

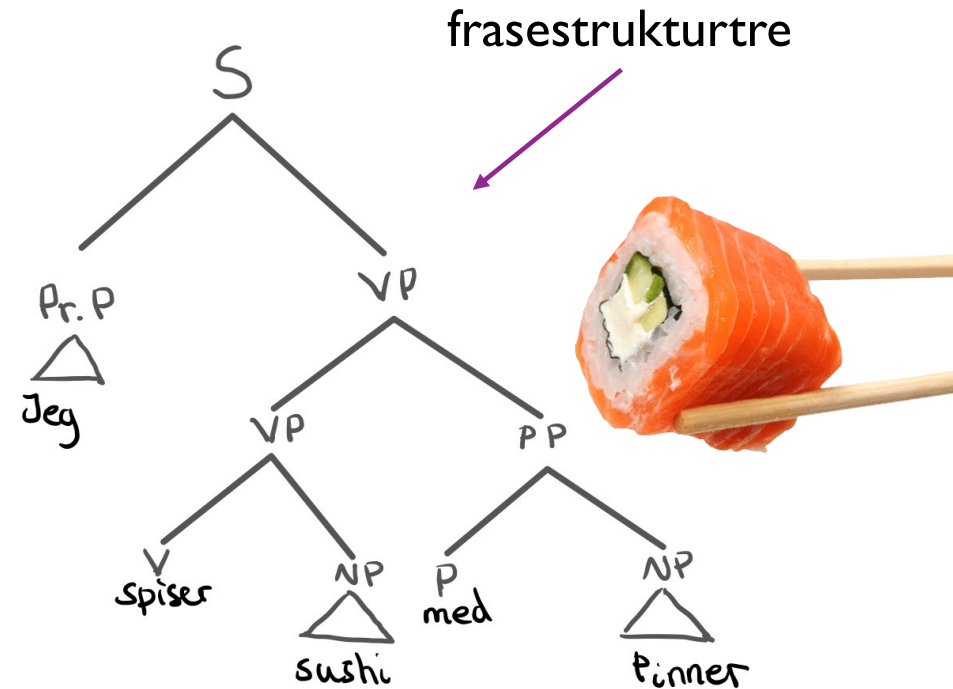
DEPENDENSSYNTAKS

SYNTAKS

Studiet av hvordan ord settes sammen til større enheter, f.eks. en setning eller en frase

Vi kan gruppere ord i *konstituent*er: bestanddel av en setning som fungerer som en enhet

- Jeg spiser sushi med pinner
- Sushi spiser jeg med pinner



SYNTAKTISK FORM

Konstituenten **beskrives med ordklasser og fraser**

- Fraser: større konstituenten over ordninvå
- Fraser navngitt av *hodet*- et sentralt, obligatorisk medlem
 - F.eks.VP(verbfrase), NP(substantivfrase)

