Tree Parsing with Synchronous Tree-Adjoining Grammars

Matthias Büchse Mark-Jan Nederhof Heiko Vogler

International Conference on Parsing Technologies

October 2011

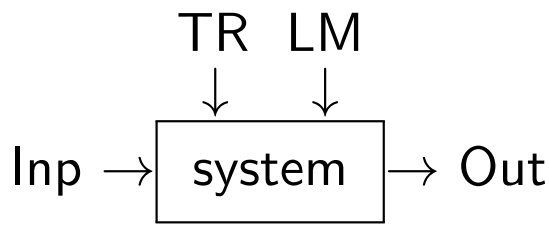
Translation System Example

$\mathsf{TR}\colon \mathsf{Tree} \times \mathsf{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ translation probabilities (French, English)

$\mathsf{LM}\colon \mathsf{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ probabilities for English parse trees

$\mathsf{Inp}\colon \mathsf{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ input weighted tree language

$\mathsf{Out}\colon \mathsf{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ output weighted tree language



$$\mathsf{Out}(e) = \sum_{f \in \mathsf{Tree}} \mathsf{Inp}(f) \cdot \mathsf{TR}(f, e) \cdot \mathsf{LM}(e)$$

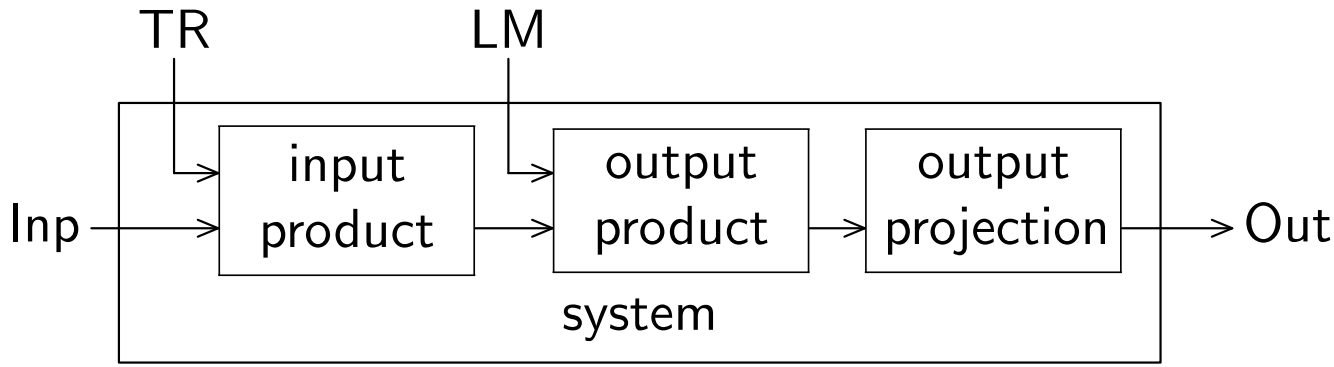
Translation System Example

$\mathrm{TR}\colon \mathrm{Tree} \times \mathrm{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ translation probabilities (French, English)

$\mathrm{LM}\colon \mathrm{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ probabilities for English parse trees

$\mathrm{Inp}\colon \mathrm{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ input weighted tree language

$\mathrm{Out}\colon \mathrm{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ output weighted tree language



$$\mathrm{Out}(e) = \sum_{f \in \mathrm{Tree}} \mathrm{Inp}(f) \cdot \mathrm{TR}(f, e) \cdot \mathrm{LM}(e)$$

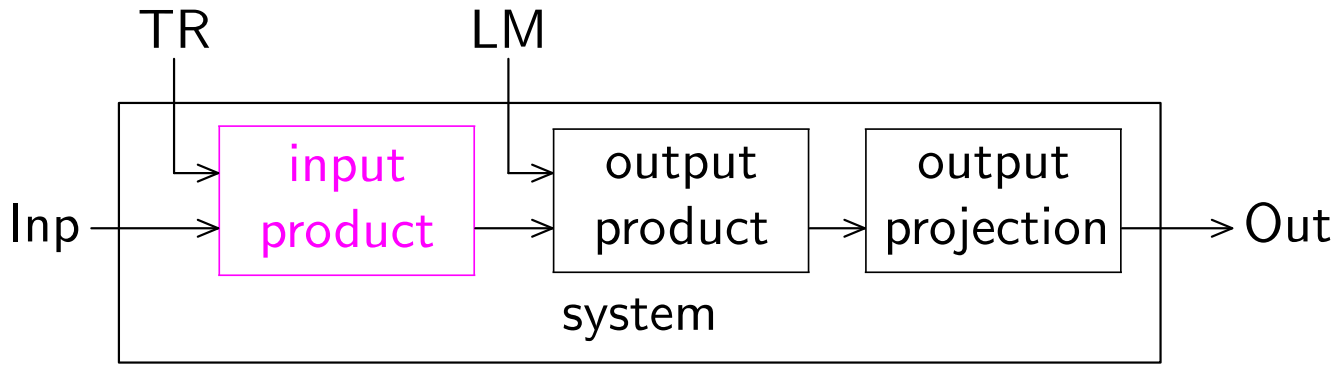
Translation System Example

$\mathsf{TR}\colon \mathsf{Tree} \times \mathsf{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ translation probabilities (French, English)

