Lexicon and grammar. Lexicon (categories, predicates and arguments, theta-roles, subcategorization of complements)

> **Today's class in a nutshell:**
> - Lexicon and grammar
> - Lexicon:
>   - Word categories are defined based on morphological and distributional criteria.
>   - Word categories of English:
>   - **thematic**: verbs (V), nouns (N), adjectives (A), prepositions (P)
>   - **functional** categories: inflections (I), determiners (D), degree adverbs (Deg), complementisers (C)
>   - distinguished by binary features:
>   - [±F] = functional/non-functional
>   - [±N] = nominal/non-nominal
>   - [±V] = verbal/non-verbal
>   - Predicates have arguments, and predicates distribute theta-roles (semantic roles) to their arguments → theta grid: part of a predicate's lexical entry
>   - Predicates determine the syntactic category of their complements → subcategorization frame: part of a predicate's lexical entry

- **Aim** of the course: to provide a description of some aspects of the syntax of English using the **generativist** theory
  - mainly the **Government and Binding** theory
  - The generativist theory is just one of the existing theories for the description of a language. It is a **formal** approach to language (other, e.g., cognitive-functional approaches also exist)
- Language is seen in this approach as a **system** that enables people to produce and understand linguistic expressions:
  - How is it possible that we understand utterances that probably **have never been produced before**, e.g. *The bishop was wearing a flowing red dress with matching high heeled shoes and singing the Columbian national anthem*?

  → it is possible to have knowledge of an infinite set of things without actually storing them in our heads: see, e.g., the set of numbers is infinite → this is possible because we know a **small number of simple rules**

    - **lexicon** = what is **stored** in our minds
    - **grammar = a finite set of rules**, which allows us to produce and understand an **infinite** set of linguistic expressions

- **I-language:** internal to the mind; a finite system; this is what generativists try to model with grammars
- **E-language:** an infinite set of expressions defined by the I-language – the actual production, the utterances

2 Word Categories
2.1 The Lexicon
- a kind of mental dictionary: information about words which is stored in our head: **arbitrary** facts about particular languages:
    - how a given word is pronounced (phonetic information):
    e.g. *macska* (Hungarian), *chat* (French), *Katze* (German), *gatto* (Italian), *kot* (Russian), *kissa* (Finnish), *neko* (Japanese), *mao* (Chinese), *paka* (Swahili))
    - what it means (semantic information)
    - its word category
    - if it's a predicate: its theta-grid (cf. 3.2) and subcategorization frame (3.4)

2.2 Categories/Word categories ('szófajok')
- traditional approach: defining categories based on meaning: e.g., verbs are that part of speech (category) which express an action or an event, e.g., *to eat* → however, certain nouns also express an action or event (*eating*)

    → We'll define categories based on morphological and distributional criteria:
    - what morphemes a word combines with
    - what positions it can occupy

2.2 Morphological criteria for determining category
- only words of certain categories can host morphemes of certain types: plural forms, in English, are restricted to nouns and other categories do not have them:

    idea → *ideas* ✓
    dog → *dogs* ✓

    > An asterisk (*) marks an ungrammatical form or sentence

    nice → **nices:* ungrammatical → not a noun (adjectives do not have a plural form in English)

    > Could you name a few languages in which adjectives do have a plural form?

    warm → *warms:* not a plural but a present tense form → not a noun but a verb

    → if it has a plural form, it is a noun and if it has a present tense form, it is a verb

2.4 Distribution
- the **distribution** of a category = the set of positions that the grammar determines to be possible for a given category
    o there are certain positions in a sentence that some words can occupy and other words cannot – this is determined by the words' category:

(2)   a. the **cat** slept
      b. he fed Pete's **cat**
      c. I tripped over a **cat**

→ these positions can be occupied by nouns and not by, e.g., complementisers – complementisers have a different distribution

a' *the **if** slept
b' *he fed Pete's **if**
c' *I tripped over an **if**

Also, nouns cannot occupy certain positions:

(3)   a. *the dog **cat** the mouse
       b. ***cat** dog howled
       c. *the dog slept **cat** a kennel

> Exercise: Replace "cat" in (3a-c) with other words so that the sentences be grammatical. What is the category of these new words?

