



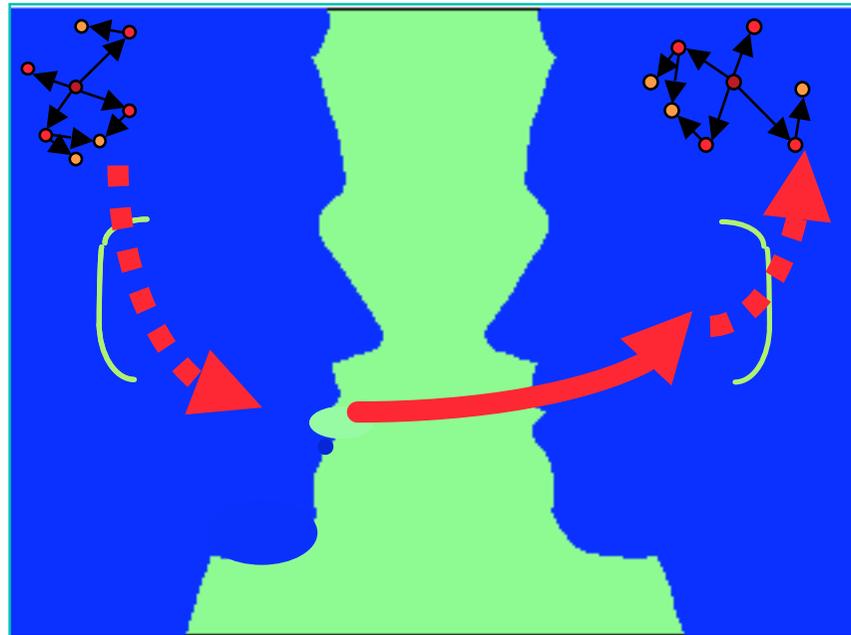
Vorlesung Grammatikformalismen:

Die Rolle der Grammatik

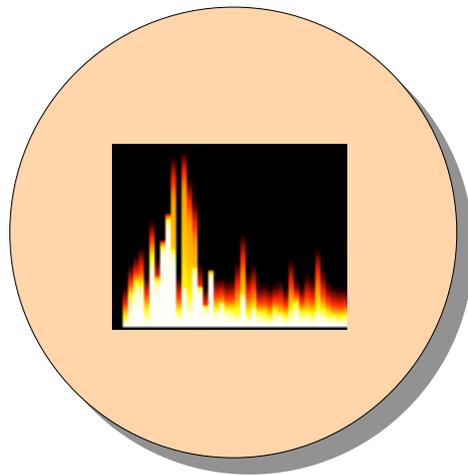
Ersetzungsregelgrammatiken

Hans Uszkoreit

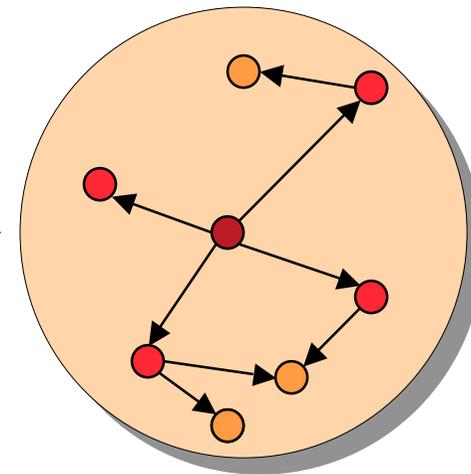
WHAT HAPPENS IN BETWEEN?



WHAT HAPPENS IN BETWEEN?

