Intro to Linguistics – Syntax 1

Jirka Hana – November 7, 2011

Overview of topics

- What is Syntax?
- Part of Speech
- Phrases, Constituents & Phrase Structure Rules
- Ambiguity
- Characteristics of Phrase Structure Rules
- Valency 1

What to remember and understand:

Syntax, difference between syntax and semantics, open/closed class words, all word classes (and be able to distinguish them based on morphology and syntax)

Subject, object, case, agreement.

1 What is Syntax?

Syntax – the part of linguistics that studies sentence structure:

- word order:
  I want these books.
  **want these I books.*
- agreement – subject and verb, determiner and noun, ... often must agree:
  He wants this book.
  **He want this book.*
  I want these books.
  **I want this books.*
- How many complements, which prepositions and forms (cases):
  I give Mary a book.
  **I see Mary a book.*
  I see her.
  **I see she.*
- hierarchical structure – what modifies what
  We need more (intelligent leaders). (more of intelligent leaders)
  We need (more intelligent) leaders. (leaders that are more intelligent)
- etc.

Syntax is not about meaning! Sentences can have no sense and still be grammatically correct:

Colorless green ideas sleep furiously. – nonsense, but grammatically correct

**Sleep ideas colorless furiously green.* – grammatically incorrect

Syntax: From Greek *syntaxis* from *syn* (*together*) + *taxis* (*arrangement*).
Cf. *symphony, synonym, synthesis; taxonomy, tactics*

2 Parts of Speech

- Words in a language behave differently from each other.
- But not each word is entirely different from all other words in that language.

⇒ Words can be categorized into **parts of speech** (lexical categories, word classes) based on their morphological, syntactic and semantic properties.

Note that there is a certain amount of arbitrariness in any such classification. For example, should *my* be classified as a pronoun or as a determiner, should numerals/participles/auxiliary verbs be a separate category. If *he* is a pronoun, should *do* be a pro-verb?

Open versus closed classes:

- **Open class** – new and new items are added to the class over the time – nouns, verbs, adjectives, and adverbs.
- **Closed class** – contains small number of words, new items are added very rarely – determiners, pronouns, prepositions, conjunctions

2.1 Open Classes

2.1.1 Nouns (N)

- morphology – most nouns form plural
  cat-s, house-s, kiss-es, men, sheep
- syntax – usually has a determiner (except proper names like *John*) and can be modified by an adjective:

  Determiner (Adjective) __

  a cat, many kisses, few men, several sheep
  a small cat, many exciting kisses, few clever men, several bored sheep
- semantics – name of a person, thing or place.
  But: problem with abstract nouns (*beauty, anger, aspect*) and actions (*a thump*).

2.1.2 Verbs (V)

- morphology – form third person, past tense, past participle, present participle:

walk	*walks*	*walked*	*walked*	*walking*
go	*goes*	*went*	*gone*	*going*
buy	*buys*	*bought*	*bought*	*buying*
run	*runs*	*ran*	*run*	*running*

- syntax:
  1. can combine with an auxiliary: Aux __
     will go, have seen, should run, must leave, is swimming
  2. can be modified by an adverb: Adverb __ or __ Adverb
     usually sleep, read carefully

- semantics – usually describes an action, a process or a state of being

  But: problem with some verbs (*know*, *remember*)

  But: hard to distinguish from nouns describing actions (*a thump*)

2.1.3 Adjectives (A, Adj)

- morphology:
  1. form comparative and superlative forms:
     cool – cooler – coolest, successful – more successful – most successful, good – better – best
  2. many can be changed into adverbs by the *-ly* suffix:
     sad – sadly, funny – funnily, nice – nicely, beautiful – beautifully
- syntax:
  1. Can modify a noun: Determiner __ N
     a tall man, a cool day, a wonderful trip
  2. Can be modified by an adverb: Adverb __
     very clever, extremely clever, unusually hot
- semantics – usually describes a quality or attribute

2.1.4 Adverbs (Adv)

quickly, soon, morally, today, here, very, before

- morphology – often formed from adjectives by the *-ly* suffix.

