notice that the strings DP, VP, V and so on are considered to be symbols; but they are not symbols of English. They are therefore nonterminal symbols. Only nonterminal symbols may occur on the left of a rule. But nonterminal symbols may occur on the right, too.

The role of the start symbol is the following. A string of terminal symbols is a **sentence** of the grammar if there is a derivation of it beginning with the start symbol. The **language** generated by the grammar is the set of all sentences (using only terminal symbols!) that it generates. Ideally, a grammar for English should generate exactly those sentences that are proper English. Those grammars are incredibly hard to write.

Notes on this section. Let's return to our definition of constituent occurrence. Context free grammars provide strings with a constituent analysis. However, not every such analysis conforms to the definition of constituents. To make everything fall into place we require that the grammars of English have the following

universal rule schema. For every category symbol X they contain a rule

(151) $X \rightarrow X \_ \textsf{and} \_ X$

Furthermore, for every X they must have a rule

(152) $X \rightarrow \vec{w}$

for some word. If that is so, the criteria ① – ④ will be met by all constituents assigned by the grammar.

Syntax II: Argument Structure

> We shall introduce a general schema for phrase structure, called X-bar syntax. In X-bar syntax, every word projects a phrase consisting of up to two arguments, and any number of adjuncts.

Lexical Categories

There are several tens of thousands of categories in a language, maybe even millions. Thus the number of rules that we have to write is far too large to be written one by one. Thus, while in phonology the desire for general rules could still be thought of as not so urgent, here it becomes absolutely central. We shall put to use our notation of attribute value structures (AVSs). We start off with a few general purpose rules and then refine them as we go along.

First, words fall into roughly two handful of so-called **lexical** or **major categories**. The ones we shall be using are: **noun** (N), **verb** (V), **adjective** (A), **adverb** (Adv), **preposition** (P), **complementizer** (C), **determiner** (D), and **tense** (T). Not all classes have single words in them, but most of them do:

(153)

N	`car`, `house`, `storm`, `insight`
V	`run`, `eat`, `hasten`, `crawl`
A	`greedy`, `raw`, `shiny`, `cheerful`
Adv	`very`, `steadily`, `allegedly`, `down`
P	`in`, `for`, `about`, `below`
C	`that`, `which`, `because`, `while`
D	`a`, `the`, `this`, `those`

Our first attribute is CAT, and it has the values just displayed (so far, they are N, V, A, Adv, P, C, D, T).

Subject and Object

The next distinction we want to make is that between a **word** and a **phrase**. We have made that distinction earlier, when we called certain words determiners and certain constituents DPs (= determiner phrases). The distinction is intimately

connected with the notion of argument structure. The argument structure tells us for each word what kinds of constituents it combines with to form a phrase. Take the verbs /run/ and /accuse/. The difference between the two is that the first is happy to combine with just one DP, say /the sailors/ to form a sentence, while the latter is not:

(154) The sailors run.

(155) *The sailors accuse.

To make the latter into the sentence, you need to supply two more DPs, one for the one who is accused and one for what he is accused of:

(156) The sailors accuse the captain of treason.

We say, /run/ takes **one argument**, /accuse/ takes **three arguments**. It means that in order to get a grammatical sentence, /run/ only needs to combine with one DP, /accuse/ needs three. (There are differences in the degree to which /accuse/ needs any one of its arguments; this is a matter we shall not go into.)

As for verbs, they always have one argument in a sentence, and this is the **subject**. The subject is that argument which is in the nominative case. You can recognize it by the form of pronoun that you have to use. The subject is /she/, /he/, /it/, /I/, /you/, /they/ and not /her/, /him/, /me/, /us/ and /them/. (The form /you/ is both nominative and accusative, so we cannot use it for our test.) If instead you have to use the accusative forms, you are looking at the (**direct**) **object**.

(157) They run./*Them run.

(158) They accuse the captain of treason.

(159) *Them accuse the captain of treason.

(160) The sailors accuse us of treason.

(161) *The sailors accuse we of treason.

Unless adverbs are present, English puts the subject right before the verb, and the object right after it. So, this actually is another diagnostic for subject and object. The third argument is /of treason/. This is formed by using a preposition (/of/), so it is called a PP (= preposition phrase). There are as many preposition phrases as there are prepositions. We shall deal with them a little later.

A verb is called **intransitive** if it has no direct object. Otherwise it is called **transitive**. Hence, /run/ is intransitive, /accuse/ is transitive. Verbs may also be both; /break/ and /eat/ can occur both with and without a direct object. Such a verb is both transitive and intransitive; we say that it is *used* transitively if there is a direct object (as in (162)), and intransitively otherwise (as in (163)).

(162) The children are eating the cake.

(163) The children are eating.

The distinction between transitive and intransitive tells us whether or not the verb needs a direct object as an argument.

We observe that intransitive verbs are distributionally equivalent with the combination of transitive verb+direct object:

(164)
$$\left\{\begin{array}{l}\text{The child}\\ \text{Napoleon}\\ \text{My neighbour's dog}\end{array}\right\}\left\{\begin{array}{l}\text{ran.}\\ \text{lost the battle of Waterloo.}\\ \text{ate the cake.}\\ \text{is beautiful.}\end{array}\right\}$$

This tells us two things: transitive verbs plus (direct) objects form a constituent, and this constituent can be replaced by an intransitive verb.

Linguists have therefore proposed the following scheme. Constituents have two attributes: a major category and a **projection level**. The levels are 0, 1, and 2. They are also called **bar levels** (because they used to be denoted by overstrike bars). Other notation is: D^0 for determiner, 0 level; D^1, $\overline{D}$ or D′ for determiner first or **intermediate** projection; and D^2, $\overline{\overline{D}}$ or D″ for determiner level 2, or **determiner phrase**. Words are uniformly assigned level 0, and phrases are level 2. So, what we used to call a determiner (D) is now a D^0, and the representation of it and a determiner phrase (DP) is like this:

(165)
$$D^0 = \left[\begin{array}{l}\textsc{cat}:\text{D}\\ \textsc{proj}:0\end{array}\right] \qquad \text{DP} = \left[\begin{array}{l}\textsc{cat}:\text{D}\\ \textsc{proj}:2\end{array}\right]$$

The need for an intermediate level will become clear below. The following rules are proposed for English (to the right you find a more user-friendly notation *of the*

same rule):

(166) $\begin{bmatrix} \text{CAT} & :\text{V} \\ \text{PROJ} & :2 \end{bmatrix} \rightarrow \begin{bmatrix} \text{CAT} & :\text{D} \\ \text{PROJ} & :2 \end{bmatrix} \begin{bmatrix} \text{CAT} & :\text{V} \\ \text{PROJ} & :1 \end{bmatrix}$ $\quad \text{VP} \rightarrow \text{DP} \quad \text{V}'$

(167) $\begin{bmatrix} \text{CAT} & :\text{V} \\ \text{PROJ} & :1 \end{bmatrix} \rightarrow \begin{bmatrix} \text{CAT} & :\text{V} \\ \text{PROJ} & :0 \end{bmatrix}$ $\quad \text{V}' \rightarrow \text{V}^0$

(168) $\begin{bmatrix} \text{CAT} & :\text{V} \\ \text{PROJ} & :1 \end{bmatrix} \rightarrow \begin{bmatrix} \text{CAT} & :\text{V} \\ \text{PROJ} & :0 \end{bmatrix} \begin{bmatrix} \text{CAT} & :\text{D} \\ \text{PROJ} & :2 \end{bmatrix}$ $\quad \text{V}' \rightarrow \text{V}' \quad \text{DP}$

The start symbol is momentarily VP. Now, the rule (166) says that a sentence (= VP) is something that begins with a determiner phrase and then has a level 1 V. (167) is for intransitive verbs, (168) is for transitive verbs. Thus, we further add an attribute TRS with values + and −, to prevent an intransitive from taking a direct object:

(169) $\begin{bmatrix} \text{CAT} & :\text{V} \\ \text{PROJ} & :1 \end{bmatrix} \rightarrow \begin{bmatrix} \text{CAT} & :\text{V} \\ \text{PROJ} & :0 \\ \text{TRS} & :- \end{bmatrix}$

(170) $\begin{bmatrix} \text{CAT} & :\text{V} \\ \text{PROJ} & :1 \end{bmatrix} \rightarrow \begin{bmatrix} \text{CAT} & :\text{V} \\ \text{PROJ} & :0 \\ \text{TRS} & :+ \end{bmatrix} \begin{bmatrix} \text{CAT} & :\text{D} \\ \text{PROJ} & :2 \end{bmatrix}$

(171) $\begin{bmatrix} \text{CAT} & :\text{V} \\ \text{PROJ} & :0 \\ \text{TRS} & :- \end{bmatrix} \rightarrow \texttt{sit} \mid \texttt{walk} \mid \texttt{talk} \mid \cdots$

(172) $\begin{bmatrix} \text{CAT} & :\text{V} \\ \text{PROJ} & :0 \\ \text{TRS} & :+ \end{bmatrix} \rightarrow \texttt{take} \mid \texttt{see} \mid \texttt{eat} \mid \cdots$

Notice that we did not specify transitivity for V^1. This is because we have just said that a V^1 is a transitive verb plus direct object or an intransitive verb. Thus we say that a V^1 is intransitive, and there is simply no transitive V^1. (Cases like /`They call him an idiot.`/ seem like involving a verb with two direct objects. This is not so, but to make our case would take us too far afield.) Thus, in the above rules with can simply drop mentioning transitivity for V^1s.

We should mention here the fact that some verbs want sentences as arguments,

not DPs. The sentence must be opened with a complementizer, usually /that/.

(173) John believes that the earth is flat.

(174) John knows that two plus two is four.

These sentences are also called objects. They can be replaced by /this/ or /that/, showing that they are either subjects or objects. They follow the verb, so they must be objects. In the following examples they are subjects:

(175) That the earth is flat is believed by John.

(176) That the cat was eating his food annoyed the child.

Oblique Arguments and Adjuncts

Verbs can have other arguments besides the subject and the object, too. These are called **oblique**. Let us look at the verb /accuse/ again. In addition to a direct object it also wants a PP expressing subject matter. This PP must begin with the preposition /of/. There are verbs that require a PP with /on/ (for example /count/), others a PP with /about/ (for example /think/), and so on. Let us add a new attribute PREP whose value can be any of the prepositions of English. Then /accuse/ will get the following syntactic category:

(177)
$$\begin{bmatrix} \text{CAT} & :\text{V} \\ \text{PROJ} & :0 \\ \text{TRS} & :+ \\ \text{PREP} & :\texttt{of} \end{bmatrix}$$

This tells us that /accuse/ wants a subject (because all verbs do), a direct object (because it is transitive) and a PP opened by /of/ (because this is how we defined the meaning of PREP:of). To get the positioning of the phrases right, we propose to add the following rules:

(178)
$$\begin{bmatrix} \text{CAT} & :\text{V} \\ \text{PROJ} & :1 \\ \text{PREP} & :\texttt{of} \end{bmatrix} \rightarrow \begin{bmatrix} \text{CAT} & :\text{V} \\ \text{PROJ} & :1 \\ \text{PREP} & :\texttt{of} \end{bmatrix} \begin{bmatrix} \text{CAT} & :\text{P} \\ \text{PROJ} & :2 \\ \text{PREP} & :\texttt{of} \end{bmatrix}$$

(179)
$$\begin{bmatrix} \text{CAT} & :\text{P} \\ \text{PROJ} & :2 \\ \text{PREP} & :\texttt{of} \end{bmatrix} \rightarrow \begin{bmatrix} \text{CAT} & :\text{P} \\ \text{PROJ} & :1 \\ \text{PREP} & :\texttt{of} \end{bmatrix}$$

(180) $\begin{bmatrix} \text{CAT}:\text{P} \\ \text{PROJ}:1 \\ \text{PREP}:\texttt{of} \end{bmatrix} \rightarrow \begin{bmatrix} \text{CAT}:\text{P} \\ \text{PROJ}:0 \\ \text{PREP}:\texttt{of} \end{bmatrix} \begin{bmatrix} \text{CAT}:\text{D} \\ \text{PROJ}:2 \end{bmatrix}$

(181) $\begin{bmatrix} \text{CAT}:\text{P} \\ \text{PROJ}:0 \\ \text{PREP}:\texttt{of} \end{bmatrix} \rightarrow \texttt{of}$

Similarly for every other preposition. Notice that we have used the feature [PREP: `of`] also for prepositions; this makes sure that the right preposition appears at last in the structure!

The rules say that the PP is to be found to right of the direct object, if there is one. This is generally the case:

(182) `They found the gold in the river.`

(183) *`They found in the river the gold.`

(184) `The pilot flew the airplane to Alaska.`

(185) *`The pilot flew to Alaska the airplane.`

Indeed, we can observe that the above rules hold for arbitrary prepositional phrases, so we simply replace /on/ by a placeholder, say α. The rules then look like this. We show the Rule (178) only, which becomes (186).

(186) $\begin{bmatrix} \text{CAT}:\text{V} \\ \text{PROJ}:1 \\ \text{PREP}:\alpha \end{bmatrix} \rightarrow \begin{bmatrix} \text{CAT}:\text{V} \\ \text{PROJ}:1 \\ \text{PREP}:\alpha \end{bmatrix} \begin{bmatrix} \text{CAT}:\text{P} \\ \text{PROJ}:2 \\ \text{PREP}:\alpha \end{bmatrix}$

The idea is that in this rule schema, α may be instantiated to any appropriate value. In this case the appropriate values are the English prepositions. If we choose α, every occurrence of α must be replaced by the same value. For example, the following is *not* a correct instance of (186):

(187) $\begin{bmatrix} \text{CAT}:\text{V} \\ \text{PROJ}:1 \\ \text{PREP}:\texttt{about} \end{bmatrix} \rightarrow \begin{bmatrix} \text{CAT}:\text{V} \\ \text{PROJ}:1 \\ \text{PREP}:\texttt{on} \end{bmatrix} \begin{bmatrix} \text{CAT}:\text{P} \\ \text{PROJ}:2 \\ \text{PREP}:\texttt{about} \end{bmatrix}$

Notice that the PP does not pass directly to P^0+DP, there is an intermediate level 1 projection. (The reason is not apparent from the data given so far, and some syntacticians dispute whether things are this way. However, for our purposes it makes

the syntax more homogeneous.) Notice also that the level does not decrease. A verb plus PP has the same level as the verb itself, namely 1. Therefore the PP is called an adjunct. Adjuncts do not change the level of the projection, but direct arguments do.

X-Bar Syntax

The structure of PPs looks similar to that of transitive verb phrases, except that the subject is generally missing. This similarity exists across categories. We take as an example NPs. There exist not only nouns as such, but nouns too take arguments. These arguments are mostly optional, which is to say that nouns can be used also without them.

(188) `the counting of the horses`

(189) `some president of the republic`

(190) `a trip to the Philippines`

(191) `this talk about the recent events`

In English, nouns cannot take direct arguments. The verb /`count`/ is transitive (/`count the horses`/), but the gerund wants the former object in a PP opened by /`of`/.

We can account for this in the same way as we did for verbs. We allow nouns to additionally have a feature [PREP : `of`], or [PREP : `in`], and so on. And if they do, they can (but need not) add a PP opened by the corresponding preposition.

(192) $$\begin{bmatrix} \text{CAT} & :\text{N} \\ \text{PROJ} & :1 \\ \text{PREP} & :\texttt{of} \end{bmatrix} \rightarrow \begin{bmatrix} \text{CAT} & :\text{N} \\ \text{PROJ} & :0 \\ \text{PREP} & :\texttt{of} \end{bmatrix} \begin{bmatrix} \text{CAT} & :\text{P} \\ \text{PROJ} & :2 \\ \text{PREP} & :\texttt{of} \end{bmatrix}$$

(193) $$\begin{bmatrix} \text{CAT} & :\text{N} \\ \text{PROJ} & :1 \\ \text{PREP} & :\texttt{of} \end{bmatrix} \rightarrow \begin{bmatrix} \text{CAT} & :\text{N} \\ \text{PROJ} & :0 \\ \text{PREP} & :\texttt{of} \end{bmatrix}$$

These two rules are abbreviated as follows.

(194) $$\begin{bmatrix} \text{CAT} & :\text{N} \\ \text{PROJ} & :1 \\ \text{PREP} & :\texttt{of} \end{bmatrix} \rightarrow \begin{bmatrix} \text{CAT} & :\text{N} \\ \text{PROJ} & :0 \\ \text{PREP} & :\texttt{of} \end{bmatrix} \left(\begin{bmatrix} \text{CAT} & :\text{P} \\ \text{PROJ} & :2 \\ \text{PREP} & :\texttt{of} \end{bmatrix} \right)$$

The brackets around an item say that it is optional. For completeness, let us note that we have the following rules.

(195) $\begin{bmatrix} \text{CAT}:\text{N} \\ \text{PROJ}:0 \\ \text{PREP}:\texttt{of} \end{bmatrix} \rightarrow \texttt{counting} \mid \texttt{brother} \mid \texttt{election} \mid \texttt{chief} \mid \cdots$

(196) $\begin{bmatrix} \text{CAT}:\text{N} \\ \text{PROJ}:0 \\ \text{PREP}:\texttt{about} \end{bmatrix} \rightarrow \texttt{doubt} \mid \texttt{question} \mid \texttt{talk} \mid \texttt{rumour} \mid \cdots$

We note in passing the following. A PP opened by /about/ has (mostly) the meaning "concerning". It indicates the person or thing in the centre of attention. With nouns denoting feelings, opinions, and so on it feels natural that the preposition /about/ is used, so we are inclined to think that the grammar does not have not tell us this. One indication is that /about/ can be replaced by /concerning/. Unfortunately, this does not always give good results:

(197) `doubts concerning the legitimacy of the recall`

(198) `questions concerning the statement by the police officer`

(199) *`talk concerning the recent book release`

Thus, although the intuition seems generally valid, there also is a need to record the habits of each noun as to the PP that it wants to have.

Now we take a bold step and abstract from the categories and declare that English has the following general rules:

(200a) $\text{XP} \rightarrow (\text{YP}) \quad \text{X}'$

(200b) $\text{XP} \rightarrow \text{XP} \quad \text{YP}$

(200c) $\text{X}' \rightarrow \text{YP} \quad \text{X}'$

(200d) $\text{X}' \rightarrow \text{X}' \quad \text{YP}$

(200e) $\text{X}' \rightarrow \text{X}^0 \quad (\text{YP})$

Translated into words this says: phrasal adjuncts are always on the right, while first level adjuncts are either on the right or on the left. The projection of X to the right is called the **head**. The YP in Rule (200a) is called the **specifier** of the phrase, the one in Rule (200e) is called the **complement**; finally, YP in Rules (200b), (200c) and (200d) is called an **adjunct**. Subjects are specifiers of verbs, direct objects complements of verbs. We note the following general fact about English.

> In English, specifiers are on the left, complements on the right of the head.

All these terms are *relational*. An occurrence of a phrase is not a complement per se, but the complement of a particular phrase XP in a particular construction.

If we look at other languages we find that they differ from English typically only the relative position of the specifier, complement and adjunct, not the hierarchy. So, they will also make use of the following rules absent from the list above:

(201a) $\mathrm{XP} \to \mathrm{X}' \quad \mathrm{YP}$

(201b) $\mathrm{XP} \to \mathrm{YP} \quad \mathrm{XP}$

(201c) $\mathrm{X}' \to \mathrm{YP} \quad \mathrm{X}^0$

Moreover, the direction changes from category to category. German puts the verb to the end of the clause, so the complement is to its left. The structure of nouns and PPs is however like that of the English nouns. Japanese and Hungarian put prepositions at the end of the PP. (That is why they are called more accurately **postpositions**. In general, and **adposition**) is either a preposition or a postposition. Sanskrit puts no restriction on the relative position of the subject (= specifier of verb) and object (= complement of the verb).

> **Sentence Structure I.** (*Universal*) All syntactic trees satisfy X-bar syntax. This means that Rules (200a) – (201c) hold for all structures.

It is important to be clear about the sense of "holds" in this principle. It is take to be here in the sense of . A set of rules R of the form $X \to YZ$ or $X \to Y$ is said to **hold** for a labelle tree (or a set thereof) if whenever a node in a tree has category X, then it either has no daughter, or only one daughter with label Y and the rule $X \to Y$ is in R, or it has two daughters with labels Y and Z (in that order) and the rule $X \to YZ$ is in R. This looks like the definition of last chapter, but it differs because here we are not saying how we get the trees (important when we study movement) and we also do not require that there exists a context free grammar for the trees (the language can be far more complex, but still the rules are said to hold).

Phrases and Pro-Forms

I began the lecture by citing different words and their categories. It emerged from this list that words invariably a lexical and project a phrase. This is true in general, but there are important exceptions. The first class is the **pro-forms**. The pro-forms are such words that actually stand in for an unspecified word or phrase. One class are the pronouns, which fact replace a DP. So we should actually call them pro-DPs (the name *pronoun* was coined before one thought of DPs). Other examples are /one/ (pro-NP), /do/ (pro-VP).

(202) `John wrote a good essay. Paul also wrote a good one.`

(203) `John wrote an essay and Paul did as well.`

In that connection we should also mention the question words: /who/, /what/, /where/, /why/ and so on, and the deictics (/this/, /thus/, /so/ and so on) which also are phrasal. /who/ and /what/ are DPs, /why/ and /where/ are PPs.

The Global Structure of Sentences

We shall now come to an important fact.

> **Sentence Structure II.** (*Universal*) The start symbol is CP.

Constituents that are CPs are often called **clauses**. Sentences have a global structure. In English, it looks like this:

(204) [XP [C^0 [YP [T^0 VP]]]]

This means the following. The leftmost element is the specifier of CP. Then follows the complementizer (C^0), and then the complement of the complementizer. It in turn has the following structure. First comes the specifier of T(ense). Then the T^0 and then the VP. We will see later why we need all these parts in the sentence. The largest CP, which contains the entire sentence is called the **matrix clause**.

Notes on this section. Since PPs are adjuncts, they can be repeated any number of times. But the PPs that are selected by the verb cannot be repeated:

(205) *`The sailors accused the captain of treason of cruelty.`

The feature system of this section allows for arbitrary repetition, however. This can be prevented. What is really needed is a feature system that adds the requirements for each occurring object and then checks them against those of the verb. I shall not work out the details here, as they are not revealing (and not pretty either). This ultimately leads to a system called **Generalised Phrase Structure Grammar** (**GPSG**), proposed in [Gazdar *et al.*, 1985], which later developed into **Head Driven Phrase Structure Grammar** (**HPSG**, see [Pollard and Sag, 1994]).

Syntax III: Local Dependencies and Constraints: Selection, Agreement and Case Marking

> The organisation of phrases is as much a matter of placement as a matter of morphological marking. When a head (for example a verb) wants an argument it also determines some of its morphological features. Selection is the fact that it wants and argument; case is the feature that identifies a phrase as a particular argument of the head. And finally agreement is something that both head and argument share.