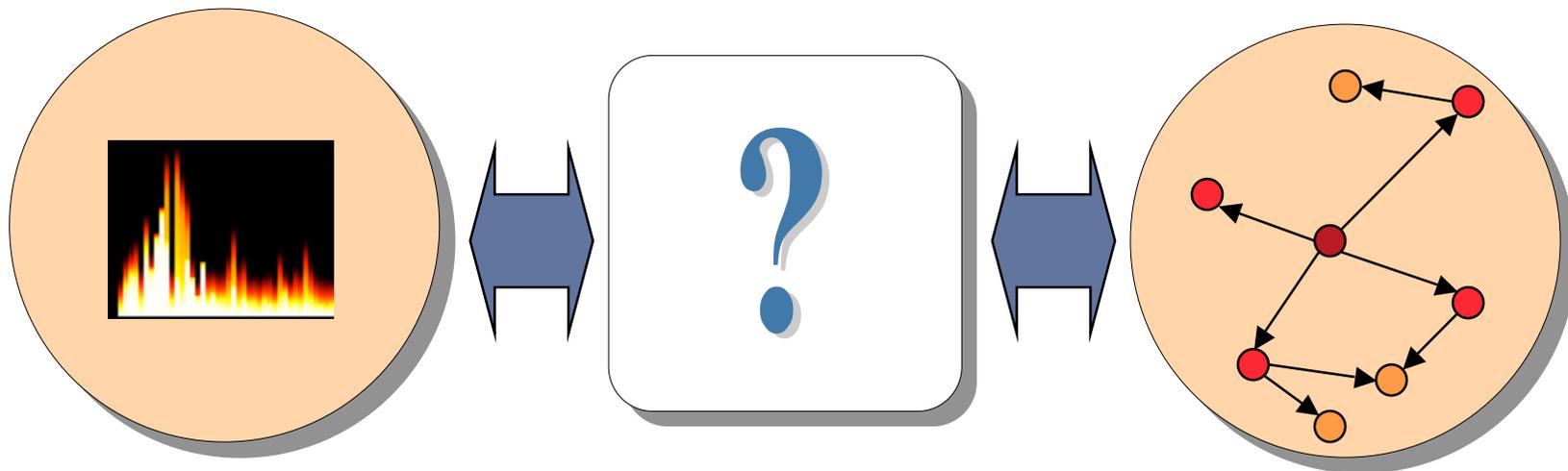


sound waves



activation of concepts

WHAT HAPPENS IN BETWEEN?

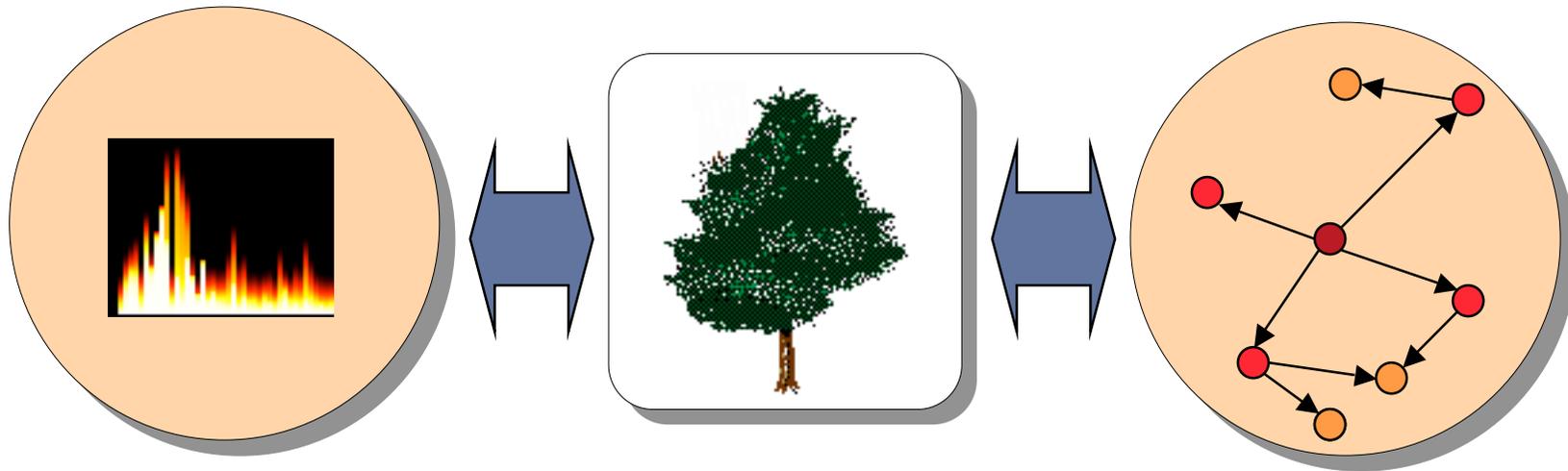


sound waves

Grammar

activation of concepts

WHAT HAPPENS IN BETWEEN?