$\mathsf{LM}\colon \mathsf{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ probabilities for English parse trees

$\mathsf{Inp}\colon \mathsf{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ input weighted tree language

$\mathsf{Out}\colon \mathsf{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ output weighted tree language



$$\mathsf{Out}(e) = \sum_{f \in \mathsf{Tree}} \mathsf{Inp}(f) \cdot \mathsf{TR}(f, e) \cdot \mathsf{LM}(e)$$

$$= \sum_{f \in \mathsf{Tree}} (\mathsf{Inp} \lhd \mathsf{TR})(f, e) \cdot \mathsf{LM}(e)$$

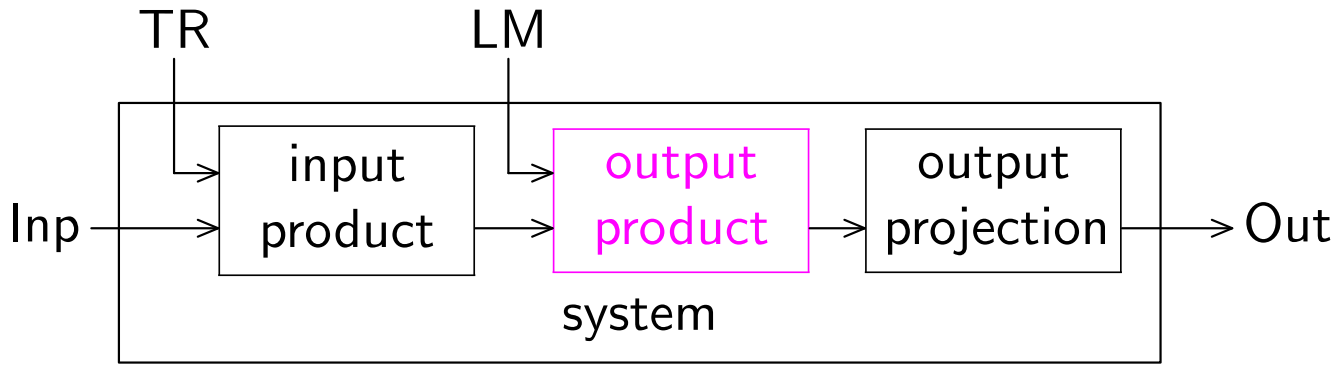
Translation System Example

$\mathrm{TR}\colon \mathrm{Tree} \times \mathrm{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ translation probabilities (French, English)

$\mathrm{LM}\colon \mathrm{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ probabilities for English parse trees

$\mathrm{Inp}\colon \mathrm{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ input weighted tree language

$\mathrm{Out}\colon \mathrm{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ output weighted tree language



$$\mathrm{Out}(e) = \sum_{f \in \mathrm{Tree}} \mathrm{Inp}(f) \cdot \mathrm{TR}(f, e) \cdot \mathrm{LM}(e)$$

$$= \sum_{f \in \mathrm{Tree}} ((\mathrm{Inp} \lhd \mathrm{TR}) \rhd \mathrm{LM})(f, e)$$

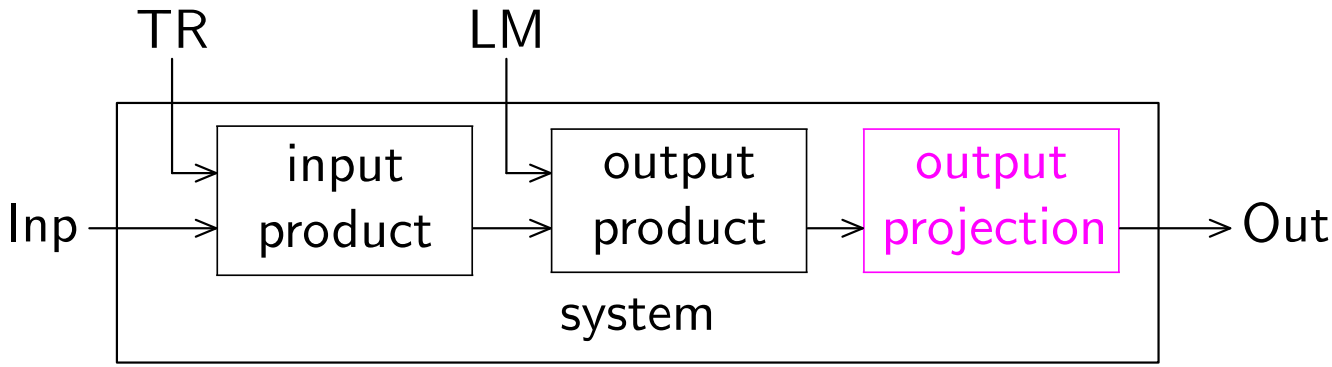
Translation System Example

$\mathsf{TR}\colon \mathsf{Tree} \times \mathsf{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ translation probabilities (French, English)