	VP		
Pr. P	VP	PP	
Jeg	Spiser sushi	med pinner	

SYNTAKTISK FUNKSJON

Konstituenten **beskrives etter rollen deres i setningen** som

helhet

-Subjekt

-Objekt

-Adverbial/predikat

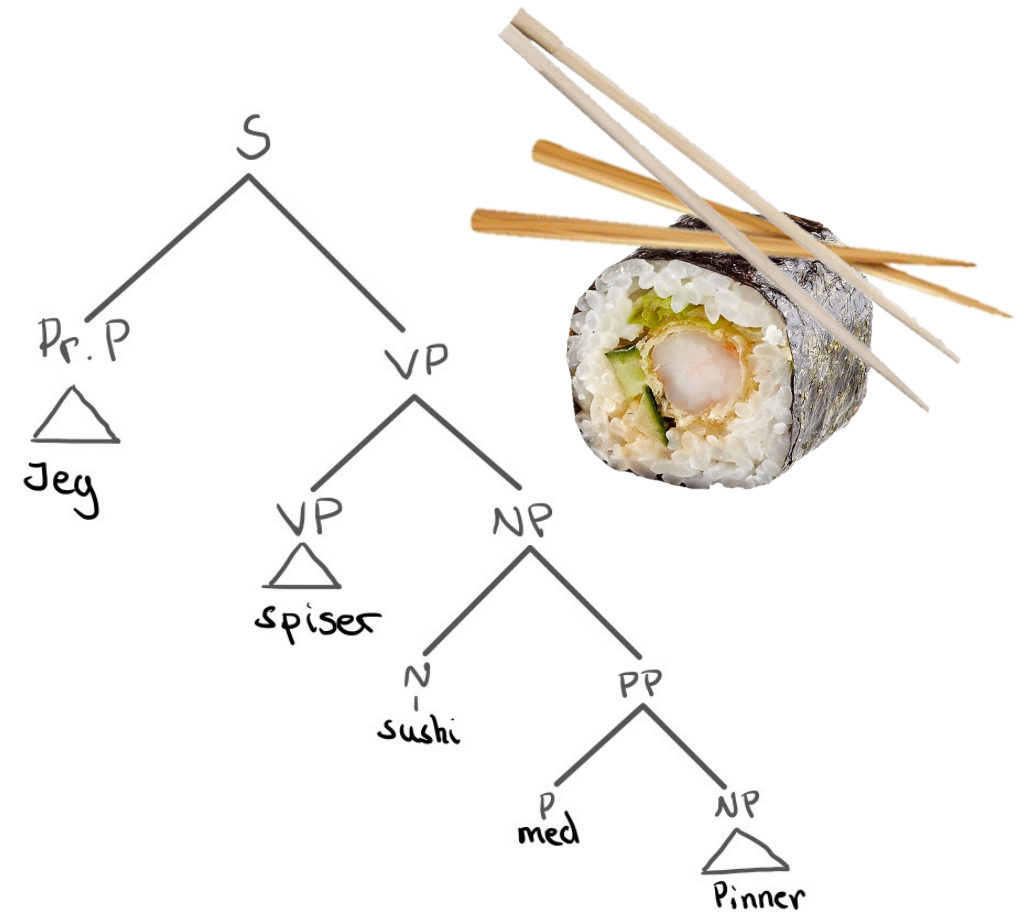
	Predikat		
Subjekt		Objekt	
Jeg	Spiser sushi	med pinner	

ALTERNATIV TOLKNING

- Jeg spiser sushi med pinner
- jeg spiser de
- Sushi med pinner spiser jeg

VP		
Pr. P	VP	NP
Jeg	spiser	sushi med pinner

Predikat		
Subjekt		Objekt
Jeg	spiser	sushi med pinner



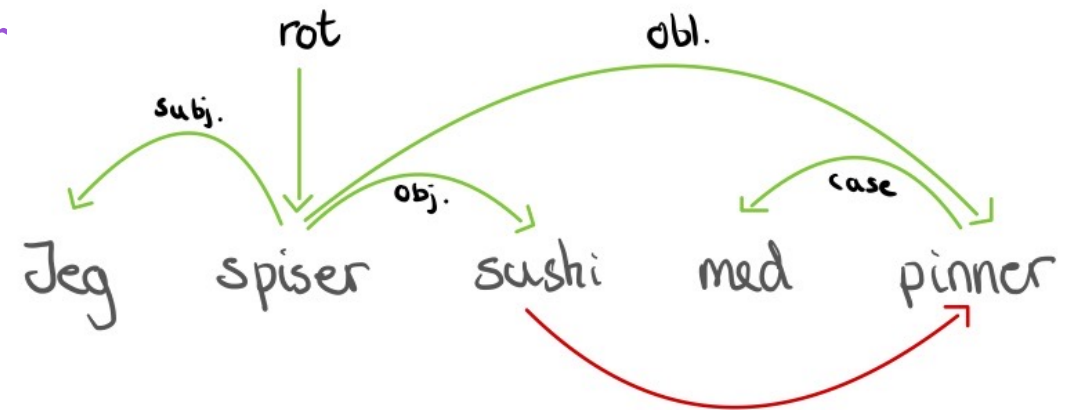
DEPENDENSSYNTAKS

En alternativ representasjon til frasestruktur-representasjoner

Består av leksikalske enheter som er koblet sammen med binære asymmetriske relasjoner- *dependenser*

Basert på *relasjoner* mellom ord, og *ikke* konstituentgrupperinger slik som i frasestrukturtær