Grammatical Features

Nouns come in two varieties: there are singular and plural nouns. Singular and plural are called **numbers**. We have already mentioned a few ways in which the singular and plural of nouns are formed. For syntax the different ways of forming the plural are not relevant; only the fact whether a noun is singular or plural is relevant. A way to represent the number of a noun is by adding an attribute NUM, with values sing and pl (in other languages there will be more...). So we have that the **syntactic representation** of 'mouse' and 'mice' is

(206) $$\texttt{mouse}:\begin{bmatrix}\text{CAT}:\text{N}\\ \text{PROJ}:0\\ \text{NUM}:\text{sing}\end{bmatrix},\texttt{mice}:\begin{bmatrix}\text{CAT}:\text{N}\\ \text{PROJ}:0\\ \text{NUM}:\text{pl}\end{bmatrix}$$

This we rephrase by using the following rules:

(207) $$\begin{bmatrix}\text{CAT}:\text{N}\\ \text{PROJ}:0\\ \text{NUM}:\text{sing}\end{bmatrix}\rightarrow \texttt{mouse}$$

(208) $$\begin{bmatrix}\text{CAT}:\text{N}\\ \text{PROJ}:0\\ \text{NUM}:\text{pl}\end{bmatrix}\rightarrow \texttt{mice}$$

An alternative notation, which is used elsewhere and should now be self-explanatory, is DP[sing] and DP[pl]. (The attribute NUM is omitted, because sing and pl are only values of that attribute not of any other.) We shall use this type of notation without further warning. It is to be thought of as abbreviatory only.

Also noun phrases in English, like nouns, can be both singular and plural. Whether or not an NP is singular or plural can be seen by looking at the head noun.

(209) `fearful warrior` (singular)

(210) `fearful warriors` (plural)

(211) `brother of the shopkeeper` (singular)

(212) `brothers of the shopkeeper` (plural)

In the first example the head noun is /`warrior`/, in the second it is /`warriors`/. This indicates that in the first example the NP is singular, and plural in the second. Thus, the number feature is passed up, so to speak, to the phrase from the head. Notice that the determiner is different depending on whether the NP is singular or plural.

(213) `this fearful warrior`

(214) *`these fearful warrior`

(215) *`this fearful warriors`

(216) `these fearful warriors`

We say that like nouns, determiners have a singular and a plural form and that the determiner **agrees** with the NP complement in number. This will be put into the rules as follows. The letter α may be instantiated to any legitimate value, in this case sing or pl. Notice that within a single rule, each occurrence of α must be replaced by the same value. However, whether or not the letter α also occurs in another rule is of no significance.

(217) $$\begin{bmatrix} \text{CAT}:\text{D} \\ \text{PROJ}:1 \\ \text{NUM}:\alpha \end{bmatrix} \rightarrow \begin{bmatrix} \text{CAT}:\text{D} \\ \text{PROJ}:0 \\ \text{NUM}:\alpha \end{bmatrix} \begin{bmatrix} \text{CAT}:\text{N} \\ \text{PROJ}:2 \\ \text{NUM}:\alpha \end{bmatrix}$$

(218) $$\begin{bmatrix} \text{CAT}:\text{D} \\ \text{PROJ}:0 \\ \text{NUM}:\text{sing} \end{bmatrix} \rightarrow \texttt{this} \mid \texttt{the} \mid \texttt{a} \mid \cdots$$

(219) $$\begin{bmatrix} \text{CAT}:\text{D} \\ \text{PROJ}:0 \\ \text{NUM}:\text{pl} \end{bmatrix} \rightarrow \texttt{these} \mid \texttt{the} \mid \varnothing \mid \cdots$$

Here, $\varnothing$ is the empty string. It is needed for indefinite plurals (the plural of /`a car`/ is /`cars`/).

The choice of the determiner controls a feature of DPs that is sometimes syntactically relevant: definiteness. DPs are said to be **definite** if, roughly speaking, they refer to a specific entity given by the context or if they are uniquely described by the DP itself. Otherwise they are called **indefinite**. The determiner of definite DPs is for example /this/, /that/ and /the/. The determiner of indefinite DPs is for example /a/ in the singular, /some/ in singular and plural, and ∅ in the plural.

In other languages, nouns have more features that are syntactically relevant. The most common ones are **case** and **gender**. Latin, for example, has three genders, called **masculine**, **feminine** and **neuter**. In English the three genders have survived in the singular 3rd person pronouns: /he/, /she/, /it/. The Latin noun /homo/ ('man') is masculine, /luna/ ('moon') is feminine, and /mare/ ('sea') is neuter. The adjectives have different forms in each case (notice that the adjective likes to follow the noun, but it does not have to in Latin):

(220) homo ruber (red man)

(221) luna rubra (red moon)

(222) mare rubrum (red sea)

Nouns have many different declension classes (morphology deals with them, syntax does not), and there are rules of thumb as to what declension class nouns have which gender, but they can fail. Nouns ending in /a/ are generally feminine, but there are exceptions (/nauta/ 'the seafarer', /agricola/ 'the farmer' are masculine). Similarly, adjectives have many declension paradigms, so the forms can vary. But for each adjective there are forms to go with a masculine noun, forms to go with a feminine noun, and forms to go with a neuter noun. We therefore say that the adjective **agrees** in gender with the noun. It is implemented by installing a new attribute GENDER, whose values are 'm(asculine)', 'f(eminine)' and 'n(euter)'. English is said not to have any forms of gender agreement. Moreover, English gender is said to be semantic, that is, based on what a noun denotes. To know whet.her you should use /he/, /she/ or /it/ you have to ask yourself whether the referent is male, female, or neither. (The case where sex is unknown (or is not to be revealed by the speaker) is another case, and it is different from the case whether the sex is known, or in fact the notion of sex does not apply, for example in the case of cars. In some variants of English, for example, /they/ is used to avoid communicating the sex, as opposed to /it/ when the notion does not apply.) The semantics of gender is not fully congruent with sex; ships and other vessels are an exception in that you must use /she/ to refer to them, even though they have no sex. This does not mean, though, that gender is not semantic; it means

that its meaning is not fully determined by sex. A different case is presented by German, where gender is demonstrably not fully semantic. All diminutives are neuter, regardless of what they refer to. In referring to a small man, for example, you can either say /der kleine Mann/ ('the-masc.nom.sg small man') or /das Männchen/ ('the-neut.sg.nom man-dimin').

Latin nouns also have different **cases**. There are five cases in Latin: nominative (for subjects), accusative (for direct objects), dative, genitive and ablative. Just as verbs select PPs with a particular preposition in English they can also select a DP with a particular case. If it is accusative the verb is transitive; but it can be dative (/placere/+DP[dat] 'to please someone'), ablative (/frui/+DP[abl] 'to enjoy something'), and genitive (/meminisse/+DP[gen] 'to remember someone/something'). There exist verbs that take several DPs with various cases. For example, /inferre/ 'to inflict' (with perfect /intuli/) wants both a direct object and a dative DP.

(223) Caesar Gallis bellum intulit.
Caesar Gauls-dat war-acc inflict.upon-perf
Caeser inflicted war on the Gauls.

This is just like English /inflict/, that wants a direct object and a PP[on].

For us, a **grammatical feature** is anything that defines a syntactic category. You may think of it as the syntactic analogue of a phonemic feature. However, beware that elsewhere the usage is a little different. Grammatical category and projection level are typically discarded, so that grammatical feature refers more to the kinds of things we have introduced in this section above: number, gender, definiteness, and case. A fifth one is *person*. (This list is not exhaustive.)

The position of subject can be filled with DPs like /the mouse/, /a car/, but also by so-called **pronouns**. Pronouns are distinct from nouns in that they express little more than that they stand in for a DP. However, they have many different forms, depending on the grammatical feature. In addition they show a distinction in **person**. Across languages, there is a pretty universal system of three persons: 1st, 2nd, and 3rd. First person means: includes speaker. Second person means: includes hearer, and third person means: includes neither speaker nor hearer. So, /I/ is first person (it includes me and no one else); it is also singular, because it refers to just one thing. The plural /we/ is used when one refers to several people, including speaker. By contrast, /you/ is used for individuals or groups

including the hearer (there is no distinction between singular and plural). The third person pronouns distinguish also gender in the singular (/he/, /she/, /it/), but not in the plural (/they/). Moreover, as we explained above, pronouns distinguish nominative from accusative. Thus, there are more morphological distinctions in the pronominal system in English than there is in the ordinary DPs.

More on Case

Cases have many different functions. One function is to indicate the nature of the argument. A verb has a subject, and the case of the subject is referred to as **nominative**. The direct object has a case that is referred to as **accusative**. Some people believe that English has cases (because pronouns still reflect a distinction between subject and object: /she/ : /her/, /he/ : /him/, and so on). On the other hand, this is confined to the pronouns, and nouns show no distinction in case whatsoever. This is why we say that English has *no case*. Strictly speaking, it means only that there is no distinction in case. (One might say: there is one case and only one. This is useful. For example, there is a famous principle of syntactic theory which states that nouns need case. If there is no case, this principle fails.) Chinese is another example of a language that has no cases. These languages make no distinction between subject and object in form; nevertheless, one can tell the difference: the subject precedes the verb, and the object follows it (both in English and in Chinese).

Verbs can have more than two arguments, and many more adjuncts. To distinguish between them some kind of marking is needed. In English this is done by means of prepositions. For example, there often is an argument towards which the action is directed or for which it is performed (the 'goal') and it is given by a PP opened by /to/ (= PP[to], for example /talk to someone/). The goal is also called **indirect object**. Latin has a case for this, the **dative**. There is from a global viewpoint not much of a difference whether the goal is encoded by a case or by a PP. Languages can choose which way to go.

Another important case is the **genitive**. It marks possession. English has basically two ways to mark possession (apart from obvious ones like /which belongs to/). One is the so-called **Anglo-Saxon genitive**, formed by adding an /'s/ (/my neighbour's car/). The other is a PP opened by /of/ (/the car of my neighbour/). The genitive is used a lot in English. Nouns that have ar-

guments that are not PPs put them in the genitive:

(224) the election of the chancellor

(225) the peak of the mountain

(226) Napoleon's destruction of the city

Notice that two of these nouns have been obtained from transitive verbs. The rule in English is that the noun can take any of the arguments that the verb used to take (though they are now optional). However, the subject and the object must now appear in the genitive. The PPs on the other hand are taken over as is. For example, /destroy/ is transitive, so the noun /destruction/ can take two genitives, one for the subject and one for the object. The verb /talk/ takes a subject, an indirect object and subject matter, expressed by a PP headed by /about/. The latter two are inherited as is by the noun /talk/, while the subject is put in the genitive. Alternatively, it can be expressed by a PP[by].

(227) John talked to his boss about the recent layoffs.

(228) John's talk to his boss about the recent layoffs

(229) talk by John to his boss about the recent layoffs

Subject-Verb Agreement

English displays a phenomenon called **subject-verb agreement**. This means that the form of the verb depends on the grammatical features of the subject. The agreement system is very rudimentary; the only contrast that exists is that between singular 3rd and the rest:

(230) She runs.

(231) They run.

Notice that since the verb does agree in person with the subject it has to make a choice for DPs that are not pronominal. It turns out that the choice it makes is that ordinary DPs trigger 3rd agreement:

(232) The sailor runs.

This applies even when the DP actually refers to the speaker! So, agreement in person is (at least in English) not only a matter of what is actually talked about, but it is also a syntactic phenomenon. There are rules which have to be learned.

Other languages have more elaborate agreement systems. Let us look at Hungarian. The verbal root is /lát/ 'to see'.

(233)	Én látok.	Mi látunk.
	I see	We see
	Te látsz.	Ti látatok.
	You(sg) see	You(pl) see
	Ö lát.	Ök látuk.
	He/she/it sees	They see

Hungarian has no distinction whatsoever in gender (not even in the pronominal system; /ö/ must be rendered by 'he', or 'she', or 'it', depending on what is talked about). However, it does distinguish whether the direct object is definite or not. Look at this (translation is actually word by word):

(234) Én látok egy madarat.
I see a bird

(235) Én látom a madarat.
I see the bird

The subject is the same in both sentences, but the object is indefinite in the first (a bird) and definite in the second (the bird). When the object is indefinite, the form /látok/ is used, to be glossed roughly as 'I see', while if the object is definite, then the form /látom/ is used, to be glossed 'I see it'. Hungarian additionally has a form to be used when subject is first person singular and the direct object is 2nd person:

(236) Én látlak.
I see-sub:1sg.ob:2sg

(The dot is used to say that the form /lak/ is a single morpheme expressing to syntactic agreement facts.)

Who Agrees with Whom in What?

Agreement is pervasive in some languages, and absent in others. Chinese has no agreement whatsoever, English has next to none. The most common type of agreement is that of verbs with their subjects. Some languages even have the verb agree

with the direct object (Hungarian, Mordvin (a language spoken in Russia, related to Hungarian), Potawatomi (an American Indian language)). Other languages have the verb agree in addition with the indirect object (Georgian). Agreement is typically in person and number, but often also in gender. Above we have seen that definiteness can also come into the picture. Adjectives sometimes agree with the nouns they modify (Latin, German, Finnish), sometimes not (Hungarian). There is no general pattern here. This is one of the things that one has to accept as it is.

Obligatoriness

Morphology is obligatory. A noun has a case in a language with cases, it has a number in a language with numbers, and so on. This creates a certain predicament about which I have spoken earlier in connection with gender. Let me here discuss it in connection with number. The intuitive approach is this. An English noun is either singular or plural. If it is singular it means 'one of a kind' and if it is plural it means 'several of a kind'. So, /`car`/ means 'one car', and /`cars`/ means 'several cars'. But this immediately raises questions.

① Are there remaining cases to be dealt with?

② What is the notion of number does not apply?

③ What if the speaker cannot (or does not want to) communicate the number of things involved?

I deal with these questions in turn.

①. The most obvious case that we left out is 'zero of a kind'; but others also come to mind, such as fractions, and negative numbers.

②. Mass nouns (/`water`/, /`iron`/), which we shall discuss later, have no obvious plural, since we cannot count what they refer to.

③. The most evident situation is a question; but there are many circumstances in which we do not know how many, such as if there crime and we talk about the people who did it.

In all of these cases, the language comes up with more or less explicit recipes as to how to solve the problem. These fixes are to some degree arbitrary. Consider

the first point:

(237) `They gave me a quarter point for my answer.`

(238) `They gave me 0.25 points for my answer.`

(239) `They gave me 1 1/4 points for my answer.`

(240) `They gave me 1.25 points for my answer.`

What we see is that decimals expansions are treated differently from fractions (less semantically).

When you ask for the identity of some person, you often do not know how many there are. In English you do not have to commit yourself, you use /who/. However, even though the number is unknown, /who/ itself triggers singular agreement in the verb.

(241) `Who drank my coffee?`

In Hungarian, there are two question words, a singular /ki/ 'who-sg' and a plural /kik/ 'who-pl'. If you use the latter you indicate you expect the answer to be several. If you are not using a question, the plural sounds less committal than the singular, for example (242) commits you just one, while (243) does not.

(242) `The murderer of the shopkeeper.`

(243) `The murderers of the shopkeeper.`

The important lesson to learn is that all categorisations leak. There simply is no overarching system to classify everything. Consequently, languages must implement a few strategies of dealing with the unknown cases.

Feature Percolation

When we look at the distribution of number in English, we find that we get the following rules:

(244) $\text{DP}[\text{NUM} : \alpha] \rightarrow (\text{YP}) \quad \text{D}'[\text{NUM} : \alpha]$

(245) $\text{DP}[\text{NUM} : \alpha] \rightarrow \text{YP} \quad \text{DP}[\text{NUM} : \alpha]$

(246) $\text{DP}[\text{NUM} : \alpha] \rightarrow \text{DP}[\text{NUM} : \alpha] \quad \text{YP}$

(247) $\text{D}'[\text{NUM} : \alpha] \rightarrow \text{D}[\text{NUM} : \alpha] \quad \text{YP}[\text{NUM} : \alpha]$

(248) $\text{D}'[\text{NUM} : \alpha] \rightarrow \text{D}[\text{NUM} : \alpha]$

(249) $\text{D}'[\text{NUM} : \alpha] \rightarrow (\text{YP}) \quad \text{D}'[\text{NUM} : \alpha]$

(250) $\text{D}'[\text{NUM} : \alpha] \rightarrow \text{D}'[\text{NUM} : \alpha] \quad \text{YP}$

This is the same also for the projection of N, and V. There is in general no connection between the features on the YP and that of the D, with one exception: the determiner passes the number feature on to the NP complement. It is the same also with gender features, and with case. Might this therefore be a general feature of the X-bar syntax?

To answer the question let me first draw attention to a case that is not of the form. Recall the rules (169) and (170). What we did not establish then was the identity of the V′ with respect to its transitivity. We use the following diagnostic: in a coordination only constituents which have the same features can be coordinated. If this is so then the following sentence would be ungrammatical if we called /`kicked the ball`/ as a transitive V′.

(251) `John kicked the ball and fell.`

The solution is to class both as intransitive V′. Thus the rules are as follows.

(252) $\begin{bmatrix} \text{CAT} : \text{V} \\ \text{PROJ}: 1 \\ \text{TRS} : - \end{bmatrix} \rightarrow \begin{bmatrix} \text{CAT} : \text{V} \\ \text{PROJ}: 0 \\ \text{TRS} : - \end{bmatrix}$

(253) $\begin{bmatrix} \text{CAT} : \text{V} \\ \text{PROJ}: 1 \\ \text{TRS} : - \end{bmatrix} \rightarrow \begin{bmatrix} \text{CAT} : \text{V} \\ \text{PROJ}: 0 \\ \text{TRS} : + \end{bmatrix} \begin{bmatrix} \text{CAT} : \text{D} \\ \text{PROJ}: 2 \end{bmatrix}$

The full story is to distinguish two kinds of features: **selectional features** (like TRS and PREP) and **agreement features**. Selectional features are not passed from head to phrase, while agreement features are.

It is to be noted, though, that coordination does *not* follow X-bar syntax (we have discussed the schema above). Furthermore, agreement features are *exempt* from the identity requirement in coordination:

(254) `John and Mary danced.`

(255) `William and the firemen saluted the governor.`

Syntax IV: Movement and Non-Local Dependencies

> Even though there is a way to account for the syntax of questions in terms of context free rules, by far the most efficient analysis is in terms of *transformations*. Typically, a transformation is the movement of a constituent to some other place in the tree. This lecture explores how this works and some of the conditions under which this happens.

Movement

We have learned that in English the transitive verb requires its direct object immediately to its right. This rule has a number of exceptions. The first sentence below, (256), displays a construction known as **topicalisation**, the second, (257), is a simple question using a question word.

(256) `Air pilots Harry admires.`

(257) `Which country have you visited?`

We could of course give up the idea that the direct object is to the right of the verb. It would be possible to argue for topicalisation that the sentence has the structure Object-Subject-Verb. But the facts are quite complex, especially when we turn to questions. For example, no matter what kind of constituent the question word replaces (subject, object, indirect object and so on), it is at the first place even if it is not the subject.

(258) `Alice has visited Madrid in spring to learn Spanish.`

(259) `What has Alice visited in spring to learn Spanish?`

(260) `Who has visited Madrid in spring to learn Spanish?`

(261) `When has Alice visited Madrid to learn Spanish?`

(262) `Why has Alice visited Madrid in spring?`

We see that the sentences involving question words differ from (258) in that the question word is in first place and the verb in second place. There is a way to arrive at a question in the following way. First, insert the question word where it ought to belong according to our previous rules. Next, take it out and put it in first

position. Now move the auxiliary (/has/) into second place. We can represent this as follows, marking removed elements in red, and newly arrived ones in blue:

(263) Alice has visited what in spring to learn Spanish?

(264) What Alice has visited what in spring to learn Spanish?

(265) What has Alice has visited what to learn Spanish?

A more standard notation is this:

(266) Alice has visited what in spring to learn Spanish?

(267) What Alice has visited __ in spring to learn Spanish?

(268) What has Alice __ visited __ to learn Spanish?

(The underscore just helps you to see where the word came from. It is typically neither visible nor audible.) The first notation is more explicit in showing you which element came from where (assuming they are all different). However, neither notation reveals the order in which the movements have applied. It turns out, though, that this is irrelevant anyhow. (The tree structures do not reveal much about the derivation of a sentence either, and that is mostly considered irrelevant detail anyway.)

The good side about this proposal is that it is actually simple. However, we need to be clear about what exactly we are doing. For it seems that what we do is derive one surface sentence of English from another surface sentence of English. (This was in fact the way transformations were originally thought of.) Here however we pursue a different conception, namely that every surface sentence of English is derived in a two stage process. First, we generate a structure in which every head projects its phrase, satisfying its requirements (such as verbs projecting a VP consisting of a subject and, if transitive, an object, and the necessary PPs). After that is taken care of, we shall apply a few transformations to the structure to derive the actual surface string. This is what we have done with phonological representations, too. First we have generated deep representations and then we have changed them according to certain rules. Thus, we say that the context free grammar generates **deep syntactic representations**, but that the rules just considered operate on them to give a final output, the **surface syntactic representation**. The rules are also referred to as (**syntactic**) **transformations**.

Wh-Movement

Let us investigate the properties of the so-called **Wh-Movement**. This is the transformation which is responsible to put the question word in front of the sentence. Question words are also referred to as **wh-words**, since they all start with /wh/ (/who/, /what/, /where/, /why/, etc.). At first blush one would think that syntactic transformations operate on strings; but this is not so. Suppose the original sentence was not (258) but

(269) Alice has visited which famous city in Mexico to wait for her visa?

Then the output we expect on this account is (270). But it is ungrammatical. Instead, only (271) is grammatical.

(270) *Which has Alice __ visited __ famous city in Mexico to wait for her visa?

(271) Which famous city in Mexico has Alice __ visited __ to wait for her visa?

It is the entire DP that contains the question word that goes along with it. There is no way to define that on the basis of the string, instead it is defined on the basis of the tree. To see how, let us note that the sentence (258) has the following structure. (Some brackets have been omitted to enhance legibility.)

(272) Alice [has [[visited [which famous city in Mexico][to wait for her visa]]]]?

Now, /which famous city in Mexico/ is a constituent (it passes for example the tests ① – ④). Moreover, it is the object of the verb /visited/. The word /which/ is a determiner, and the smallest phrase that contains it is the one that has to move.

> **Wh-Movement I.** (*Preliminary*)
> Only *phrases* can be moved by Wh-Movement. What moves is the least phrase containing a given wh-word. It moves to the beginning of a clause (= CP).

This specification is imprecise at various points. First, what happens if there are several wh-words? In English only one of them moves and the others stay in place; the choice of the one to move is a bit delicate, so we shall not deal with that question here. In other languages (Rumanian, Bulgarian, Hungarian are examples) *all of them move*. Second, what happens if the wh-word finds itself inside a sentence that is inside another sentence? Let us take a look.

