



Faculty of Philosophy General Linguistics

# Syntax & Semantics WiSe 2022/2023 Lecture 2: Basic Concepts I

27/10/2022, Christian Bentz



## Overview

Section 1: Recap of Lecture 1

#### Section 2: Constituency Definition

- Section 3: Constituency Tests
  - Substitution Pronominalization Question Formation Permutation Fronting Coordination Problems
- Section 4: Parts of Speech (POS)
  - Definition Determining POS





Faculty of Philosophy General Linguistics

## Section 1: Recap of Lecture 1



## Syntax

(1) /baʊm/ Baum tree.NOM.SG NOUN "tree" *tree*′ Section 1: Recap of Lecture 1

Section 2: Constituency

Section 3: Constituency Tests

Section 4: Parts of Speech (POS)

References

Syntax typically deals with the morphosyntactic and POS level.



## Interlude: Chomsky's Colorless Green Ideas

(2) Farblos-e grün-e Idee-n schlaf-en furios. colorless-NOM.PL green-NOM.PL ideas-NOM.PL sleep-PRS.3PL furiously [[[ADJ ADJ NOUN] VERB] ADV] "Colorless green ideas sleep furiously." Meaning ???

Chomsky (1957, p. 15) has famously argued that syntax and semantics are independent of one another, since sentences as the one above can be syntactically analyzed and are grammatically correct, while deriving a meaning is impossible. Is it? ...

Section 1: Recap of Lecture 1

Section 2: Constituency

Section 3: Constituency Tests

Section 4: Parts of Speech (POS)



## The Design Features of Human Language

"A set of 13 design-features is presented in the illustration on the opposite page. There is solid empirical justification for the belief that all the languages of the world share every one of them."

Hockett (1960). The origin of speech, p. 90.



#### Section 1: Recap of Lecture 1

Section 2: Constituency

Section 3: Constituency Tests

Section 4: Parts of Speech (POS)



## **Duality of Patterning**

"Language is structured on at least two levels (Hockett, 1960). On one level, a small number of meaningless building blocks (phonemes, or parts of syllables for instance) are combined into an unlimited set of utterances (words and morphemes). This is known as combinatorial structure. On the other level, meaningful building blocks (words and morphemes) are combined into larger meaningful utterances (phrases and sentences). This is known as compositional structure."

Little et al. (2017), p. 1.





## **Compositionality in Syntax**

(3) Kim sieht einen großen Baum [PROPN [VERB [DET [ADJ NOUN]]]]"Kim sees a big tree"

In the example above, the elements of the sentence which combine to larger phrases (e.g. adjective and noun, determiner and noun phrase, etc.) are *adjacent* to one another. However, this is not always the case... Section 1: Recap of Lecture 1

Section 2: Constituency

Section 3: Constituency Tests

Section 4: Parts of Speech (POS)



## **Problem: Non-Adjacency**

(4) Kim hat den Wald vor lauter Bumen nicht gesehen
... [AUX ... ... ... ... ... VERB]
"Kim hasn't seen the forest because of all the trees."

The term *non-adjacency* refers to the fact that elements of a sentence which depend on each other, do not necessarily occur next to each other in linear order. This is also sometimes referred to as the *non-linearity* of syntax, or *long-distance dependency*.

Section 1: Recap of Lecture 1

Section 2: Constituency

Section 3: Constituency Tests

Section 4: Parts of Speech (POS)



## **The Elements of Combinations**

(5) Kim sieht einen großen Baum [PROPN [VERB [DET [ADJ NOUN]]]]"Kim sees a big tree"

In the example above, **the elements** of the sentence which **combine to larger phrases** (e.g. adjective and noun, determiner and noun phrase, etc.) are *adjacent* to one another.

How do we know which elements combine?

Section 1: Recap of Lecture 1

Section 2: Constituency

Section 3: Constituency Tests

Section 4: Parts of Speech (POS)





Faculty of Philosophy General Linguistics

## **Section 2: Constituency**





## **Definition:** Constituents

Both the **basic elements**/units of a sentence – often orthographic words – as well as **combinations of those**, i.e. **phrases**, count as constituents.

