

Syntax Seminar

Revision. Grammar and rules. Categories and phrases

Instructor: Irina Burukina

irina.burukina@btk.elte.hu

The terms you are likely already familiar with:

sentence, noun, verb, subject, object, prepositional object, predicate.

The notions you are expected to be **already** familiar with from the Foundations of Syntax/Introduction to Linguistics courses: recursion, constituent, phrase, head, word category, functional categories, lexical categories.

At home, for the next class:

- Read BESE, 1.1-1.3.1 to remind yourself on the notions mentioned above.
- Read this handout, do the exercises. Make sure that everything is clear. If you have any questions – send me an email or ask me next time in class (do not hesitate to do that!)

1 Language and grammar

Language is a system of communication. We build sentences that express propositions and that we can utter.

? Why do we need syntax? Can't we jump straight from form to meaning or from meaning to form?

! Syntax relates form and meaning.

Cf. the **T/Y-model**: **PF** (phonological form) – Syntax – **LF** (logical form)

Note: PF, LF = interfaces.

Syntax – a set of rules, principles, and processes that govern the structure of sentences in a given language.

A relatively finite number of syntactic rules and lexical items → An infinite number of grammatical sentences.

→ Language is not about enumeration, it is about rules and constituents.

An example of a universal syntactic rule: **Recursion** – the ability to use output of a rule as input of the same rule. A linguistic rule can be applied to the result of the application of the same rule.¹

- (1) Yesterday Mary told me, that John had said, that Bill had promised ...
[SENTENCE 1 [SENTENCE 2 [SENTENCE 3 [...]]]]

? Why do we study syntax?

¹Daniel Everett: Pirahã language – no recursion. See his book *Don't sleep, there are snakes: Life and language in the Amazonian jungle*. However, his claim was opposed by many linguists, including David Pesetsky, Andrew Nevins, and Cilene Rodrigues. See their paper *Pirahã exceptionality: A reassessment* <pdf>.

Language (with a capital L) = Human Language Capacity = the ability to speak. Language is a psychological or cognitive property of humans. Syntax is one of the most important components of Language.

Sentences are generated by a set of procedures, which are part of our cognitive abilities

→ **Generative grammar** – we model the rules that allow us to build all grammatical sentences and that reject all ungrammatical sentences.

! We come up with a group of rules (i.e. grammar) based on the actual language data (empirical observations). Our rules are *hypotheses* – they must make predictions and be falsifiable. Our *theory* must be descriptively and explanatory adequate.

* Exercise: Consider the distribution of reflexive pronouns and their *antecedents* (the word that determines the reference of the pronoun), illustrated in (2) and (3). What kind of rule(s) can capture it?

Note: we use indices to show that a pronoun and a noun refer to the same person/object.

(2) Good/acceptable/well-formed examples

- a. Frodo_i saw himself_i in the mirror.
- b. Galadriel showed Frodo_i to himself_i.
- c. The friends_i of Frodo saw themselves_i in the mirror.

(3) Bad/unacceptable/ill-formed examples

- a. Himself_i saw Frodo_i in the mirror.
 - b. Galadriel showed himself_i to Frodo_i.
 - c. Merry_i asked Éowyn to hug himself_i.
 - d. The friends of Frodo_i saw himself_i in the mirror.
-

2 Universal rules

? Do you know what Universal Grammar is? Can you guess from its name?

! Language is an instinct → innate **Language faculty** → **Universal Grammar**.

Support for the idea of Universal Grammar comes from:

1. The species specificity.
2. The localization of (some) language knowledge in the brain: Broca's area (production) and Wernicke's area (comprehension).
3. Plato's Problem: Any human child can acquire any language with native proficiency if placed in the appropriate speech community.
4. (related) The logical problem of language acquisition: As a rule-governed infinite system syntax should be unlearnable.
5. (related) The poverty of stimulus argument: children have knowledge about their language which is not in the input data that they hear:
 - Children know that (4-a) and (4-c) are well-formed and (4-b) and (4-d) are ill-formed even if they are not taught that explicitly.
 - Children know that (5) are ill-formed even if they occasionally hear such examples.

6. (related) Universal patterns of language acquisition.
- (4) a. Good: Who do you think ___ will question Seamus first?
 b. Bad: Who do you think that ___ will question Seamus first?
 c. Good: Is the woman who is singing ___ happy?
 d. Bad: Is the woman who ___ singing is happy?
- (5) a. What do you? (instead of What do you think?)
 b. I don't want. (instead of I don't want that.)

Examples of universal properties of Language → *principles*:

- (6) a. All languages make a distinction between nouns and verbs.
 b. In all languages sentences must have a subject (overt or covert).
 c. All languages have pronouns.
 d. More abstract:
 (i) All languages build sentences by putting pairs of items together.
 (ii) More abstract: All languages have recursion.
 (iii) More abstract: Grammatical operations are local.