  But: *good – well, fast – fast, friendly – friendly, eastward*
- syntax – can modify verbs (hence ad-verbs), adjectives & often other adverbs.
  1. V __ or __ V
     stop abruptly, usually eat
  2. __ Adj
     amazingly cheap, very bad (very cannot modify a verb), *quite nice*
  3. __ Adv
     very quickly, quite soon

2.1.5 Summary of open classes

	Typical Morphology	Typical Syntax	Typical Semantics
Noun	plural *house – houses*	D (Adj) __ *the big house*	thing, person, place
Verb	tenses, ... *walk – walked*	combines with an Aux *would walk*	action
Adj	comparative, superlative *big – bigger – biggest*	D __ N *the big house*	quality, property
Adverb	often has *-ly* suffix *really*, but: *well*	modifies V, Adj, Adv *a really big house*	manner, degree, ...

2.2 Closed Classes

2.2.1 Determiners (D, Det)

articles (*a, the*), quantifiers (*many, any, all, several*), possesives(*my, your, his, her*)

Syntax – come before nouns: __ (Adj) N

2.2.2 Auxiliary verbs (Aux)

will, may, must, shall, would, can, have
Syntax:

1. Is followed by a verb: __ V
   It will rain. You must be quiet.
2. Is negated directly: __ *not*
   He cannot swim. She would not come.
   **He doesn't can swim. *She doesn't would come.*

Normal verbs vs. Auxiliary verbs:

	Verbs	Auxiliary verbs
negation	needs aux *do* *I don't want it.* **I want not it.*	directly **I don't will come.* *I will not come.*
question	needs aux *do* *Do you know it?* *Know you it?*	inversion **Do you will come?* *Will you come?*
agreement	yes *He knows it.* **He know it.*	no **He wills come* *He will come*

Be, have and *do* are sometimes normal verbs and sometimes auxiliary verbs. (*I don't have it.* vs. *I have not seen it.*)

2.2.3 Pronouns (Pron)

Words that stand for a noun or a whole noun phrase.
I, you, he, she, it, we, they, me, him, her, us, them

Note: It makes sense to classify possessives (traditionally called possessive pronouns) as determiners. Syntactically, pronouns and possessives behave differently – pronouns act as nouns, but possessives modify nouns:
pronoun: *I run.* – **My run.*
possessive: *John likes my house.* – **John likes I house.*

pronoun: Based on Latin *pro* (*for*) + *noun*

2.2.4 Prepositions (P)

in, on, about, with, at, to, of, under

Syntax: stand before noun phrases (see later, simply NP = Det (Adj) N)
Semantics: usually express spacial, temporary, etc. relations.
on the table, with nice colors, about mammals

2.2.5 Conjunctions (Conj)

and, or, but, ...
Syntax: connect two words or phrases on the same level

1. N __ N (*women and men*)
2. V __ V (*run or walk*)
3. Adj __ Adj (*warm but rainy*)
4. S __ S (*I will talk and he will write.*)
5. etc.

2.3 Words belonging to more than one lexical category

Some words belong to several categories:

- They have similar meaning:
  - *She* talks *very much.* vs. *She is giving three* talks.
  - *It's* cold. vs. *I got a* cold.
- They have completely different meanings:
  - *I can't* bear *the noise.* vs. *There is a* bear *in the wood.*
  - *tree* barks vs. *the dog* barks

Virtually all English nouns can be turned into verbs.

3 Phrases, Constituents & Phrase Structure Rules

3.1 Describing Noun Phrases

In English, a noun phrase a determiner followed by a noun, or determiner followed by an adjective followed by a noun, or a single noun, or ...