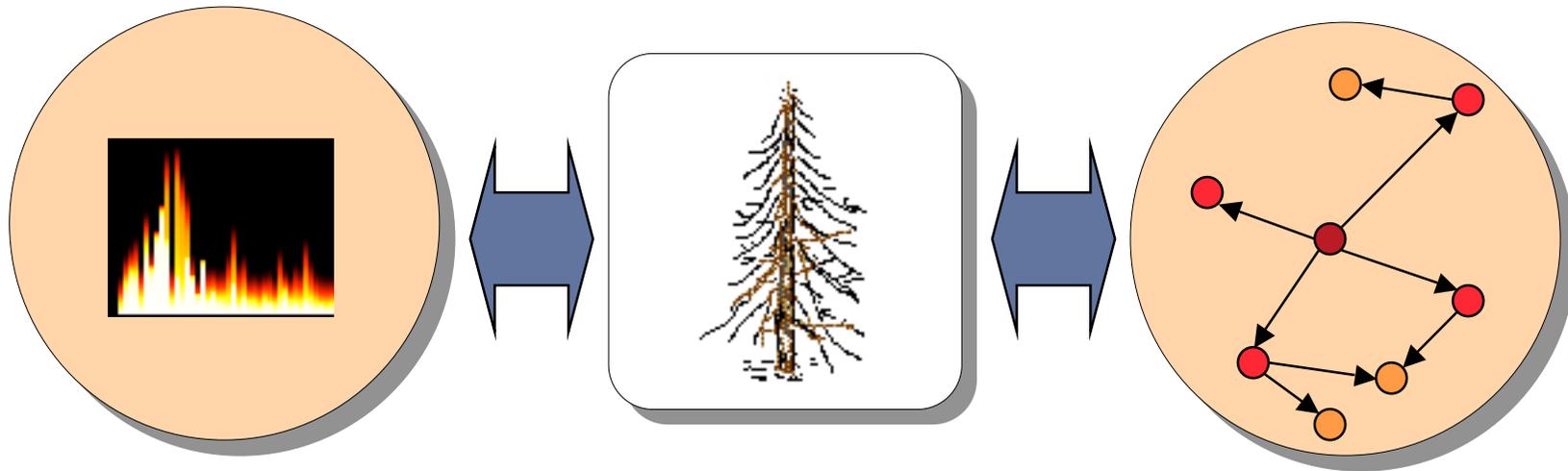


sound waves

Grammar

activation of concepts

WHAT HAPPENS IN BETWEEN?

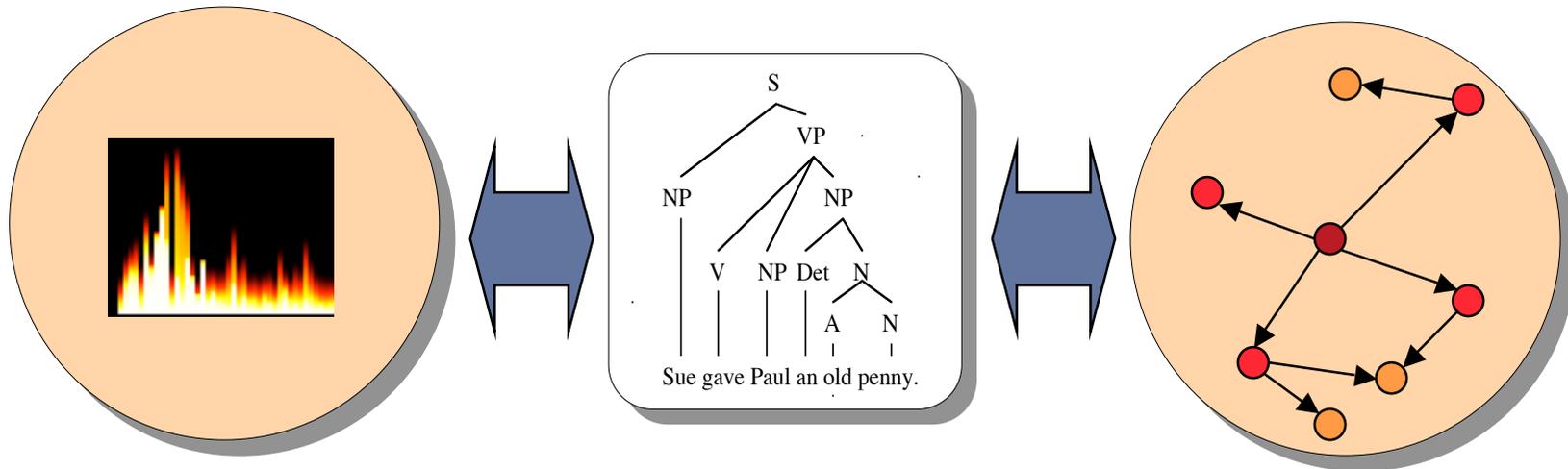


sound waves

Grammar

activation of concepts

WHAT HAPPENS IN BETWEEN?

