

# THE HISTORY OF THE CONCEPT OF THE PHRASE

*Part I: Why early linguists had no  
idea what a phrase was*



# A BRIEF INTRODUCTION TO THE PHRASE

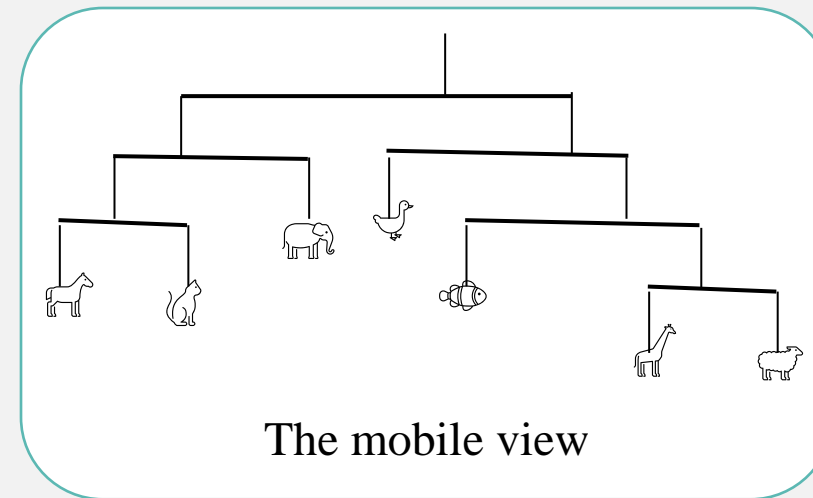
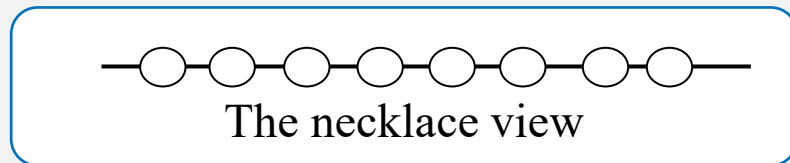
- The phrase is a syntactic unit bigger than a word:

[<sub>NP</sub> The cowboy] [<sub>VP</sub> reached [<sub>PP</sub> for [<sub>NP</sub> his gun]]]

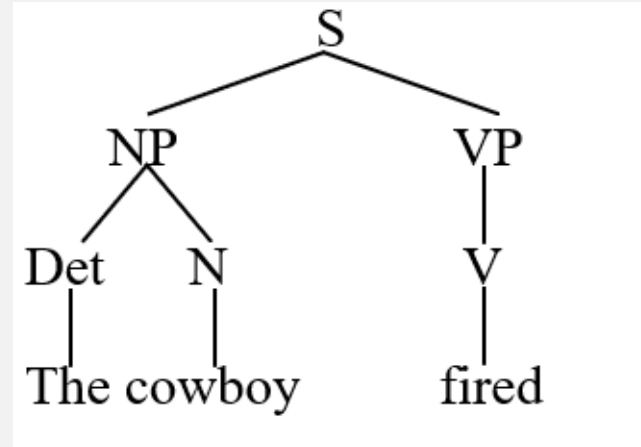
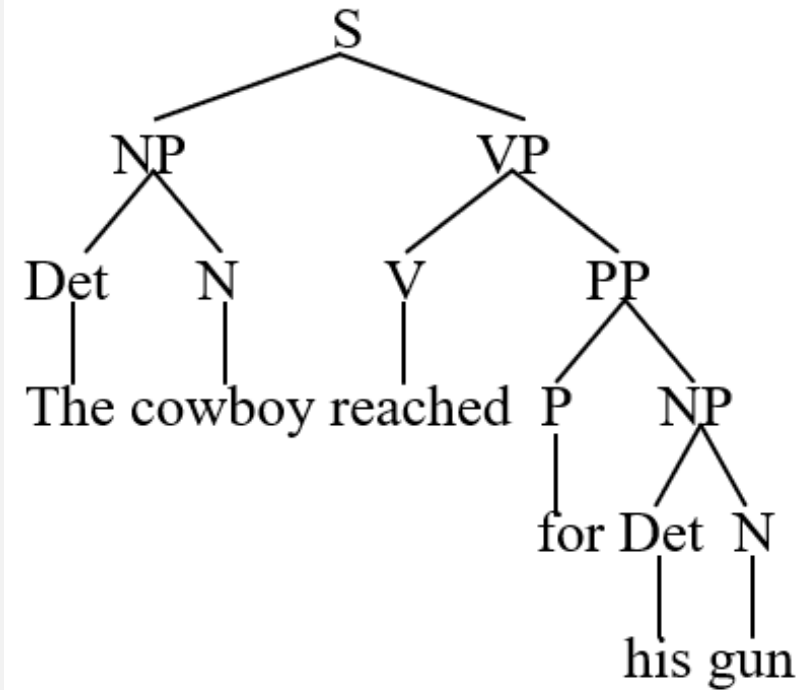
- How many words make a phrase?