(273) `Mary thinks you ought to see what city?` (deep structure)

Here the wh-phrase moves to end of the higher sentence (and notice that something strange happens to the verb too):

(274) `What city does Mary think you ought to see?`

However, some verbs dislike being passed over. In that case the wh-phrase ducks under; it goes to the left end of the lower sentence.

(275) `*What city does Mary wonder you have seen?`

(276) `Mary wonders what city you have seen.`

So, let us add another qualification.

> **Wh-Movement II.** (*Preliminary*)
> The wh-phrase moves to the beginning of the leftmost phrase possible.

We shall see further below that this is not a good way of putting things, since it refers to linear order and not to hierarchical structure.

Verb Second

Wh-movement is a movement of phrases. In addition to this there is also movement of zero level projections, or **heads**. (It is therefore called **head movement**.) A particular example is verb movement. Many languages display a phenomenon called **Verb Second** or **V2**. German is among them. Unlike English, the verb is

not always in second place. Here is a pair of sentences with word-to-word translation.

(277) `Hans geht in die Oper.`
Hans goes into the opera

(278) `Der Lehrer ist erfreut, weil Hans in die Oper geht.`
the teacher is pleased, because Hans into the opera goes

The main verb is /`geht`/ ('goes'). In the first example it is in second place, in the second example it is at the end of the sentence. Notice that in the second example there is a CP which is opened by /`weil`/ ('because'). It is called **subordinate**, because it does not display the same kind of order as a typical clause. Now one may suspect that the verb simply occupies a different place in subordinate clauses. However, if we look at an auxiliary plus a main verb, matters start to become more complex.

(279) `Hans will in die Oper gehen.`
Hans wants into the opera go

(280) `Der Lehrer ist erfreut, weil Hans in die Oper gehen will.`
the teacher is pleased, because Hans into the opera go wants

Only the auxiliary (/`will`/) is found in the second place in the main clause. More facts can be adduced to show the following. The verb is at the end of the clause in deep structure. In a subordinate clause it stays there. Otherwise it moves to second position.

Now, what exactly is 'second position'? It cannot be the second word in the sentence. In the next example it is the fifth word (the dot in the transcription shows that /`im`/ is translated by two words: 'in' and 'the').

(281) `Der Frosch im Teich ist kein Prinz.`
the frog in.the pond is no prince

Obviously, it is not the second word, it is the second *constituent* of the sentence. Once again we find that the operation of movement is not described in terms of strings but in terms of the structure.

English has a phenomenon similar to this. Notice that I said the subject is in the specifier of VP. That works fine for a simple tense but creates word order

problems for complex tenses. For we would like to posit that /will/ is a tense head (so here is one of the words that go into T^0). If that is so, the basic structure is this.

(282) [[∅ [[will [John [see Alice]$_{V'}$]$_{VP}$]$_{T'}$]$_{TP}$]$_{C'}$]$_{CP}$

This gives us the following base order.

(283) will John see Alice

This time it is the subject, however, that moves to specifier of TP, so that we get

(284) [[∅ [[John will [__ see Alice]$_{VP}$]$_{T'}$]$_{TP}$]$_{C'}$]$_{CP}$

This is now extended to all kinds of clauses as follows. We propose that T^0 consists of tense and subject agreement. How can this be? In the future tense we say, for example, /will see/, while in the present tense we say /sees/. Blurring the distinction between syntax and morphology, we segment /sees/ into /see/ and /s/, the latter the exact equivalent of /will/. We propose therefore that /will/ as well as /s/ are T^0 heads.

Thus, by analogy, the original present tense sentence is this:

(285) [[∅ [[s [John [see Alice]$_{V'}$]$_{VP}$]$_{T'}$]$_{TP}$]$_{C'}$]$_{CP}$

Thus we discover another reason for movement: morphological integrity. The agreement morpheme /s/ cannot survive by itself, and the verb moves up to attach to it.

(286) [∅ [∅ [[see+s [John [__ Alice]$_{V'}$]$_{VP}$]$_{T'}$]$_{TP}$]$_{C'}$]$_{CP}$

We shall concern ourselves with the exact structure of /see+s/.

Thus, we propose that across the board the V^0 is void of any tense and inflectional feature; while the future tense and negation require a separate T^0 in the overt sentence, present tense and past tense only have one word. For example, since /saw/ is the past tense of /see/, we analyse it as /see+PAST+3Sg/. The sequence /PAST+3Sg/ is hosted in T^0, and the root /see/ moves up to compose with it.

(287) [[PAST+3Sg [John see Mary]$_{VP}$]$_{T'}$]$_{TP}$

(288) [[see+PAST+3Sg [John __ Mary]$_{VP}$]$_{T'}$]$_{TP}$

(289) [John [see+PAST+3Sg [__ __ Mary]$_{VP}$]$_{T'}$]$_{TP}$

There are various other reasons why this is a good analysis, having to do with placement of negation and adverbs. On the other hand, it raises questions about the compatibility of categories; for we require that after movement the categorial structure is still in accordance with X-bar syntax, so V must still project a VP, not a CP. I spell this principle out in full.

> **Categorial Transparency**
> The structure obtained after movement must conform to X-bar syntax.

Since the movement of the verb into T^0 does not change any projections (it is just movement of a constituent), we will retain a T' node, and so the head of this phrase must still be a T. There is a fix for this problem, which I shall discuss below (but it is not part of this lecture). Notice that that the T^0 contains material in it. Moreover, when the verb moves, it does not replace the material but rather adds itself on the left of it. The resulting string is therefore still of the same category. We shall at the details of this later.

The dance now continues. When a question is formed, not only does the question word (or the phrase containing it) move to first position, also the auxiliary does move next to it. For rather than just getting (291) by moving /what/, we get (292), which results in a second movement, of the auxiliary /have/.

(290) [[∅ [you will [__ see what]$_{VP}$]$_{TP}$]$_{C'}$]$_{CP}$

(291) [what [∅ [you will [__ see __]$_{VP}$]$_{TP}$]$_{C'}$]$_{CP}$

(292) [what [will+∅ [you __[__ see __]$_{VP}$]$_{TP}$]$_{C'}$]$_{CP}$

The evidence is clear (on the basis of what we have said so far) that the tense head /will/ must also move. If it does, we need to supply a reason, as we did for the movement of the verb to T^0. It is believed that the empty element in C^0 actually is a question morpheme, call it Q. The derivation therefore looks more like this:

(293) [[Q [you will [__ see what]$_{VP}$]$_{TP}$]$_{C'}$]$_{CP}$

(294) [what [Q [you will [__ see __]$_{VP}$]$_{TP}$]$_{C'}$]$_{CP}$

(295) [what [will+Q [you __[__ see __]$_{VP}$]$_{TP}$]$_{C'}$]$_{CP}$

In relative clauses, however, this movement does not happen. Calling the head of a relative clause RC, this is the structure:

(296) ...the man [[RC [you will [__ see who]$_{VP}$]$_{TP}$]$_{C'}$]$_{CP}$

(297) ...the man [who [RC [you will [__ see __]$_{VP}$]$_{TP}$]$_{C'}$]$_{CP}$

We can say that the difference between Q and RC is that although both are empty, only RC can be on its own, while Q wants to attach to a host, and so triggers movement.

How Movement Works

We have already seen that movement must respect X-bar syntax. Whatever the end result, it must conform again to X-bar syntax. We have to begin with an important definition. Recall that a tree is a pair $\langle T, <\rangle$, where T is a set, the set of **nodes** and $<$ the relation 'is (properly) dominated by'. (I use 'is dominated by' synonymously with 'is properly dominated by'. This relation is never reflexive: no node dominates itself.) x and y are said to be **comparable** if $x = y$ or $x < y$ or $x > y$. (Thus, x can be found from y by either following the lines upwards or following them downwards.)

Definition 12 *Let $\langle T, <\rangle$ be a tree and $x, y \in T$ be nodes. x **c-commands** y if and only if x and y are not comparable and the mother of x dominates y.*

This can be phrased differently as follows. x c-commands y if x is sister to z and y is below z (or identical to it). Thus, c-command is with any sister and its heirs. Recall the tree from Lecture 8, repeated here as Figure 5. Here is a complete list of which nodes c-command which other nodes:

(298)

	c-commands
1	–
2	3, 6, 7, 8, 9
3	2, 4, 5
4	5
5	4
6	7, 8, 9
7	6
8	9
9	8

The relation of c-command is inherited by the strings that correspond to the nodes. For example, /`this villa`/ c-commands /`costs a fortune`/ and its subconstituents.

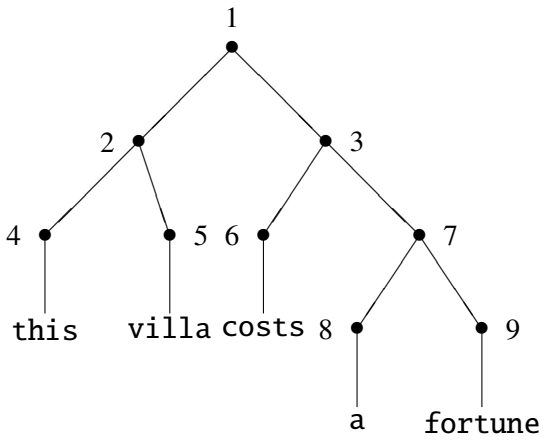


Figure 5: A Syntactic Tree

Now, movement is such that the constituent that is being moved is moved to a place that (i) is empty, and (ii) c-commands its original place. To make this proposal work we assume that instead of lexical items we can also find ∅ below a terminal node. Trees 6 and 7 describe a single movement step, for the movement that is otherwise denoted using strings by

(299) `Simon likes which cat?`

(300) `Which cat Simon likes __?`

We have also added the labellings. The constituent consisting of /`which cat`/ (which contains the node dominating these words plus everything below it) is moved into a c-commanding position. It can go there only if the node is absent. Thus, movement creates its own constituent, but it must to so in agreement with X-bar syntax. In place of the earlier constituent a node dominating ∅ is left behind. (We shall say a little more about this element in the next lecture.)

For the whole story to go through it is assumed that the X-bar grammar produces a number of choices for positions that can be used for elements to move into. (You do not move into nodes that are already there.) One such node is typically the specifier of CP. The positions that an element moves into must also match in label (category). For example, the C^0-position (which we have not shown above) is also often empty, but a phrase cannot go there, otherwise the labels do not match. It

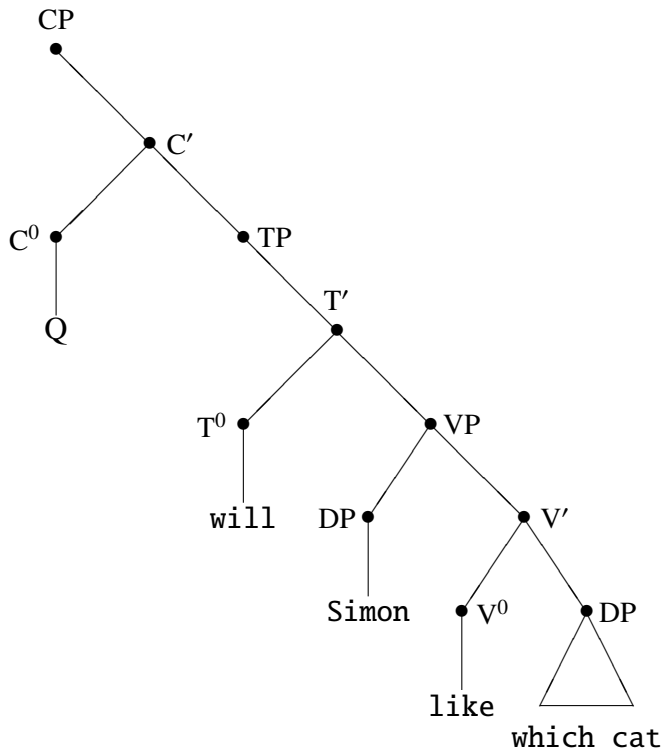


Figure 6: Before Movement (Step 0)

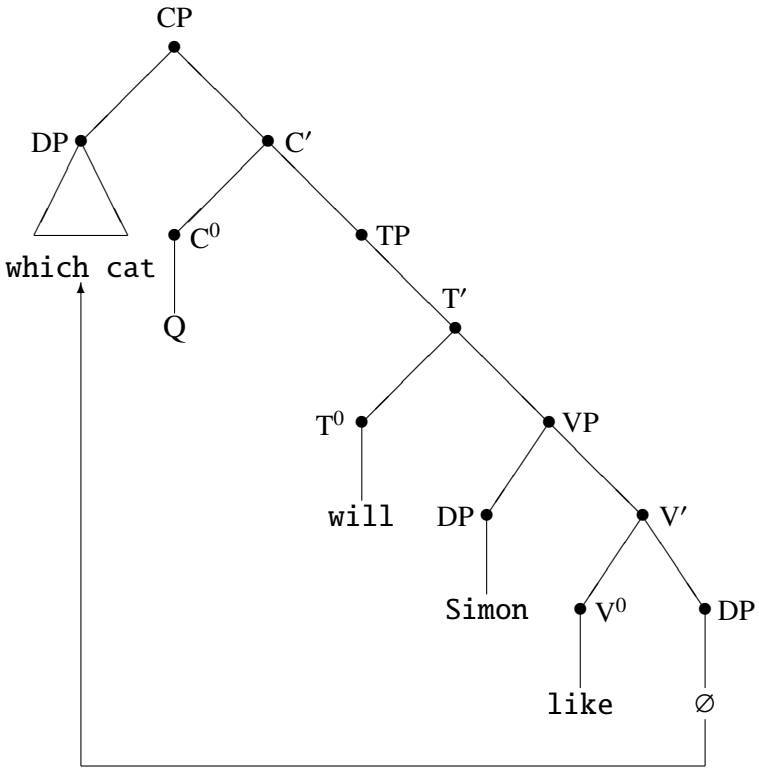


Figure 7: After Movement (Step 1)

is believed that the verb moves there in the relevant examples from (258). You can check that X-bar syntax places no restriction on the category of the specifier except that it must be a phrase. Thus, wh-phrases can go there. The solution has an advantage worth mentioning. CP has only one specifier, and therefore only one of the wh-phrases can go there. The others have to stay in place.

Now, why do we also need the condition of c-command? To answer this, one should realise that without restrictions the moving constituent may be put into any position that is free. For example, take this deep structure

(301) which salesperson wonders should he promote which product

The surface structure is

(302) Which salesperson wonders which product he should promote?

There are two specifiers of CP:

(303) [$_{CP}$[$_{VP}$which salesperson wonders [$_{CP}$[$_{TP}$[$_{T'}$should [$_{VP}$he promote which product]]]]]]

Many possibilities are open now including a movement of the phrase /which salesperson/ downwards to the right of /wonders/.

(304) *Wonders which salesperson he should promote which product?

As we have said, if the CP is filled by one wh-phrase, the other stays in place, so this should be grammatical. But it is not. And the reason for that is that movement has to be upwards, and into c-commanding position.

Other possibilities (mentioned above that still need to be excluded) is that the lower wh-phrase moves to the higher specifier (and /do/-insertion):

(305) *which product does which salesperson wonder he should promote

To exlucde this, we require first that every wh-phrase can only move to the next CP. If it is unavailable, the wh-phrase stays in place. Second, we shall say that /wonders/ blocks the movement of the wh-phrase.

(For Addicts:) Head Movement

I shall briefly discuss the problem of the movement of the verb to T^0, and that of the auxiliary to C^0. We have said that the moving head combines with the already existing head in the structure. But how can we make this compatible with X-bar syntax? First of all, by the facts we need to assume that /sees/, which is the same as /see+s/, actually is again a T^0. On the other hand, /see/ is a V^0, and /s/ is a T^0. Thus we seems to have this structure:

(306) $[\texttt{see}_{V^0}\ \texttt{s}_{T^0}]_{T^0}$

This is exactly what we shall propose. We shall supplement our X-bar with one more rule, which covers the above case. Namely, we shall allow adjunction also to zero level projections; however, adjunction is only on the left, and only zero level projections can adjoin. The result is again a zero level projection (since this is adjunction):

(307) $X^0 \rightarrow Y^0 \quad X^0$

Additionally, we require that when a head is moved it must move to the next c-commanding head and adjoin there. Adjunction is a movement that splits up a single node of category X into two nodes with category X, the upper one having the moved constituent and the lower X node as its daughters. It is clear that the rule (307) makes adjunction to heads an option. Let us perform this step with Figure 7. Notice that we should now consider that all heads in the initial structure are filled by either some lexical material or by a phonetically empty head (as we have effectively required). This then gives the structure in Figure 8. In order for everything to fall into place we now need to adjust our notion of c-command. We shall say that a within a structure $[Y^0 \quad X^0]_{X^0}$, all heads have the same c-command domain, which is the one of the enture constituent.

The derivation is complete after we have moved the subject to spec of TP.

Notes on this section. It should be clear that a proper formulation of such an initially simple idea as moving a wh-phrase to the front of a sentence needs a lot of attention to detail. And this is not only because the facts are unpredictable. It is also because what looks like a simple proposal can become quite complex once we start looking harder. This applies for example to programming, where we starts out as a simple program can become quite long because we need to make sure it works properly in all cases.

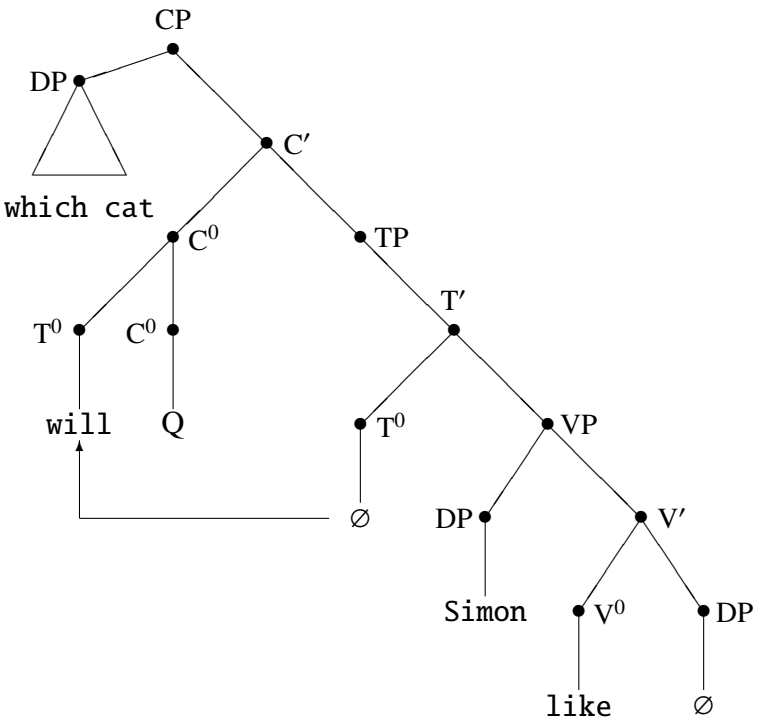


Figure 8: After Head Movement (Step 2)

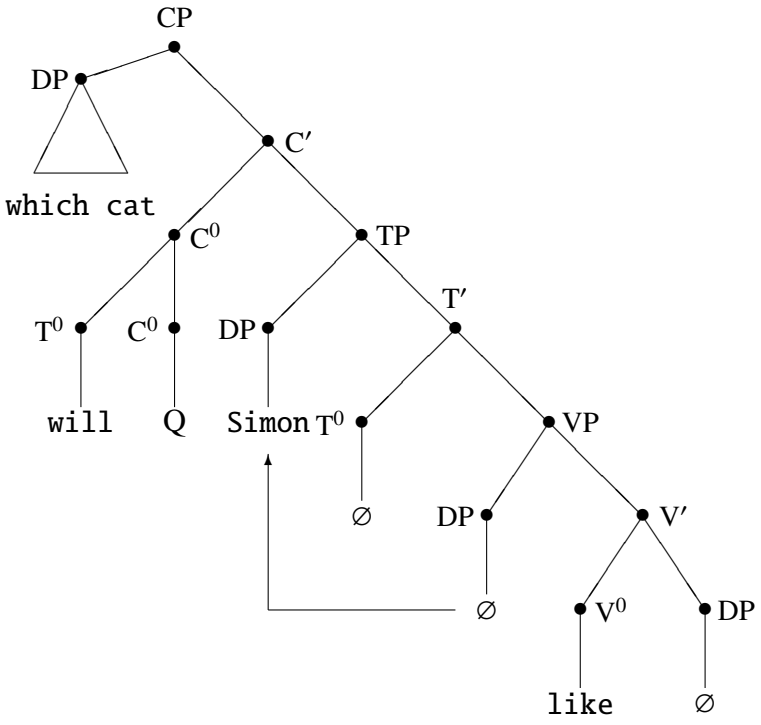


Figure 9: After Movement of Subject (Step 3)

Syntax V: Binding

> Pronouns do not only refer to a particular person or thing, they often refer to some constituent. This provides coherence to a text. Binding theory is about the interpretation of pronouns as well as the way in which they can be linked to other parths of the sentence. Key concepts are *Principles A*, *B* and *C* in addition to *c-command* and *binding*.

Pronouns

In this chapter we shall look at a phenomenon that is most intimately connected with c-command, namely **binding**. Binding is, as we shall see, as much of a semantic phenomenon as a syntactic one, but we shall ignore its semantic aspects as much as we can. What is at issue is that there are several different kinds of DPs: ordinary DPs, **names** and **pronouns**. Pronouns can be either **reflexive** (like /myself/, /yourself/ etc.) or not (/he/, /she/, etc.). In addition, there are also **reciprocals** (/each other/). In generative grammar, the term **anaphor** is used to refer to both reciprocals and reflexives but this is not the usage elsewhere, where an anaphor is an expression that refers to something that another expression already refers to (and therefore also nonreflexive pronouns generally counts as anaphors).

In English, the reflexive pronouns exist only in the accusative. There is no **/hisself/*, for example. (The pronoun /her/ can both be genitive and accusative, so /herself/ is not a good text case. On the other hand, it is /ourselves/ and not **/usselves/*. English has not been complelety consistent in arranging its case system.) There additionally are **demonstratives** (/this/, /that/), **relative pronouns** (/who/, /what/, /which/), which are used to open a relative clause.

(308) I could not find the CD [which you told me about].

(309) There is a new film [in which no one ever talks].

The enclosed constituents are called **relative clauses**. They are used to modify nouns, for example. In that they are like adjectives, but they follow the noun in English rather than preceding it.

(310) *I could not find the which you told me about CDs.

(311) *There is a new in which no one ever talks film.

The relative clause is opened by a relative pronoun, /which/, together with the preposition that takes the pronoun as its complement. One should strictly distinguish the relative pronoun from the complementiser /that/ as found in sentential complements (/believe that/). The complementiser occupies a different position (C^0). Some relative clauses can be opened by /that/.

(312) I could not find the CD [that you told me about].

