ENG 701: Principles of Linguistic Analysis — Fall Semester 2004: TUE 18.00-22.00
MA in Applied Linguistics, Department of English Studies — University of Cyprus
Kleanthes Grohmann (Room M 004, Phone x2106, Email kleanthi@ucy.ac.cy)

October 26, 2004

CLASS 8: CATEGORIES AND FUNCTIONS

This discussion is largely based on Adger (2002: chaps. 2-4); see PLA-homepage for more.

WHAT IS A FEATURE?

Features are our shorthand notation for **morphosyntactic features**, the abstract properties of individual classes to which particular **morphological forms** belong (a given word shape).

They are **morphological** in being related to the morphological form, and **syntactic** in playing a role for **the computational system** that assembles these forms from the **lexicon** and creates the syntactic **structure** which is then interpreted:

- for meaning by the **conceptual-intentional system** in the **semantic component** (also known as **Logical Form / LF**), and

- for articulation meaning by the **articulatory-perceptual / sensorimotor system** in the **phonetic component** (viz. **Phonetic Form / PF**).

A classic example of a feature is **number** expressing singular, plural, or dual numbers across languages. There are **at least five ways** to integrate this expression:

1. a **unique feature** approach (for each property): [SINGULAR], [PLURAL], [DUAL]
2. a **feature bundle** approach: [SG], [PL], [SG, PL]
3. a **privative feature** approach ("default"): [ ], [PL] (depending on one's view, [DL])
4. a **binary feature** approach: [+SG, –PL], [–SG, +PL], [+SG, +PL] — *[-SG, –PL]
5. a **feature class** approach: [NUM: SG], [NUM: PL] — [NUM: SG, PL] / [NUM: DL]

Adger (and us): use the **privative feature approach** for the time being and develop it.

WHAT ARE THE FEATURES?

This is a good question with a bad answer: anything that seems necessary to **relate form and meaning** (incl. form without meaning, meaning without form) — and sometimes beyond (no form and no meaning, but a **necessity for the syntactic computation**).

- categorial features: [N], [V], [A], …
- semantic features: [mass], [count], …
- phonological features: [voice], [nasal], …

But these arguably don't play as big a **role in the syntax** as the following:

- phi-features: person ([1], [2], [3] / [ ]), number ([SG], [PL]), gender ([MASC], [FEM], [NEUT])
- case-features (nominal): nominative [NOM], accusative [ACC], genitive [GEN], …
- tense-features (verbal): present [PRES], past [PAST], …
- (other) V-features: [INF], [PART], …

(Morpho)syntactic features are **interpretable** or **uninterpretable** at the semantic interface.

CONSTITUENCY

(1) The structure of this nice sentence should have concerned everyone.

A string of words is any number of words that are **(string) adjacent** to one another. However, not every string of words carries the same meaning: *the structure of this nice sentence* vs. *sentence should*. Those strings of words that do carry some meaning or function, or form a unit of sorts, are called **constituents**. For constituency, there are a number of **tests**.

One constituent can be replaced by another of the same type. Before going into different types, let's play with a sequence of words that can be replaced by a single word; such tests are interchangeably known as **Replacement Test**, **Substitution Test**, or **Pronominalization**:

(2) a. **It** should have concerned everyone.
b. * The structure of this nice **it** have concerned everyone.

Constituents can be represented by **bracketing**. And since a sentence is also a constituent, it carries brackets as well. In fact, constituents may be made up of further constituents, which we can also test for (more tests to consider are ***one*-substitution** or ***do*-substitution**):

(3) a. [The structure of this nice sentence should have concerned everyone]
b. [[The structure of this nice sentence] should have concerned everyone]
c. [[The structure of this [nice sentence]] should have concerned everyone]
d. [[The structure of this [nice sentence]] should have [concerned everyone]]

(4) The structure of this **one** should have concerned everyone.

(5) The structure of this nice sentence should have **done**.