To save words, we can use the so called Phrase Structure Rules capture this:

(1)

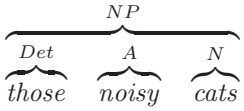
a.	NP → Det N	*the cat*
b.	NP → Det A N	*those noisy cats*
c.	NP → N	*cats*
d.	NP → A N	*noisy cats*

We can mark optional subphrases with parentheses and save even more words:

(2) NP → (Det) (A) N *cats*, *noisy cats*, *the cat*, *those noisy cats*

A phrase structure rule tell us two things:

- Which smaller phrases (Det, A, N) use to build a bigger phrase (NP).
- How to order the smaller phrases – the rule (2) allows *noisy cats*, but not *cats noisy*

(3) 

In addition, a pronoun can be a noun phrase:

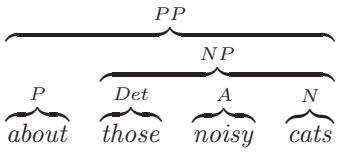
(4) NP → Pron *she, you, ...*

3.2 Describing Prepositional phrases

In English, preposition is usually followed by a noun phrase (let's ignore the prepositions at the end of the sentence).

(5) PP → P NP *about those noisy cats*

Now we can put that together and say things like:

(6) 

3.3 Describing Sentences

In English, a sentence consists of a subject (usually a noun phrase) followed by a verb which is sometimes followed by an object (another noun phrase), prepositional phrases etc.

(7) a. *Alphons slept.* – Subject + V
b. *Alphons saw his dog.* – Subject + V + Object
c. *Alphons asked for a beer.*
d. *Alphons begged beer from his dog.*

(8) a. S → NP V – *Alphons slept*
b. S → NP V NP – *Alphons saw his dog*
c. S → NP V PP – *Alphons asked for a beer*
d. S → NP V NP PP – *Alphons asked his dog for a beer*

We can abbreviate these rules as:[1]

[1] Of course, we ignored many other sentences like:

(i) a. S → NP V NP NP – *Alphons offered his dog some beer*
b. S → NP V PP PP – *Alphons argued with his dog about beer*
c. S → NP V NP InfP – *Alphons persuaded his dog to buy some beer*
d. S → NP V NP that S – *Alphons persuaded his dog that it would be wise to bring beer*

Linguists often distinguish between sentences and verb phrases (VP). A verb phrase is a sentence without a subject (e.g. *saw his dog*). Then you have to describe sentence in two steps: First, S → NP VP and then VP → V (NP) (PP).

(9) S → NP V (NP) (PP)

This rule says: Sentence is a noun phrase followed by a verb and possibly some other noun phrase and/or prepositional phrase. For example:

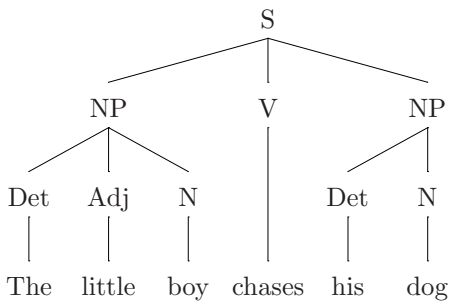
(10) $\overbrace{\overbrace{\textit{Alphons}}^{NP}\ \overbrace{\textit{saw}}^{V}\ \overbrace{\textit{his dog}}^{NP}}^{S}$

3.4 Phrase structure trees

- Phrases are created from other phrases or words.
- Sentence is the biggest phrase.

We can depict the fact that a sentence is built from smaller parts by a diagram:

(11)



A tree diagram represents several aspects of "how words are put together" in a sentence:

- the order of the words in a sentence.
- the word class (Part of Speech) of each word.
- the hierarchical structure of a sentence – the grouping of words into phrases, and the grouping of phrases into larger phrases.
- the centers of phrases that other words group around (e.g. N in NP, V in S)

3.5 Phrase/Constituent

Intuitively, a **constituent** (phrase) is a group of words which "belong together" in a sentence. They are usually coherent by themselves (i.e., when taken out of the context of the sentence) and make a coherent contribution to the meaning of the sentence as a whole.

(12) a. The dog ate the bone.
b. The president of the company likes to see big profits.
c. My stupid kid brother told my mom about my F in algebra.

Always relative to a given sentence. What is a constituent in one sentence is not necessarily a constituent in another sentence.

(13) a. I reminded *the president of the company* policy regarding smoking which he himself had instated.
b. When she was *told, my mom* broke into uncontrolled sobbing.