[<sub>NP</sub> He] [<sub>VP</sub> fired]

- As these words have the same distribution as the phrases, they are phrases themselves
- Thus, a phrase can be as small as one word
- So, how is a phrase *bigger* than a word?
- The answer has to do with the notion of structure



# THE REPRESENTATION OF SYNTACTIC STRUCTURE: TREES



The phrase may only contain one word – but it can ‘potentially’ contain more.

Words cannot contain any more words

Therefore, the phrase must be bigger.

# HOW BIG CAN A PHRASE BE?

- Phrases are contained in sentences
  - But is a sentence something different from a phrase?
- Sentences contain a subject and a predicate

I consider [that he is foolish] – finite

I consider [him to be foolish] – non-finite

I consider [him foolish] - ???

- Current wisdom has it that the last case involves an adjectival phrase with a subject.
- Sentences therefore are no different from phrases.
  - They are considered to be a certain kind of phrase.

# REASONS TO ASSUME PHRASES

- Intuitions
  - Might be more semantic based

The cowboy

reached for his gun

The cowboy reached

for his gun

# REASONS TO ASSUME PHRASES

- Distribution

The cowboy

reached for his gun

- Distribution is determined by grammatical rules
- If something has a distribution, it must be something the grammar recognises

# REASONS TO ASSUME PHRASES

- Distribution **The stranger in the ten gallon hat** **reached for his gun**
- Distribution is determined by grammatical rules
- If something has a distribution, it must be something the grammar recognises

# REASONS TO ASSUME PHRASES

- Distribution

He

reached for his gun

- Distribution is determined by grammatical rules
- If something has a distribution, it must be something the grammar recognises



# EARLY LINGUISTIC WORKS: INDIAN GRAMMARIANS

- Indian grammarians studying Sanskrit from about 600 BCE
- Sanskrit was the language of religious texts and ceremonies
  - It was spoken 1700 – 1200 years ago
  - It wasn't anyone's mother tongue at the time of the Indian grammarians (dead language)
  - It needed to be preserved – texts and incantations needed to be pronounced properly and grammatically

- Sanskrit was a morphologically complex language

- 3 persons
- 3 numbers
- 5 tenses
- 3 voices
- 4 moods

- In principle, this gives 540 different verb forms!

- Morphemes could be suffixes, infixes or reduplications:

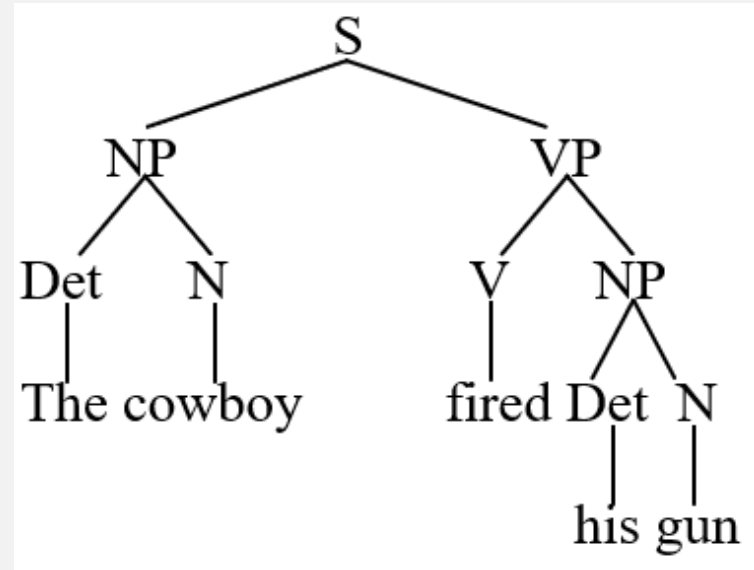
a	naya-nti lead-3.pl (=they lead)
b	naya-si lead-1.sing (=you lead)
c	naya-ami lead-1.sing (=I lead)
d	naya → neṣya-nti lead+fut(infix)-3.pl (=they will lead)
e	naya-ya-nti lead-caus-3.pl (they make someone lead)

- Because of rich verbal and nominal morphology, word order was fairly free
  - Involved nuanced meaning differences, like Hungarian