(6) **It** should have **done**.

Related to such tests, there is the phenomenon of **ellipsis** (deletion of a constituent):

(7) a. The structure of this nice sentence should have concerned everyone.
b. The structure of this nice sentence should have ~~concerned everyone~~.
c. The structure of this nice sentence should ~~have concerned everyone~~.

More tests include **movement** (e.g. to the front) and **clefting** (of the form "*It is X that…*"):

(8) a. **Everyone**, the structure of this nice sentence should have concerned.
b. *It is* **the structure of this nice sentence** *that* should have concerned everyone.

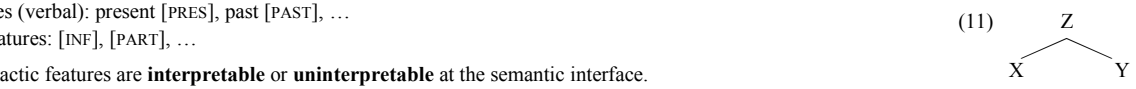
(9) a. John kissed Mary on the cheek with pleasure.
b. **With pleasure**, John kissed Mary on the cheek.
c. **On the cheek**, John kissed Mary with pleasure.

(10) a. *It is* **with pleasure** *that* John kissed Mary on the cheek.
b. *It is* **on the cheek** *that* John kissed Mary with pleasure.
c. * *It is* **kissed Mary** *that* John on the cheek with pleasure.

MERGE

The operation that puts words together to build up a tree is **Merge**. Merge takes two items *X* and *Y* (words, phrases, constituents, etc.), **joins** them, and gives the new **object** a **label** *Z*.

(11)



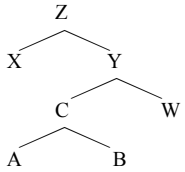
Note that X and Y can each be more **complex**, of course (i.e. have **internal structure**). The label of the new object is derived from the features of the items joined together, or Merged. The labels of **lexical items** are their **categorial labels** (D, A, N, P, V, and so on).

And since Merge joins together two items by definition, it is **binary**:

(12) a. * Z
X

b. * Z
X Y W

Some **structural relations**: *mother, daughter, sister*
contain, immediately contain
branch, node, root node, terminal node

(13)


Merge does not specify **linear order**. We'll address the issue of **linearization** later (Class 9).

Simplest assumption: **Merge only applies to root nodes**.

HEADS

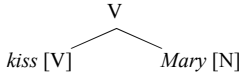
Which syntactic object (string of words, sequence of words, constituent, or other terms) determines the labels of the new object created by Merge then? This will be the **head**. And the name of the object that the head lends its label to is **phrase** — the phrase is a **projection** of the head, or the other way around: **the head projects its label**.

We test for heads through **distribution** (where it can appear within the sentence) and its **sentence frame** (the distribution of the phrase with respect to other elements/constituents — essentially, this is something like its **function** in the sentence: subject, object, and so on).

Two criteria for determining a head:

(i) *the head of a constituent conditions the distribution of a constituent of which it is a part*
(ii) *the head of a constituent if the most important element of the constituent semantically*

This leads us to **a first characterization** of phrases, projections, and the like:

(14)


We will explore and **expand** this approach next.

θ-ROLES

Thematic or theta-/θ-roles express the role that an argument plays with respect to its predicate — each 1-, 2-, or 3-place predicate assigns a unique θ-role to each of its arguments.

0-place: weather-verbs — *rain, snow*
1-place: intransitive V — *walk, jump* (unergative) / *appear, arrive* (unaccusative)
2-place: (mono)transitive V — *kiss, like*
3-place: ditransitive V — *give, donate*

Remember that not only verbs are predicates (adjectives can be, nouns to a lesser extent).

θ-roles: Agent, Theme, Goal, Experiencer, Source…

(15) ***Unique θ-Generalization***
Each θ-role must be assigned but a constituent cannot be assigned more than one θ-role.