3.5.1 Constituency Tests

- Ability to stand alone
- Constituents can be replaced by pro-forms (pronouns, *do*)

(14) a. HE ate the bone.
b. The dog ate IT.
c. HE likes to see THEM.
d. The president likes to see SUCH profits.
e. My stupid kid brother DID.

(15) a. * The dog it/such/that bone.
b. * The that/such/did company likes to see big profits.

- Movement – (some) constituents can be moved, non-constituents cannot.

(16) a. The bone ,
the dog ate.
b. Big profits ,
the president of the company likes to see.

- Clefts – only constituents can form cleft sentences

(17) a. What the dog did was [eat the bone].
b. What the dog ate was [the bone].
c. It is [the president of the company] who likes to see big profits.

Not always clear. Some strings pass some tests but not all.

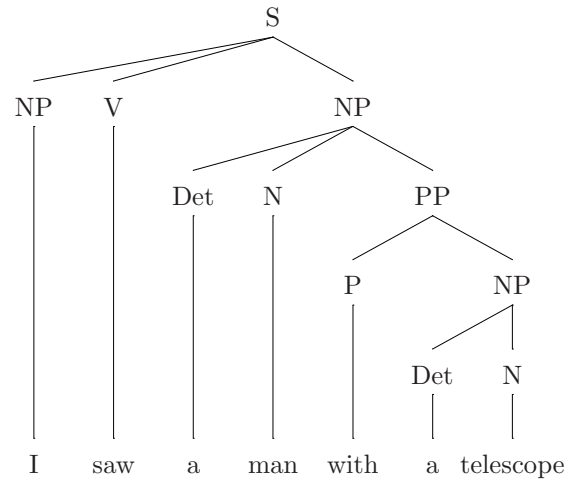
4 Ambiguity

Ambiguity:

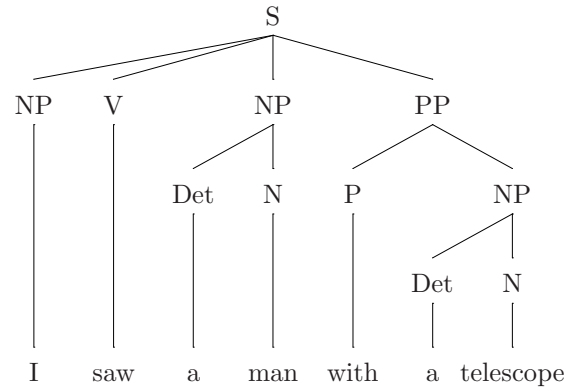
- syntactical – more than one possible structure for the same string of words.
I saw a man with a telescope. (Who has the telescope, me or the man?)
We need more intelligent leaders. (*need more* or *more intelligent*?)
- morphological – a form has more than one morphological interpretations
Napětí vyvolalo zdražení.
- lexical (homonymity) – a word has more than one meaning.
Did you see the bat?
Where is the bank?

All languages have expressions which have more than one possible interpretation.

(18)



(19)



Another example:

(20) Old men and women are exempt from the new tax.

Sometimes, world knowledge can help you to select the right interpretation:

(21) a. I saw a policeman with a gun.
b. I saw a dog with a telescope.

Syntactically, these examples are ambiguous, however your knowledge of the world helps you to choose the most probable interpretation.

5 Characteristics of Phrase Structure Rules

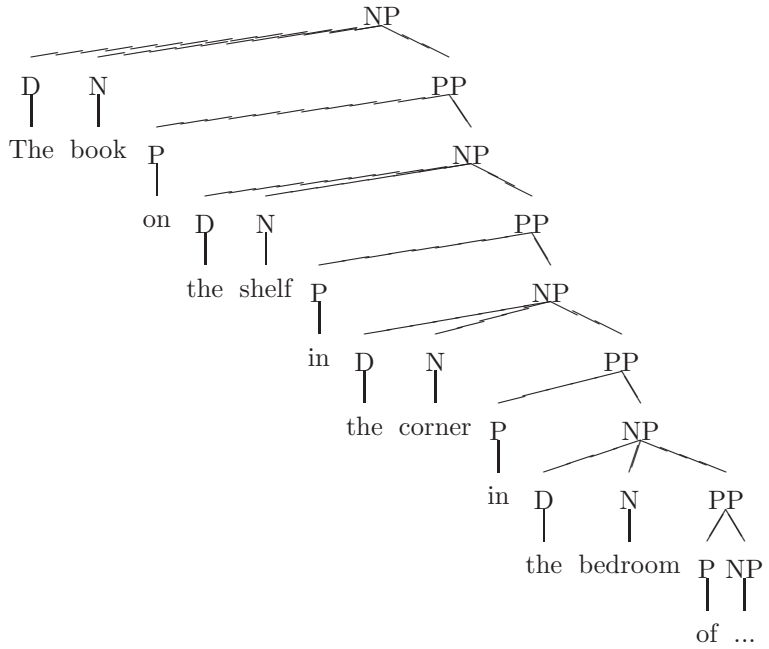
A simple grammar:

(22) a. S → NP V (NP) (PP)
b. NP → (Det) (A) N (PP)
c. PP → P NP
d. N → {dog, boy, ...}, P → {on, in, ...}, ...

This grammar describes a simple language (similar to English). It has several characteristics, which it shares with grammars of real languages:

- **Generativity**:
  It does not *list* the sentences of the language, it describes the way how to *build* them. This is important, since languages contain infinite number of sentences.
- **Ambiguity**:
  Some sentences can be build in more than one way (starting with the S rule and ending with the words in the sentence) These sentences have more than one syntactic structure – they are syntactically ambiguous (for example the *telescope* sentence). This also shows that sentences are more than just simple strings of words.
- **Infinite Recursion**:
  The grammar is recursive and thus allows to produce an infinite number of sentences using a finite (very small) number of rules.

(23)



Questions

Some things to think about:

- Is it similarly easy to write a simple grammar of Czech?
- How would one capture agreement? (In English?, In Czech?)
- How is Phrase Structure Grammar different from Dependency Grammar (the type of syntax taught at Czech schools)

5.1 Formal grammars and Chomsky hierarchy (a refresher)

Formally, a grammar is: $G = [N, T, P, S]$ where $P \subseteq (N \cup T)^+ \rightarrow (N \cup T)^*$; $S \in N$

Chomsky hierarchy of grammars:

	Languages	Complexity	Automaton	Type of rules
0	Recursively enumerable	Undecidable	Turing machine	any
1	Context-sensitive	Exponential	Linear-bounded	$\alpha A \beta \rightarrow \alpha \gamma \beta$
2	Context-free	Polynomial	pushdown	$A \rightarrow \gamma$
3	Regular	Linear	Finite state	$A \rightarrow a \vert aB \vert Ba$

6 Valency

6.1 Subject & Object/Complement

(24) a. ***John*** *snores loudly.*
b. ***John*** *wrote a long homework.*
c. ***A long homework*** *was written by John.*

Roughly: **Subject** is the active participant in the active sentence.

(25) a. ***John*** *writes a letter.*
b. However: ***John*** *underwent torture (at the hands of the terrorists).*

Note: We are defining a syntactic notion using semantics. It works in most of the cases, however not always (e.g. 2b)
Objects (complements) are the other participants.

Most of the active sentences can be transformed into passive sentences – an object becomes the subject, the subject becomes the by-PP or can be omitted.

(26) a. ***John*** *loves Mary.* (active)
b. ***Mary*** *is loved by John.* (passive)
c. ***Mary*** *is loved.* (passive)

All English sentences (except imperative) have a **subject**. Many (but not all) have an object. In some languages (e.g. Spanish), the subject can be omitted if it is understandable from the context.

The pronoun in tag questions refers to the subject:

(27) a. *The boy wrote a book, didn't he?*
b. *The girl wrote a book, didn't she?*
c. *John underwent torture at the hands of the terrorists, didn't he?*
d. **John underwent torture at the hands of the terrorists, didn't they?*

6.2 Transitive and intransitive verbs

- **Intransitive** verb – a verb with a subject and no objects: *sleep, snore*

  (28) a. ***John*** *snores.*
  b. ****John*** *snores* *a book*.