This use of /that/ seems to be a hybrid between the complementiser and a relative pronoun (recall that /that/ is also a demonstrative). (And syntacticians are divided on the question whether to call it a relative pronoun or a complementiser in this usage.) The relative pronoun, for example, can move together with its preposition (/the film in which I played piano/), but you cannot say /the film in that I played piano/; instead, you have to say /the film that I played piano in/. Typically, a relative clause either has a relative pronoun or a complementiser; this has become an iron rule of generative grammar ("Doubly filled COMP filter"). But facts are different. Some dialects of German give us evidence that the relative pronoun and the complementiser really can cooccur. In Suebian, /wo/ is used as a complementiser for relative clauses, accompanied by a relative pronoun.

(313) die Leit die wo beim Daimler schaffet
the people who that at.the Daimler work
the people who work at the Daimler Benz factory.

As with questions we like to think of the relative clauses as being derived from a structure in which the relative pronoun is in the place where the verb expects it.

(314) the CD [∅ you told me about which]

(315) a new film [∅ no one ever talks in which]

The position of specifier of CP is vacant, and the relative pronoun wants to move there. (The position of C is also empty, but we like to think that there is a silent C that sits there. Anyway, phrases can never go there.) Sometimes the relative pronoun goes alone (being a DP, hence a phrase, it can do that), sometimes it drags the P along. The latter is known as **Pied-Piping**, from the fairy tale of the piper who promised the city to get rid of the rats...

Different languages have different systems. Latin does not distinguish reflexive and irreflexive pronouns in the 1st and 2nd person. It has however two 3rd

pronouns, /is/ ('he'), and /se/ ('himself'). The reflexive exists in all cases but the nominative. It is the same both in the singular and the plural.

(316)

nom	–
gen	sui
dat	sibi
acc	se
abl	se

There is another set of pronouns called *possessive*. They are used to denote possession. They are just like adjectives. For example,

(317) equus suus
horse-nom his-nom

(318) equum suum
horse-acc his-acc

Binding

These different kinds of expressions each have their distinct behaviour. Look at the following three sentences.

(319) John votes for John in the election.

(320) John votes for himself in the election.

(321) John votes for him in the election.

There is an understanding that the two occurrences of the word /John/ in (319) point to two different people. If they do not, we have to use (320) instead. Moreover, if we use (320) there is no hesitation as to the fact that there is an individual named /John/ which casts a vote for the same individual (also named /John/). If we use (321), finally, we simply cannot mean the same person, /John/, by the word /him/. John cannot be self-voting in (321), he votes for someone else.

There are several ways one may try to understand the distribution of full DPs, pronouns and reflexives. First of all, however, let us notice that a reflexive pronoun expects the thing it refers to be the same as something else in the sentence. The expression that denotes this is called the **antecedent** of the pronoun. In (320), for

example, the antecedent of /himself/ is /John/. To express that some constituent refers to the same object we give them little numbers, called **indices**, like this.

(322) *John$_1$ votes for John$_1$ in the election.

(323) John$_1$ votes for himself$_1$ in the election.

(324) *John$_1$ votes for him$_1$ in the election.

We have already assigned grammaticality judgments. Needless to say, any other number (say, 112, 7, 56 or 34) would have done equally well, so that as far as syntax is concerned (322) is the same as

(325) John$_{34}$ votes for John$_{34}$ in the election.

(In the books you often find letters *i*, *j* and *k* in place of concrete numbers, but this not a good idea since it suggests that plain numbers are not abstract enough. But they in fact are.) In (319) the two occurrences of /John/ are supposed to point to the same individual. If they are supposed to point to different individuals, we write different numbers:

(326) John$_1$ votes for John$_2$ in the election.

The numbers are devices to tell us whether some constituent names the same individual as some other constituent, or whether it names a different one.

Pronouns seem to encourage a difference between subject and object in (320), and similarly with names (321). However, things are tricky. In (327) the pronoun /his/ can be taken to refer not only to someone different from John, but also to John himself. And similarly in (328). (Just an aside: the reflexives cannot be used in genitive, they only have accusative forms. This may be the reason why we do not find them here, but the theory we are going to outline here tries a different line of argumentation.)

(327) His lack of knowledge worried John.

(328) He looks at himself in the mirror.

So, it is not really the case that when we have a pronoun that it must be used to talk about a different person than the others in the sentence (= that it must have a different index).

The conditions that regulate the distribution of these expressions are at follows. First, we define the notion of binding.

Definition 13 *A constituent X **binds** another constituent Y if X c-commands Y and X and Y have the same index.*

Binding is an interesting mixture between semantical conditions (carrying the same index, hence talking about the same individual) and purely structural ones (c-command). A note of warning is in order. Constituents are strings, but we talk here as if they are nodes in a tree. This confusion is harmless. What we mean to say is this: suppose that x and y are nodes and the corresponding constituents are X and Y. Then if x c-commands y and has the same index, then x binds y, and X binds Y. So, X binds Y if there are nodes x and y such that x binds y and X is the constituent of x and Y the constituent of y. Now we are ready to say what the conditions on indexing are.

Principle A.
A reflexive pronoun must be bound by a constituent of the same CP (or DP).
Principle B.
A pronoun must not be bound by a constituent inside the same CP.
Principle C.
A name must not be bound.

For example, (319) can be used successfully to talk about two individuals named John. So, the expression alone is not sufficient to rule out a sentence. But the rules do tell us sometimes what the possible meanings are.

To understand the role of c-command, we need to look at the structure of some sentences. The subject c-commands the object of the same verb, since smallest constituent containing the subject is the VP, which also contains the object. However, the object does not c-command the subject, since the smallest constituent containing it is only the V′. So the following is illegitimate no matter what indices we assign:

(329) *`Himself voted for John.`

The subject also c-commands all other arguments. This accounts for the correctness of the following.

(330) `The queen was never quite sure about herself.`

(331) `The students were all talking to themselves.`

Notice that the principles not only say that these sentences are fine, they also tell us about the assignment of indices. They claim that (332) is fine but (333) is not.

(332) The queen$_1$ was never quite sure about herself$_1$.

(333) *The queen$_1$ was never quite sure about herself$_2$.

This is because the reflexive (/herself/) must be bound inside the sentence; this means there must be some antecedent c-commanding it having the same index. In (332) it is /the queen/, but in (333) there is no such constituent. Thus, the overall generalization seems to be good. Notice that Principle A says that the binder must be in the same clause. Thus the following contrast is explained.

(334) [John thinks [that Alice likes herself.]$_{CP}$]$_{CP}$

(335) *[John thinks [that Alice likes himself.]$_{CP}$]$_{CP}$

In both (334) and (335), the reflexive pronoun must be bound inside the lower CP, by Principle A. However, the only binder can be /Alice/, and not /John/, and thus (335) is ungrammatical.

Now we look at pronouns. The situation where a pronoun should not be used in the same sentence is when actually a reflexive would be appropriate according to Principle A. For example, the following sentences are ruled out if meant to be talking about the same individual(s) in the same sentence (that is, if /her/ refers to /the queen/ and /them/ refers to /the students/).

(336) *The queen$_1$ was never quite sure about her$_1$.

(337) *The students$_1$ were all talking to them$_1$.

By the same token, if we use numbers we can also write:

(338) *The queen$_1$ was never quite sure about herself$_2$.

(339) *The students$_1$ were all talking to themselves$_2$.

Here, different numbers mean that the expressions are meant to refer to different individuals or groups. Principle B talks about binding inside the same CP. It does not exclude that a pronoun is bound in a higher CP. This is borne out.

(340) [John$_1$ thinks [that Alice$_3$ likes him$_1$.]$_{CP}$]$_{CP}$

(341)

Notice that the contrast between a reflexive and a nonreflexive pronoun only matters when the antecedent is c-commanding the pronoun. We repeat (327) below:

(342) $\texttt{His}_{1/2}$ lack of knowledge worried $\texttt{John}_1$.

Here, /John/ is the antecedent. We can take the sentence to mean that John is worried about his own lack of knowledge, or that he is worried about someone else's lack of knowledge. In none of the cases would a reflexive pronoun be appropriate.

Let us now change the position of the two:

(343) $^*\texttt{John}_1$'s lack of knowledge worried $\texttt{himself}_1$.

Here, /John/ does not c-command the pronoun. But the reflexive must be bound by something that c-commands it inside the clause. This can only be /John's lack of knowledge/. But if we think that, we would have to say /itself/ rather than /himself/. Next, why is (344) fine?

(344) $\texttt{John}_1$'s lack of knowledge worried $\texttt{him}_1$.

Even though it has the same index as /John/, it is not c-commanded, hence not bound by it. Having the same index nevertheless means that they refer to the same thing (John), but for syntax binding takes place only if c-command holds in addition. Hence, the next sentence is also fine for syntax:

(345) $\texttt{John}_1$'s lack of knowledge worried $\texttt{John}_1$.

Admittedly, we would prefer (346) over (345), but it is agreed that this is not a syntactic issue.

(346) $\texttt{His}_1$ lack of knowledge worried $\texttt{John}_1$.

Movement Again

We have argued earlier that an element can move only into a position that c-commands the earlier position. It seems that binding theory could be used in this situation. However, in order to use binding theory, we need to have an element that is being bound. Recall from previous discussions that when an element

moves it leaves behind an empty element. But why is that so? To see this, look at movement from out of a subordinate clause.

(347) Who do you think answered which question?

(348) *Who do you think which question answered?

In these sentences, it is the subject of the subordinate clause which moves to the specific of CP of the main clause. If that was all that happened, we would expect that the specifier of the subordinate CP would be empty and so the object /which question/ would have to move into it. But this is not the case. Thus we conclude that the specifier of the lower CP is not empty. It is filled by something that we cannot hear. That thing is the element that /who/ left behind, when it moved to the higher specifier. So the full proposal is this: movement of a wh-phrase must be first to the specifier of CP of the same clause. After that it is to the specifier of CP of the next higher clause; and so on. Each time it moves, it leaves behind an empty element, called a **trace**. The constituent in the new position and the trace will be coindexed. Finally, we declare that traces are some sort of pronouns, and hence must be bound.

> **Condition on Traces.**
> All traces must be bound.

If we assume this much it follows that movement must inevitably be into a position that c-commands the original position. It does something else, too. It ensures that whatever thing we choose to interpret the moved element by, it will be used to interpret the trace.

(349) Air pilots$_1$ Harry admires t_1.

The relevant structure of (347) is now as follows.

(350) [Who$_1$ do you think [t_1 [answered$_2$ [t_1 t_2 which question]$_{TP}$]$_{C'}$]$_{CP}$]$_{CP}$

The t_1 in specifier of CP blocks movement of /which question/ into. However, since movement there is impossible, it is also not required, so the sentence is grammatical.

One may wonder why this is a good theory. It postulates empty elements (traces), so how can we be sure that they exist? We do not see them, we do not

hear them, so we might as well assume that they do not exist. Furthermore, (349) is not the sentence we shall see in print or hear, so this actually adds structure that is seemingly superfluous.

Opinions on this diverge. What is however agreed is that empty elements are very useful. We have used them occasionally to save our neck. When nouns get transformed into verbs, no change is involved. Other languages are not so liberal, so it is not the nature of nouns and verbs that allows this. It is, we assume, an empty element which English has (and other languages do not) which can be attached to nouns to give a verb. At another occasion we have smuggled in empty complementizers. You will no doubt find more occasions on which we have made use of empty elements.

Binding And Agreement

Pronouns distinguish not only case and number but also gender in English. Crucially, when a pronoun becomes bound it must agree in number and gender with its binder (not in case though).

(351) John voted for himself.

(352) *John voted for herself.

(353) *John voted for themselves.

(354) The committee members voted for themselves.

(355) Mary voted for herself.

The fact that agreement holds between binder and pronoun means that certain indexations are not licit. Here is an example.

(356) *Her$_1$ lack of knowledge worried John$_1$.

(357) *Their disinterest in syntax$_1$ bothered John$_1$.

Also, we have said that a pronoun cannot be bound inside the same CP. But it can be bound outside of it:

(358) John$_1$ told his boss$_2$ that he$_{1/2}$ looked good.

In the previous example, both indices are licit, since binding by /John/ or by /boss/ is compatible with agreement. However, in the following sentence one of

the options is gone.

(359) John$_1$ told Mary$_2$ that she$_{*1/2}$ looked good.

For if the pronoun is bound by /John/, it must agree with it in gender, which it does not. So it can only be bound by /Mary/.

Morphology II: Similarities and Dissimilarities to Syntax. Representational Issues

> We return to morphology. We shall investigate the possible shapes of morphs and then turn to the question in what ways morphology is different from syntax. At the end we shall return to the issue of head movement and the structure of complex heads.

Morphs and Morphemes

Before we begin, a few words on terminology. As I already said in the introduction, we distinguish between a **morph** and a **morpheme**. This conflicts to some degree with the terminology of Definition 3. I shall not attempt to resolve the conflict here. Suffice it to say that we think in this chapter of morphs and morphemes as units in the same way as phones (= sounds) and morphemes. As in phonology and syntax, we shall distinguish a deep structure form a surface structure. At deep structure, we only have morphemes. At surface structure, we have morphs. Morphs of the same morpheme are called **allomorphs**. What unites the morphs of a morpheme is the common syntactic and semantic features. An example is the plural morpheme. Syntactically, it turns a root noun into a plural noun; semantically, it changes from denoting properties to denoting groups of things. The plural morpheme takes different shapes in English. Among other there are the following ways to form the plural:

① Add [z] at the end: [dɔg] → [dɔgz]

② Add [s] at the end: [kʌp] → [kʌps]

③ Add [əz] at the end: [bʌs] → [bʌsəz]

④ Change the root vowel from [ʊ] to [ɪ]: [wʊmən] → [wɪmən]

⑤ Do not change anything: [ʃip] → [ʃip]

⑥ Add [ʃn]: [ɔks] → [ɔksən]

Notice however that we have used the square brackets: [·], and not the slashes: /·/. This means that we have lost part of the abstraction that we gained in deep phonology. This means that we listed more types of changes than necessary. Namely, at the deep phonology the first three cases (① – ③) become one. From the standpoint of deep phonology the surface forms [z], [s] and [əz] are just the surface manifestations of the deep form /z/. Thus, althout we factually deal with written form, we should remind ourselves of the fact that we have already reached a pretty abstract level of representation.

Nevertheless, even if this is taken into account, there remain some other forms that we cannot treat as phonological surface variants any more. The change of root vowel, the suffix [ən] and the zero suffix belong to this category. We therefore get the following revised list.

❶ Add /z/ at the end: /dɔg/ → /dɔgz/; /kʌp/ → /kʌps/; /bʌs/ → /bʌsəz/.

❷ Change the root vowel from /ʊ/ to /ɪ/: /wʊmən/ → /wɪmən/

❸ Do not change anything: /ʃip/ → /ʃip/

❹ Add [ʃn]: /ɔks/ → /ɔksən/

Now let us do the following. The plural morpheme has (at least) four different morphs. The first is /z/, the second is root vowel change, the third is ∅ (the empty morph) and the fourth /ʃn/. Which morph is used in a particular case depends on the noun to which we apply the morpheme. At the deep morphology we only see a combination of a noun root with the one plural morpheme. At the surface however we see that the root we choose influences what morph is to be used in the construction.

Kinds of Morphological Processes

In syntax we have said that words or constituents are simply concatenated. The only choice is therefore in which order we concatenate them. In morphology this is not always so. We shall review a few ways in which two morphs can be composed. The general term for grammatical morphs is **affix**. Generally, an affix to a string $\vec{x}$ is a string $\vec{y}$ that puts itself somewhere in $\vec{x}$. Given the terminology below, it is either a prefix, a suffix or an infix. Some writers use it in a more general

sense, but we shall not do that. Morphs are not always affixes, however. A morph need not be a piece (= string) that we add somewhere; it may be several pieces (transfix, circumfix), or simply a certain kind of change effected on the string in question (like vowel change). The general term for all of these is **morphological change**. We shall give a few examples of what kinds of morphological changes there are in the languages of the world.

Suffixes and Prefixes

A **suffix** is a string that is added at the end, a **prefix** is a string that is added at the beginning. English has mostly suffixes (derivational ones like /ation/, /ize/, /ee/; inflectional ones like /s/ and /d/). But it also has prefixes: /de/, /re/, /un/ are prefixes. It is generally agreed that—if we use an analogy with syntax here—the affix is the head, and that it either expects the string on its right (prefix) or on its left (suffix).

If there were only suffixes and prefixes, morphology would look like syntax. Unfortunately, this is not the case.

Circumfixes

A **circumfix** consists of a part that is being added before the word and another that is added thereafter. It is thus a combination of prefix and suffix. The German perfect is a circumfix. It consists in the prefix /ge/ and a suffix which changes from stem to stem (usually it is /en/ or /t/). The infinitive is a suffix /en/ as can be seen from the table below.

(360)

Infinitive	Root	Perfect
sehen	seh	gesehen
back	back	gebacken
filmen	film	gefilmt
hausen	haus	gehaust

The two parts (the prefix and the suffix) are historically of different origin (the prefix /ge/ did not exist originally, and in English you do not have it). It present day German, however, there is no sense in taking the two separate.

The superlative in Hungarian also is a circumfix.

(361)

nagy	'great'	legnagyobb	'greatest'
fehér	'white'	legfeherebb	'whitest'

Here as in the German perfect, the circumfix has two identifiable parts. The suffix is found also in the comparative:

(362)

nagy	'great'	nagyobb	'greater'
fehér	'white'	feherebb	'whiter'

The same situation is found in Rumanian. We have /frumos/ ('beautiful') in the positive, /mai frumos/ ('more beautiful') in the comparative, and /cel mai frumos/ ('most beautiful'). However, whether or not the superlative can be seen as being formed from the comparative by adding a prefix depends on the possibility to decompose the meaning of the superlative in such a way that it is derived from the meaning of the comparative. To my knowledge this has not been proposed.

Infixes

Infixes insert themselves inside the string. Look at the following data from Chrau (a Vietnamese language).

(363)	vŏh	'know'	vanŏh	'wise'
(364)	căh	'remember'	canăh	'left over'

The string /an/ is inserted after the first consonant! The string is cut into two pieces and the nominaliser inserts itself right there.

(365) vŏh → v + ŏh → v + an + ŏh → vanŏh

Transfixes

A **transfix** is an even more complex entity. We give an example from Egyptian Arabic. Roots have three consonants, for example /ktb/ 'to write' and /drs/ 'to study'. Words are formed by adding some material in front (prefixation), some

material after (suffixation) and some material in between (infixation). Moreover, all these typically happen *at the same time*. Let's look at the following list.

(366)

[katab]	'he wrote'	[daras]	'he studied'
[ʔamal]	'he did'	[naʿal]	'he copied'
[baktib]	'I write'	[badris]	'I study'
[baʔmil]	'I do'	[banʿil]	'I copy'
[iktib]	'write!'	[idris]	'study!'
[iʔmil]	'do!'	[inʿil]	'copy!'
[kaatib]	'writer'	[daaris]	'studier'
[ʔaamil]	'doer'	[naaʿil]	'copier'
[maktuub]	'written'	[madruus]	'studied'
[maʔmuu]	'done'	[manʿuul]	'copied'

It requires some patience to find out how the different words have been formed. There are variations in the patterns (just like in any other language). One thing however is already apparent: the vowels get changed from one form to the other. This explains why the vowel is not thought of as being part of the root.

Other Kinds of Changes

Reduplication is the phenomenon where a string is copied and added. For example, if /abc/ is a string, then its (re)duplication is /abcabc/. Austronesian languages like to use reduplication or sometimes even triplication for various purposes (to form the plural, to intensify a verb, to derive nouns and so on). The following is from Indonesian.

(367)

orang	'man'	orang-orang	'child'
anak	'man'	anak-anak	'children'
mata	'eye'	mata-mata	'spy'

An example of triplication is provided in Pingelapese:

(368)

kɔul	'to sing'	kɔukɔul	'singing'	kɔukɔukɔul	'still singing'
mejr	'to sleep'	mejmejr	'sleeping'	mejmejmejr	'still sleeping'

Reduplication need not copy the full word. For example, in Latin some verbs form the perfect in the following way. The first consonant together with the next vowel

is duplicated and inserted:

(369)

`pendit`	'he hangs'	`pependit`	'he has hung'
`tendit`	'he stretches'	`tetendit`	'he has stretched'
`currit`	'he runs'	`cucurrit`	'he has run'
`spondet`	'he promises'	`spopondit`	'he has promised'

The last example shows that the /s/ is exempt from reduplication.

The Difference Between Morphology and Syntax

There is no way to predict whether some piece of meaning is expressed by a morpheme or by a separate lexeme or both. There exist huge differences across languages. Some languages pack all kinds of meanings into the verb (Inuit, Mohawk), some keep everything separate (Chinese). Most languages are somewhere in between. Morphology is more or less important. However, even within one language itself the means to express something change. Adjectives in English have three forms: **positive** (normal form), **comparative** (form of simple comparison) and **superlative** (form of absolute comparison). Now look at the way they get formed.

(370)

positive	comparative	superlative
`high`	`higher`	`highest`
`fast`	`faster`	`fastest`
`common`	`more common`	`most common`
`good`	`better`	`best`
`bad`	`worse`	`worst`

The first set of adjectives take suffixes (zero for the positive, /`er`/ for the comparative and /`est`/ for the superlative). The second set of adjectives take a separate word, which is added in front (/`more`/ and /`most`/). The third set is irregular. The adjective /`good`/ changes the root in the comparative and superlative before adding the suffix (/`bett`/ in the comparative and /`b`/ in the superlative), while /`bad`/ does not allow an analysis into root and suffix in the comparative and superlative. (We may define a suffix, but it will be a one-time-only suffix, since there is no other adjective like /`bad`/. It therefore makes not much sense to define a separate comparative suffix for /`worse`/. However, /`worst`/ is a debatable case.)

It is not far fetched to subsume the difference between PPs and case marked DPs under this heading. Finnish has a case to express movement into a location, and a case for movement to a location:

(371) `Jussi menee taloon.`
Jussi goes house-into = Jussi goes into the house.

(372) `Jussi menee talolle.`
Jussi goes house-to = Jussi goes to the house.

When it comes to other concepts—like 'under' and 'over'—Finnish runs out of cases and starts to use adpositions:

(373) `Jussi menee talon alle.`
Jussi goes house-gen under = 'Jussi goes under the house.'

(374) `Jussi menee talon yli.`
Jussi goes house-gen over = 'Jussi goes over the house.'

Hungarian has a case for being on top of, but the situation is quite analogous to Finnish.