C-SELECTION

C(ategorial)-selection features express a predicate's subcategorization frame. Uninterpretable c-selectional features on a predicate thus determine the category that its argument must have (note that categorial features expressing an expression's category are always interpretable).

C-features: N, V, A, P (and later on we'll encounter more: *any* grammatical category)

(16) a. *destruction* [N] *kiss* [V] *fond* [A] *of* [P]
b. *destruction* [N, *u*P] *kiss* [V, *u*N] *fond* [A, *u*P] *of* [P, *u*N]

(17) ***Full Interpretation***
The structure to which the semantic interface rules apply contains no uninterpretable features.

Hence: In the course of the derivation, all uninterpretable features must be deleted.

(18) ***Checking Requirement***
Uninterpretable c-selectional features must be checked and can then be deleted.

(19) ***Checking under Sisterhood***
An uninterpretable c-selectional feature F on a syntactic object X is checked when X is sister to another syntactic object Y which bears a matching feature F.

S-SELECTION

S(emantic)-selection features pertain to the type of argument required by a predicate.

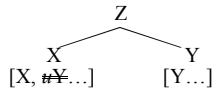
S-features: (i) proposition [describes a state of affairs that holds in the world]
(ii) entity [some kind of object or thing]
(iii) property [some kind of attribute of an entity]

Merge (as a structure-building operation within a "dumb" system) cannot see s-selectional properties — s-selectional features are semantic in nature and do not drive the syntactic derivation. Violations of s-selection only will result in a syntactically well-formed linguistic expression that simply doesn't make sense (cf. *Colourless green ideas sleep furiously*).

DERIVATIONS

Constituency
Binary Merge interacting with the checking requirement of c-selectional features.

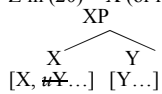
(20)



Headedness
Unchecked features of the selector project to the mother node.

(21) **Headedness**
The item that projects is the item that selects.

(22) Z in (20) = X (or rather "XP" where *P* stands for phrase/projection, which we'll get to):



Predicate
θ-roles are associated with the selectional features via the θ-grid.

give: V	c-selectional F	θ-grid	s-selectional F	θ-roles
	N	x	entity	Agent
	N	y	entity	Theme
	P	z	entity	Goal

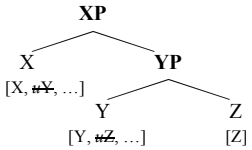
Selection
C-selectional (= uninterpretable categorial) features have to be checked under sisterhood.

(23) ***Extension Condition***
A syntactic derivation can only be continued by applying operations to the root projection of the tree (i.e. by extending the existing tree structure).

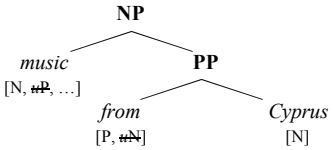
FIRST MERGE

Syntactic objects (= elements in the phrase structure, *not* an "object" as you might know!) that have all c-selectional features checked are **maximal projections** (because they don't project any further) — such a projection level is called a **phrase** (of X), generally **XP**.

(24) a.



b.



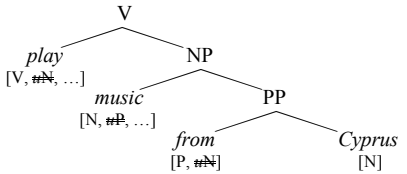
In contrast, simple lexical items are always **minimal projections** (also known as **X** or $\mathbf{X^0}$) — all they project (at that point) is into the phrase structure by virtue of being Merged.

Sometimes, a minimal projection can also be a maximal projection at the same time: **X = XP**. (For now, this goes on in cases where one lexical item doesn't have any c-selectional features and is ready for further Merge: e.g. bare nouns like *freedom* or proper names like *John*.)

In (24b), we have a noun plus PP — the noun *music* is the **head** that c-selects a PP, which consists of the head P (*from*) which c-selects an N (*Cyprus*). The c-selected element is called the **complement** (of the head that c-selects it). We thus yield a **head-complement structure**.