- **Transitive** verb – a verb with a subject and an object: *buy, brush, write, catch*

  (29) a. ***John*** *buys a candy.*
  b. ****John*** *buys.*

Some verbs are both intransitive and transitive: *dance* – *John dances* × *John dances samba*.

- **Ditransitive** verbs – a subclass of transitive verbs, take two objects (direct & indirect).

  (30) a. ***John*** *give* *a book* *to his friend*.
  b. ***John*** *gives* *his friend* *a nice book*.
  c. ***A nice book*** *is given* *to Mary* *by John.*
  d. ***Mary*** *is given* *a nice book* *by John.*

6.3 Form

Typical realization of subjects and objects:

- Noun phrase (NP)

  (31) a. ***The cat*** *caught* *the mouse*.
  b. ***He*** *goes to Chicago.*

- A sentence (S)

  (32) a. ***That I lost the tickets*** *annoys me.*
  b. *I know* *that this is true*.

6.4 How to know what is what?

6.4.1 Word order

English:

- declarative sentences – subject goes before the verb, object after it:
  A small dog *chases* *a big cat*. Subj V Obj
- interrogative sentences – subject after the auxiliary verb, object after the verb:
  Does ***a small dog*** *chase* *a big cat?* Aux Subj V Obj

6.4.2 Case

Case – morphological marking of a word suggesting its syntactic function. Usually, words have a special suffix at the end indicating the case.

English: very sporadic, only some pronouns distinguish it:

(33) a. ***He*** *sees Mary.*
b. *Mary sees him.*

If a language has cases, then this is the typical situation:

- **nominative** case is used for subjects (e.g. Latin *deus* God-subject)
- **accusative** case is used for (direct) objects (e.g. Latin *deum* God-object)

Some languages have more cases (Old English – 4, German – 4, Latin – 6, Finnish – 15) serving generally the same function as English prepositions.

Czech (the same word order, different cases, different meaning):

(34) a. *Mal-á kočk-a vidí Honz-u.* Subj V Obj
Small-nom cat-nom sees John-acc nom acc
The small cat sees John.
b. *Mal-ou kočk-u vidí Honz-a.* Obj V Subj
Small-acc cat-acc sees John-nom acc nom
John sees the small cat.

The suffix *a* in *kočka* says it is nominative, therefore it is the subject. The suffix *u* in *Honzu* says it is accusative, therefore it is the object.

6.4.3 Agreement

In many languages, subject and verb agree – they share certain morphological properties (number, person, gender, etc.)

English: Limited only to the 3 sup *rd* person singular and everything else distinction:

(35) a. *I write a letter.*
b. **I writes a letter.*

Spanish:

	Singular			Plural		
1	*yo*	*habl-o*	I speak	*nosotros*	*habl-amos*	we speak
2	*tú*	*habl-as*	you speak	*vosotros*	*habl-áis*	you (guys) speak
3	*él*	*habl-a*	he speak	*ellos*	*habl-an*	they speak

Russian:

	Singular			Plural		
1	*ya*	*vizh-u*	I see	*my*	*vid-im*	we see
2	*ty*	*vid-ish'*	you see	*vy*	*vid-ite*	you (guys) see
3	*on*	*vid-it*	he sees	*oni*	*vid-yat*	they see

on vid-el he saw *ona vid-ela* she saw

Some languages (Bantu in Africa, etc.) have object-verb agreement.

6.5 Adjuncts

(36) a. John eats [often].
b. John eats [loudly].
c. John eats [in the morning].
d. John eats [when he gets hungry].
e. John eats [in a restaurant] [on Sunday].

- Can combine with any verb (mostly)
- Can be repeated
- Have meaning on their own
- Usually expressed by adverbs, PPs

7 Syntax vs. Semantics

Compare:

(37) a. *I ate a cake.*
b. *I entered a room.*

The sentences have the same syntactic structure – both *a cake* and *a room* are (direct) objects, but the cake disappeared after I ate it, while the room did not after I entered it.
Compare:

(38) a. *I took my friend to Tesco.*
b. *I took D5 to Tesco.*

Both *my friend* and *D5* are (direct) objects, but ...