Of course, there is more than one human language and there are also language-specific rules → *parameters*:

- (7) Word order:
 a. Subject-Verb/Verb-Subject: SV – English, VS – Welsh (Celtic), Kaqchikel (Mayan)
 b. Verb-Object/Object-Verb: VO – English, OV – Mari (Uralic)
 c. Prepositions or postpositions: prepositions – Indo-European, postpositions – Uralic

3 Prescription vs description

Prescriptive rules – how a language should be used (for instance, how to speak ‘proper’ English) → ‘correct’ sentences

Descriptive rules – how a language is actually used → grammatical sentences

? What rules are we interested in as linguists?

! We focus on description and (mostly) examine spoken data (not written texts).

An example of the difference between prescription and actual language data: split infinitives.

- (8) [Carnie:36] “The linguist Geoff Pullum reports that he heard Alex Chadwick say the sentence below on the National Public Radio Show “Day to Day”:

But still, the policy of the Army at that time was not to send – was specifically to not send – women into combat roles.

Here, Mr. Chadwick corrects himself from not splitting an infinitive (was not to send) to a form where the word not appears between to and send, thus creating a classic violation of this prescriptive rule. One might wonder why he would correct the sentence in the wrong direction. Pullum observes that the two versions mean quite different things. The policy was not to send women into combat means that it was not the policy to send women into combat (i.e. negating the existence of such a

policy). The sentence with the split infinitive by contrast, means that there was a policy and it was that they didn't send women into combat. It's a subtle but important distinction in the discussion."

4 Data

? How do we get language data? → *introspection, elicitation, corpora*

Corpora (sg: corpus) – collections of naturally produced texts (written or spoken, monolingual or multilingual (parallel corpora), of different genres, typically annotated).

? What kind of data can we get by studying corpora? What kind of data are not available?

! A corpus can give us *positive* data (what is possible) but not *negative* data (what is prohibited).

Elicitation – collecting language data directly from a native speaker (usually via grammaticality judgment tasks). By eliciting grammaticality judgments from native speakers we get **negative** information about the language.

I-languages – internal languages, speakers' tacit knowledge of their language (the ability to judge sentences as (un)acceptable; e.g. English of an individual speaker.

Cf. **E-languages** – external languages, public languages used by populations; e.g. English, Japanese.

! We focus on I-languages and individual language competence.

Competence – what we know about our language.

Performance – what is actually produced and heard; influenced by external factors.

- (9)
- The horse raced past the barn fell. – a *garden path sentence*
 - After Bill drank the water proved to be poisoned.
 - The man who hunts ducks out on weekends.
 - Cheese mice cats catch love stinks. – well-formed but impossible to comprehend and not produced
 - The man the boy the woman saw heard left.

Note: acceptability and grammaticality are not exactly the same. Grammatical sentences are not always perfectly acceptable. Ungrammatical sentences can sometimes be accepted by native speakers.

- (10) Types of grammaticality judgments:
- John does not know anything. – *well-formed*
 - %John does not know nothing. – *variation*
 - *John not know anything. – *syntactically ill-formed*
 - ?Peter John does not know. – *marginal*
 - #The table does not know anything. – *semantically ill-formed*

5 Constituents

Last time we said that syntax is about rules that allow us to put lexical items (and some functional elements) together into sentences.

? Should/can we limit ourselves to 'word'-related rules?

- (11) Cats eat fish. ← cats + eat + fish

Consider the following more complex examples:

- (12) a. My cats eat baked fish.
b. My two cats eat baked sea fish.
c. My two fluffy cats eat baked sea fish from this market.
d. My two fluffy cats, who I adore, eat every day baked sea fish from this market.

In (12) "my" modifies "cats" but has nothing to do with "eat" or "fish". Similarly, "baked" adds some information about "fish" but not about "cats". These words together form intermediate units (smaller than a complete sentence) – **constituents**. In (13) constituents are marked with [].

- (13) a. [My cats] eat [baked fish].
b. [My two cats] eat [baked sea fish].
c. [My two fluffy cats] eat [baked sea fish [from [this market]]].
d. [My two fluffy cats, [who I adore]], eat [every day] [baked sea fish [from [this market]]].

Constituent – a group of words that function together as a unit.

! Syntactic rules "see" the structure, that is, the way items in the sentences are organized into constituents. Studies confirmed that babies are also aware of the notion of constituency when they acquire a language.

- (14) Example 1: subject-verb order
Anna likes hamsters. – [The girl] likes hamsters. – *The likes girl hamsters.
- (15) Example 2: formation of yes-no questions
a. The girl can dance. – Can the girl ___ dance?
b. The girl who can sing can dance.
c. *Can the girl who ___ sing can dance? – Can [the girl who can sing] ___ dance?
- (16) Example 3: agreement
a. Anna likes hamsters.
b. [Anna and John] like/*likes hamsters.
c. [All friends of Anna] likes/*likes hamsters.