And once we have a maximal projection, we can keep Merging — in particular, if an element enters the derivation which has a c-selectional feature matching that maximal projection:

(25)



We know that this structure is **pretty much complete** — at least as complete as to consider it a constituent. Recall from imperatives (which lack subjects) that these structures do occur:

(26) a. Play music from Cyprus!
b. Play music from there!
c. Play it!

SECOND MERGE

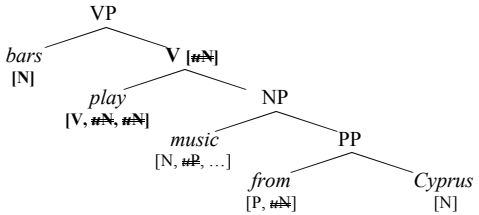
Head-complement structures are only one type of structure (still barring further discussion of **linearization**). If we don't want to express (26) as an imperative, we do need something else.

(27) Bars play music from Cyprus.

Questions: **Where** does *bars* come from? **What** does it do? **How** does it enter the derivation?

Answers: Just as NP *music from Cyprus* is a c-selected complement of V *play*, N *bars* is a c-selected element — we call it a **specifier**. The NP Merged with the predicate V must receive a **θ-role** (THEME) and intuitively, we need someone who does the playing, the AGENT: *bars*.

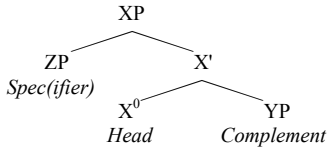
(28)



First Merge joins together a head and its c-selected complement. **Second Merge** then adds another syntactic object c-selected by the head, but because Merge is binary, it must be in a slightly different position: to the sister of an **intermediate projection** (also known as **X'**).

(29) a. minimal projection: head — X^0 / X^{min}
b. intermediate projection: bar-level — X' / X-bar
c. maximal projection: phrase — XP / X^{max}

(30)



Remember:
$X^0 = X$

ADJOIN

Yet other elements won't find a place in the structure as either complement or specifier — they won't for the reason that they are not c-selected, hence bear no θ-role: they are **adjuncts**.

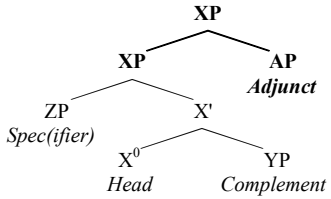
(31) Bars play music from Cyprus **at night**.

(32) Bars play music from Cyprus **loudly**.

(33) Bars **regularly** play music from Cyprus.

We represent adjuncts as **adjoined to XP** (which can usually be **to the left or to the right**):

(34)



Adjuncts are all the guys without a θ-role and without which the sentence would still be good — **adverbs** (Adv) / **adverbial phrases** (AdvP) / **other adverbials** (NPs, full sentences, etc.).

Rather than call this an instance of Merge, we name this operation simply **Adjoin**.

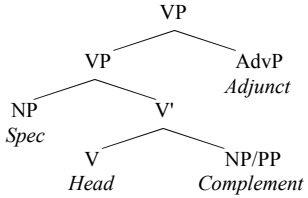
Merge *(from Adger 2003: pp. 90-91)*
1. Merge applies to **two** syntactic objects to form a new syntactic object.
2. The new syntactic object is said **contain** the original syntactic objects, which are **sisters** but which are not linearized.
3. Merge only applies to the **root** nodes of syntactic objects.
4. Merge allows the checking of an uninterpretable c-selectional feature on a head, since it creates a sisterhood syntactic relation.

Adjoin *(based on Adger 2003: pp. 112-113)*
1. Adjoin applies to **two** syntactic objects and thereby expands the old syntactic object.
2. The old syntactic object is said **contain** the new syntactic object and be **sister** with it, but they are not linearized.
3. Adjoin only applies to the **root** nodes of syntactic objects.
4. Adjoin doesn't allow any type of checking nor is it triggered by some feature since it doesn't create a sisterhood syntactic relation.