# PANINI (ABOUT 400 BCE)

- Panini's grammar was highly developed
  - Contained notions (e.g. the phoneme) which did not show up in European grammars until late 19<sup>th</sup> C
  - Very formal (not part of European grammars until mid 20<sup>th</sup> C)
- It has been claimed that the grammar contained rules similar to those used in the 1960s to produce phrase structures:
  - $S \rightarrow NP VP$
  - $VP \rightarrow V NP$
  - $NP \rightarrow D N$
- $A \rightarrow AA / C-D$ 
  - This is a morphological rule (reduplication)
  - The form is not the one that Panini used

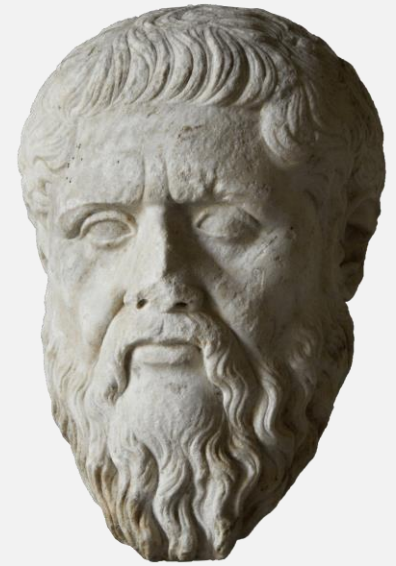


# THE LACK OF SYNTAX IN INDIAN GRAMMARS

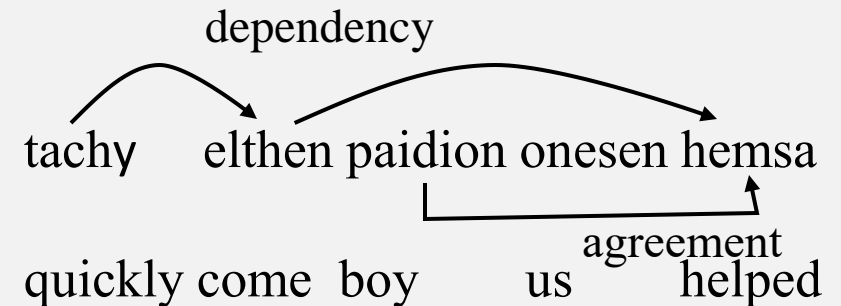
- The main focus was on pronunciation (phonetics and phonology) and word form (morphology)
  - These things are obvious features of the language and important for the purpose of the grammar
- Free word order obscured syntactic phenomena
  - It is hard to see phrases if words do not have to ‘stick together’ as they do in a language like English

# GREEK LINGUISTICS

- The Greeks developed phonetic based writing systems from about 800 BCE
  - This showed an awareness of phonetic aspects of the language
  - The term ‘grammar’ originally meant ‘the study of letters’
- But they didn’t study language as a specific topic until much later
  - It was all part of philosophy, rhetoric, logic, epistemology, etc.
- Most original works have been lost and we only know about them through secondary sources
- Early work by Plato, Socrates and Aristotle
  - First proposal of subject – predicate distinction
    - But part of logic, not linguistic analysis per se
    - Defined as words, not phrases (idea persists in ‘school grammars’)
  - Aware of a notion of ‘subordination’
    - But this was ‘semantic’ in nature – dependency
    - Doesn’t necessarily lead to the idea of a phrase
- Later work by Thrax and Apollonius
  - Thrax’s grammar (100 BCE) is the earliest surviving text on language
    - Concerned phonology and morphology – no syntax
  - Apollonius wrote about syntax
    - But only about agreement, concord and semantic dependency



Plato



‘Quickly coming up, the boy helped us.’

# LACK OF PHRASES IN GREEK GRAMMAR

- Like Sanskrit, Greek was morphologically complex and had free word order
  - These focus attention on the word and away from syntax
- A more philosophical approach focussed on meaning
- Thus, relations between words in a sentence were looked on in terms of morphology and semantics
  - There didn't seem to be any need to involve syntax

# LATIN

- The Romans were good soldiers and engineers.
  - They relied on the Greeks for ‘academic’ thinking.
- Early works by Varro did attempt to further debates by Greek philosophers, but mainly in relation to epistemology and etymology.
- The most influential Latin grammar was written by Priscian (500 CE)
  - Based almost entirely on Thrax’s grammar
    - Mainly concerned with forcing an analysis of Latin into a framework designed for Greek
  - Had two volumes on syntax
    - But nothing interesting – wrongheaded ideas about word order reflecting the natural world
  - Introduced the concept of ‘subordinate clause’
    - But again, more to do with dependency than structural hierarchy



