

L2 Syntax Lecture 2: Lexical entries and sentence structure

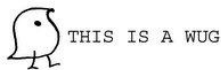
Robert Truswell

Generative grammar

"A generative grammar is an algorithm for generating **all** and **only** the grammatical sentences in a language"

S → NP VP
VP → V (NP)
NP → (Det) N
Det → {a | the}
N → {dog | cat | monkeys}
V → {surprised | chased | slept}

The importance of rules



"... now there are two of them.
There are two..."

"Wugs!"

Descriptive grammar

We are interested in observations and generalizations about how people actually use language

- We don't know which side the police are on
- ? We don't know on which side are the police
- ?? We don't know on which side the police are
- * We don't know are on which side the police

We are interested in **description**, not prescription

Generative grammar

- Description, not prescription
- Explicit description of a language
- Cognitive basis
- Aim for maximally accurate, simple, and general descriptions

The impossibility of memorization

- How many (morphologically simple) words does an average speaker of English know?
- How long is the longest sentence in English?
- How many sentences are there in English?
- How long would it take you to rote-learn all the sentences of English?

Chunks

- [[The boy who is sleeping] is happy]
- *[Is [the boy who ___ sleeping] is happy]?
- [Is [the boy who is sleeping] ___ happy]

The substitution test

- (1) The man chased the dog
 - (2) The man chased it
 - (3) He chased it
- "the" and "dog" form a constituent
 - "the" and "man" form a constituent

The coordination test

- (1) The man chased the dog
 - (2) [The man] and [the woman] chased the dog
 - (3) The man chased [the dog] and [the cat]
 - (4) The man [chased the dog] and [stroked the cat]
- "the" and "man" form a constituent
 - "the" and "dog" form a constituent
 - "chased", "the" and "dog" form a constituent

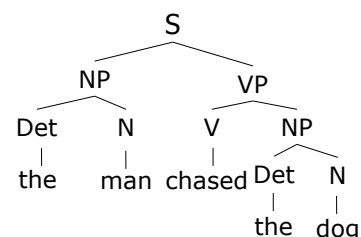
The question test

- (1) The man chased the dog
 - (2) Who chased the dog? The man
 - (3) What did the man chase? The dog
 - (4) What did the man do? Chase the dog
- "the" and "man" form a constituent
 - "the" and "dog" form a constituent
 - "chase", "the" and "dog" form a constituent

Bracket representation

the man chased the dog
[the man] chased [the dog]
[the man] [chased [the dog]]
[[the man] [chased [the dog]]]
[_S [NP the man] [VP chased [NP the dog]]]

A clearer representation: trees



The constituents are all subtrees

Today: moving towards lexicalized theories

- Syntactic categories
- Verbs and subcategorization
- Thematic roles

Today's reading

- Haegeman (1994), sections 1-4 of ch.1 (pp.33-55)
- Copies in the DSB ground floor resource room, on WebCT, and in the library

What is a syntactic category?

Noun:	table, fight, emptiness, ...
Verb:	jump, sleep, exist, ...
Preposition:	on, of, since, despite, ...
Determiner:	a, the, no, ...
Adjective:	tall, former, nonexistent ...

How do we work out syntactic categories for a word? How do kids?

Working out syntactic categories

- (1) The **piblo** was 6 feet wide
- (2) That guy was **snarping** in my bed!
- (3) A **gont** nested in Hamish's beard
- (4) It's over there, **asmast** the ganyions
- (5) **Breft** dogs savaged our best ewe
- (6) **Gonts snarp asmast breft piblos**, usually

Working out syntactic categories

Syntactic categories are determined by the **distribution** of words and constituents (not semantics or surface form)

- words and phrases of the same category appear in the same positions
- that's why the substitution test works

When you learn a language, you learn the categories and the words in them.

The lexicon

Lexicon: mental dictionary, full of memorized facts about words

dog: N
slept: V
piblo: N
snarped: V

...

Grammar and lexicon

Syntactic information in the lexicon is visible to the grammar.

S	→	NP VP
VP	→	V (NP) (NP)
NP	→	Det N
Det	→	{a the}
N	→	{man woman book dog}
V	→	{slept chased sold}

Grammar and lexicon

Grammar	Lexicon
S → NP VP	the: Det
VP → V (NP)(NP)	a: Det
NP → Det N	slept: V
	chased: V
	sold: V
	man: N
	...

Testing the grammar again

S	→	NP VP
VP	→	V (NP) (NP)
NP	→	Det N
Det	→	{a the}
N	→	{man woman book dog}
V	→	{slept chased sold}

Testing the grammar again

- [S] (S → NP VP)
- [NP VP] (NP → Det N)
- [[Det N] VP] (VP → V NP)
- [[Det N] [V NP]] (NP → Det N)
- [[Det N] [V [Det N]]] (Lexical insertion)
- [[A dog] [chased [the man]]]
- [[The man] [sold [a book]]]
- ...