- distributional positions are not defined in terms of linear order: if it was so, then we could define, e.g., the 1st position as a position for nouns, the 2nd position for verbs and the 3rd position for nouns, as in (17):

(17)   dogs chase cats

→ but this would predict the sentence in (19) to be ungrammatical:

(19)   obviously dogs chase cats

→ distributional positions are defined but in terms of **structural positions**

3 A Typology of Word Categories
- **thematic** (or lexical) categories: verbs (V), nouns (N), adjectives (A), prepositions (P)
- **functional** categories: inflections (e.g., modal auxiliaries) (I), determiners (*a, the*) (D), degree adverbs (*so, too*) (Deg), complementisers (*if, that*) (C)

- distinguished by binary **features**: each category can be defined in terms of a unique collection of these features, but they may share some of the features with other categories, accounting for similarities between them

    1) [±F] = functional/non-functional
    2) [±N] = nominal/non-nominal
    3) [±V] = verbal/non-verbal

(32)
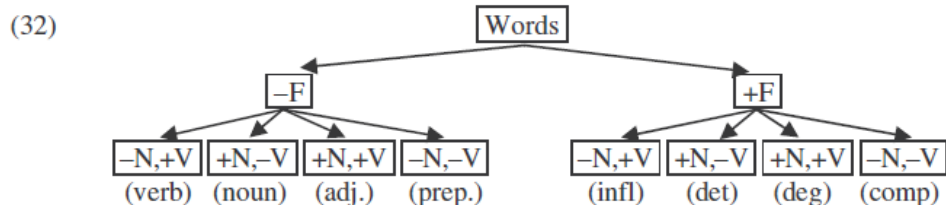
- nouns and verbs are diametrically opposed:
  nouns = [–F, +N, –V]
  verbs = [–F, –N, +V]

- How can we capture the sense that determiners (*the, a*) have something in common with nouns, and that modal auxiliary verbs (*can, must* etc.) have something in common with verbs?
  determiners = [+F, +N, –V] → determiners are the functional equivalents to nouns
  modals = [+F, –N, +V] → modals are functional verbs

3.2 Predicates and arguments
- all thematic categories can function as predicates

(33) Peter chased Mary → *chased*: **predicate**
                    → *Peter, Mary*: **arguments** – they are required by the predicate in order for the sentence to be grammatical – their omission results in ungrammaticality:

(33') *Chased Mary.
(33") *Peter chased.

(34)   a. Selena slept → *sleep* is a **one-place predicate** involving one argument
        b. Tom is tall → *see* is a **two-place predicate** involving two arguments
        c. Percy placed the penguin on the podium → *place* is a **three-place predicate**

Exercise 2

Identify the arguments in the following sentences.

(1)   a Peter left his family.
        b Peter left after dinner.
        c Peter and Mary met in the park.
        d Mary suddenly noticed that her purse was missing.
        e Before leaving the house she checked her bag.
        f The purse was on the kitchen table.
        g Peter considers Mary beautiful.
        h John knew that Peter and Mary met in the park in the afternoon.
        i John knows Mary.
        j Peter wanted John out of the room.
        k They treated their guests kindly during their stay.
        l Peter wrote a letter to Mary the other day.
        m He sent her a box of chocolate, too.
        n Peter called Mary yesterday.
        o John called Peter a liar.

- the meaning of a predicate determines the semantics of the arguments: → arguments have **thematic roles (= Θ-roles/theta-roles/semantic roles)**

(35)   a. Harold hit Henry
          Harold: the one who deliberately performs an action = **AGENT**
          Henry: the one who is acted upon = **PATIENT**
        b. Sam saw Simon
          Sam: the one who does the seeing = **EXPERIENCER**
          Simon: the one who gets seen = **THEME**

→ **theta-roles organise into theta-grids** – part of a predicate's lexical entry**:**

(36) *sleep* **Θ-grid:** <agent>
       hit **Θ-grid:** <agent, patient>
       see **Θ-grid:** <experiencer, theme>
       place **Θ-grid:** <agent, patient, location>

- adjectives (38-39), nouns and prepositions (46) can also be predicates, in which case they have their arguments and theta-grids:

(38) a. Fred is **fond** of Fiona
      b. Kevin is **keen** on karate

(39) *fond* **Θ-grid:** <experiencer, theme>
      keen **Θ-grid:** <experiencer, theme>

(45) the house is **on** the hill
(46) *on* **Θ-grid:** <theme, location>

→ it is **[-F] categories that can have theta-grids**

Exercise 3

Here is a list of definitions of theta roles. Given the definitions, label the arguments in the sentences below.