Most basic constituents: [Kim] [sees] [a] [big] [tree]

Higher level constituents: [big[tree]], [a[big[tree]]], etc.

Müller (2019). Grammatical theory, p. 7.

Note: Most of the time when the term *constituent* is used in linguistic articles, it will refer to higher level constituents, rather than single words. Section 1: Recap of Lecture 1

Section 2: Constituency

Section 3: Constituency Tests

Section 4: Parts of Speech (POS)





Faculty of Philosophy General Linguistics

## Interlude: Wordhood



## What is a word anyways?

The general distinction between morphology and syntax is widely taken for granted, but it crucially depends on a cross-linguistically valid concept of '(morphosyntactic) word'. I show that there are no good criteria for defining such a concept. I examine ten criteria in some detail [...] and I show that none of them is necessary and sufficient on its own, and no combination of them gives a definition of 'word' that accords with linguists' orthographic practice.

Haspelmath (2011). The indeterminacy of word segmentation and the nature of morphology and syntax, p. 31.

Section 1: Recap of Lecture 1

Section 2: Constituency

Section 3: Constituency Tests

Section 4: Parts of Speech (POS)



## Word Criterion: Free Occurrence

"Bloomfield (1933: 160) called utterance segments that can occur on their own **free forms**, and he famously defined the word as "a free form which does not consist entirely of (two or more) lesser free forms; in brief, a word is a minimum free form"."

Haspelmath (2011), p. 39 citing Bloomfield (1933).

## Example

(6) Where are you? - *Here*.What do you need? - *Money*.

Section 1: Recap of Lecture 1

Section 2: Constituency

Section 3: Constituency Tests

Section 4: Parts of Speech (POS)



## Word Criterion: Free Occurrence

"But this definition does not single out forms that correspond to our intuition of grammatical words. On the one hand, it is too strict, because by this definition compounds [...] would not be words, but phrases, because they have constituents that are themselves free forms. On the other hand, it is much too loose, because many phrases [...] would count as words [...]"

Haspelmath (2011), p. 39-40.

## Example

- (7) *firewater* (two separate free forms): fire water
- (8) *the flower* (one single free form): \*the *al-zahra* 'the flower' (Standard Arabic)

Section 1: Recap of Lecture 1

Section 2: Constituency

Section 3: Constituency Tests

Section 4: Parts of Speech (POS)



			395	~					
	1983		& Mchombo 1995	7661 p.				900	
	ıllum		<b>Ichon</b>	eSour	1999		05	lise 2	2007
	Zwicky & Pullum	Kanerva 1987	nan & N	Ackema & LeSourd	Monachesi 1	Harris 2000	Milićević 2005	Lieber & Scalise 2006	Bickel et al. 2
	Zwio	Kane	Bresnan	Acke	Mon	Harr	Milid	Lieb	Bick
ee occurrence				+			+		
ernal mobility and internal fixedness	+			+	+	+			
interruptibility				+					+
n-selectivity	+	+			+	+	+		+
n-coordinatability			+	+	+		+	+	+
phoric islandhood			+					+	
nextractability			+					+	
orphophonological idiosyncrasies	+	+			+	+	+		
viations from biuniqueness									+

#### *Table 1.* Nine studies that examine wordhood using test batteries

Haspelmath (2011), p. 60.



## **Back to Constituency**

tree NOUN big [tree] ADJ [NOUN] a [big [tree]] DET [ADJ [NOUN]] sees [a [big [tree]]] VERB [DET [ADJ [NOUN]]] Kim [sees [a [big [tree]]]] PROPN [VERB [DET [ADJ [NOUN]]]] Section 1: Recap of Lecture 1

Section 2: Constituency

Section 3: Constituency Tests

Section 4: Parts of Speech (POS)



## But why not:

sees VERB [sees] a [VERB] DET [[sees] a] big [[VERB] DET] ADJ [[[sees] a] big] tree [[[VERB] DET] ADJ] NOUN [Kim [[[sees a] big] tree]] [PROPN [[[VERB DET] ADJ] NOUN]] ?