* Exercise: The following sentences are all structurally ambiguous. For each sentence, define the two interpretations. Determine which constituent is 'responsible' for the ambiguity. Group constituents together using [] to show the difference between the structures corresponding to the different interpretations.

- (17) a. John threatened the boy with a book.
b. The dog of the woman that had an umbrella was brown.
c. Susan met very tall boys and girls.
d. Susan met boys and girls who were holding hands.
e. The boy and the girl's uncle stayed.
f. The professor said on Monday he would give an exam.
g. For this salad we need three tomatoes and onions.

! Constituents **can include** one another but they **cannot overlap**: if a word belongs to two constituents than one of these constituents must include the other.

- (18) a. {The professor said [on Monday] he would give an exam].
– cannot mean "said on Monday" AND "give on Monday" at the same time
b. {The professor said on Monday [(that) he would give an exam on the same day/on Monday]}.

6 Constituency tests

6.1 Substitution (= replacement)

As a single unit, a constituent can be substituted by a single pro-form or a wh-expression.

- Substitution by a **pro-form**

Note: pro-forms = pronouns, elements such as *so*, *such*, etc.

- (19) a. [The little boy] fed [the cat]. → [**He**] fed [**her**].
b. [The little boy] with a ball fed [the cat] without a tail.
→ ***[He]** with a ball fed [**her**] without a tail.

- Substitution by a **wh-expression**

Note: *wh*-expressions – interrogative pronouns, relative pronouns: e.g. **Which** book did you like? *The book, which he really liked, was there.*

- (20) a. Susan saw John [in a garden]. → **Where** did she see him? ***Where** did she see him in a?
b. Mary bought [a red dress]. → **What** did she buy? ***What** did Mary buy a?

- *Do-so/did-too* replacement

- (21) Anna [hugged John] in a garden.
→ Bill *did so* in a garden. Bill *did it* in a garden too.

6.2 Movement (= dislocation, separation)

As a single unit, a constituent can be separated from the rest of the clause.

- Fragment answers

- (22) a. Susan kissed John [in a garden].
Where did she kiss him? → [In a garden].
b. Mary bought [a red dress].
What did she buy? → [A red dress].

- Topic- and focus-fronting²

² *Topic* – a piece of old information that the sentence is about.

- (23) a. Ali Baba returned home [wiser than before].
 [Wiser than before], Ali Baba returned home ____.
- b. Mary bought [a red dress].
 [A red dress] Mary bought ____, and not a white bag.

- Clefting

It-cleft = It is [constituent] that/wh-expression

- (24) [Ordinary cats] like the smell of meat.
 It is [ordinary cats] that like the smell of meat.

Specificational pseudoclefts = Wh-expression ... is [constituent].

- (25) The cats will [play in the garden].
 What the cats will do is [play in the garden].

6.3 Coordination

Only constituents of the same type (category; more on these in the next handout) can be coordinated.

- (26) a. [The fluffy cat] and [the very big dog] are best friends.
 b. *[The fluffy cat] and the very big are best friends.

***** Exercise: Apply the constituency tests to determine whether a string in brackets is a single constituent.

- (27) a. The postman lost [his key] yesterday.
 b. Peter met Mary [in the park yesterday].
 c. Ordinary cats [detest the smell] of citrus fruits.

Focus – a piece of new information that we want to emphasize, highlight.

- (i) a. Context: And what about Mary? What did she buy yesterday?
 Mary_{TOP} bought A BOOK_{Foc} yesterday.
- b. Context: Did Mary buy a book?
 No, Mary bought A PENCIL_{Foc}. – *contrastive focus*
- c. Context: What did each of the kids (Mary, Bill and Bob) buy?
 Mary_{TOP} bought a book. John_{TOP} bought a pencil. Bob_{TOP} bought a pen. – *contrastive topics*

*** At home** (the answers to the exercises marked as "optional" will be at the end of the next handout)

1. *Optional:* In (28) you will find several Sanskrit verbs (a) and their translations into English, written in a different word order (b). Match the Sanskrit verbs with the correct translations. Write down the syntactic rules of Sanskrit that these data allowed you to come up with. Are there any similarities between Sanskrit and English? [The answers are at the end of HO2]

- (28) a. Sanskrit: nayasi, icchati, anayam, nayāmi, icchasi, icchāmi, anayat
b. English: I want, you lead, he wants, I lead, I led, you want, he led.

2. *Optional:* (29) is structurally ambiguous. Paraphrase the sentence in two different ways to show the two possible interpretations. Apply movement tests and provide some sentences that show that there is an ambiguity in structure.

- (29) Susan accidentally pushed the button twice.