Agent: the participant who deliberately initiates the action denoted by the verb (usually animate).
Theme: the participant (animate or inanimate) moved by the action.
Patient: an affected participant (animate or inanimate) undergoing the action (the roles 'theme' and 'patient' are often collapsed).
Experiencer: the participant (animate or inanimate) that experiences some (psychological, emotional, etc.) state.
Beneficiary/Benefactive: the participant that gains by the action denoted by the verb.
Goal: the participant towards which the activity is directed.
Source: the place from which something is moved as a result of the action.
Location: the place in which the action or state denoted by the verb is situated.
Propositional: clausal arguments have the propositional theta role.

(1) a Peter loves Mary.
     b Peter knows Mary well.
     c The door opened.
     d The purse was stolen.
     e Mary wrote a letter to John the following day.
     f John received a letter from Mary.
     g Mary cut the cake with a knife.
     i There arrived some visitors.
     j Mary was cooking dinner when they entered.
     k Peter has broken his leg.

l Peter has broken a vase.
m It surprised everyone that the visitors arrived.
n They wondered what to do.
o Mary is beautiful.
p John is in Paris.
q That the purse was stolen shocked everyone.

3.4 The Thematic categories

3.4.1 Verbs

- complements: the arguments which **follow** the verb (!! the subject is an argument, but not a complement)
  the verb determines i) the number and ii) the category of its complements

(67)   the mayor gave [the hero] [a reward] – **2 nominal complements**
(68)   a. the villain awaited **his trial – 1 nominal complement**
       b. the villain waited **for his trial – 1 prepositional complement**

- the **category of the complement** (nominal/prepositional/adjective/adverb/sentence) is stated as a separate piece of information in a verb's lexical entry → **subcategorization frame**:

(70)   *await*   **category:** [–F, –N, +V]
                 Θ-grid: <agent, goal>
                 subcat: [nominal]
       wait   **category:** [–F, –N, +V]
                 Θ-grid: <agent, goal>
                 subcat: [prepositional]

- **intransitive verbs:** verbs without a nominal complement:
  - **"true intransitives":** no complement at all – *to fly, to sleep. to laugh:* (*The dragon flew; Susan slept; Vicky laughed*):

    (72)   *laugh*   **category:** [–F, –N, +V]
                     Θ-grid: <agent>
                     subcat: [∅]

  - *to wait* **for**. **prepositional verb**

- **transitive Vs**: verbs with a nominal complement (*to await sg*)
  - verbs with one nominal complement (73a)
  - verbs with two nominal complements (73b): **ditransitives**

(73)   a. the hero fought [the dragon]
       b. the king gave [the hero] [half the kingdom]

  - **complex transitive verbs:** take both a nominal and a prepositional complement:

(75)   a. Percy placed [the penguin] [on the podium]

  -

- verbs with **adjectival or adverbial complements**:

(76) a the judge looked mean
     b *look* **category:** [–F, –N, +V]
              Θ-grid: <theme, attribute>
              subcat: [adjectival]
(77) a the pianist performed passionately
     b *perform* **category:** [–F, –N, +V]
              Θ-grid: <agent, manner>
              subcat: [adverbial]

- verbs with **sentences as their complements**:

(78) b. Theodore thinks [Larry left]

(79)   *think* **category:** [–F, –N, +V]
              Θ-grid: <experiencer, proposition>
              subcat: [sentence]

Homework:

To read Chapter 1 in BESE: p. 1–51., paying special attention to: 1.3.4–1.3.6 (The Thematic Categories; Functional categories; Functionally underspecified categories)

+ **Exercises** from BESE. You don't have to send them to me, but we will check them in class and you might be asked to write/draw on the blackboard, so please come prepared:
- p. 53, Ex.6
- p. 54, Ex. 9, from a) to e)
- p. 55, Ex. 12

For ex. 12, remember that the lexical entry of a predicate consists of
- the category of the predicate
- the theta-grid
- the subcategorization of the complements (the arguments but the subject)