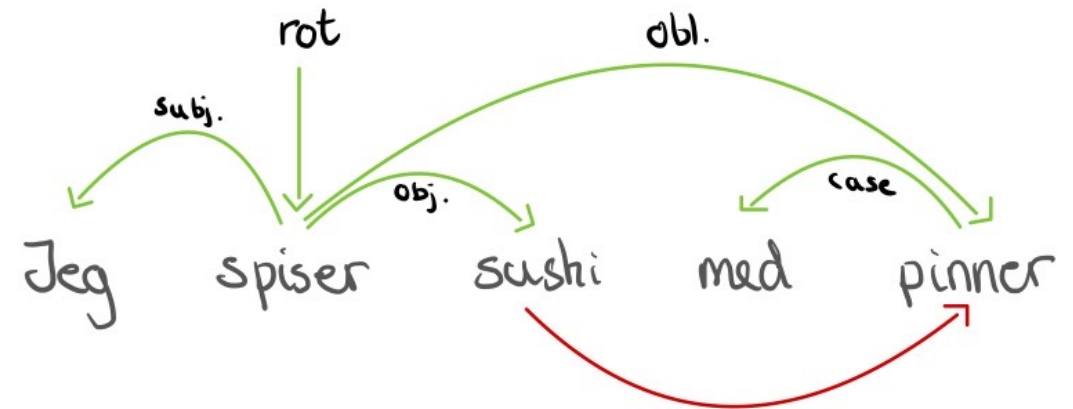
- $A \rightarrow B$: **A** styrer over **B**, og **B** avhenger av **A**
- $\text{spiser} \rightarrow \text{pinner}$: **spiser** styrer over **pinner**, og **pinner** avhenger av **spiste**
- $\text{spiser} \rightarrow \text{jeg}$: **spiser** styrer over **jeg**, og **jeg** avhenger av **spiser**
- Siden både subjekt og objekt avhenger av **spiser**, er **spiser**



KRITERIER FOR HODER OG DEPENDENSER

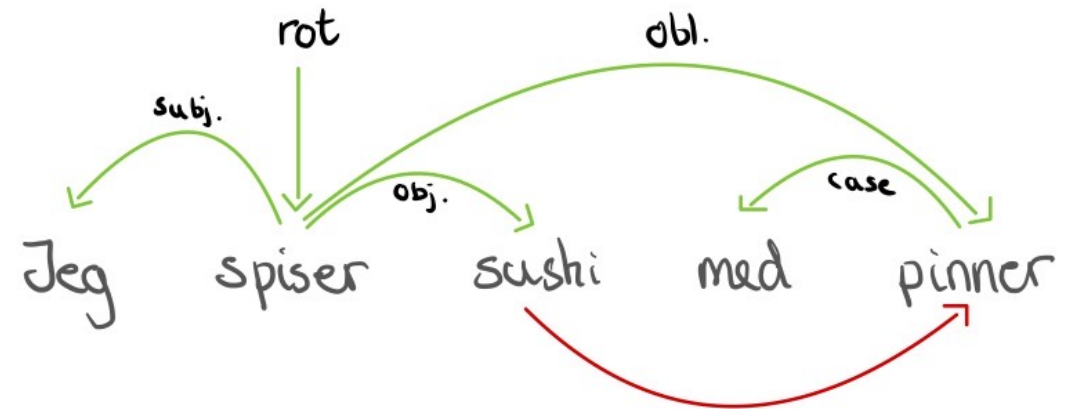
Hode H , dependens D , konstruksjon K

1. H avgjør den syntaktiske kategorien til K : H kan erstatte K
2. H avgjør den syntaktiske kategorien til K : D spesifiserer H
3. H er obligatorisk, D kan være valgfri
4. Formen av D avhenger av H (*agreement* eller *government*)
5. Den lineære posisjonen til D er definert med utgangspunkt i H



FORMELLE EGENSKAPER VED DEPENDENSGRAFER

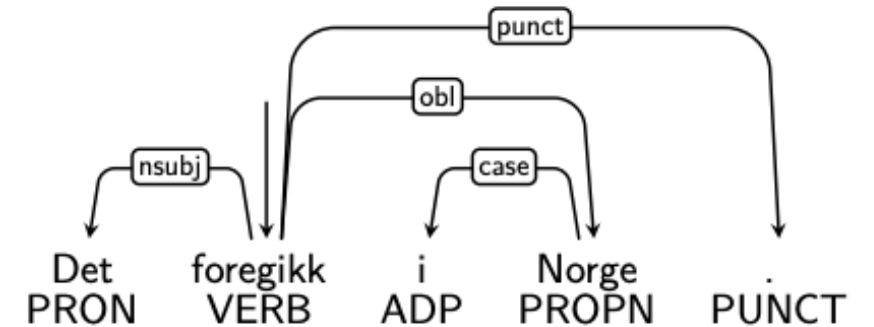
- Antisymmetriske – Hvis A styrer over B, styrer ikke B over A
- Irrefleksive – ingen ord styrer over seg selv
- Ikke-transitiv – dependensrelasjonene markerer direkte avhengighet
- «Labeled» - hver relasjon har en merkelapp (label)