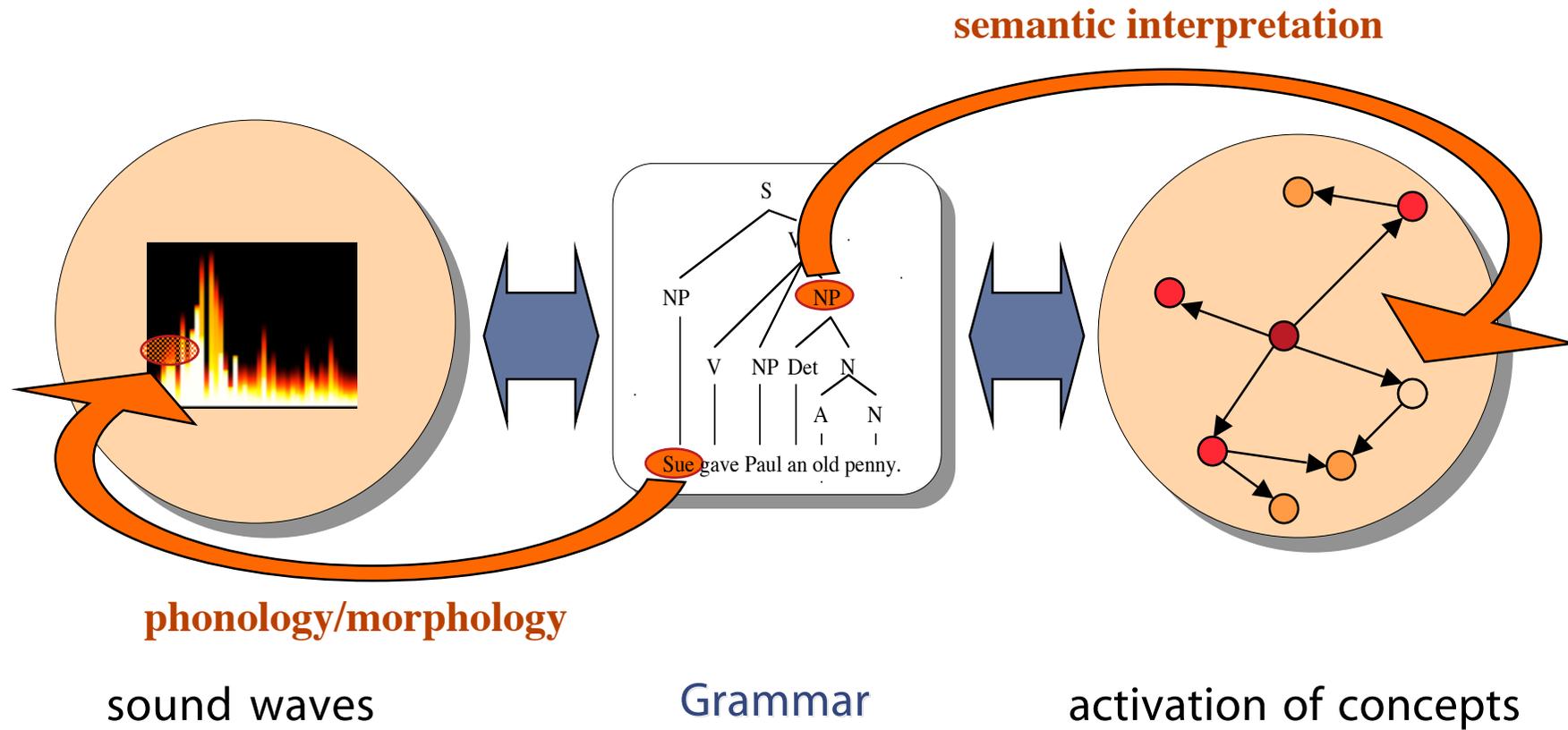


sound waves

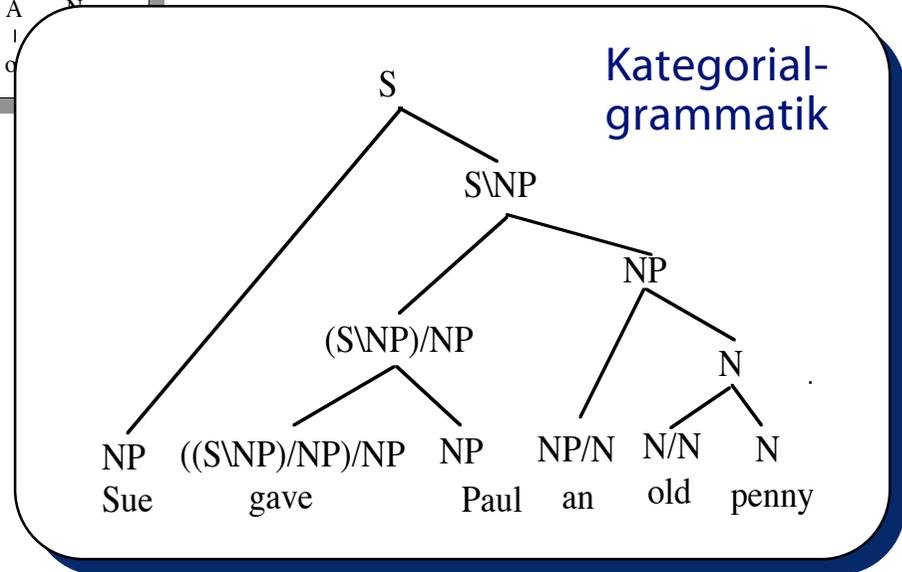
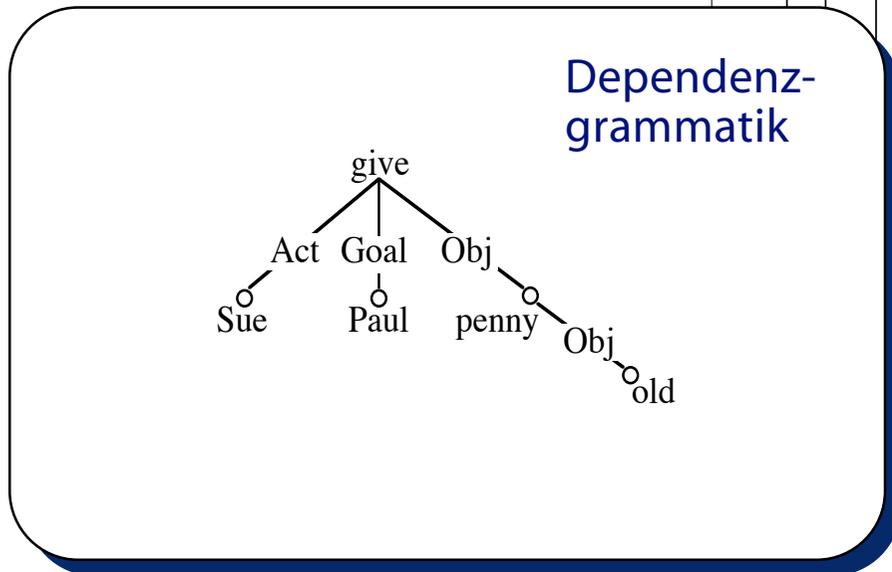