Thus, the cutoff point between morphology and syntax is arbitrary. However, the two may show to behave differently in a given language so that the choice between morphological and syntactical means of expression has further consequences. We have seen, for example, that the comparative morpheme /`er`/ is a suffix—so it is added after the word. However, the comparative lexeme /`more`/ wants the adjective to its right. The latter is attributable to the general structure of English phrases. The complement is always to the right. Morphemes are exempt from this rule. They can be on the other side, and generally this is what happens. For example, verb+noun compounds in English are formed by placing the verb after the noun: /`goalkeeper`/, /`eggwarmer`/, /`lifesaver`/ and so on. If these were two words, we should have /`keeper goal`/, /`warmer egg`/, and /`saver life`/. The reason why we do not get that is interesting in itself. English used to be a language where the verb follows the object (as is the case in German). It then changed into a language where the verb is to the left of the object. This change affected only the syntax, not the morphology. French forms compounds the other way around (/`casse-noix`/ lit. cracker-nut = 'nutcracker', /`garde-voie`/ lit. guard-way = 'gatekeeper'). This is because when French started to form compounds, verbs already preceded their objects.

The Latin Perfect—A Cabinet of Horrors

This section shall illustrate that the same meaning can be signaled by very different means; sometimes these are added independently. To start, there is a large group of verbs that form the perfect stem by adding /v/. (3rd person does not signal gender; we use 'he' instead of the longer 'he/she/it'. The ending of the 3rd singular perfect is /it/. The ending in the present is /at/, /et/ or /it/, depending on the verb.)

(375)

amat	'he loves'	amavit	'he has loved'
delet	'he destroys'	delevit	'he has destroyed'
audit	'he hears'	audivit	'he has heard'

We also find a group of verbs that form the perfect by adding /u/.

(376)

vetit	'he forbids'	vetuit	'he has forbidden'
habet	'he has'	habuit	'he has had'

The difference is due to modern spelling. The letters /u/ and /v/ were originally not distinct and denoted the same sound, namely [u], which became a bilabial approximant in between vowels. The difference is thus accounted for by phonemic rules. (You are however warned that we do not really know that what is nowadays written /u/ is the same sound as what is now written /v/; this is an inference, partly based on the fact that the Roman did use the same letter. Probably pronunciation was different; however, the difference was for all we know not phonemic.)

Other verbs add an /s/. The combination /gs/ is written /x/.

(377)

regit	'he reigns'	rexit	'he has reigned'
augit	'he fosters'	auxit	'he has fostered'
carpit	'he plucks'	carpsit	'he has plucked'

There are verbs where the last consonant is lost before the /s/-suffix.

There are verbs where the perfect is signaled by lengthening the vowel (lengthening was not written in Roman times, we add it here for illustration):

(378)

iuvit	'he helps'	iūvit	'he has helped'
lavat	'he washes'	lāvit	'he has washed'

Sometimes this lengthening is induced by a loss of a nasal:

(379)

rumpit	'he breaks'	rūpit	'he has broken'
fundit	'he pours'	fūdit	'he has poured'

Finally, there are verbs which signal the perfect by reduplication:

(380)

currit	'he runs'	cuccurrit	'he has run'
tendit	'he stretches'	tetendit	'he has stretched'

Now, certain verbs use a combination of these. The verb /frangere/ ('to break') uses loss of nasal, accompanied by vowel lengthening, plus ablaut:

(381) frangit 'he breaks' frēgit 'he has broken'

The lengthening of the vowel could be attributed to the loss of the nasal (in which case it is called **compensatory lengthening**). However, the next verb shows that this need not be the case.

The verb /tangere/ uses a combination of reduplication, loss of nasal and ablaut (and no lengthening of the vowel):

(382) tangit 'he touches' tetigit 'he has touched'

The root is /tang/. The nasal is dropped, yielding /tag/. Reduplication yields /tetag/ (actually, it should give */tatag/.) Finally, the vowel changes to give /tetig/. The change of vowel is quite common in the reduplicating verbs since the reduplication takes away the stress from the root. We have /ˈtangit/ and /ˈtetigit/. The reason is that the root vowel is actually still short (so no compensatory lengthening). If it were long, we would have to pronounce it /teˈtīgit/.

What Syntax Gets To See

We have seen above that the comparative of adjectives is sometimes formed by adding a morpheme and sometimes by using a separate word. It is unfortunate, however, that there is such an asymmetry. For from a syntactic perspective, there is not much of a difference between the two. It is for this reason—and others—that one has sought to expand the scope of syntax downwards into morphology. Linguists one used to think of the form /thinks/ as a verb form with some features, thus a syntactically indecomposable whole. Nowadays, however, one sees /thinks/ as syntactically complex. This brings us back to Lecture 11, Page 131. There we have already said that the reason for head movement was that the verb form /thinks/ is syntactically complex, but that its parts are distributed in the

sentence in the wrong way. Given that the string /s/ wants to be added to the end of a verb, and that the verbal root cannot be pronounced as such, the solution is to move the verb upwards, and form a complex head [$_{C}$[$_{V}$`think`][$_{C}$`s`]]. Similarly, at the deep level all the surface complications of the Latin perfect disappear. There is a single morpheme, and it is added to the stem.

The analysis of words as complex heads has proved to be a very fruitful idea. With verbs actually hosting quite a number of other morphemes (agreement (in person, in number, in gender), voice (active or passive), aspect (completed or ongoing activity), definiteness, to name a few), the scope of syntax expanded massively at the expense of morphology.

Semantics I: Basic Remarks on Representation

> Semantics studies meanings. It is intimately connected with logic, the study of reasoning. To see whether we have the correct meaning it is sometimes illuminating to check whether the purported meaning carries the correct logical consequences.

So far we have been concerned only with the *form* of language expressions, and not with their meaning. Ultimately, however, language is designed to allow us to communicate to each other how the world is like, to make each other do or believe something, and so on. The part of linguistics that deals mainly with the question of what is meant by saying something is called **semantics**. Meaning is not just some aspect of the form in which expressions are put by the language. If I tell you that Paul has pestered at least three squirrels this morning you can conclude that he has pestered at least two squirrels this morning. That follows not by virtue of the form the words /three/ and /two/ have, but in virtue of what these words mean. Likewise, if I get told that exactly half of my students do not like syntax, and I have 180 students, then I conclude that 90 students do not like syntax. This reasoning works independently of the language in which the sentences are phrased. The same information can be conveyed in English, French, Swahili and so on. The form will of course be much different.

We introduce some bits of terminology. First, a **statement** is a sentence which can be said to be either true or false. In distinction to a statement, a *question* is not true or false; it is a request for information. Likewise, a *command* is a request on the part of the speaker that he wants something done. We shall be concerned here exclusively with statements.

Statements express **propositions**. We think of propositions as existing independently of the language. A sentence is a faithful translation of another sentence if both express the same proposition (if speaking about statements). So, the French sentence

(383) `Marc a vu la Tour Eiffel.`

is translated into English by

(384) `Marc has seen the Eiffel Tower.`

just because they express the same proposition. A principal tool of semantics is to study the logical relations that hold between sentences in order to establish their meanings. Let us look at the following reasoning.

(385)
```
   Peter gets married.
   Sue gets married.
∴  Peter and Sue get married.
```

This is less than easy. For there is a certain subtlety that might easily get overlooked. Let's change the wording somewhat. The following is a more straightforward case.

(386)
```
   Peter gets married.
   Sue gets married.
∴  Peter marries Sue.
```

There is a clear sense in which the second reasoning does not go through. Suppose Peter marries Joan and Sue marries Alec. Then the first two sentences are true, but the conclusion fails. Thus (386) does not go through. However, (385) still seems to be OK. And this is because the conclusion is true even in the case that Peter does not marry Sue. (Though one is likely to read it that way; I shall return to that problem below.) It is the task of semantics to explain this.

Definition 14 *Let $A_1, \ldots, A_n$ and B be propositions. We say that B is a **logical consequence** of A_1 to A_n if whenever A_1, A_2, … A_n are true, so is B. A_1 to A_n are called the **premises** and B the **conclusion**. We write $A_1, \ldots, A_n \vDash B$ in this case; and $A_1, \ldots, A_n \nvDash B$ otherwise.*

An alternative notation has been used above:

(387)
$$\begin{array}{l} A_1 \\ \vdots \\ A_n \\ \hline \therefore \quad B \end{array} \qquad \text{or} \qquad \begin{array}{l} A_1 \ldots A_n \\ \hline \therefore \quad B \end{array}$$

We shall use this sometimes, with or without ∴. This definition talks about propositions, not about statements. However, we still need to use some language to denote the statements that we want to talk about. Often enough linguists are content in using the statements in place of the proposition that they denote. So we

write

(388) `Peter gets married., Sue gets married.`
$\models$ `Peter and Sue get married.`

Also the following is true

(389) `Peter marries Sue.` $\models$ `Peter gets married.`

But this often is an oversimplification. It so turns out that sentences do not always express just one proposition, they can express various propositions, sometimes depending on the context. We say that they are **ambiguous**. Look again at the example (385). There are two ways to understand

(390) `Peter and Sue get married.`

One way is to understand it as them getting married to each other; another is to understand it expressing that Peter gets married to someone, and Sue does. Even though the latter interpretation covers the first one, intuitively they are distinct. And it is in fact this stronger reading that is the preferred one. This explain that some people would actually think that (385) is in fact false since the premisses do not support the (stronger, but preferred) reading where Peter marries Sue. Thus, judgements on logical conclusion are difficult to elicit in presence of ambiguity.

There are other instances where one and the same sentence can have different meanings without being called ambiguous. The easiest example is provided by pronouns. Suppose I say

(391) `I love Sachertorte.`

Then the proposition expressed by this sentence is different from the proposition expressed by the same sentence uttered by Arnold Schwarzenegger. This is so since I can truly utter (391) at the same time when he can truly deny (391), and vice versa. Moreover, (391) uttered by me is roughly equivalent to

(392) `Marcus Kracht loves Sachertorte.`

This of course would not be so if it were an utterance by Arnold Schwarzenegger, in which case the utterance expressed the same proposition as

(393) `Arnold Schwarzenegger loves Sachertorte.`

Another case is the following. Often we use a seemingly weaker sentence to express a stronger one. For example, we say

(394) `Peter and Sue got married.`

when we want to say that Peter married Sue. There is nothing wrong in doing so, as (394) is true on that occasion, too. But in fact what we intended to convey was that Peter married Sue, and we expect our interlocutor to understand that. Semantics does not deal with the latter problem; it does not investigate what we actually *intend* to say by saying something. This is left to **pragmatics**. (Some would also relegate the Sachertorte-example to pragmatics, but that view is not shared by everyone.) Still, even if such things are excluded, natural language meanings are not always clear beyond doubt. Let's look at the following sentence.

(395) `John is a bright UCLA student.`

Suppose it is agreed that in order to be a UCLA student one has to be particularly bright in comparison to other students. Now, is (395) saying that John is bright even in comparison to other UCLA students, so that he is far brighter than average? Or does it only say that John is bright as a student, and that in addition he is a UCLA student? There is no unequivocal answer to this.

Here is another case.

(396) `Six companies sent three representatives to the trade fare.`

How many representatives got sent? Three or eighteen? I think the intuitive answer is 18, but in other sentences intuitions are less clear.

(397) `Six students visited three universities.`

The example (395) may just as well involve in total six students and in total three universities. But it may also involve four universities, or five, or even 18. (397) is formally identical to (396), so why is there a difference?

It turns out that the difference between (396) and (397) is not due to the particular meanings of the expressions. Rather, it is the way the world normally is that makes us choose one interpretation over another. For example, we expect that one

is employed by just one company. So, if a company sends out a representative, it is by default a representative only of this company and not of another. We can imagine things to be otherwise. If the companies strike a deal by which they all pool together and send a three employees to represent them all, (396) would also be true. However, our expectations are different, and this implies that we think that (396) speaks about in total 18 representatives. In (397) on the other hand there is no expectation about the sameness or difference of the universities that the individual students visit. Each students visits six universities, but it is quite possible that they visit partly or totally the same universities.

There are two ways of making clear what one wants to say. The first is to use a sentence that is clearer on the point. For example, the following sentences make clear on the point in question what is meant.

(398) `Six companies sent three representatives each to the trade fare.`

(399) `Six companies sent in total three representatives to the trade fare.`

The other option is to use a formal language that has well-defined meanings and into which propositions are rendered. It is the latter approach that is more widespread, since the nuances in meaning are sometimes very difficult to express in natural language. Often enough, a mixture of the two are used. In the previous section we have used indices to make precise which DPs are taken to refer to the same individual. We shall do the same here. For example, we shall use the following notation.

(400) $\mathsf{run}'(x)$

Here, x is a variable which can be filled by an individual, say john' and it will be true or false. So,

(401) $\mathsf{run}'(\mathsf{john}')$

is a proposition, and it is either true or false. This looks like a funny way of saying the same thing, but the frequent use of brackets and other symbols will actually do the job of making meanings more precise. What it does not do, however, is explain in detail what it means, for example, to *run*. It will not help in setting the boundary between *walking* and *running*, even though semantics has a job to

do there as well. It is felt, though, that this part of the job falls into the lexicon and is therefore left to lexicographers, while semantics is (mostly) concerned with explaining how the meanings of complex expressions are made from the meanings of simple expressions.

Semantics II: Compositionality

> The thesis of compositionality says, roughly, that what is in the meaning of a sentence is all in the meaning of its parts and the way they were put together.

Truth Values

The thesis of compositionality says that the meaning of a complex expression is a function of the meaning of its parts and the way they have been put together. Thus, in order to understand the meaning of a complex expression we should not need to know how exactly it has been phrased as long as the two expressions are synonymous. In fact, we have made compositionality a design criterion of our representations. We said that when two constituents are merged together the meaning of the complex expression is the result of applying a function to the meaning of its parts. Rather than study this in its abstract form, we shall see how it works in practice.

We shall use the numbers 0 and 1 for saying whether or not a given sentence is true. The numbers 0 and 1 are therefore also called **truth values** (they are also numbers, of course). 0 represents the 'false' and 1 represents the 'true'. A given sentence is either false or true, so its truth value is either 0 (when it is false) or 1 (when it is true). We imagine that there is a ghost, or machine, or oracle which tells us for every possible statement whether or not it is true or false. In mathematical jargon we call this ghost a **function**. We represent this function by a letter, say g. g needs a string and returns 1 or 0 instead of 'yes' or 'no'. Now, g must abide by the conventions of the language. For example, a sentence of the form S `and` T is true if and only if both S and T are true. So,

(402) $g(S \texttt{ and } T) = 1$ if and only if $g(S) = 1$ and $g(T) = 1$

This can be rephrased as follows. We introduce a symbol $\cap$, which has exactly the following meaning. It is an operation which needs two truth values and returns a truth value. Its table is the following.

(403)

$\cap$	0	1
0	0	0
1	0	1

Now we can write

(404) $g(S \texttt{ and } T) = g(S) \cap g(T)$

This suggests that we take the meaning of /and/ to be exactly $\cap$ so that one may wonder what it is I am saying here. Notice however that $\cap$ is a formal symbol which has the meaning that we have given it; /and/ on the other hand is a word of English whose meaning we want to describe. Thus, to say that /and/ means $\cap$ is actually to say something meaningful! In the same way we can say that the meaning of /or/ is the following function

(405)

$\cup$	0	1
0	0	1
1	1	1

and that the meaning of /if/$\cdots$/then/ is the function

(406)

$\rightarrow$	0	1
0	1	1
1	0	1

This takes a while to understand. We claim that a sentence of the form /if S then T/ is true if either S is false or T is true. Often, people understand such a sentence as saying that S must be true exactly when T is true. But this is not so. (It is an error of fact, to be sure; I am not claiming the meaning of /if·then/ ought to be different from what it is. I am claiming that people do factually use /if$\cdots$then/ in that way. However, it is known that the meaning of /if$\cdots$then is a complex affair.) For example, suppose I order a book at the bookstore and they say to me

(407) `If the book arrives next week, we shall notify you.`

Then if the book indeed arrives and they do not notify me they have issued a false promise. On the other hand, if it does not arrive and still they notify me (saying that it hasn't arrived yet), that is still alright. To give another example: in mathematics there are lots of theorems which say: 'if the Riemann hypothesis is true then...'. The point is that nobody really knows if Riemann's hypothesis is actually true. (If you have a proof, you will be famous in no time!) But the theorems will remain true no matter which way the hypothesis is eventually decided. To see how this follows in our setup notice that we have said that

(408) $g(\texttt{if } S \texttt{ then } T) = g(S) \rightarrow g(T)$

This means that `if` S `then` T is true if either S is false (that is, $g(S) = 0$) or T is true (that is $g(T) = 1$). Why it is that people consider a sentence of the form /`if` S `then` T/ to say the same as /S `if and only if` T/ still needs to be explained. Part of the explanation is pragmatic. It comes from analysing two things: the utterance situation, and the language system. The utterance situation defines in which way we talk and how we actually use language. If talking 'mathematics' the meaning of /`if`$\cdots$`then`/ has a meaning that is fixed by convention in a way it is not in ordinary talk. Thus the example of Riemann conjecture was misleading: it showed us not what /`if`$\cdots$`then`/ ordinarily means, but rather what it means in mathematical discourse. The second aspect comes down to looking at communication as a strategic game using language. The actual utterance meanings are derived from the standard meanings, they are parasitic on the latter. This has the unfortunate consequence that the utterance meaning may confuse us about the semantic meaning. We shall not discuss the any further, though.

Let us see how this works in practice.

(409) `Pete talks and John talks or John walks.`

The question is: does this imply that Pete talks? Our intuition tells us that this depends on which way we read this sentence. We may read this as being formed from the sentences /`Pete talks`/ and /`John talks or John walks`/ using /`and`/; or we can read it as being formed from /`Pete talks and John talks`/ and /`John walks`/ using /`or`/. The string is the same in both cases.

(410) g(`Pete talks and John talks or John walks`)
$=g$(`Pete talks`) $\cap$ (g(`John talks or John walks`))
$=g$(`Pete talks`) $\cap$ (g(`John talks`) $\cup$ g(`John walks`))

(411) g(`Pete talks and John talks or John walks`)
$=g$(`Pete talks and John talks`) $\cup$ g(`John walks`)
$=$(g(`Pete talks`) $\cap$ g(`John talks`)) $\cup$ g(`John walks`)

The results are really different. Suppose for example that Pete does not talk, that John talks and that John walks. Then we have

(412) g(`Pete talks`) $\cap$ (g(`John talks`) $\cup$ g(`John walks`))
$=0 \cap (1 \cup 1)$
$=0$

(413) $(g(\texttt{Pete talks}) \cap g(\texttt{John talks})) \cup g(\texttt{John walks})$
$=(0 \cap 1) \cup 1$
$=1$

So, the different interpretations can give different results in truth!

The True Picture

We have started out by assuming that there is a function g telling us for each sentence whether or not it is true. This function is in a predicament with respect to (409), for it may be that it is both true and false. Does this mean that our picture of compositionality is actually wrong? The short answer is: no! To understand why this is so, we need to actually look closer into the syntax and semantics of /and/. First, we have assumed that syntactically everything is binary branching. So the structure of /Pete talks and John walks/ is

(414) $[[\texttt{Pete talks}]_{\text{CP}}\ [\texttt{and}\ \ [\texttt{John walks}]_{\text{CP}}]]_{\text{CP}}$

Thus, /and/ forms a constituent together with the right hand CP (the latter is also called **conjunct**). And the two together form a CP with the left hand CP (also called **conjunct**). The argument in favour of this is that it is legitimate to say

(415) `Pete talks. And John walks.`

So, we want to have a semantics that gives meanings to all constituents involved. We shall first give the semantics of the expression and T. This is a function which, given some truth value x will return the value $x \cap g(T)$. In mathematics, one writes $x \mapsto x \cap g(T)$. There is a more elegant notation saying exactly the same:

(416) $g(\text{and } T) = \lambda x.x \cap g(T)$

The notation $\lambda x.\cdots$ is similar to the set notation $\{x|\cdots\}$ (or $\{x : \cdots\}$). Given an expression, it yields a function. For example, given the term x^2 it gives the *function* $\lambda x.x^2$, which outputs the square of its input. In ordinary mathemantics one writes this as $y = x^2$, but this latter notation (which is intended to say exactly the same) is actually not useful for the purpose at hand. To avoid the shortcomings of this 'usual' notation, the λ got introduced. In practice, if you know what the answer is in each case, say it is x^2, then $\lambda x.x^2$ is a function (or machine, oracle,

algorithm, whatever you find more instructive) that gives you the answer n^2 when you give it the number n. For example, you give it the number 3, it gives you back 9; you give it 7, it answers 49. And so on. But why is $\lambda x.x^2$ different from x^2? The difference can be appreciated as follows. I can say: *Let* $x^2 < 16$. Because depending on what x is, x^2 is smaller than 16 or it isn't. It is like saying *Let* $-4 < x < 4$. But about the function $f(x) = x^2$ I cannot simply say: *Let* $f < 16$. A function cannot be smaller than some number, only numbers can be. So, there is an appreciable difference between numbers and functions from numbers to numbers. It is this difference that we shall concentrate on.

By our conventions we write $\lambda x.x^2$ instead of f. This is *not* a number, it is a function. You call it, you give it a number, and it gives you back a number. In and of itself, it is not a number. So why use the notation with the λ's? The reason is that the notation is so useful because it can be iterated. We can write $\lambda y.\lambda x.(x-y)$. What is this? It is a function that, when given a number m, calls another function; so you may call it a second order function. The function it calls is the function $G := \lambda x.(x-m)$. G on its turn waits to be given a number, same or different. Give it a number n and it will return the number $n-m$. Notice that $\lambda x.\lambda y.(x-y)$ also is a second order function, but a different one. Give it the number m and it will call the function $H := \lambda y.(m-y)$. Give H the number n and it will give you $m-n$, which is not the same as $n-m$. (For example, $3-5=-2$, while $5-3=2$.)

Let us see about $\lambda x.x \cap g(T)$. Suppose that T is true. Then the function is $\lambda x.x \cap 1$. x has two choices, 0 and 1:

(417) $\quad (\lambda x.x \cap 1)(0) = 0 \cap 1 = 0, \qquad (\lambda x.x \cap 1)(1) = 1 \cap 1 = 1$

Suppose next that T is false. Then $\lambda x.x \cap g(T) = \lambda x.x \cap 0$.

(418) $\quad (\lambda x.x \cap 0)(0) = 0 \cap 0 = 0, \qquad (\lambda x.x \cap 0)(1) = 1 \cap 0 = 0$

Now we are ready to write down the meaning of /and/. It is

(419) $\quad \lambda y.\lambda x.x \cap y$

How does this help us? We shall assume that if two expressions are joined to form a constituent, the meaning of the head is a function that is applied to the meaning of its sister. In the construction of S and T, the head is /and/ in the first step, and and T in the second step. Suppose for example that Pete talks but John does not

walk. Ignoring some detail (for example the morphology), the signs from which we start are as follows:

(420)
$$\begin{bmatrix} 1 \\ \text{CP} \\ \texttt{Pete talks} \end{bmatrix} \quad \begin{bmatrix} \lambda y.\lambda x.x \cap y \\ \text{C} \\ \texttt{and} \end{bmatrix} \quad \begin{bmatrix} 0 \\ \text{CP} \\ \texttt{John walks} \end{bmatrix}$$

Let's put them together.