Section 1: Recap of Lecture 1

Section 2: Constituency

Section 3: Constituency Tests

Section 4: Parts of Speech (POS)

References

## How do we know which elements combine?

19 — Syntax & Semantics, WiSe 2022/2023, Bentz





Faculty of Philosophy General Linguistics

## Section 3: Constituency Tests



## Constituency Tests

There is a set of tests for "sequences of words" in order to establish whether these constitute *higher order* constituents.

Substitution
Pronominalization
Question Formation
Permutation
Fronting
Coordination

Note: While Müller (2019, p.7) does not explicitly define the length of such sequences of words as > 1, it seems clear that this is meant here.

Section 1: Recap of Lecture 1

Section 2: Constituency

Section 3: Constituency Tests

Section 4: Parts of Speech (POS)



## Substitution Test

"If it is possible to **replace a sequence of words in a sentence with a different sequence of words** and the acceptability of the sentence remains unaffected, then this constitutes evidence for the fact that each sequence of words forms a constituent."

Müller (2019). Grammatical theory, p. 7.

## Example:

(9) he knows [the man] he knows [a woman]

## Problem:

(10) he [knows the] man he [sees a] man

Does this mean that "[knows the]" and "[sees a]" are constituents?

Section 1: Recap of Lecture 1

Section 2: Constituency

Section 3: Constituency Tests

Section 4: Parts of Speech (POS)



## **Pronominalization Test**

"Everything that can be **replaced by a pronoun** forms a constituent."

Müller (2019). Grammatical theory, p. 8.

- (11) Peter versucht, [das Buch zu lesen].Peter tries the book to read"Peter is trying to read the book."
- (12) Klaus versucht das auch.
   Klaus tries that also
   "Klaus is trying to do that as well."

Section 1: Recap of Lecture 1

Section 2: Constituency

Section 3: Constituency Tests

Section 4: Parts of Speech (POS)





## **Question Formation Test**

## "A sequence of words that can be elicited by a question forms a constituent."

Müller (2019). Grammatical theory, p. 9.

**Example:** Spanish (spa, Indo-European)

- (13) [La mujer] trabaj-a. the woman work-PRS.3SG "The woman works."
- (14) **Quién** trabaj-a? who work-PRS.3SG "Who works?"

Section 1: Recap of Lecture 1

Section 2: Constituency

Section 3: Constituency Tests

Section 4: Parts of Speech (POS)



## **Permutation Test**

"If a **sequence of words can be moved** without adversely affecting the acceptability of the sentence in which it occurs, then this is an indication that this word sequence forms a constituent."

Müller (2019). Grammatical theory, p. 9.

## Example:

- (15) [...] dass keiner [dieses Kind] kennt[...] that nobody this child knows
- (16) [...] dass [dieses Kind] keiner kennt[...] that this child nobody knows"[...] that nobody knows this child"

**Problem 1:** Look at the English glossing. The equivalent permutations in English are ungrammatical. Hence, this test only works if word order is flexible.

Section 1: Recap of Lecture 1

Section 2: Constituency

Section 3: Constituency Tests

Section 4: Parts of Speech (POS)





## Permutation Test

**Problem 2:** Latin (lat, Indo-European)

- nemo cognosc-it (17)hunc puer-um nobody know-PRS.3SG this.ACC.SG child-ACC.SG "Nobody knows this child."
- nemo hunc puerum cognoscit (18)hunc cognoscit nemo puerum hunc puerum cognoscit nemo etc.

**Question:** do languages that *freely permute words* not have constituents beyond the most basic level (i.e. words)?

Section 1: Recap of Lecture 1

Section 2: Constituency

Section 3: Constituency Tests

Section 4: Parts of Speech (POS)





## Fronting Test

"The possibility for a sequence of words to be **fronted** (that is to occur in front of the finite verb) is a strong indicator of constituent status."