Testing the grammar again

S	→	NP VP
VP	→	V (NP) (NP)
NP	→	Det N
Det	→	{a the}
N	→	{man woman book dog}
V	→	{slept chased sold}

Does this generate **only** grammatical sentences?

Verbs differ in transitivity

- *The man chased
The man chased the dog
- *The man chased the woman the book

“chased” is a **transitive** verb

- it takes exactly 1 **complement**

Verbs differ in transitivity

The man slept

*The man slept the dog

*The man slept the woman the book

“slept” is an **intransitive** verb

- it takes no complements

Verbs differ in transitivity

*The man sold

The man sold the dog

The man sold the woman the book

“sold” is either **transitive** or **ditransitive**

- it takes one or two complements

Fixing our grammar

S → NP VP
VP → V (NP) (NP)
NP → Det N
Det → {a | the}
N → {man | woman | book | dog}
V → {slept | chased | sold}

We need to add some information to this:

- “slept” is intransitive, “chased” is transitive, “sold” is transitive or ditransitive

Subcategorization frames

- The standard approach: **subcategorization frames** in lexical entries.

slept: V, []
chased: V, [NP]
sold: V, [NP]
sold: V, [NP, NP]

Subcategorization frames

- The standard approach: **subcategorization frames** in lexical entries.
- Although this information seems semantic, it is at least partially syntactic
- John ate/devoured the beetroot
- John ate/*devoured

Subcategorization frames

- The standard approach: **subcategorization frames** in lexical entries.
- Although this information seems semantic, it is at least partially syntactic
- **Eat:** V, []
- **Eat:** V, [NP]
- **Devour:** V, [NP]

Grammar and lexicon again

Grammar

S	→	NP VP
VP	→	V (NP)(NP)
NP	→	Det N

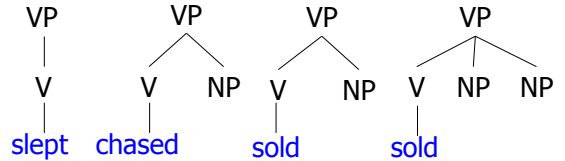
Lexicon

slept:	V, []
chased:	V, [] NP
sold:	V, [] NP
sold:	V, [] NP NP
...	

Same info specified multiple times:
redundant

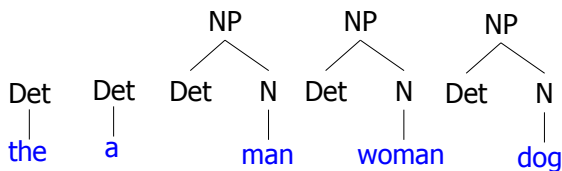
Enriching lexical entries: elementary trees

One solution: replace phrase structure rules for VP with lexical entries using elementary trees



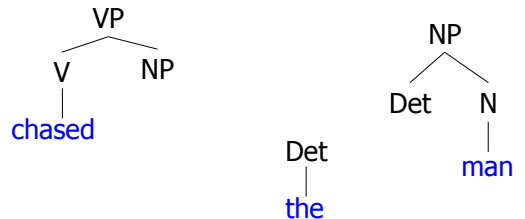
More elementary trees

Similar trees for nouns and determiners:



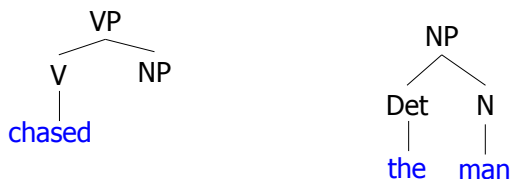
Building a tree

The elementary trees fit together by substitution



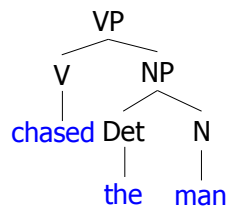
Building a tree

The elementary trees fit together by substitution

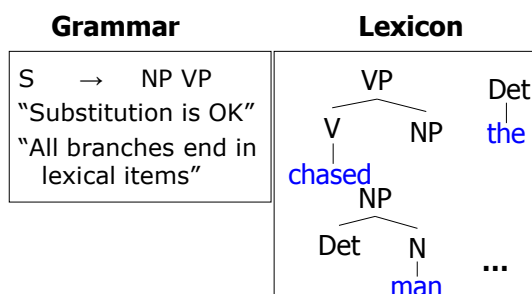


Building a tree

The elementary trees fit together by substitution



Grammar and lexicon again



Summary so far

- Lexical entries contain information on
 - syntactic category
 - subcategorization
- Having the same info in two places (grammar + lexicon) is redundant
 - Replace phrase structure rules with enriched lexical entries
- S → NP VP still needed (for now...)

Summary so far

- The intuition: grammatical relations are local
 - Everything to do with the verb is determined within VP
 - The verb doesn't care about more distant material
- Another example: agreement
 - The boy snores/*snore
 - The parents snore/*snore
 - [[The boy]'s parents] snore/*snore
 - He says his parents snore/*snore

Thematic roles

The boy ate the apple

There are two participants here – the boy and the apple.