(421)
$$\begin{bmatrix} 1 \\ \text{CP} \\ \texttt{Pete talks} \end{bmatrix} \bullet \left(\begin{bmatrix} \lambda y.\lambda x.x \cap y \\ \text{C} \\ \texttt{and} \end{bmatrix} \bullet \begin{bmatrix} 0 \\ \text{CP} \\ \texttt{John walks} \end{bmatrix} \right)$$
$$= \begin{bmatrix} 1 \\ \text{CP} \\ \texttt{Pete talks} \end{bmatrix} \bullet \begin{bmatrix} \lambda x.x \cap 0 \\ \text{C}' \\ \texttt{and John walks} \end{bmatrix}$$
$$= \begin{bmatrix} 0 \\ \text{CP} \\ \texttt{Pete talks and John walks} \end{bmatrix}$$

It is not hard to imagine that we can get different results if we put the signs together differently.

Types of Ghosts

We have said that there are functions, and that there are second order functions. They are higher up, because they can call ordinary functions to fulfill a task for them. To keep track of the hierarchy of functions, we shall use the following notation. Ghosts have *types*. Here are the types of the functions introduced so far.

(422)
$$\begin{aligned} 0, 1 &: t \\ \lambda x.x \cap 1 &: t \to t \\ \lambda y.\lambda x.x \cap y &: t \to (t \to t) \end{aligned}$$

The general rule is this:

Definition 15 *A truth value is of* ***type*** *t. A function from objects of type α into objects of type β is a function of* ***type*** *$\alpha \to \beta$. A function of type $\alpha \to \beta$ can only apply to an object of type α; the result is an object of type β.*

Something of Type 1 is a truth value; by definition, it is either 0 or 1. So, it is known to us directly. Everything else needs some computation on our side. A function of type $\alpha \to \beta$ is asking for an object of type α to be given. It will then call an object of type β for an answer. Typically we had $\alpha = t$. An object of type $t \to \beta$ is waiting to be given a truth value and it will then return an object of type β.

Is there more to it than just truth values? There is! We shall assume, for example, that there is a type e that comprises all objects of the world. Every physical object belongs to e, including people, animals, stars, and even ideas, objects of the past like dinosaurs, objects of unknown existence like UFOs, they are all of type e. Ordinary names, like /John/ and /Pete/ can denote objects, and so their meaning is of type e, we say. A function of type $e \to t$ is an object that waits to be given an object and it will return ... a truth value. It says 1 or 0 (or 'yes' or 'no', however we like). These functions are called (**unary**) **predicates**. Intransitive predicates denote unary predicates. For example, /run/, /talk/ and /walk/. We shall truly treat them as functions. We do not know how they work. This is part of the lexicon, or if you wish, this part of the knowledge of the language that we must acquire what it means that someone talks, or walks, or runs. In the absence of anything better we write run' for the function of type $e \to t$ that tells us if someone is running. And likewise we write walk' for the function that tells us if someone is talking. And so on. Then we do however know that the meaning of

(423) John talks.

is $\mathsf{talk}'(\mathsf{john}')$. Big deal. But things get trickier very soon. There are also transitive predicates like /scold/. Since they form a category of intransitive verb together with a name, we conclude that they must be functions of second order: they are of type $e \to (e \to t)$. And so on. What this means is that semantics mirrors syntax by way of types. If something is a transitive verb, not only does it need two syntactic arguments, also its meaning is a second order function, waiting to be given two objects before it will answer with a truth-value.

Notes to this section. I have earlier (at the end of Lecture 8) required that grammars of English contain the rule

(424) CP → CP␣and␣CP

but the present rules do not have it. Nevertheless, it still satisfies the properties ① – ④ on Page 92. And this is because the rule (424) is what one calls a derived rule

of the grammar, obtained by doing three steps: from CP we step to CP C′ and in a second step to CP C CP and finally to CP␣and␣CP.

Semantics III: Basic Elements of Interpretation

> The things that exist come in different forms. There are objects, time points, events, truth values, and so on. These properties are fundamental. A truth value can never be a number regardless of the fact that we use numbers (or letters) to stand in for them.

On What There Is

With language we can seemingly talk about everything that exists. While this can (at least abstractly) be shown to be impossible, the range of things that we can talk about is immense. However, this vast expressive power creates its own challenges. For to the human mind the things around us are of different nature. There are physical objects, people, places, times, properties, events and so on. Languages reflect this categorisation of things in different ways. For example: many languages classify things into various groups, called *genders* or *classes*. This classification proceeds along several lines. Many languages distinguish animate from inanimate, for example. And within the animate between people and animals, and within the people the male and the female, and so on. This division is in no way objectively given. Why should we care to distinguish humans from animals, or animals from stones? Clearly, the answer is that language is shaped to serve the needs of humans, not dogs or stones. And further: language answers to the need of a society and culture.

The division of things into genders or classes is realised in different ways. For some languages it is part of the morphology (French, for example) for other it is irrelevant (Hungarian). One has made a lot out of that difference; the so-called **Sapir-Whorf-Thesis** states that the language we speak has a profound influence on the way we think. The reasoning is that if gender distinctions are obligatory then we need to pay more attention to it and consequently take this to be a feature worth attending to. Nowadays one tends to think that the influence of language on perception has been much overstated. The fact that the Hopi language does not distinguish past from furture does not mean they cannot tell the difference or are in any way less likely to understand it. Yet there seem to be areas where the thesis has its merits, be it only in categorical perception of sounds. Another example spatial language. It has been claimed that there are languages where there

are only absolute spatial prepositions of the type /north/, /south/, and that the speakers of these languages are far better at tracking absolute directions.

To the extent that the classification of things is relevant to language, it is going to be reflected in the basic semantic *types*. We shall review a few basic types of things that are relevant to natural languages.

Number: Individuals and Groups

Look at the following contrast.

(425) John and Mary met.

(426) The senators met.

(427) The committee met.

(428) *John met.

Evidently, in order to meet you need to meet with somebody. But apparently several people can meet in the sense that they meet with each other. One way to account for this difference is to say that the verb /meet/ needs a group of people as its subject. Since committees, lines and other things can also meet, it is not simply a fact of grammar, in other words, it cannot be decided by syntactic means alone whether a sentence is well formed, witness the distinction between (427) and (428). So we say instead that /meet/ needs a group of things, without further qualification. There are basically two ways to form groups: you use the plural or you use /and/.

Thus there are basic DPs that denote individuals, and those that denote groups. It is however to be noted that singular DPs rarely denote groups directly; rather, they denote some entity that is constituted (among other things) by its members (like a committee or a parliament). They therefore can be used as group denoting DPs only in a derived sense. The primary source of group denoting expression is something else. It is plural DPs and DPs coordinated with the help of /and/. Groups can evidently be subjects of verbs, and sometimes verbs specifically require groups as subjects. Other verbs take both.

(429) Paul is eating.

(430) Paul and the neighbouring cat are eating.

In the case of /eat/ we may think of it as being an activity that is basically performed by each individual alone. If it is understood to be this way, the verb is said to be **distributive**. It means in the present case that the fact that Paul is eating and that the neighbouring cat is eating is enough to make the sentence (430) true. We can also say that the following inference is correct.

(431)
Paul is eating.
The neighbouring cat is eating.
∴ Paul and the neighbouring cat are eating.

In general, if a verb V is distributive, then the following inferences go through:

(432)
A Vs.
B Vs.
∴ A and B V.

A and B V.
∴ A Vs.

A and B V.
∴ B Vs.

Now, if there are groups of people, are there also groups of groups? There are! Here is an interesting pattern. You can get married as a group of two people by getting married to each other. This involves a one-time event that makes you married to each other. You can get married as a group by each marrying someone else. The sentence (433) can be read in these two ways. The reason why this is so lies in the fact that /get married/ actually also is a verb that takes individuals.

(433) John and Sue got married.

(434) John got married.

Moreover, the above test of distributivity goes through in that case. But if we understand it in the non-distributive sense, the inference does not go through, of course. Now, let's form a group of groups:

(435) John and Alison and Bill and Sue got married.

There is a reading which says that John marries Alison, and Bill marries Sue. This reading exists in virtue of the following. /John and Alison/ denotes a group, so does /Bill and Sue/. The verb applies to such groups in the meaning 'marry each other'. By our understanding of marriage, if several groups get married to each other, this means that all groups get married separately.

There are also verbs that encourage a group of groups reading.

(436) The ants were marching eight by eight.

Here, we think of a group of groups of ants, each group consisting of eight ants in a line.

Note that the same pattern can be observed with /meet/:

(437)
The senators meet.
The congressmen meet.
∴ The senators and the congressmen meet.

However, reader beware: the conclusion has a reading where the subject is a single group which meets in a one-time event. This is not what can be concluded from the premises. All that follows is that the senators met (with each other) and that the congressmen met (with each other).

Time

Properties can change over time. I can be tired in the evening, but maybe next morning I am not. Some properties have an inbuilt time dependency. For example, a cat is a kitten only through certain stages of its life. When it is old enough it ceases to be a kitten, even though it never ceases to be a cat. We picture time as consisting of a continuum of time points on a line. We write t, t', u for time points, and we write $t < t'$ to say that t is prior to t', and $t > t'$ to say that it is after t'. For any two time points t and t', either $t = t'$ (they are the same) or $t < t'$ or $t > t'$. This trichotomy translates into the three basic tenses of English. The present is used for something that happens now, the past is used for something that has happened before now, and the future is used for something that will happen later.

(438) John runs. (present)

(439) John ran. (past)

(440) John will run. (future)

We make reference to time through various ways. One is the words /now/ and /yesterday/. /now/ refers to the very moment of time where it is uttered; textt-tyesterday/ refers to any time point that is on the day before today. Today on the other hand is the day of 'now'. Other words require some calculation.

(441) John realized on Monday that he had lost his wallet the day before.

We do not know exactly when things happened. We know that John's realising he had lost his wallet happened in the past (because we used past tense); and it happened on a Monday. His losing the wallet happened just the day before that Monday, so it was on a Sunday. Suppose we replaced /`the day before`/ by /`yesterday`/:

(442) `John realized on Monday that he had lost his wallet yesterday.`

Then John's realizing is in the past, and it is on a Monday. His losing his wallet is prior to his realizing (we infer that among other from the phrase /`had lost`/ which is a tense called **pluperfect**). And it was yesterday. So, today is Monday and yesterday John lost his wallet, and today he realizes that. Or he realized yesterday that on that day he had lost his wallet. (Actually, the phrase /`on Monday`/ is dispreferred here. We are not likely to say exactly what day of the week it is when it is today or yesterday or tomorrow. But that is not something that semantics concerns itself with. I can say: *let's go to the swimming pool on Thursday* even when today is Thursday. It is just odd to do so.)

That time is linear and transitive accounts for the following inference patterns.

(443)
$$\frac{\begin{array}{l}\text{A Ved.}\\ \text{B Ved.}\end{array}}{\therefore\ \text{A Ved before B or A Ved after B or A Ved at the same as B.}} \qquad \frac{\begin{array}{l}\text{A Ved before B.}\\ \text{B Ved before C.}\end{array}}{\therefore\ \text{A Ved before C.}}$$

Location

Space is as important in daily experience as time. We grow up thinking spatially; how to get furniture through a door, how to catch the neighbour's cat, how to not get hit by a car, all these things require coordination of actions and thinking in time and space. This is the reason why language is filled with phrases that one way or another refer to space. The most evident expressions are /`here`/, which functions the same way as /`now`/, and /`there`/ (analogous to /`then`/). It denotes the space that speaker is occupying at the moment of utterance. It is involved in the words /`come`/ and /`go`/. If someone is approaching me right now, I can say

(444) `He is coming.`

But I would not say that he is going. That would imply he is moving away from me now. German uses verbal prefixes (/hin/ and /her/) for a lot of verbs to indicate whether movement is directed towards speaker or not.

Space is not linear, it is organized differently, and language reflects that. Suppose I want to say where a particular building is located on campus. Typically we phrase this by giving an orientation and a distance (this is known as 'polar coordinates'). For example,

(445) `200 m southwest of here`

gives an orientation (southwest) and a distance (200 metres). The orientation is either given in absolute terms, or it can be relative to the way one is positioned, for example /`to the right`/. To understand the meaning of what I am saying when I say /`Go to the right!`/ you have to know which way I am facing.

Worlds and Situations

We have started out by saying that sentences are either true or false. So, any given sentence such as the following is either true or false.

(446) `Paul is chasing a squirrel.`

(447) `Napoleon lost the battle of Waterloo.`

(448) `Kitten are young cats.`

We can imagine with respect (446) that it is true right now or that it is false. In fact, we do not even know. With (447) it is the same, although if we have learned a little history we will know that it is true. Still, we find ourselves thinking 'what if Napoleon had actually won the battle of Waterloo ...'. Thus, we picture a situation that is contrary to fact. The technical term is **world**. Worlds decide every sentence one way or another. There are worlds in which (446) is true and (447) is false, others in which (446) is false and (447) is true, others in which both are false, and again others in which both are false. And there is one (and only one) world we live in.

Seemingly then, any combination of saying this and that sentence is true or false is a world. But this is not quite true. (448) is different. It is true. To suppose otherwise however would be tantamount to violating the rules of language. If I

were to say 'suppose that kittens are not young cats but in fact old rats ...' what I ask you is to change the way English is understood. I am not talking about a different world. Worlds have an independent existence from the language that is being used. We say then that (448) is **necessarily true**, just like /4+7=11/. If you do not believe either of them you are just not in the picture.

The denotation of a word like /cat/ in this world is the set of all beings that are cats. They can change from world to world. We can imagine a world that has absolutely no cats. (If we go back in time, there was a time when this was actually true.) Or one that has no mice. But we do not suppose that just because there are different sets of cats in different worlds the meaning of the word changes—it does not. That's why you cannot suppose that kittens are old rats. We say that the meaning of the word /cat/ is a function cat' that for each and every world w gives some set, $\mathsf{cat}'(w)$. We of course understand that $\mathsf{cat}'(w)$ is the set of all cats in w. (Some people use the word **intension** for that function.) Likewise, the intension of the word /rat/ gives for each world w the set of all rats in w, and likewise for the word /kitten/. It is a fact of English that

(449) $\mathsf{kitten}'(w) \subseteq \mathsf{cat}'(w), \qquad \mathsf{kitten}'(w) \cap \mathsf{rat}'(w) = \varnothing$

There are plenty of words that are sensitive not just to the denotation but to the meaning.

(450) John doubts that Homer has lived.

(451) Robin thinks that Napoleon actually won Waterloo.

Nobody actually knows whether or not Homer has existed. Still we think that the sentence 'Homer has lived.' has a definite answer (some ghost should tell us...). It is either true or not. Independently of the answer, we can hold beliefs that settle the question one way or the other, regardless of whether the sentence is factually true or not. Robin, for example, might be informed about the meaning of all English words, and yet is a little weak on history. So she holds that Napoleon won Waterloo. John might believe the opposite, and Robin might believe that Homer has lived. Different people, different opinions. But to disagree on the fact that kittens are cats and not rats means not using English anymore.

Events

When I sit behind the computer typing on the keyboard, this is an activity. You can watch me do it and describe the activity in various ways. You can say

(452) `He is typing on the keyboard.`

(453) `He is fixing the computer.`

(454) `He is writing a new book.`

Both (452) and (453) can be manifested by the same physical activity: me sitting behind the computer and typing something in. Whether or not I am fixing the computer by doing so, time will tell. But in principle I could be fixing the computer just by typing on the keyboard. The same goes for writing a book. Thus, one and the same physical activity can be a manifestation of various different things. In order to capture this insight, one has proposed that verbs denote particular objects called **events**. There are events of typing, as there are events of fixing the computer, and events of writing a book. Events are distinct from the processes that manifest them.

Events will figure in Lecture 18, so we shall not go into much detail here. There are a few things worth knowing about events. First, there are two kinds of events: **states** and **processes**. A state is an event where nothing changes.

(455) `Lisa knows Spanish.`

(456) `Harry is 8 feet tall.`

These are examples where the truth of something is asserted at a moment of time, but there is no indication that something changes. By contrast the following sentences talk about processes.

(457) `Paul is running.`

(458) `The administrator is filling out the form.`

(459) `The artist is making a sculpture.`

In the first example, Paul is changing place or posture. In the second the administrator is for example writing on some piece of paper which changes that very piece of paper. In the third example a statue comes into existence from some lump of material. Events have participants whose number and contribution can

vary greatly. A process always involves someone or something that undergoes change. This is called the **theme**. In (457) the theme is Paul, in (458) the theme is the form, in (459) the theme is the sculpture. Events usually have a participant that makes the change happen; in (457) the actor is again Paul, in (458) it is the administrator, in (459) it is the artist. There need not be an actor, just as there need not be a theme; but mostly there is a theme. Some events have what is called an **experiencer**. In the next sentence, Jeremy is the experiencer of hate.

(460) `Jeremy hates football.`

Notice that experiencer predicates express states of emotion, so they fall into the category of verbs that express states rather than processes. Another class of participants are the **beneficiaries**; these are the ones for whose benefit an action is performed, like the wife of the violinist in the following example.

(461) `The violinist composed a sonata for his wife.`

The list of participant types is longer, but these ones are the most common ones.

Processes are sometimes meant to finish in some state and sometimes not. If you are said to be running no indication is made as to when you will stop doing so. If you are said to be running a mile then there is an inherent point that defines the end of the activity you engage in. Some verbs denote the latter kind of events: /`arrive`/, /`reach`/, /`pop`/, /`finish`/. The process they denote is finished when something specific happens.

(462) `Mary arrived in London.`

(463) `The composer finished the oratorio.`

In (462) the arriving of Mary happens at a more or less clearly defined time span, say when the train gets close to the station up to when it comes to a halt. Similarly for (463), where the finishing is the last stretch of the event of writing the oratorio. The latter is a preparatory action. So, you can write an oratorio for a long time, maybe years, but you can only finish it in a considerably shorter time, at the end of your writing it.

Notes on this section. There are most likely to be a few more sorts, for example, degrees. Notice that there is a big difference between the classification here and what is nowadays in computer science called an ontology, which is a rather rich classification of things.

Semantics IV: Scope

> Different analyses of sentences give rise to different c-command relations between constituents. These in turn determine different interpretations of sentences. Thus one of the reasons why sentences can mean different things is that they can have different structures.

Let us return to example (409), repeated here as (464).

(464) `Pete talks and John talks or John walks.`

We have said that under certain circumstances it may turn out to be both true and false, depending on how we read it. These interpretations are also called **readings**. In this lecture we shall be interested in understanding how different readings may also be structurally different. (Structural differences are not the only reason for different interpretations.) The syntactic notion that is pivotal here is that of **scope**. We shall say that every constituent has a scope; the scope is that string part (constituent) that serves as a semantic argument to it. This definition is semantic by intention. However, we can also give syntactic correspondences for the scope and thereby eliminate reference to semantics. For example, for a head the scope is exactly its complement. We shall fill this definition with life right away. In (464) we find two logical connectives, /and/ and /or/. Each of them takes two CPs, one to the right and one to the left. This is all the requirements they make on the syntactic side. This means that syntactically the sentence can be given two different structures. They are shown in Figure 10. The structure of the individual CPs is not shown, to save space. Let us look at (a). We can tell from its structure what its meaning is. We shall work it out starting at the bottom. The complement of /or/ is the CP /John walks/. So, /or John walks/ is a constituent formed from /or/ and /John walks/. This means that its meaning is derived by applying the meaning of /or/ to the meaning of /John walks/. Notice that the latter is also the constituent that the node labelled C just above /or/ c-commands (recall the definition of c-command form Lecture 12). The meaning of the lower C′ node is therefore

(465) $(\lambda y.\lambda x.x \cup y)(g(\texttt{John walks}))$
$= \lambda x.x \cup g(\texttt{John walks})$

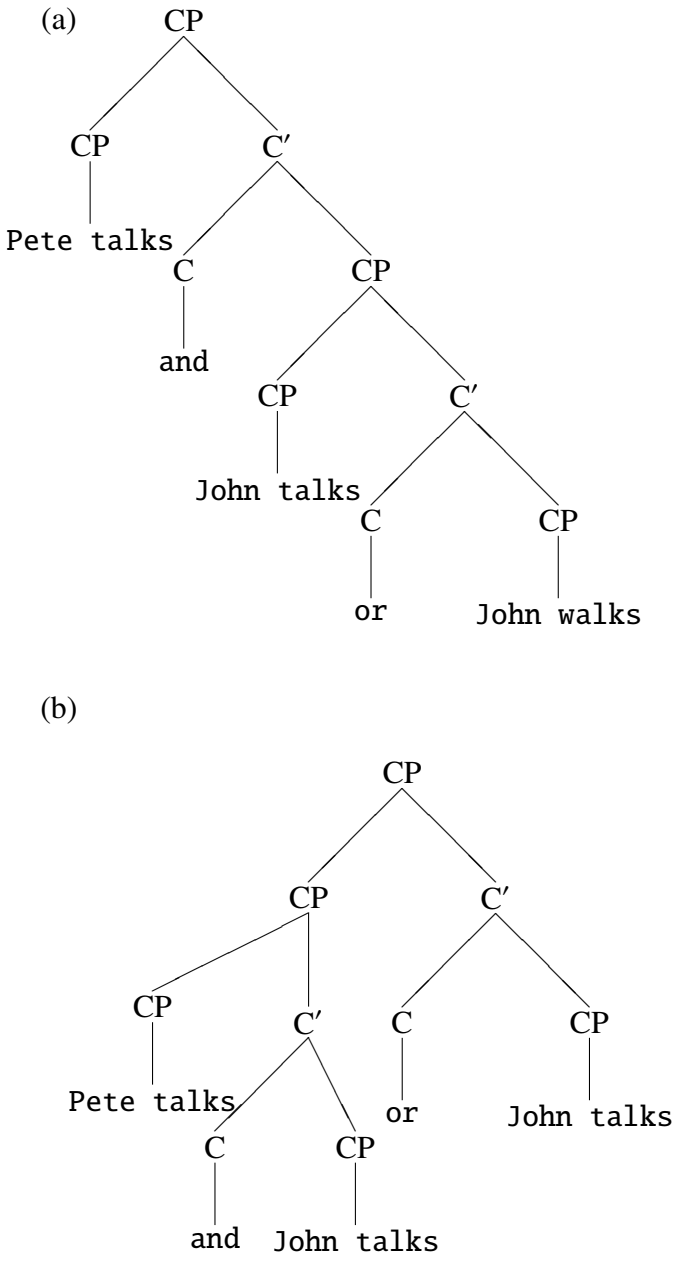


Figure 10: Two Analyses of (464)

This node takes the CP node above /John talks/ as its sister, and therefore it c-commands it. The meaning of the CP that the two together form is

(466) $(\lambda x.x \cup g(\texttt{John walks}))(g(\texttt{John talks}))$
$=g(\texttt{John talks}) \cup g(\texttt{John walks})$

This is the meaning of the lower CP. It is the complement of /and/. The two form a constituent, and its meaning is

(467) $(\lambda y.\lambda x.x \cap y)(g(\texttt{John talks}) \cup g(\texttt{John walks}))$
$=\lambda x.x \cap (g(\texttt{John talks}) \cup g(\texttt{John walks}))$

Finally, we combine this with the specifier CP:

(468) $(\lambda x.x \cap (g(\texttt{John talks}) \cup g(\texttt{John walks})))(g(\texttt{Pete talks}))$
$=g(\texttt{Pete talks}) \cap (g(\texttt{John talks}) \cup g(\texttt{John walks}))$

This is exactly the same as the interpretation (410).