Müller (2019). Grammatical theory, p. 10.

**Example:** Danish V2 (Verb Second) Order (dan, Indo-European)

- [...] at Bo har ikke læst [denne bog] (19) [...] that Bo has not read this book '[...] that Bo has not read this book.'
- [...] at [denne bog] har Bo ikke læst (20) book has Bo not read [...] that this '[...] that Bo has not read this book.'

Adopted from Thrainsson (2007). The syntax of Icelandic, p. 42.

Section 1: Recap of Lecture 1

Section 2: Constituency

Section 3: Constituency Tests

Section 4: Parts of Speech (POS)



## Coordination

"If two sequences of words can be **conjoined** then this suggests that each sequence forms a constituent."

Müller (2019). Grammatical theory, p. 10.

Example: French (fra, Indo-European)

(21) [la femme] et [l'homme] travaill-ent.the woman and the=man work-PRS.3PL"The woman and the man work."

Also works for whole phrases (e.g. infinitive constructions):

(22) Er hat versucht, [das Buch zu lesen] und [es dann unauffällig he has tried the book to read and it then secretly verschwinden zu lassen].

disappear to let

"He tried to read the book and then make it quietly disappear."

Section 1: Recap of Lecture 1

Section 2: Constituency

Section 3: Constituency Tests

Section 4: Parts of Speech (POS)



## **Problems with Constituency Tests**

"It would be ideal if the tests presented here delivered clear-cut results in every case, as the empirical basis on which syntactic theories are built would thereby become much clearer. Unfortunately, this is not the case. There are in fact a number of problems with constituent tests, [...]"

Müller (2019). Grammatical theory, p. 11.

Section 1: Recap of Lecture 1

Section 2: Constituency

Section 3: Constituency Tests

Section 4: Parts of Speech (POS)



## **Universality of Constituency (?)**

"Many discussions presume that constituency is an absolute universal, exhibited by all languages. But in fact constituency is just one method, used by a subset of languages, to express constructions which in other languages may be coded as dependencies of other kinds."

Evans & Levinson (2009), p. 440.

## Note:

- If we count "basic elements/units of a sentence" as constituents (see definition above), then constituency is trivially a universal.
- However, the criticism is leveled towards the universal existence of *higher level constituents*.

Section 1: Recap of Lecture 1

Section 2: Constituency

Section 3: Constituency Tests

Section 4: Parts of Speech (POS)



## **Universality of Constituency (?)**

Thalanyji (?, Pama-Nyungan(?))

(23) Kupuju-lu **kaparla-nha** yanga-lkin **wartirra-ku-nha** child-ERG dog-ACC chase-PRES woman-DAT-ACC "The child chases the woman's dog."

"Note how possessive modifiers – coded by a special use of the dative case – additionally pick up the case of the noun they modify, as with the accusative -nha on "dog" and "woman-Dat" [...] It is this **case-tagging**, rather than **grouping of words into constituents**, which forms the basic organizational principle in many Australian languages."

#### Evans & Levinson (2009), p. 441.

**Note however:** We don't know what the different constituent tests above would say about the constituency of *kaparla-nha wartirra-ku-nha*. This is only possible with a detailed knowledge of how the language is used.

Section 1: Recap of Lecture 1

Section 2: Constituency

Section 3: Constituency Tests

Section 4: Parts of Speech (POS)





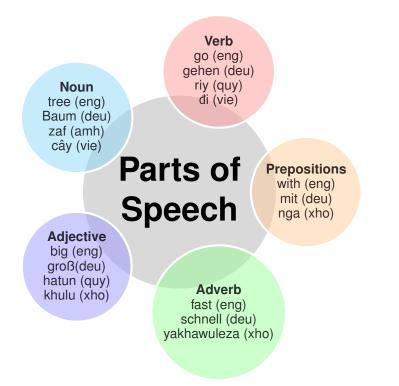
Faculty of Philosophy General Linguistics

## Section 4: Parts of Speech (POS)



## Definition

**Parts of Speech** are classes of words that each lexical item is assigned to according to its *morphosyntactic* properties. According to Müller (2019: 18) the basic POS are *Verb*, *Noun*, *Adjective*, *Adverb*, *Prepositions*.