TREBANKER

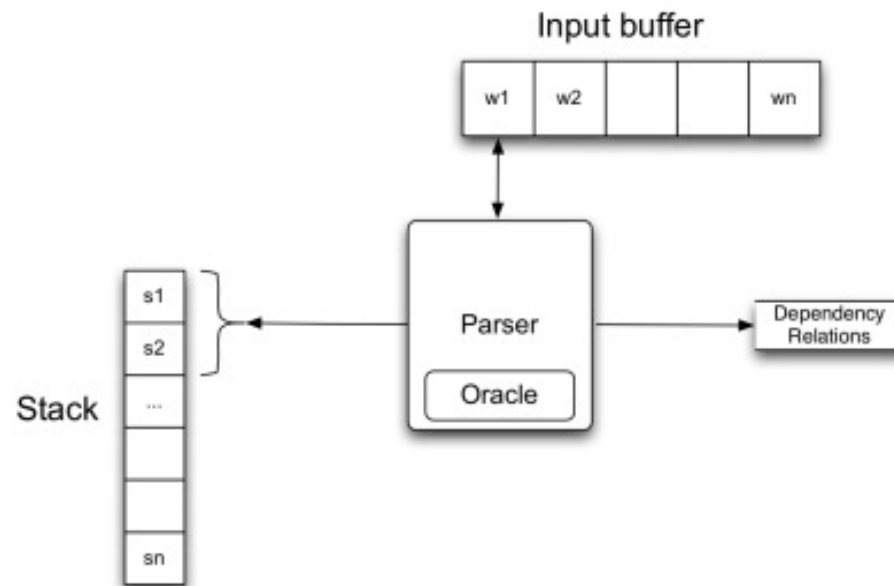
- En samling av setninger manuelt annotert med syntaktisk analyse
- Brukes til trening av datadrevne NLP-verktøy (tagging, parsing)
- CoNLL-U-format:

1	Det	det	PRON	Gender=Neut ...	2	nsubj
2	foregikk	foregå	VERB	Mood=Ind ...	0	root
3	i	i	ADP	—	4	case
4	Norge	Norge	PROPN	—	2	obl
5	.	.	PUNCT	—	2	punct



SYNTAKTISK PARSING

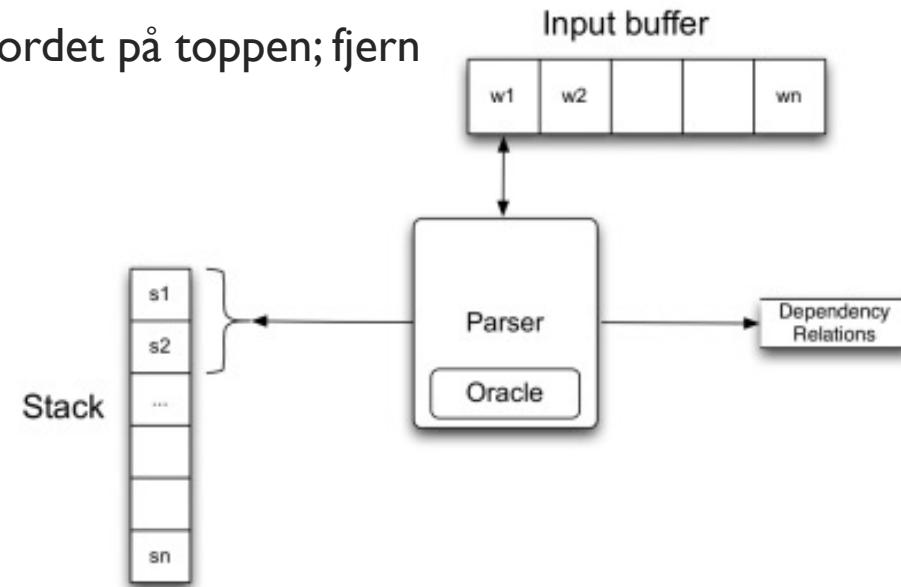
Automatisk avgjøre syntaktisk struktur for en setning.
Modellene kan være enten *grafbaserte*, eller *transisjonsbaserte*. Vi fokuserer på sistnevnte.