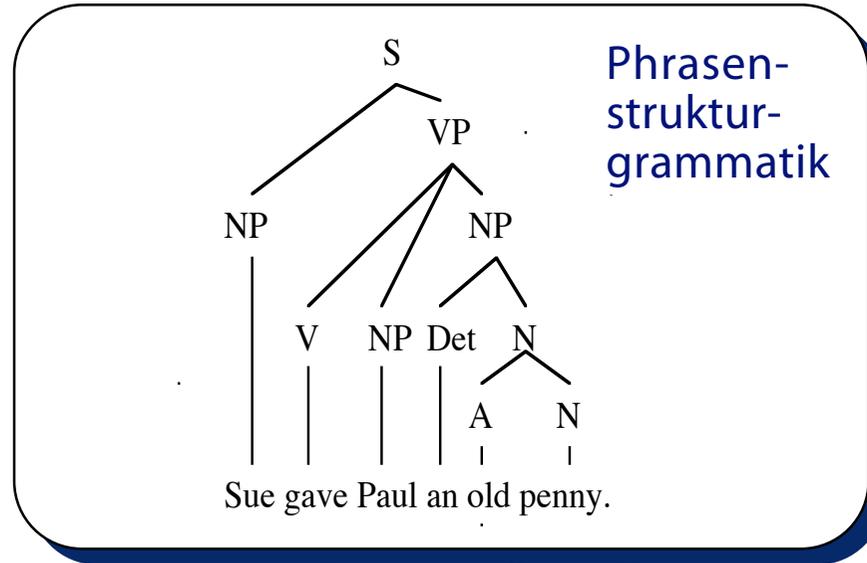
Grammar

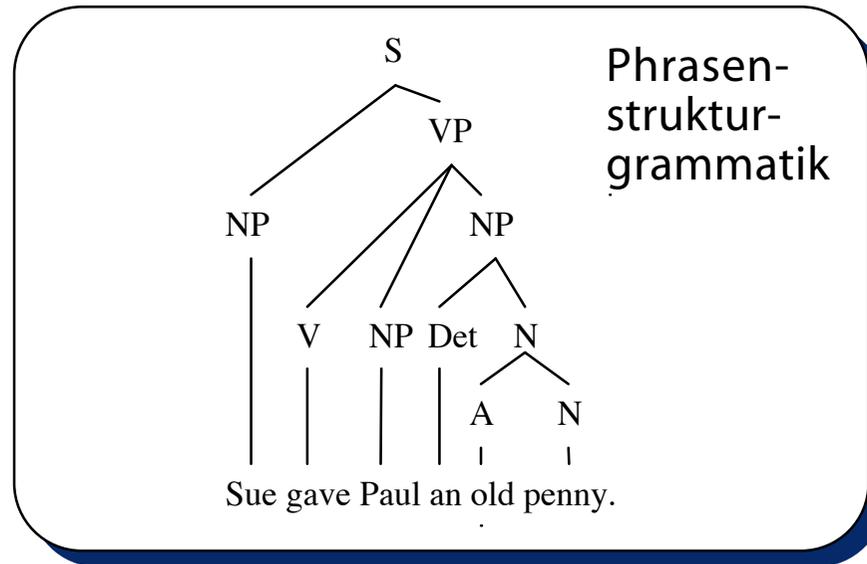
activation of concepts

WHAT HAPPENS IN BETWEEN?