Section 1: Recap of Lecture 1

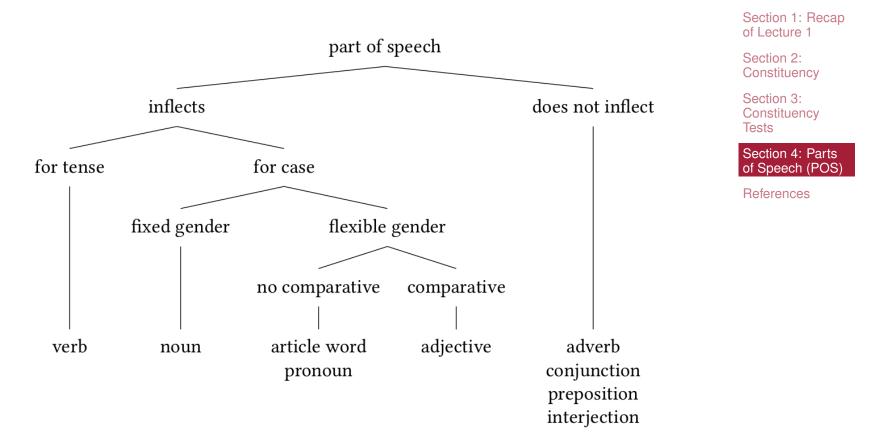
Section 2: Constituency

Section 3: Constituency Tests

Section 4: Parts of Speech (POS)



## **Decision Tree**



Müller (2019). Grammatical theory, p. 24. Based on Duden Grammar by Eisenberg et al. (2005).



## **Example: Determining POS**

- (24) Das ist ein schön-er Baum
   This is a beautiful-M.NOM.SG tree-M.NOM.SG
   "This is a beautiful tree."
- (25) Ich seh-e ein-en **schön-en** Baum I see-1SG DET-M.ACC.SG beautiful-M.ACC.SG tree "I see a beautiful tree."
- (26) Das ist ein-e **schön-e** Blume This is DET-F.NOM.SG beautiful-F.NOM.SG flower-F.NOM.SG "This is a beautiful flower."
- (27) Der Baum ist schön-er als die Blume The tree is beautiful-more than the flower"The tree is more beautiful than the flower."

#### **POS inference:**

"schön"  $\rightarrow$  inflects  $\rightarrow$  for case (i.e. (24) vs. (25))  $\rightarrow$  has flexible gender (i.e., (25) vs. (26))  $\rightarrow$  has a comparative form (i.e. (27))  $\rightarrow$  adjective **Beware:** nouns have fixed grammatical gender (e.g. Baum (M), Blume (F)), but additional morphology might reflect biological gender (e.g. Student (M), Student-in (F)).

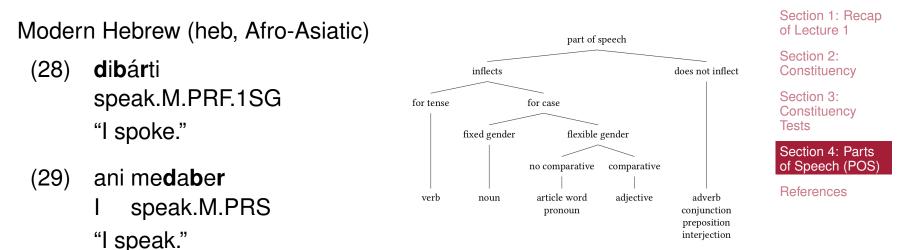
Section 2: Constituency

Section 3: Constituency Tests

Section 4: Parts of Speech (POS)



## **Example: Determining POS**



(30) a**d**aber speak.FUT.1SG "I will speak."