This means that adjoined objects ("adjuncts")
- are always **hierarchically outside** complements and specifiers;
- may appear on **either side** of the phrase they adjoin to;
- don't receive θ-roles;
- **are** phrasal.

PHRASE STRUCTURE

(35)



The *specifier* is more often than not the **subject** (θ-role: AGENT), the *complement* the **(direct) object** (θ-role: e.g. THEME), the *adjunct* an **adverbial of sorts**. More later on this station.

C-COMMAND

Sisterhood does a lot for us already: it is the relation under which features are checked. Beyond that, **c-command** is a new relation which will become increasingly important.

(36) *C-Command*
A node A c-commands a node B if, and only if [= *iff*] A's sister either:
a. is B, or
b. contains B.

As you can see, c-command is not as **local** as sisterhood: it can stretch quite some distance.

While many nodes may c-command others in a given (large) tree, there is nothing inherently important to this relation — but **some grammatical phenomena require c-command**. In these, we speak of the requirement of **an antecedent to c-command the element it licenses**.

(37) ***The Reflexive Generalization***
A reflexive must be coreferential with a c-commanding expression.

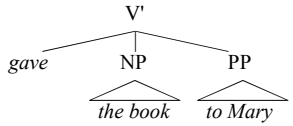
(38) ***The NPI Generalization***
NPIs must c-commanded by a negative element.

THREE ARGUMENTS — WHAT TO DO?

Ditransitive verbs have three arguments: usually, an **NP-subject** (AGENT, should be in Spec), an **NP-object** (THEME, should be the verb's complement), and a **PP-(indirect) object** — but where does that go (GOAL, BENEFICIARY, SOURCE, etc.)??? It should also be a *complement*…

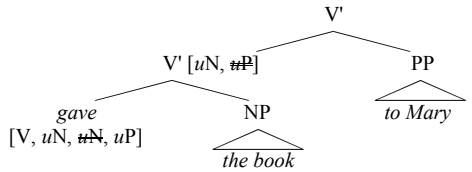
Let's focus on the position of the **two objects**, the intermediate projection called V':

(39)



But the structure in (1) is **ternary branching**, which isn't possible with Merge as defined.

(40)



An **intermediate projection** arises if the head still has unchecked (c-selectional) features — so there shouldn't be anything wrong with having an additional intermediate projection: V'.

However, the only instance we've seen such **recursion** so far was with adjuncts, which don't get a θ-role, so this might be conceptually worrying nevertheless. Moreover, what determines the **ordering of complements**: why couldn't PP merged first, checking [*u*P], and then NP?

SOME ARGUMENTS FOR BINARITY WITH THREE ARGUMENTS

But the big question is: apart from **conceptual considerations** (binarity), are there **empirical reasons** to favour a binary-branching analysis over a ternary-branching one for ditransitives?

(41) a. John said he would **give the book to Mary** and [**give the book to Mary**] he did.
b. *John said he would **give the book** to Mary and [**give the book**] he did to Mary.

(42) A: Who gave the book to Mary? — B: * John **{ Ø / did }** to Mary.

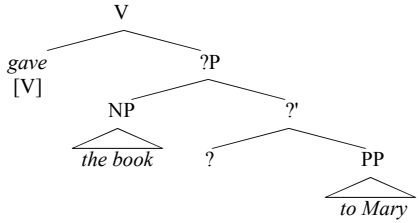
(43) a. John [gave the book] and [sent the paper] to Mary. *PAUSE, INTONATION* ⇨
b. John [gave the book **Ø**] and [sent the paper to Mary]. ☑

We have to look for alternatives. Let's **play with the pause/intonational funniness** point:

(44) a. John [gave the book to Mary] and [Ø the paper to Jane] *NO SUCH PATTERN* ⇨
b. John gave [the book to Mary] and [the paper to Jane] ☑

NP- & PP-objects seem to form **one constituent** (say **?P**), with the **verb somewhere higher**:

(45)



The so-called **double-object construction** gives us further support for something like (7):

(46) a. John gave the book **to Mary**.
b. Mary sent some flowers **to Peter**.
c. Peter showed a picture **to Jane**.