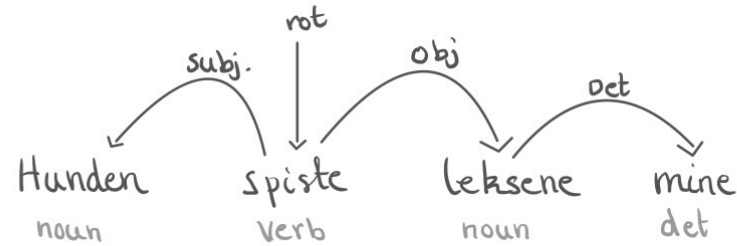
ARC STANDARD

TRANSISJONER

- SHIFT: flytte fra toppen av bufferen til toppen av stacken
- LEFT-ARC: relasjon fra ordet på toppen av stacken til det andre ordet; fjern det andre ordet fra stacken
- RIGHT-ARC: relasjon fra det andre ordet på stacken til ordet på toppen; fjern ordet fra toppen av stacken



ARC STANDARD



Stack	Ordliste	Transisjon	Relasjon
[rot]	[hunden, spiste, leksene, mine]	SHIFT	
[rot, hunden]	[spiste, leksene, mine]	SHIFT	
[rot, hunden, spiste]	[leksene, mine]	LEFT-ARC	(hunden \leftarrow spiste)
[rot, spiste]	[leksene, mine]	SHIFT	
[rot, spiste, leksene]	[mine]	SHIFT	
[rot, spiste, leksene, mine]	[]	RIGHT-ARC	(leksene \rightarrow mine)
[rot, spiste, leksene]	[]	RIGHT-ARC	(spiste \rightarrow leksene)
[rot, spiste]	[]	RIGHT-ARC	(rot \rightarrow spiste)
[rot]	[]		

ARC EAGER

- Alternativ transisjonsbasert modell
- Kan sette RIGHT-ARC-relasjoner mye tidligere enn arc standard
 - RIGHT-ARC i arc standard fjerner dependenten med en gang de er tildelt et hode
 - RIGHT-ARC i arc eager flytter dependenten til starten av stacken

TRANSISJONER

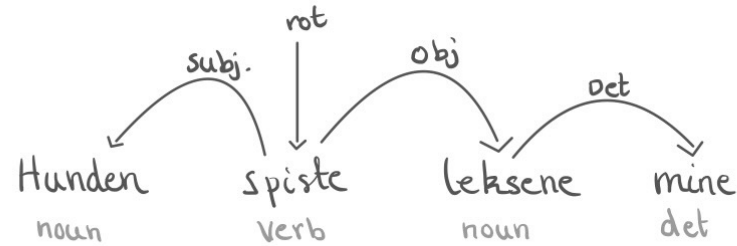
LEFTARC: tildel en relasjon fra ordet på starten av bufferen til ordet på starten av stacken; **reduce**

RIGHTARC: tildel en relasjon fra ordet på toppen av stacken til ordet på toppen av bufferen; **shift** ordet på toppen av bufferen til toppen av stacken

SHIFT: Fjern ordet fra toppen av input bufferen til stacken

REDUCE: pop stacken

ARC EAGER



Stack	Ordliste	Transisjon	Relasjon
[rot]	[hunden, spiste, leksene, mine]	SHIFT	
[rot, hunden]	[spiste, leksene, mine]	LEFT-ARC	(hunden \leftarrow spiste)
[rot]	[spiste, leksene, mine]	RIGHT-ARC	(rot \rightarrow spiste)
[rot, spiste]	[leksene, mine]	RIGHT-ARC	(spiste \rightarrow leksene)
[rot, spiste, leksene]	[mine]	RIGHT-ARC	(leksene \rightarrow mine)
[rot, spiste, leksene, mine]	[]	REDUCE	
[rot, spiste, leksene]	[]	REDUCE	
[rot, spiste]	[]	REDUCE	
[rot]	[]	FERDIG	