Now we take the other structure. This time we start with the interpretation of the middle CP. It is now c-commanded by the C above the word /and/. This means that the two form a constituent, and its interpretation is

(469) $(\lambda y.\lambda x.x \cap y)(g(\texttt{John talks}))$
$=\lambda x.x \cap g(\texttt{John talks})$

This constituents take the specifier /Pete talks/ into a constituent CP. Thus, it applies itself to the meaning of /Pete talks/:

(470) $(\lambda x.x \cap g(\texttt{John talks}))(g(\texttt{Pete talks}))$
$=g(\texttt{Pete talks}) \cap g(\texttt{John talks})$

This constituent is now the specifier of a CP, which is formed by /Pete talks and John talks/ and /or John walks/. The latter has the meaning ($\lambda x.x \cup g(\texttt{John walks})$), which applies itself to the former:

(471) $(\lambda x.x \cup g(\texttt{John walks}))(g(\texttt{Pete talks}) \cup g(\texttt{John talks})))$
$=(g(\texttt{Pete talks}) \cup g(\texttt{John talks})) \cap g(\texttt{John walks})$

And this is exactly (411).

Thus, the two different interpretations can be seen as resulting from different structures. This has motivated saying that our function g does not take sentences as inputs. Instead, it wants the entire syntactic tree. Only when given a syntactic tree the function can give a satisfactory answer. The way the meaning is computed is by working its way up. We assume that each node has exactly two daughters. Suppose we have a structure

(472) $\gamma = [\alpha\, \beta]$

We assume that the semantics is arranged in such a way that if two constituents are merged into one, the interpretation of one of the nodes is a function that can be applied to the meaning of its sister. The first node is then called the **semantic head**. If the meaning of α and β is known and equals $g(\alpha)$ and $g(\beta)$, respectively, then the meaning of γ is

(473) $$g(\gamma) = \begin{cases} g(\alpha)(g(\beta)) & \text{if } \alpha \text{ is the semantic head} \\ g(\beta)(g(\alpha)) & \text{if } \beta \text{ is the semantic head} \end{cases}$$

The only thing we need to know is: is α the semantic head or is it β? The general pattern is this: a zero level projection is always the semantic head, and likewise the first level projection. Adjuncts are semantic heads, but they are not heads (the latter notion of head is a syntactic notion and is different, as this case shows). Notice that by construction, the semantic head 'eats' its sister: its meaning is a function that applies itself to the meaning of the sister. And the sister is the constituent that it c-commands. This is why c-command has become such a fundamental notion in syntax: it basically mirrors the notion of scope, which is the one needed to know what the meaning of a given constituent is.

We shall discuss a few more cases where scope makes all the difference. Look at the difference between (474) and (475).

(474) `This is the only car we have, which has recently been repaired.`

(475) `This is the only car we have which has recently been repaired.`

The part /`which has recently been repaired`/ is a clause that functions like an adjective; it is called a **relative clause**. (Recall that relative clauses are opened by relative pronouns in English.) Unlike adjectives, relative clauses follow the

noun. Notice that /we have/ also is a relative clause, though it is somewhat shortened (we could replace it /which we have/).

Suppose you go to a car dealer and he utters (474). Then he is basically saying that he has only one car. Moreover, this car has been under repair recently. Suppose however he says (475). Then he is only claiming that there is a single car that has been under repair recently, while he may still have tons of others. It is clear from which dealer you want to buy. To visualize the difference, we indicate the scope of the operator /the only/:

(476) This is the only [car we have], which has recently been repaired.

(477) This is the only [car we have which has recently been repaired].

In the first sentence, /the only/ takes scope over /car we have/. It has scope over /car we have which has recently been repaired/ in the second sentence. It is not necessary to formalize the semantics of /the only/. We need only say that something is *the only P*, if and only if (a) it is a *P*, and (b) nothing else is a *P*. So, if /the only/ takes scope only over /car we have/, then the car we are talking about is a car the dealer has (by (a)), and there is no other that he has (by (b)). So he has only one car. If the scope is /car we have which has recently been repaired/, then the car we are talking about has recently been repaired (by (a)), it is one of the dealer's cars (also by (a)), and there is no other car like that. So, there may be other cars that the dealer has but they have not been repaired, and there may be other cars that were not the dealer's but have been repaired.

The difference in structure between the two is signaled by the comma. If the comma is added, the scope of /the only/ ends there. The relative clause is then said to be **non-restrictive**; if the comma is not there, the relative clause is said to be **restrictive**. If we look at $\overline{X}$-syntax again we see that non-restrictive relative clauses must be at least D′-adjuncts, because they cannot be in the scope of /the/. In fact, one can see that they are DP-adjuncts. Let us see how this goes. First we notice that /the/ can be replaced by a possessive phrase (which is in the genitive):

(478) Simon's favourite bedtime story

(479) Paul's three recent attacks on squirrels

The possessives are phrases, so they are in specifier of DP, quite unlike the determiner /the/ itself. Notice that even though the possessive and the determiner cannot co-occur in English this is not due to the fact that they compete for the same position. In Hungarian they can co-occur:

(480) Mari-nak a cipője
Mary-DAT the shoe-POSS3SG
Mary's shoe

Literally translated, this means *Mary's the her shoe*. The possessive /Marinak/ (in the dative!) occurs before the actual determiner. Now, the actual structure we are proposing is this (notice that the 's is analysed as the D-head; there are other possibilities, for example taking it to be a case marker (which I prefer)):

(481) $[_{DP}[_{DP}$Paul$][_{D'}[_{D}$'s$][_{NP}$three recent attacks on squirrels$]]]$

Now, take a DP which has non-restrictive relative clause.

(482) Simon's only bedtime story, which he listens to carefully

(483) Paul's only attacks on squirrels, which were successful

These DPs are perfect, but here we have /Simon's only/ and /Paul's only/. We shall not go into the details of that construction and how it differs from /the only/. In the first DP, /Simon's only/ takes /bedtime story/ in its scope. It can only do so if the relative clause is an adjunct to DP. I remark here that I have not given an analysis of the construction /A's only B/ just of /the only B/, though this can easily be done. Moreover, there are some problems in making the string /the only/ and /Simon's only/ a constituent, so as to make the above semantic interpretation mechanism work. Inasfar as the string /the only/ is concerned we can make /only/ an adjective, thus forming part of the NP. /the only/ is then no longer a constituent. For possessors, we should then first make them possessors of the NP: /Paul recent attack on squirrels/, make /only/ take this into its scope, and then finally add the head /'s/. How to reconcile this with the surface syntax is a moot point. (It costs us some movement steps. Alternatively, we need to postulate some infix operations for syntax so as to infix /only/ into /Paul recent attack on squirrels/ to yields /Paul only

recent attack on squirrels/, and another to give /Paul's only recent attack on squirrels/.)

Likewise one may wonder about the place of restrictive relative clauses. It is clear that they can be neither adjuncts to DP nor adjuncts to D′ (because then /the only/ cannot take scope over them). The restrictive relative clauses is therefore adjunct to either N′ or NP. We shall not go into more detail.

So far we have seen that differences in interpretation manifest themselves in differences in structure. The next example is not of the same kind—at least at first sight. This example has to do with quantification. Suppose that the professors complain about office space and the administrator says

(484) Every professor has an office.

He might be uttering a true sentence even if there is a single office that is assigned to all professors. If this is to be understood as a remark about how generous the university is, then it is probably just short for

(485) Every professor has his own office.

in which case it would be clear that each professor has a different office. The first reading is semantically stronger than the second. For there is a single office and it is assigned to every professor then every professor has an office, albeit the same one. However, if every professor has an office, different or not, it need not be the same that there is just a single office. We can use 'stilted talk' to make the difference visual:

(486) There is an office such that it is assigned to every professor.

(487) Every professor is such that an office is assigned to him.

In the first sentence, /there is an office/ takes scope over /every professor/. In the second sentence /every professor/ takes scope over /an office/.

Returning to the original sentence (485), however, we have difficulties assigning different scope relations to the quantifiers: clearly, syntactically /every professor/ takes /his own office/ in its scope. This problem has occupied

syntacticians for a long time. The conclusion they came up with is that the mechanism that gets the different readings is syntactic, but that the derivation at some point puts the object into c-commanding position over the subject.

There are other examples that are not easily accounted for by syntactic means. Let us give an example.

(488) `John searches for the holy grale.`

There are at least two ways to understand this sentence. Under one interpretation it means that there is something that John searches for, and he thinks it is the holy grale. Under another interpretation it means that John is looking for something that is the holy grale, but it may not even exist. This particular case is interesting because people are divided over the issue whether or not the holy grale existed (or exists, for that matter). Additionally, it is not clear what it actually was. So, we might find it and not know that we have found it. We may paraphrase the readings as follows.

(489) `There is the holy grale and John searches for it as that.`

(490) `There is something and John searches for it as the holy grale.`

(491) `John is searching for something as the holy grale.`

Here, the meaning difference is brought out as a syntactic scope difference. The first is the strongest sentence: it implies that both speaker and John identify some object as the holy grale. The second is somewhat weaker: according to it, there is something of which only John believes that it is the holy grale. The third is the weakest: John believes that there is such a thing as the holy grale, but it might not even exist. Such differences in meaning are quite real. There are people who do in fact search for the holy grale. Some of them see in it the same magical object as which it is described in the epics, while others do not think the object has or ever had the powers associated to it. Rather, for them the importance is that the knights of King Arthur's court possessed or tried to possess it. So they interpret the medieval stories as myths that nevertheless tell us about something real (as Schliemann believed that Homer's Iliad may not have been about the Greeks and their gods, but at least about the fall of a real city, one that he later found). If the first sentence is true then speaker and John agree in the identity of the holy

grale; if they do not agree, then either of them must think of the other that they are mistaken about the identity of the holy grale. In the second sentence the speaker concedes that John is at least concerned about the properties of an existing object, say some bowl or dish, to which John however attributes some mystical power.

Now, whether or not these differences can be related back to differences in structures corresponding to (488) remains to be seen. Such a claim is difficult to substantiate.

Semantics V: Cross-Categorial Parallelism

> There is an important distinction in the study of noun denotations between count nouns and mass nouns. An equally important distinction is between processes and accomplishments/achievements. It is possible to show that the division inside the class of nouns and inside the class of verbs is quite similar.

We have said that nouns denote objects and that verbs denote events. It has been observed, however, that some categorisations that have been made in the domain of objects carry over to events and vice versa. They ultimately relate to an underlying structure that is similar in both of them. A particular instance is the distinction between mass and count nouns. A noun is said to be a **count noun** if what it refers to in the singular is a single object that cannot be conceived of consisting of parts that are also objects denoted by this noun; for example, 'bus' is a count noun. In the singular it denotes a thing that cannot be conceived as consisting of two or more busses. It has parts, for sure, such as a motor, several seats, windows and so on. But there is no part of it which is again a bus. We say that the bus is an **integrated whole** with respect to being a bus. Even though some parts of it may not be needed for it to be a bus, they do not by themselves constitute another bus. Ultimately, the notion of integrated whole is a way of looking at things: a single train, for example, may consist of two engines in front of several dozens of wagons. It may be that it has even been obtained by fusing together two trains. However, that train is not seen as two trains: it is seen as an integrated hole. That is why things are not as clear cut as we might like them to be. Although we are mostly in agreement as to whether a particular object is a train or not, or whether it is two trains, an abstract definition of an integrated hole is hard to give.

The treatment of count nouns in formal semantics has therefore been this. A noun denotes a certain set of things, the integrated wholes of the kind. The word /mouse/ denotes, for example, the set of all mice in this world; call that set M. A particular object is a mouse if and only if it is in M. Groups of mice are subsets of M. Therefore, the plural /mice/ denotes the set of all subsets of M. (Or maybe just the subsets that contain more that one element, but that is not of essence here.) Let us then say this: /A is a B/ is true if what /A/ denotes is an element of the denotation of /B/. Then /Paul is a mouse./ is true if and only if Paul is a member of M, that is, if Paul is a mouse. We shall say the same about /A are B/: it is true if and only if what /A/ denotes is in /B/. Except that the agreement on the

verb has to match that of /A/ (which is not a semantic fact). Therefore /Paul and Doug are mice./ is true if and only if the denotation of /Paul and Doug/ is in the set of all subsets of M, that is, if it is a subset of M. The denotation of /Paul and Doug/ is among other things the set consisting of Paul and Doug. This is a subset of M exactly if both Paul and Doug are mice.

Let us return to the issue of dividing objects of a kind. It seems clear that the division into smaller and smaller units must stop. A train cannot consist of smaller and smaller trains. At some point, in fact very soon, it stops to be a train. There is a difference with water. Although in science we learn that things are otherwise, in actual practice water is divisible to any degree we like. And this is how we conceive of it. We take an arbitrary quantity of water and divide it as we please—the parts are still water. Thus, /water/ is not a count noun; it is a **mass noun**.

One problem remains, though. We have not talked about mass things, we have consistently talked about mass or count *nouns*. We said, for example, that /water/ is a mass noun, not that the substance water itself is a mass substance. In this case, it is easy to confuse the two. But there are occasions where the two are different. The word /furniture/ turns out to be a mass noun, even though what it denotes clearly cannot be indefinitely divided into parts that are also called furniture. But how do we know that something is a mass noun if we cannot ultimately rely on our intuitions about the world? There are a number of tests that establish this. First, mass nouns do not occur in the plural. We do not find */furnitures/ or */courages/. On the surface, /waters/ seems to be an exception. However, the denotation of /waters/ is not the same as that of a plural of a count noun, which is a group. Waters is used, for example, with respect to clearly defined patches of water (like rivers or lakes). A better test is this one. Mass nouns freely occur with so-called **classifiers**, while count nouns do not.

(492) a glass of water

(493) a piece of furniture

(494) a *glass/*piece of bus

Notice that one does not say */a glass of furniture/, nor a */piece of water/. The kind of classifier that goes with a mass noun depends on what it actually denotes. Some can be used across the board, like /lot/. Notice that classifiers must be used without the indefinite determiner. So, with respect to (494), we can say /a piece of a bus/ but then /piece/ is no longer a classifier.

There is an obvious distinction between whether or not objects of a kind are divisible into objects of the same kind, and whether a language calls the noun denoting this kind a mass noun. These must be kept separate. To a certain degree languages exercise their freedom of seeing the world in a different light. Some nouns are mass nouns when comes to using the tests, but everybody knows the kind they denote is not divisible; an example is /furniture/. A someone less clear case is /hair/. Notice furthermore that whether or not things of a kind can be divided into things of the same kind depends on what you believe the world to be like. This is clearest when we look at /water/. The scientific world view that the process of division must come to a halt. Eventually we have a single molecule of water, and here the division must stop. Yet, our own experience is different. As humans we experience water as arbitrarily divisible. The language we speak has been formed not with scientific world view in mind; there is no authority that declares that from now on /water/ must cease to be used as a mass noun. Anyhow, we see that there are example where even the naive experience tells us differently. It is the same as gender distinctions: for some languages they are not morphologically relevant, but the division into various sexes, or other classes can usually be represented one way or another.

Let us now look at verbs. Verbs denote events, as we have said. Events are like films. We may picture them as a sequence of scenes, lined up like birds on a telefone cable. For example, scene 1 may have Paul 10 feet away from some squirrel; scene 2 sees Paul being 8 feet away from the squirrel; scene 3 sees Paul just 6 feet away from the squirrel; and so on, until he is finally right next to it, ready to eat it. Assume that the squirrel is not moving at all. This sequence can then be summarized by saying:

(495) `Paul is attacking the squirrel.`

Similarly, in scene 1 someone is behind a blank paper. In scene 2, he has drawn a small line, in scene 3 a fragment of a circle. From scene to scene this fragment of a circle grows, until in the last scene you see a complete circle. You may say

(496) `He has drawn a circle.`

While you are watching the film, you can say

(497) `He is drawing.`

Or you can say

(498) `He is spreading ink on the paper.`

All these are legitimate ways of describing what is or was going on. Unfortunately, the director has decided to cut a few of the scenes at the end. So we now have a different film. Now we are not in a position to truthfully utter (496) on the basis of the film any more. This is because even though what that person began to draw looks like the beginning of a circle, that circle may actually never have been completed. Notice that the situation is quite like the one where we stop watching the movie: we are witnessing part of the story and guess what the rest is like, but the circle may not get completed. However, (497) and (498) are still fine no matter what really happens. Even if the director cuts parts of the beginning, still (497) and (498) are fine. No matter how short the film is and no matter what really happen thereafter: the description is adequate.

This is the same situation as before with the nouns. Certain descriptions can be applied also to subparts of the film, others cannot. Those events that can be divided are called **atelic**; the others being **telic**. This is the accepted view. One should note though that by definition, a telic event is one that ends in a certain state, without which the event would not be the same. In other words: if we cut out parts of the beginning, that would not hurt. But if we could out parts of the end, that would make a difference. An example is the following.

(499) `John went to the station.`

Here, it does not matter so much where John starts out from, as long it was somewhere away from the station. We can cut parts of the beginning of the film, still it is a film about John's going to the station. Telic events are directed towards a goal (that gave them their name; in Ancient Greek, 'telos' meant 'goal'.). However, as it appears, the majority of nondivisible events are telic. A different one is

(500) `John wrote a novel.`

Here, cutting out parts of the film anywhere will result in making (500) false, because John did not write the novel in its entirety.

Now, how do we test for divisibility (= atelicity)? The standard test is to see whether /`for an hour`/ is appropriate as opposed to /`in an hour`/. Divisible events can occur with /`for an hour`/, but not with /`in an hour`/. With indivis-

ible events it is the other way around.

(501) John wrote a novel in an hour.

(502) *John wrote a novel for an hour.

(503) *John was writing in an hour.

(504) John was writing for an hour.

So, divisible events can be distinguished from nondivisible events. However, let us see if we can make the parallel even closer. We have said that mass terms are divisible. But suppose also this: if you pour a little water into your glass, and then again a little bit, as a result you still have water. You cannot say you have two water(s). You can only say this if, say, you have two glasses of water, so the bits of water are separated. (Actually, with water this still sounds odd, but water in the sense of rivers allow this use.) Also, it does not make sense to divide your portion of water in any way and say that you have two pieces of water in your glasses. Likewise, suppose that John is running from 1 to 2pm and from 2pm to 3pm and did not at all stop—we would not say that he ran twice. The process of running stretches along the longest interval of time as it possibly can. There is one process only, just as there is water in your glass, without any boundary. The glass defines the boundary of water; so if you put another glass of water next to it, there are now two glasses of water. And if John is running from 1 to 2pm and then from 3 to 4pm, he ran twice; there are now two processes of running, because he did stop in between. The existence of a boundary between things or events determines whether what we have is one or two or several of them.

We must distinguish between the event type denoted by a verb and the event type denoted by the sentence as a whole. Take the verb /drinking/. This denotes a process, and is therefore atelic.

(505) *Alex was drinking in an hour.

(506) Alex was drinking for an hour.

When combined with a mass noun it remains atelic, when combined with a count noun (or a nondivisible kind) it is telic.

(507) *Alex drank water in an hour.

(508) Alex drank water for an hour.

(509) Alex drank a beer in an hour.

(510) *Alex drank a beer for an hour.

A series of individual events can make a higher level atelic event.

(511) *Alex was drinking beers in an hour.

(512) Alex was drinking beers for an hour.

Putting It All Together

Having taken a closer look at phonology, morphology, syntax and semantics, we shall revisit the big picture of the first lecture. We said that there is one operation, called 'merge' and that it operates on all of these four levels at the same time. However, we had to make concessions to the way we construe the levels themselves. For example, we argued that the English past tense marker was [d], but that it gets modified in a predictable way to [t] or [əd]. Thus we were led to posit two levels: deep phonological and surface phonological level. Likewise we have posited a deep syntactic level and a surface level (after movement has taken place), and there is also a deep morphological level and a surface morphological level. This throws us into a dilemma: we can apply the morphological rules only after we have the surface syntactical representation, because the latter reorders the lexical elements. Likewise, the deep phonological representation becomes apparent only after we have computed the surface morphological form. Thus, the parallel model gives way to a sequential model, which has been advocated for by Igor Mel'čuk (in his Meaning-to-Text theory). In this model, the levels are not parallel, they are ordered sequentially. We speak by organising first the semantic representation, then the words on the basis of that representation and the syntactic rules, then the morphological representation on the basis of the lexical, and the phonological on the basis of the morphological representation. Listening and understanding involves the converse sequence.

The paradigm of generative grammar is still different. Generative grammar assumes a generative process which is basically independent of all the levels. It runs by itself, but it interfaces with the phonology and the meaning at certain points. The transformations are not taken to be operations that are actually executed, but are ways to organize syntactic (and linguistic) knowledge. This makes the empirical assessment of this theory very difficult, because it is difficult to say what sort of evidence is evidence for or against this model.

There are also other models of syntax. These try to eliminate the distinction between deep and surface structure. For example, in GPSG the question words

are generated directly in sentence initial position, there simply is no underlying structure that puts the object first right adjacent to the verb from which it is moved to the beginning of the sentence. It is put into sentence initial position right away. Other grammars insist on special alignment rules.

There are probably as many theories as there are linguists. But even though the discussion surrounding the architecture of linguistics has been much in fashion in the last decades, some problems still remain that do not square with most theories. We mention just one very irritating fact. We have seen that Malay uses reduplication for the plural. If that is so then first of all the plural sign has no substance: there is no actual string that signals the plural (like the English /s/). What signals the plural is the reduplication. This is a function ρ that sends a string x into that string concatenated with itself in the following way:

(513) $\rho(x) := x^\frown\text{-}^\frown x$

Thus, ρ(kerani) = kerani-kerani. This function does not care whether the input string is an actual word of Malay. It could be anything. But it is *this function* which is the phonology of the plural sign. This means among other that the phonological representation of signs must be very complicated if the story of parallelism is to be upheld (recall the plural of /mouse/ with a floating segment). We have to postulate signs whose phonology is not a string but a function on strings. Unfortunately, no other theory can do better here if that is what Malay is like. Thus, the door has to be opened: there is more to the operation of merge than concatenating strings. If that is so, we can try the same for syntax; there is more to syntax than concatenating constituents. It has been claimed that Chinese has a construction that duplicates entire constituents. Even if that story is not exactly true, the news is irritating. It means that entire constituents are there just to convey a single piece of meaning (here: that the sentence is a question).