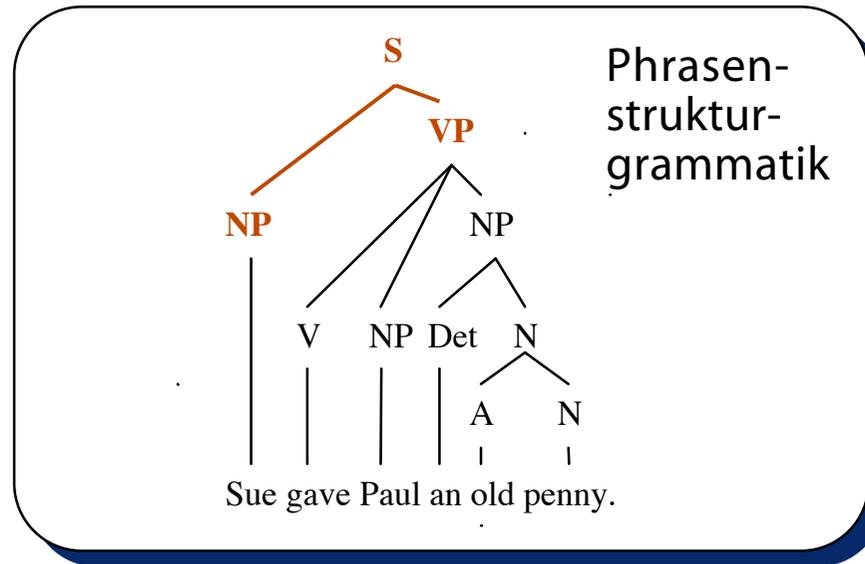


THREE TRADITIONS

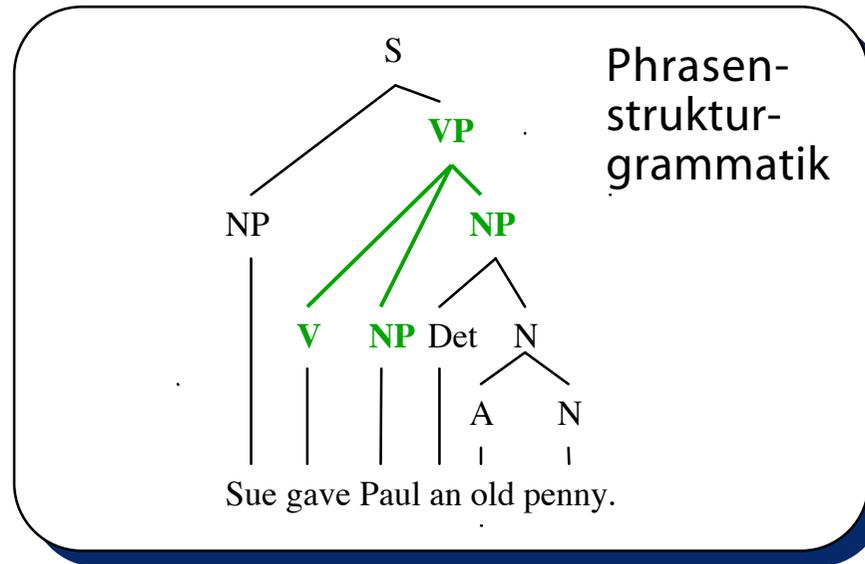




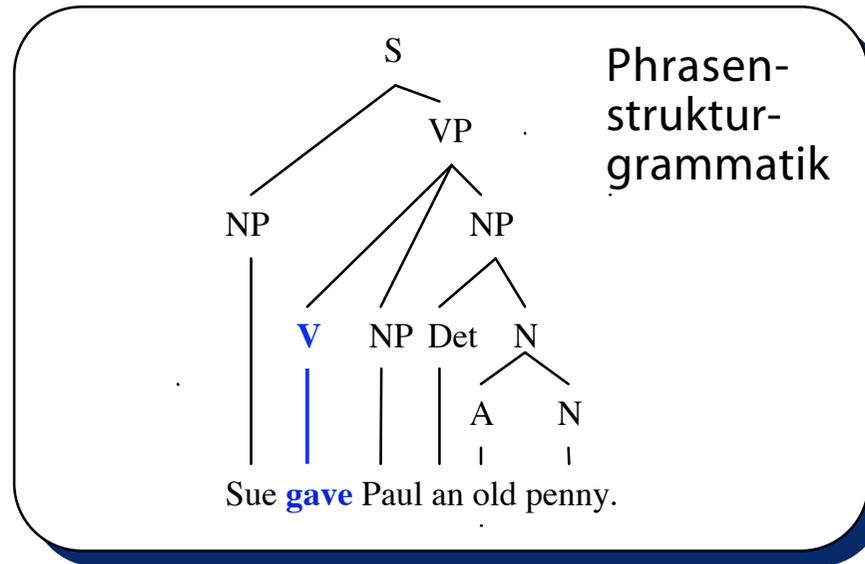
S → NP VP



S → NP VP



S \square NP VP
VP \square V NP NP



S \square NP VP
VP \square V NP NP

V \square gave



	Einzelsymbole	Ketten
nichtterminale	A, B, C, , X, Y, Z
terminale	a, b, c, , x, y, z
unspezifiziert	\square , \square , \square , \square , \square , \square , \square
Startsymbol	S	
die leere Kette	ϵ	
Zahlen	..., i, j, k, l, m, n, ...	