#### **POS inference:**

Words built on root template **dbr** (in this particular example)  $\rightarrow$  inflect  $\rightarrow$  for tense  $\rightarrow$  verb **Beware: dober** (m.)  $\rightarrow$  "speaker" (one who speaks)  $\rightarrow$  noun



## Universality of Word Classes (POS)

"Now it has often been assumed that, across all languages, the major classes – those that are essentially unlimited in their membership – will always be the same "big four": nouns, verbs, adjectives, and adverbs. But we now know that this is untenable when we consider the cross-linguistic evidence."

Evans & Levinson (2009), p. 434.

Section 1: Recap of Lecture 1

Section 2: Constituency

Section 3: Constituency Tests

Section 4: Parts of Speech (POS)



## Controversy: languages without adjectives?

"Are there adjectives in Mandarin Chinese? Most grammarians who dealt with the question of word classes or 'parts of speech' in Mandarin [...] tended to answer this question affirmatively. Their common practice was to apply the notional 'definitions' of word classes prevailing at that time in Western linguistics to the Chinese lexicon and to identify adjectives as words that 'denote properties'.

[...]

On the other hand, in more modern, structuralist-orientated approaches it has been widely agreed that those words which were traditionally called "adjectives" on semantic grounds, turned out to be, at most, a subclass of verbs, if their complex morphosyntactic behaviour was taken as a yardstick. It is this view which is taken by Chao (1968) and Li & Thompson (1981, 1990)."

Sackmann (1996). The problem of 'Adjectives' in Mandarin Chinese, p. 258.

Section 1: Recap of Lecture 1

Section 2: Constituency

Section 3: Constituency Tests

Section 4: Parts of Speech (POS)



## Controversy: languages without adjectives?

"It is widely agreed in linguistics that the supreme criterion for identifying verbs is *predicativity*. In Mandarin, 'adjectives' are able to form predicates in exactly the same way as verbs do. In predicative use, neither 'adjectives' nor verbs require, or even allow for, the use of a copula verb [...]"

Sackmann (1996), p. 261.

## Mandarin Chinese (cmn, Sino-Tibetan)

- (31) zhège xuéshēng niàn huàxué this student study chemistry"This student studies chemistry."
- (32) zhège xuéshēng [hěn] nǔlì this student [very] diligent"This student **is** diligent."

Section 1: Recap of Lecture 1

Section 2: Constituency

Section 3: Constituency Tests

Section 4: Parts of Speech (POS)



## Controversy: languages without adjectives?

"In Mandarin, both verbs and 'adjectives' can be marked for aspectual categories, either by aspectual suffixes like *-le* (perfective), *-guo* (experiential), and *-zhe* (durative), or by 'reduplication' ('delimitative'). (I tentatively adopt the position of regarding aspectual markers as (morphological) suffixes rather than (syntactic) particles. [...]"

Sackmann (1996), p. 262.

### Mandarin Chinese (cmn, Sino-Tibetan)

- (33) zhège xuéshēng shuì-le [...] this student sleep-PERF [...]
  "This student has slept [...]."
- (34) zhège xuéshēng nǔlì-le [...] this student diligent-PERF [...]
   "This student has been diligent [...]."

Section 1: Recap of Lecture 1

Section 2: Constituency

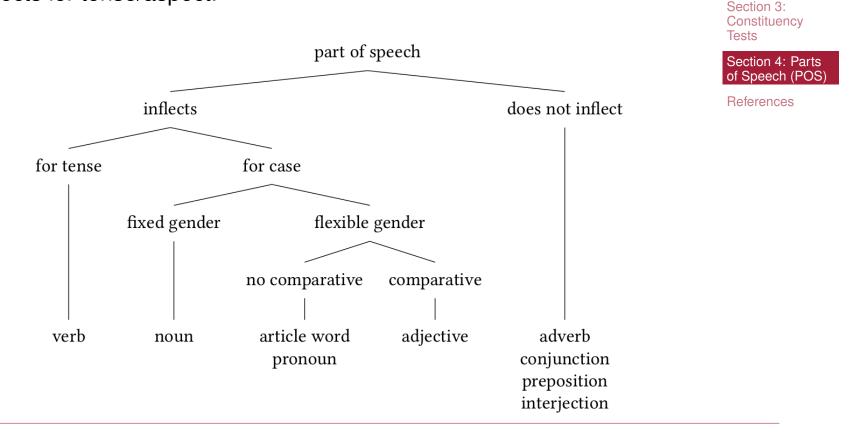
Section 3: Constituency Tests

Section 4: Parts of Speech (POS)