(9') a. John gave **Mary** the book.
b. Mary sent **Peter** some flowers.
c. Peter showed **Jane** the picture.

Above tests work here too; an interesting test for constituency comes from **reflexive binding**.

(47) a. John showed Mary **herself** in the mirror.
b. * John showed **herself** Mary in the mirror

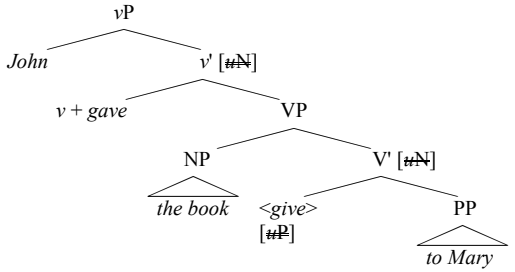
(48) *Mary showed **herself** to John.

All this suggests that **either the ternary or the '?P'-analysis are consistent with the facts**. Since ternarity would require a reformulation of Merge, let's try and **find evidence for '?P'**.

The trick will be **'little *v*'** for '?': consider evidence from **causatives** and **verb movement**.

We thus assume the following **structure with a 'little *v*'**, which we'll elaborate on next class:

(49)



LINKING: UTAH

We now have a **hierarchy of projections** (*v* >> V, with **Move**: V undergoes movement to *v*) and know which argument (AGENT, THEME, GOAL…) appears in which position. **Or do we…?**

(50) a. John gave **it** to Mary.
b. *John gave to Mary **it**.

We need a way to connect a certain type of argument to a certain type of position: **Linking**.

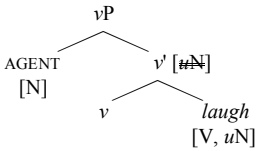
(51) ***Uniformity of Theta-Assignment Hypothesis (UTAH)***
Identical thematic relationships between predicates and their arguments are represented syntactically by identical structural relationships when items are Merged.

Now we do! Consider:

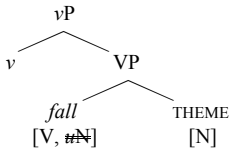
(52) a. NP daughter of *v*P → interpreted as AGENT
b. NP daughter of VP → interpreted as THEME
c. PP daughter of V' → interpreted as GOAL

This gives us also a **structural distinction** between **unergative and unaccusative** verbs:
[NB: I added the **appropriate features** that are checked, which Adger (p. 140) left out!]

(53) a.



b.



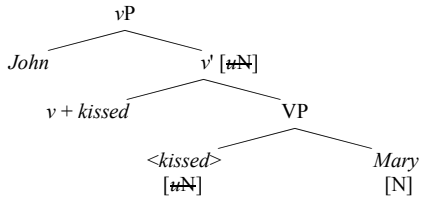
In Italian for example unergatives and unaccusatives differ **morpho-syntactically**: unergative verbs in past participle form don't **agree with the subject**, unaccusatives do, and **auxiliary selection** differs famously with the former taking a form of *have*, the latter a form of *be*.

(54) a. Molte ragazze **hanno** telefonato.
many girls have phone-PAST_PART.3SG
'Many girls have phoned.'

b. Molte ragazze **sono** arrivate.
many girls are phone-PAST_PART.3PL
'Many girls have arrived.'

And **simple transitives** can be captured by the *v*P-analysis too, of course!

(55)



ON THE PRINTED HANDOUT (NOT THE PDF), EXERCISES FOLLOW! [=16 PP. TOTAL]