Eine Sprache L über einem Alphabet (Vokabular) Σ ist eine Teilmenge von Σ^* .

Eine formale Grammatik G_L für eine Sprache L ist ein
Quadrupel $(V_N, V_T, \{S\}, P)$.

V_N - nichtterminales Vokabular, Hilfsalphabet

V_T - terminales Vokabular

$$(V_T \cap V_N = \emptyset, \quad L \subseteq V_T^*, \quad V = V_T \cup V_N)$$

$\{S\}$ - Einermenge mit dem Startsymbol, Axiomenmenge

P - Menge der Produktionen, Regelmenge

Menge von Regeln der Form $\alpha_1 \alpha_2 \rightarrow \beta_1 \beta_2$

meist geschrieben als $\alpha \rightarrow \beta$



Relation "folgt":

Wenn $G = (V_N, V_T, \{S\}, P)$, dann folgt α aus β nach G genau dann, wenn es Ketten $\alpha_1, \alpha_2, \alpha, \beta$ gibt, so daß $\alpha = \alpha_1 \alpha_2$ und $\beta = \alpha_1 \alpha_2$ und $\alpha_2 \alpha_2 \in P$.

Notation: $\alpha \xrightarrow[G]{} \beta$

Ableitung:

Eine Folge von Ketten $\alpha_1, \alpha_1, \dots, \alpha_n$ ist eine Ableitung nach G g.d.w. $\alpha_i \xrightarrow[G]{} \alpha_{i+1}$ für alle $i, 1 \leq i \leq n$.

Wenn es eine Ableitung nach G von α nach β gibt, können wir das

so schreiben: $\alpha \xrightarrow[G]^* \beta$

Die Relation ableitbar ist transitiv und wird als reflexiv definiert.



Die Sprache L: Eine Kette α ist nach G_L in der Sprache L g.d.w. die folgenden drei Bedingungen erfüllt sind:

1. $\alpha \in V_T^*$

2. $S \stackrel{*}{\Rightarrow}_G \alpha$

3. Es gibt kein β , so daß $\alpha \stackrel{*}{\Rightarrow}_G \beta$ und $\alpha \neq \beta$.

Man sagt auch G_L erzeugt die Sprache L. Die von G erzeugte Sprache L wird auch als $L(G)$ geschrieben.

Schwache Äquivalenz: Zwei Grammatiken G_1 und G_2 sind schwach äquivalent, wenn sie dieselbe Sprache erzeugen.



Typ 0 (unbeschränkte Ersetzungsregelsysteme):

Jede Grammatik, die die Definition einer formalen Grammatik erfüllt, ist vom Typ 0.

Typ 1 (kontextsensitive Grammatiken):

Jede Produktion hat die Form $\alpha A \beta \rightarrow \gamma \delta$, wobei $A \in V_T$, $\alpha \neq \epsilon$.

Typ 2 (kontextfreie Grammatiken):

Jede Produktion hat die Form $A \rightarrow \alpha$, wobei $A \in V_T$.

Typ 3 (reguläre Grammatiken):

Jede Produktion hat die Form $A \rightarrow x B$ oder $A \rightarrow x$, wobei $x \in \Sigma$.



S
NP VP
DET ADJ N VP
DET ADJ N V NP
DET ADJ N V DET ADJ N
.
.
.

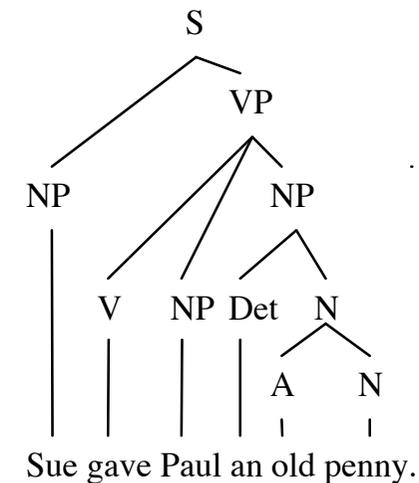
ein kleines Mädchen sucht ein kleines Mädchen



der Begriff des syntaktischen Strukturbaums

kodierte Information

1. Die hierarchische Organisation der Teile eines Satzes in Konstituenten.
2. Der Zugehörigkeit jeder Konstituente in eine syntaktische (linguistische) Klasse (Kategorie).
3. Die lineare Abfolge der Konstituenten.