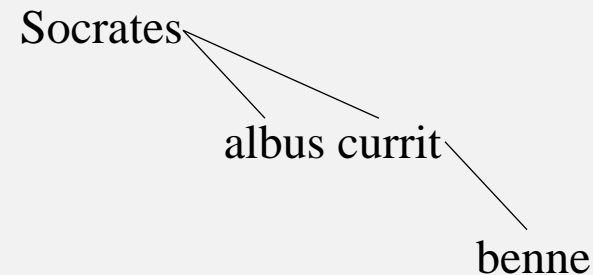
Priscian

# WHY THE ROMANS DIDN'T KNOW ABOUT PHRASES

- They were not particularly interested in scholarly development
  - Differed to the Greeks and didn't attempt to go beyond
  - As the Greeks had not discovered the notion, the Romans were not about to.
- Latin was morphologically rich and had relatively free word order,
  - So, nothing special about the language which was going to suggest a different approach to previous works

# THE DARK AGES

- Rejected ‘pagan’ Greek teachings
- Knowledge was to be based on Christian biblical teaching = Scholasticism
- Latin was the language of religion, politics and scholarly work (such as it existed)
  - Latin therefore became a lingua franca
  - But Latin, as mother tongue, had already died out
  - The teaching of Latin became very important
- The main grammar used was the Priscian one
  - Irony: Priscian grammar was based on Thrax, so Greek linguistics had more influence on language study at this time than they wanted.
- Needless to say, nothing much was learned during this time
- Speculative Grammar
  - Started towards the end of this period
  - Thomas of Erfurt (1310) reinvented Socratean analysis: subject – predicate, with dependent elements:
    - Socrates albus currit benne  
Socrates white runs well
    - The verb was dependent on the subject as it is morphologically linked to it





# WHY WAS THE PHRASE NOT DISCOVERED IN THE MIDDLE AGES?

- Very little was discovered during this time
- It wasn't called the Dark Ages for no reason.

# THE RENAISSANCE

- Classical teachings were back!
- But new ideas were also encouraged.
- Philosophically the big debate was between empiricism (British) and rationalism (French)
  - Concerned aspects of language, but nothing to do with syntax
- Port Royal Grammarians
  - Major group investigating language at the time;
  - Unfortunately moved away from the Greek morphological approach to categories towards a semantic one. This has been a problem ever since.
  - Took a step in grammatical analysis which virtually ensured they would not discover the phrase:
    - Subordination relationships were to be seen as underlying independent sentences:
      - The invisible God created the visible world.
        - God, who is invisible created the world, which is visible
        - God is visible. God created the world. The world is visible
  - Why was the phrase not discovered?
    - People were looking elsewhere – mainly meaning and so syntax was not a big part of their thinking.



# COMPARATIVE LINGUISTICS

- Main interest in classifying languages into language families
  - Partly due to Renaissance's interest in classical languages
  - But also due to European 'discovery' of Indian linguistics (Walter 1733)
- Ironically enough, the idea of 'language families' came from the scholastic based invention of Scythian:
  - According to the Bible, after the flood, Noah son Japheth moved to Europe and fathered all European nations.
    - The original language was Scythian, which turned into different languages (Tower of Babel 700 years after the flood).
- Indo-European was proposed as the ancestor to most European languages,
  - Evidence mainly on phonology (sound laws discovered at this time) and lexicology
  - Not much on syntax as there was not much available about the syntax of classical languages
- Again, the phrase was not discovered because no one was looking for it.



# EUROPEAN STRUCTURALISM



- Ferdinand de Saussure ‘Course in General Linguistics’ (1907-1911)
  - Published 1916 by his students after his death.
- ‘Structuralism’ was nothing to do with syntactic structure, but the idea that the signs (words) of a language form a system which are impossible to define outside of the system:
  - meaning of elements is determined relative to other elements
- E.g. 10.30 to Paris
- Main concentration here is on words and meaning
- Little to do with syntax.
  - So, no hope of discovering the phrase.

# CONCLUSION

- The phrase was not discovered at any point in the majority of the history of the study of language
  - From 600 BCE until 1916 CE (over 2500 years)
- The reason for this, generally, is that syntax was not a large part of linguistic investigation
  - And when it was considered, it was seen in terms of words and their semantic relation to each other
- Properties of classical languages distracted from the serious study of syntax
  - Complex morphology
  - High degree of word order freedom
- There has been a tendency to base linguistic study on previous ideas, adapting them (sometimes to the detriment of simplicity) to newer languages.
- Early studies were based on a single language and so the limits of variation were never considered.
  - When languages were compared, it was mainly based on words and pronunciation
- Ultimately, the main reason no one found the phrase was that no one was looking for it.