They are both NPs.

They play different roles:

- eater (the boy)
- eatee (the apple)

Thematic roles

The girl built a sandcastle

- "The girl" is Builder
- "A sandcastle" is Buildee

Thematic roles

The boy ate the apple
The girl built a sandcastle

- "The boy" and "The girl" are Agents: Individuals who intentionally initiate the action described by the verb;
- "The apple" and "a sandcastle" are Themes: Things affected by the action or state described by the verb.

Thematic roles

The boy ate the apple
The girl built a sandcastle

- Generalizations like these are useful because they allow us to predict the position of NPs based on the role they play.
- Agents are typically subjects
- Themes are typically objects

Thematic roles

- Agents are typically subjects
- Themes are typically objects

This means there are no verbs like *shmeat*

- "The apple shmeated the boy" means
- "The boy ate the apple"

Thematic roles

- Agents are typically subjects
- Themes are typically objects

In a passive sentence, the NPs appear to be the wrong way round:

- "The apple was eaten by the boy" means
- "The boy ate the apple"

More on this in a couple of weeks

More thematic roles

- Experiencer: Barry fears spiders.
- Beneficiary: I made this for you.
- Goal: Tom gave Jane a present.
- Location: It all happened in India.

The bad news

- There is no agreement concerning a list of thematic roles.
- Some people think there are half a dozen, some people think there are thousands (each verb's thematic roles are unique to that verb)

The good news

- For this course, we don't have to care.
- The number of thematic roles assigned by a verb is more important than their names
- It will be useful to recognize agents (eaters and builders) and themes (eatees and buildees)

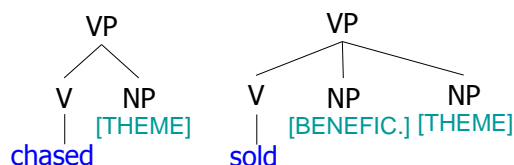
Verbs assign θ -roles

Thematic roles depend on the verb.
This information must be in the lexicon:

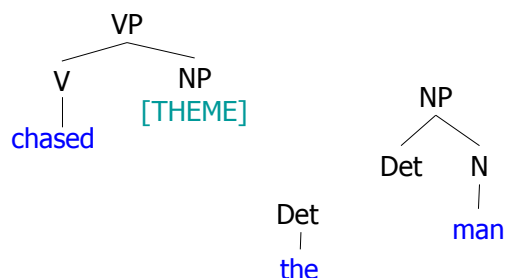
- **chased:** subject = AGENT, complement = THEME
- **danced:** subject = AGENT
- **melted:** subject = THEME
- **sold:** subject = AGENT, 1st complement = BENEFICIARY (if present), 2nd complement = THEME

Enriching elementary trees

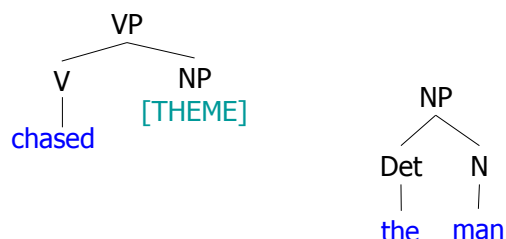
Assigning thematic roles to complements is straightforward



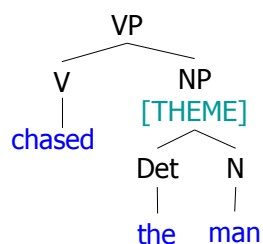
Building a tree again



Building a tree again



Building a tree again

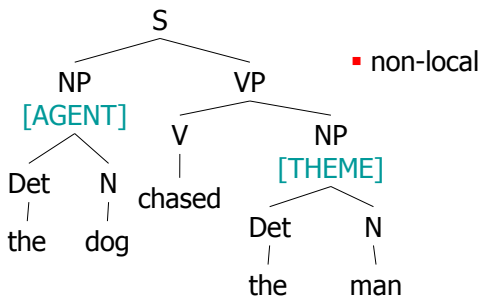


Assigning a thematic role to the subject?

- **melted:** assigns THEME to sentence subject, ...
- **chased:** assigns AGENT to subject, ...
- **sold:** assigns AGENT to subject ...

How does the **subject** of the sentence get its thematic role?

Like this?



Summary

- Lexical entries contain information on
 - syntactic category
 - subcategorization
 - thematic roles
- Enrich lexical entries at the expense of phrase structure rules
- $S \rightarrow NP VP$ is still needed, but ...
 - how does the subject get its thematic role?

Review

Pair up

- **Person 1:** explain why we got rid of (most) phrase structure rules and started using elementary trees
- **Person 2:** explain the relationship between thematic roles and elementary trees

Next Monday:

- A closer look at the structure of VPs
- Getting rid of $S \rightarrow NP VP$
- The beginnings of X' theory