Relationen: unmittelbare Dominanz - Dominanz
unmittelbare Präzedenz - Präzedenz

Konstituentenstrukturbaum: Quintupel (N, Q, D, P, L)

N - endliche Menge von Knoten

Q - endliche Menge von Etiketten

D - schwache Teilordnung in $N \times N$, die Dominanzrelation
(reflexiv, transitiv und antisymmetrisch)

P - starke Teilordnung in $N \times N$, die Präzedenzrelation
(irreflexiv, transitiv und asymmetrisch)

L - Funktion von N in Q, die Etikettierfunktion

Bedingungen:

Wurzelbedingung

Exklusivitätsbedingung

Kreuzungsfreiheit



Wurzelbedingung

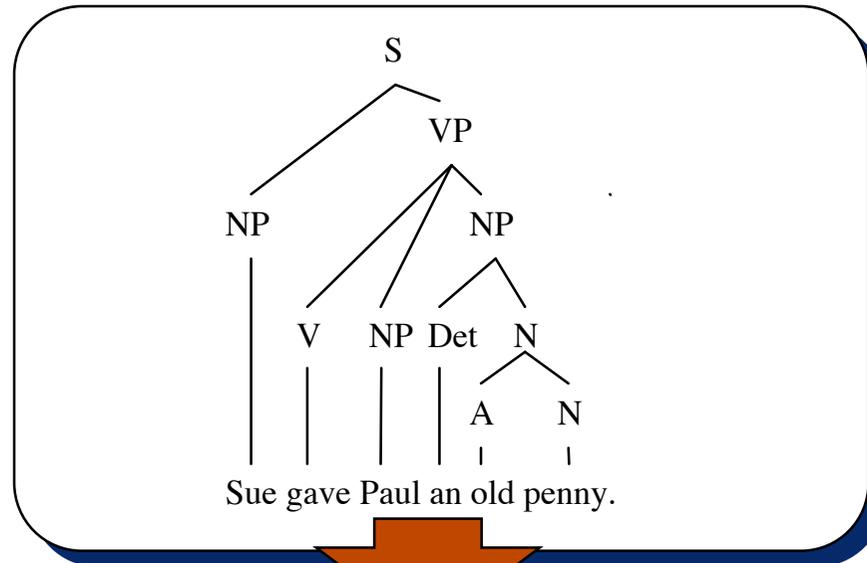
Es gibt genau einen Knoten, der alle Knoten des Baumes dominiert.

Exklusivitätsbedingung

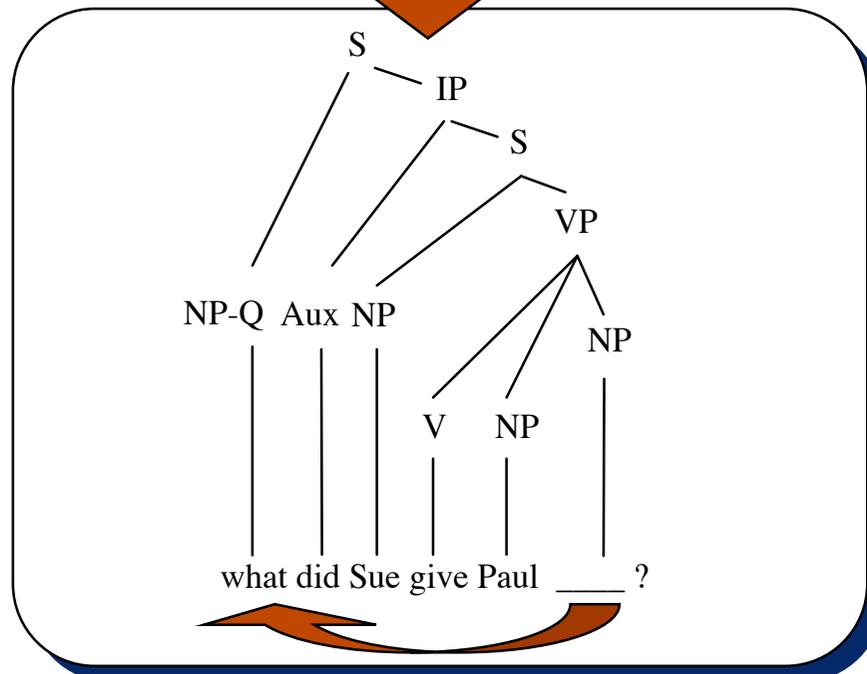
Für jegliche zwei Knoten x und y gilt, entweder $D(x,y)$ oder $D(y,x)$
oder $P(x,y)$ oder $P(y,x)$

Kreuzungsfreiheit

Wenn $P(x,y)$ dann gilt für alle x' , die von x dominiert werden $D(x,x')$ und für alle y' , die von y dominiert werden $D(y,y')$ dass auch $x' y'$ vorausgeht $[P(x',y')]$



Transformations-
grammatik





Unifikationsgrammatik