But we need not go that far. Lots of languages in Europe have agreement of one sort or another. English still has number agreement, for example, between the demonstrative and the NP (/this flag/ versus /these flags/), and with the verb. The agreement is completely formal. One says /these troops/ and not /this troops/, even though one does say /this army/. However, the number that the demonstrative carries is semantically dependent on the noun: if the latter carries plural meaning, then the whole is plural, otherwise not. The semantic contribution of 'these' is not plural irrespective of whether the noun actually specifies plural meaning. Take 'guts', whose meaning may be singular like 'courage'.

Unfortunately, it cannot really be combined with a demonstrative. (Otherwise we would expect /`those guts`/, and certainly not /`that guts`/.) A better example is Latin /`litterae`/ 'letter' (which you write to a friend), a morphological plural derived from /`littera`/ 'the letter' in the sense of 'the letter A'. The letter you write is a single object, though it is composed from many alphabetic letters. It controls plural agreement in any event. Now, if the plural morpheme appears many times in the sentence, but only once is it allowed to carry plural meaning—what are we to do with the rest of them? The puzzle has been noted occasionally, and again several solutions have been tried. Harris speaks of a 'scattered morpheme', he thought they are just one element, distributed ('scattered') over many places. The same intuition seems to drive generative grammar, but the semantics is never clearly spelled out.

Language Families and History of Languages

> It is clear even to an untrained person that certain languages, say Italian and Spanish, must somehow be related. Careful analysis has established relationships between languages beyond doubt. A language Indo-European has been proposed and argued that it is the ancestor of about half of the languages spoken in Europe and many more. The study of the history of language tries to answer (at least in part) one of the deepest questions of mankind: where do we come from?

Today, linguistics focuses on the mental state of speakers and how they come to learn language. To large parts, the investigation dismisses input that comes from an area of linguistics that was once dominant: historical linguistics. The latter is the study of the history and development of languages. The roots of historical linguistics go as far back as the late 17th century when it was observed that English, German, Dutch as well as other languages shared a lot of common features, and it was quickly observed that one could postulate something of a language that existed a long time ago, called **Germanic**, from which these languages developed. To see the evidence, let us look at a few words in these languages:

(514)

English	Dutch		German	
bring	brengen	[brɛŋən]	bringen	[bʁɪŋən]
sleep	slapen	[slapən]	schlafen	[ʃlafən]
ship	schip	[sxɪp]	Schiff	[ʃɪf]
sister	zuster	[zystɛr]	Schwester	[ʃwɛsteə]
good	goed	[xud]	gut	[gut]

This list can be made longer. It turns out that the correspondences are to a large degree systematic. It can be observed that for example word initial [p] in Dutch and English corresponds to German [pf], that initial [s] becomes [ʃ] before [t] and [p]. And so on. This has lead to two things: the postulation of a language (of which we have no record!), called **Germanic**, a set of words together with morphology and syntax for this language, and a set of rules which show how the language developed into its daughter languages. In the present case, the fact that Dutch [p] corresponds to German [pf] is explained by the fact that Germanic (not to be confused with German) had a sound *p (the star indicates reconstruction, not that the sound is illegitimate). This sound developed word initially into [p] in

Dutch and into [pf] in German. This is called a **sound law**. We may write it in the same way as we did in phonology:

(515) $^{*}p \rightarrow p \;/\; \#\_\_\_\_$ Dutch

(516) $^{*}p \rightarrow pf \;/\; \#\_\_\_\_$ German

Often one simply writes $^{*}p > pf$ for the sound change. The similarity is not accidental; in the case of phonological rules they were taken to mean a sequential process, a development from a sound to another in an environment. Here a similar interpretation is implied, only that the time span in which this is supposed to have taken place is much longer, approximately two thousand years!

The list of Germanic languages is long. Apart from the ones just listed also Danish, Swedish, Norwegian, Faroese, Icelandic, Frisian and Gothic belong there. Gothic is interesting because it is a language of which we only have written records, we do not know exactly how it sounded like.

As with the Germanic languages, similarities can be observed between French, Spanish, Italian, Rumanian and Portuguese. In fact, all these languages come from a language which we know very well: Latin. The development of Latin into these languages is well documented in comparison with others. This is important, since it allows to assert the existence of a parent language and changes with certainty, whereas in most cases the parent language has to be constructed from the daughter languages. This is so, for example, with **Celtic**, from which descended Gaelic, Irish, Welsh, and Breton. Cornish and Manx are also Celtic, but became extinct in the 19th century. Another Celtic language, Gaulish, was spoken in the whole of France, but it was completely superseded by Latin. We have records of Gaulish only in names of people and places. For example, we know of the Gaulish king Vercingetorix through the writings of Caesar. The name is a Celtic name for sure.

Throughout the 19th century it became apparent that there are similarities not only between the languages just discussed, but also between Germanic, Latin, Greek, Celtic, Sanskrit, Old Persian, Armenian, Slavic, Lithuanian, and Tocharian. It was proposed that all these languages (and their daughters, of course) descend form a single language called **Indo-European**. When people made excavations in Anatolia in the 1920s and found remains of Hittite, researchers soon realised that also Hittite belongs to this group of languages. During the last 200 years a lot of effort has been spent in reconstructing the sound structure, morphology and syntax of Indo-European, to find out about the culture and belief and the ancient homeland of the Indo-Europeans.

The time frame is roughly this: the Indo-European language is believed to have been spoken up to the 3rd millennium BC. Some equate the Indo-Europeans with people that lived in the region of the Balkan and the Ukraine in the 5th millennium BC, some believe they originate further north in Russia, other equate them with the Kurgan culture, 4400–2900 BC, in the south of Russia (near the Caspian sea). From there they are believed to have spread into the Indian subcontinent, Persia and large parts of Europe. The first to arrive in central Europe and Britain were the Celts who established a large empire only to be topped by the Romans and later by the Germans.

How the Language Looked Like

The sounds are believed to be these. Consonants are

(517)

	unaspirated		aspirated	
	voiceless	voiced	voiceless	voiced
velar	*k	*g	*kh	*gh
palatal	*k̂	*ĝ	*k̂h	*ĝh
apico-dental	*t	*d	*th	*dh
labial	*p	*b	*ph	*bh

Other people assume instead of the palatals a series of labiovelars (*k^u, *g^u, *k^uh, *g^uh). The difference is from an abstract point of view irrelevant (we do not know anyway how they were exactly pronounced...) but it makes certain sound changes more likely. Another set is *y, *w, *r, *l, *m and *n, which could be either syllabic or non-syllabic. Syllabic *y was roughly [i], nonsyllabic *y was [j]. Likewise, syllabic *w was [u], nonsyllabic *w was [w]. The nonsyllabic *r was perhaps trilled, nonsyllabic *m and *n were like [m] and [n]. Syllabic *l was written l̥, similarly m̥ and n̥. The vowels were *i (= syllabic *y), *e, *a, *o and *u (= syllabic *w).

Here are some examples of roots and their correspondences in various Indo-European languages:

*wl̥k^uos 'wolf'. In Latin we find /lupus/, in Greek /lykos/, in Sanskrit /vṛkaḥ/, Lithuanian /viĩkas/, in Germanic */wulfaz/, from which English /wolf/, and German /Wolf/ [wɔlf].

*dekm̥ 'ten'. In Sanskrit /daśa/, Latin /decem/, pronounced [dɛkɛm] or even

[dɛkɛ̃], with nasalized vowel, in Greek /deka/, Germanic *tehun/, from which Gothic /taihun/, German zehn [tse:n], and English /ten/.

Here is an example of verbal conjugation. Indo-European is believed to have had not only singular and plural but also a dual (for two). The dual was lost in Latin, but retained in Greek and Sanskrit. The root is *bher 'to carry'.

(518)

		Sanskrit	Greek	Latin
Sg	1	bhar-ā-mi	pher-ō	fer-ō
	2	bhar-ā-si	pher-eis	fer-s
	3	bhar-ā-ti	pher-ei	fer-t
Du	1	bhar-ā-vaḥ	–	–
	2	bhar-ā-thaḥ	pher-e-ton	–
	3	bhar-ā-taḥ	pher-e-ton	–
Pl	1	bhar-ā-maḥ	pher-o-mes	fer-i-mus
	2	bhar-ā-tha	pher-e-te	fer-tis
	3	bhar-ā-nti	pher-o-nti	fer-u-nt

To be exact, although Latin /fero/ has the same meaning, it is considered to belong to another inflectional paradigm, because it does not have the vowel 'i'. In Attic and Doric Greek, the 1st plural was /pheromen/. Thus, there has been a variation in the endings.

The verb *bher is also found in English in the verb /bring/ (often, root vowels become weak, giving rise in the case of *e to a so-called 'zero'-grade *bhr̥).

How Do We Know?

The reconstruction of a language when it is no longer there is a difficult task. One distinguishes two methods: comparison between languages, and the other internal reconstruction. The latter is applied in absence of comparative evidence. One observes certain irregularities in the language and proposes a solution in terms of a possible development of the language. It is observed, for example, that irregular inflection is older than regular inflection. For example, in English there are plurals in /en/ (/oxen/, /vixen/) and plurals in vowel change (/women/, /mice/). These are predicted by internal reconstruction to reflect an earlier state of the language where plural was formed by addition of /en/ and vowel change, and that the plural /s/ was a later development. This seems to be the case. Likewise, this method

predicts that the comparative in English was once formed using /er/ and /est/, but at some point got replaced by forms involving /more/ and /most/. In both cases, German reflects the earlier stage of English. Notice that the reasoning is applied to English as it presents itself to us now. The change is projected from present day English. But how can we ascertain that we are right?

First and foremost, there are written documents. We have translations of the bible into numerous languages (including medieval Georgian (a Caucasian language)), and we have an Old English bible (King Alfred's bible), and a Gothic bible, for example. Latin and Greek literature has been preserved to this day thanks to the effort of thousands of monks in the monasteries (copying was a very honorable and time consuming task in those days). Also other languages have been preserved, among which Avestan, Sanskrit and Hittite (written mostly in cuneiform). The other languages got written down from the early middle ages onwards, mostly in the form of biblical texts and legal documents. Now, this provides us with the written language, but it does not tell us how they were spoken. In the case of Hittite the problem is very obvious: the writing system was totally different from ours and it had to be discovered how it was to be read. For Sanskrit we know from the writings of the linguists of those days, among which Paṇini (500 BC) is probably one of the latest, how the language was spoken. This is because we have explicit descriptions from them of how the sounds were produced. For Latin and Greek matters are less easy. The Greeks, for example, did not realize the distinction between voiced and voiceless consonants (they knew they were different but couldn't say what the difference was). In the middle ages all learned people spoke Latin, but the Latin they spoke was very different from classical Latin, both in vocabulary and pronunciation. By that time, people did not know how things were pronounced in the classical times (= first century BC). So how come *we* know?

One answer is: for example through mistakes people make when writing Latin. Inscriptions in Pompeii and other sites give telling examples. One specific example is the fact that /m/ after vowels was either completely lost or just nasalized the preceding vowel. One infers this from the fact that there are inscriptions where one finds /ponte/ in place of what should have been /pontem/. People who made the mistakes simply couldn't hear the difference. (Which is not to say that there is none; only that it was too small to be noticeable.) Also, in verse dating from that time the endings in vowel plus /m/ counted as nonexistent for the metre. (This in turn we know for sure because we know what the metre was.) This is strong

evidence that already in classical times the final /m/ was not pronounced. The next method is through looking at borrowings into other languages. The name /Caesar/ (and the title that derived from it) was borrowed into many languages, and appears in the form of /Kaiser/ in German, taken from Gothic /kaisar/, and /césar/ [seza:r] in French. So, at the time the Goths borrowed the word the letter /c/ it was pronounced [k]. And since French descends from Latin we must conclude that the Gothic borrowing is older. Moreover, there was a diphthong. The diphthong was the first to disappear, becoming plain long [e:], and then [k] changed into [s] in French and [tʃ] in Italian and Rumanian. A third source is the alphabet itself. The Romans did not distinguish /v/ from /u/. They wrote /v/ regardless. This shows that the two were not felt to be distinct. It is unlikely that /v/ was pronounced [v] (as in English /vase/). Rather, it was originally a bilabial approximant (the nonsyllabic *w mentioned above), which became a labiodental fricative only later.

Historical explanations are usually based on a lot of knowledge. Languages consist of tens of thousands of words, but most of them are not indigenous words. Many words that we use in the scientific context, for example, come from Latin and/or Greek. Moreover, they have been borrowed from these languages at any moment in time. The linguistic terminology (/phoneme/, /lexeme/) is a telling example. These words have been artificially created from Greek source words. Learned words have to be discarded. Another problem is that words change their meaning in addition to their form. An example is German /schlimm/ 'bad', which originally meant 'inclined'. Or the word /vergammeln/ 'to rot', which is from Scandinavian /gamall/ 'old'. English /but/ derives from a spatial preposition, which is still found in Dutch /buiten/ 'outside'. From there it took more abstract meanings, until the spatial meaning was completely lost. If that is so, we have to be very cautious. If meanings would be constant, we could easily track words back in time; we just had to look at words in the related languages that had the same meaning. But if also the meaning can change—what are we to look out for? Linguists have put a lot of effort into determining in which ways the meanings of words can go and which meanings are more stable than others.

Two Examples Among Many

As an example of the beauty and danger of historical linguistics, we give the history of two words.

The first example shows that once we know of the rules, the resemblances become very striking. The English word /choose/ has relatives in Dutch and German. If we look at the verbs that these languages normally use, we might get disappointed: the Dutch word is /kiezen/, and the German is /wählen/. The relation between Dutch and the English are easier seen. First notice that the Dutch verb has a perfect participle /gekozen/ 'chosen', which has the /o/ in place of the /ie/. The change from [e] to [o], called **Ablaut**, is widely attested in the Indo-European languages. Now, [k] often becomes [tʃ] before either [e] or [i] (like Latin [ke] became [tʃe] in Italian, often with loss of /e/ in pronunciation). Now, this change occurred in English only, not in Dutch. However, we still have to see why Ablaut occured in English. The Old English word in fact was /ceōsan/. (When no star appears that means that we have written records.) The pronunciation of /c/ changed and incorporated the /e/, and the infinitive ending got lost like with other verbs. The German case seems hopeless. In fact, /wählen/ does *not* come from a related root. However, in German we do find a verb /kiesen/ in similar meaning, although it is now no longer in use. Strangely enough, the perfect passive participle (PPP) of the verb /auserkiesen/ (two prefixes added, /aus/ and /er/) is still in use: /auserkoren/ 'chosen'. The verb itself is no longer in use. Notice that the ablaut is the same as is Dutch, which incidentally also uses the circumfix /ge- -en/ as does German. Finally, in German the PPP has /r/ in place of /s/ (which would be pronounced [z]). The change from /s/ to /r/ in between vowels (called 'rhotacism') is a popular change. (There are more examples. The English /was/ is cognate to Dutch /was/, but German has /war/.) Latin has plenty of examples of this.

Now, the root from which all this derives is believed to be Germanic */keusa/ 'to try out, choose'. Once we have progressed this far, other words come into sight: Latin /gustare/ 'to taste' (from which via French English got /disgusting/), Greek /geuomai/ 'I taste', Sanskrit /juśati/ 'he likes', Old Irish /do-goa/ 'to choose'—they all look like cognates to this one. Indeed, from all these words one has reconstructed the root */geus/ 'to choose'. The Latin word presents the zero-grade */gus/. (Roots typically had /e/. Ablaut is the change of /e/ to /o/. The **zero grade** is the version of the root that has no /e/. Similarly, */dik/ is the zero grade of */deik/ 'to show'.) In West Germanic we have /kuzi/ 'vote, choice'. But beware: French /choisir/ does *not* come from Latin—there is no way to explain this with known sound laws. For example, it cannot be derived from /gustare/. Known sound laws predict /gustare/ to develop into /goûter/, and this is what we find. Instead /choisir/ was taken from—Gothic! Indeed, French has taken

loanwords from Germanic, being occupied/inhabited in large parts by Germanic tribes. The name /France/ itself derives from the name of a tribe: /Frankon/.

With respect to Dutch and German the reconstruction is actually easy, since the languages split around the 17th century. Certain dialects of North Germany still have [p] where others have [pf]. It often happens that a word is not attested in all languages. For example, English /horse/ corresponds to Dutch /paard/ and German /Pferd/, with same meaning. The sound laws allow to assert that /paard/ and /Pferd/ descend from the same word, but for English /horse/ this is highly implausible. There is a German word, /Ross/ 'horse', and a Dutch /ros/, which sound quite similar, but they are far less frequent. What we have to explain is why the words are different. Now, we are lucky to have source confirming that the word was sometimes spelt /hros/ sometimes /ros/ in Old German. In Icelandic, it is still /hross/. The loss of /h/ before /r/ is attested also in other cases, like the word /ring/. But the change happened in English, too, and the only reason why the /h/ was preserved in /horse/ is that the /r/ changed places with /o/.

Finally, where did Dutch and German get their words from? Both words come from the same source, but it is not Germanic, it is Medieval Latin /paraveredus/ 'post-horse' (200 - 900 AD), which in turn is /para/ + /veredus/. /para/ comes from Greek (!) /para/ 'aside', and /veredus/ is Celtic. It is composed from /ve/ 'under' and /reda/ 'coach'. In Kymric there is a word /gorwydd/ 'horse'. So, /paraveredus/ had a more special meaning: it was the horse that was running on the side. (The coach had two horses, one on the right and one on the left; the one on the left was saddled, and the one on the right was the 'side horse'.) English has a word /palfrey/, which denotes a kind of riding horse. The Indo-European root that has been reconstructed is *ek^{w}os. Latin /equus/, Greek /hippos/ and Sanskrit /aśvaḥ/. Had we not known that the Latin word is /equus/, we would have had to guess from French /cheval/ and Spanish /caballo/.

Thus, roots do not survive everywhere. Words get borrowed, changed, returned and so on. There are not too many roots that are attested in all languages, mostly kinship terms, personal pronouns and numbers.

Other Language Families

In addition to Indo-European, there is another language family in Europe: the **Uralic** language family. The languages that are said to belong to that family are

Finnish, Estonian, Lappish, Hungarian and a number of lesser known languages spoken in the north of Russia. The affiliation of Hungarian is nowadays not disputed, but in the 19th century it was believed to be related to Turkish. Unfortunately, the written records of these languages are at most 1000 years old, and the similarities are not always that great. Finnish, Estonian and Lappish can be seen to be related, but Hungarian is very much different. This may have to do with the fact that it was under heavy influence from Slavic, Turkish (the Turks occupied Hungary for a long time) and Germanic (not the least through the Habsburg monarchy).

The affiliation of Basque is unknown.

Other recognized language families are: **Semitic** (including Hebrew, Ethiopic, Amharic, Aramaic and Arabic), **Altaic** (Turkish, Tatar, Tungus and Mongolian), **Dravidian** (spoken in the south of India: Tamil, Telugu, Kannada and Malayalam), **Austronesian** (Malay, Indonesian, Malagassy, languages spoken on Macronesia, Micronesia and Polynesia), **Eskimo-Aleut** (Inuit (= Eskimo), indigenous languages spoken in Canada, Greenland, Western Siberia and the Aleut Islands). The list is not complete. Sometimes languages are grouped together because it is believed that they are related, but relationships are actually hard to come by. This is the case with the **Caucasian** languages (Georgian and many others). It is believed that the people who live there have not moved for several millenia. This has given rise to a dazzling number of quite distinct languages in a relatively small area in and around the southern part of the Caucasian mountains.

Probing Deeper in Time

It has been tried to probe deeper into the history of mankind. One way to do this is to classify people along genetic relations (a large project with this aim has been led by Cavalli-Sforza), another has been to establish larger groupings among the languages. Although genetic relationships need not coincide with language relationships, the two are to a large degree identical. Two hypotheses are being investigated starting from Indo-European. Joseph Greenberg proposed a macrofamily called **Eurasiatic**, which includes Indo-European, Uralic, Altaic, Eskimo-Aleut, Korean, Japanese and Ainu (spoken in the north of Japan) and Chukchi-Kamchatkan. It has been suggested on the other hand by Russian linguists (Illyč-Svityč and Dolgopolsky) that there is an even larger macrofamily

called **Nostratic**, which originally was believed to include all Eurasiatic families, Afro-Asiatic (spanning north Africa including Semitic), Dravidian and Kartvelian (= Caucasian). Greenberg did not reject the existence of Nostratic but wanted to put it even farther back in history, as a language which developed among other into Eurasiatic. Moreover, the position of modern Nostraticists has come closer to that of Greenberg's views. At a larger scale there are believed to be twelve such macrofamilies in this world, all ultimately coming from a single language ... Examples of words that are believed to have been passed to us by the single ancestor language are /kuan/ 'dog'; /mano/ 'man'; /mena/ 'think (about)'; /mi(n)/ 'what'; /ʔaq'wa/ 'water'; /tik/ 'finger, one'; and /pal/ 'two' (see Merrit Ruhlen & John Bengtson: *Global Etymologies*, in Merrit Ruhlen: *On the Origin of Languages*, Stanford University Press, 1994, 277–336). This is highly speculative, but the evidence for Eurasiatic (or even Nostratic) is not as poor as one might think. Common elements in Eurasiatic are for example: first person in m, second person in t/n. Greenberg has collected a list of some 250 words or elements that are believed to be of Eurasiatic origin. (Ruhlen and Bengtson list 27 Proto-World roots.)

Index

Languages

Bibliography

[Coulmas, 2003] Florian Coulmas. *Writing Systems. An introduction to their linguistic analysis*. Cambridge University Press, Cambridge, 2003.

[Fromkin, 2000] V. Fromkin, editor. *Linguistics: An Introduction to linguistic theory*. Blackwell, London, 2000.

[Gazdar *et al.*, 1985] Gerald Gazdar, Ewan Klein, Geoffrey Pullum, and Ivan Sag. *Generalized Phrase Structure Grammar*. Blackwell, London, 1985.

[Lass, 1984] Roger Lass. *Phonology. An Introduction to Basic Concepts*. Cambridge University Press, 1984.

[O'Grady *et al.*, 2005] William O'Grady, John Archibald, Mark Aronoff, and Janie Rees-Miller. *Contemporary Linguistics: An Introduction*. Bedford St. Martins, 5 edition, 2005.

[Pollard and Sag, 1994] Carl Pollard and Ivan Sag. *Head–Driven Phrase Structure Grammar*. The University of Chicago Press, Chicago, 1994.

[Rodgers, 2000] Henry Rodgers. *The Sounds of Language: An Introduction to Phonetics*. Pearson ESL, 2000.