## Mandarin Chinese 'adjectives'

**Note:** If we accept -le as a suffix marking perfective aspect, then we would class *nŭlì-le* "diligient-PERF" as verb on the decision tree, since it inflects for tense/aspect.



Section 1: Recap

of Lecture 1

Section 2: Constituency



## Controversy: Noun-Verb distinction

"The question of whether there is a distinction between *noun* and *verb* as lexical categories in Salish has long been a lively issue. [...] All words are *predicates*: a root plus its internal arguments, if any."

Jelinek (1990), p. 177-179.

Straits Salish (str, Salishan)

- (35) sweyqe' cə t'ilem man-3A<sup>1</sup> DET sing
   "The (one who) is singing is a man."
- (36) t'ilem cə sweyqe' sing-3A DET man"The (one who) is a man is singing."

Section 1: Recap of Lecture 1

Section 2: Constituency

Section 3: Constituency Tests

Section 4: Parts of Speech (POS)

<sup>&</sup>lt;sup>1</sup>A: agent-like argument (according to Leipzig Glossing Rules).



## **Summary: Problems with POS**

- Problem 1: The number of basic POS can differ according to the framework any particular researcher adheres to (e.g. Interjection, Conjunction, etc. might be seen as additional POS).
- Problem 2: It is controversial whether all languages even have the basic POS mentioned above.
- Problem 3: The abbreviations used for POS can also differ across frameworks.
- Problem 4: Isolating languages have very little or no inflections. According the the Decision Tree all words in these languages would be in the class of adverbs, conjunctions, etc.

Section 1: Recap of Lecture 1

Section 2: Constituency

Section 3: Constituency Tests

Section 4: Parts of Speech (POS)





Faculty of Philosophy General Linguistics





## References

Eisenberg, Peter, Jörg Peters, Peter Gallmann, Cathrine Fabricius-Hansen, Damaris Nübling, Irmhild Barz, Thomas A. Fritz & Reinhard Fiehler. 2005. *Duden: Die Grammatik*. 7th edn. Vol. 4. Mannheim, Leipzig, Wien, Zürich: Dudenverlag.

Evans, Nicholas, and Levinson, Stephen C. (2009). The myth of language universals: language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*, 32, 429-492.

Jelinek, Eloise (1990). Quantification in Straits Salish. In: *International Conference on Salish and Neighboring Languages*, p. 177-195.

Haspelmath, Martin (2011). The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica*, 45/1, 31-80.

Müller, Stefan. 2019. *Grammatical theory: From transformational grammar to constraint-based approaches. Third revised and extended edition.* **Volume I**. Berlin: Language Science Press.

Sackmann, Robin (1996). The problem of 'adjectives' in Mandarin Chinese. In: Robin Sackmann and Monika Budde (Eds.), *Theoretical Linguistics and Grammatical Description*. Amsterdam/Philadelphia: John Benjamins.

Swan, O. E. (2002). *A grammar of contemporary Polish.* Bloomington: Slavish Publishers.

Thráinsson, H. (2007). *The syntax of Icelandic*. Cambridge: Cambridge University Press.

Section 1: Recap of Lecture 1

Section 2: Constituency

Section 3: Constituency Tests

Section 4: Parts of Speech (POS)



## Thank You.

#### **Contact:**

Faculty of Philosophy General Linguistics Dr. Christian Bentz SFS Wilhelmstraße 19-23, Room 1.24 chris@christianbentz.de Office hours: During term: Wednesdays 10-11am Out of term: arrange via e-mail