$\mathsf{LM}\colon \mathsf{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ probabilities for English parse trees

$\mathsf{Inp}\colon \mathsf{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ input weighted tree language

$\mathsf{Out}\colon \mathsf{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ output weighted tree language



$$\mathsf{Out}(e) = \sum_{f \in \mathsf{Tree}} \mathsf{Inp}(f) \cdot \mathsf{TR}(f, e) \cdot \mathsf{LM}(e)$$

$$= \mathsf{Proj}((\mathsf{Inp} \lhd \mathsf{TR}) \rhd \mathsf{LM})(e)$$

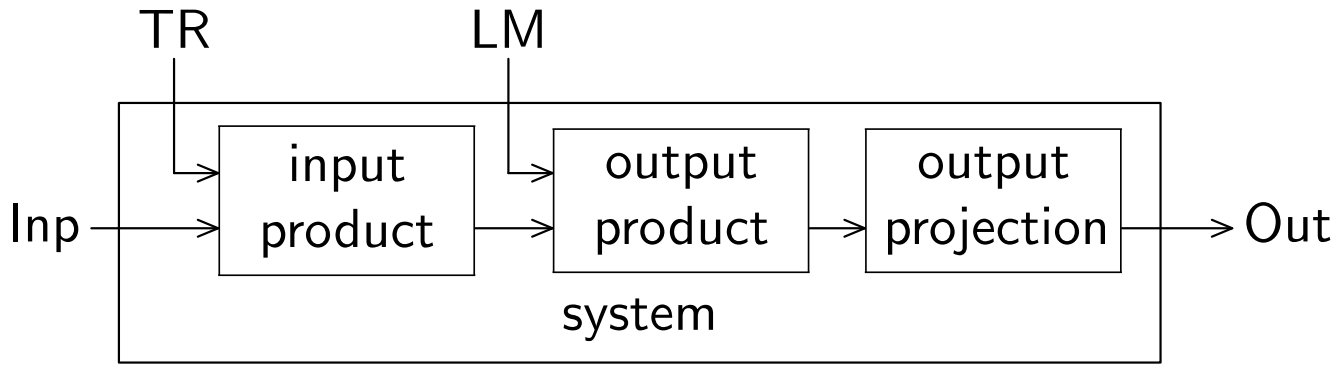
Translation System Example

$\mathrm{TR}\colon \mathrm{Tree} \times \mathrm{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ translation probabilities (French, English)

$\mathrm{LM}\colon \mathrm{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ probabilities for English parse trees

$\mathrm{Inp}\colon \mathrm{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ input weighted tree language

$\mathrm{Out}\colon \mathrm{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ output weighted tree language



$$\mathrm{Out} = \mathrm{Proj}((\mathrm{Inp} \lhd \mathrm{TR}) \rhd \mathrm{LM})$$

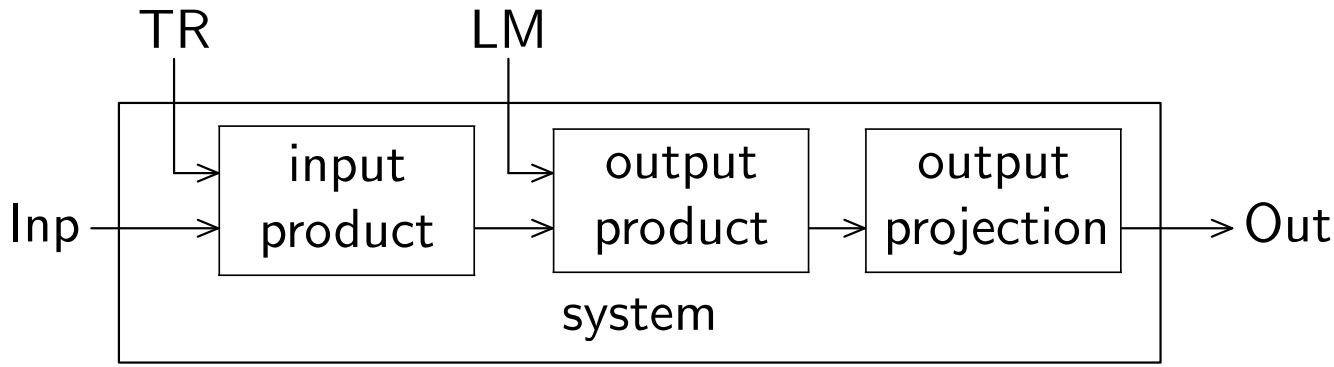
Translation System Example

$\mathrm{TR}\colon \mathrm{Tree} \times \mathrm{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ synchronous tree-adjoining grammar

$\mathrm{LM}\colon \mathrm{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ regular tree grammar

$\mathrm{Inp}\colon \mathrm{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ regular tree grammar

$\mathrm{Out}\colon \mathrm{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$???



$$\mathrm{Out} = \mathrm{Proj}((\mathrm{Inp} \lhd \mathrm{TR}) \rhd \mathrm{LM})$$

Finitely representable?

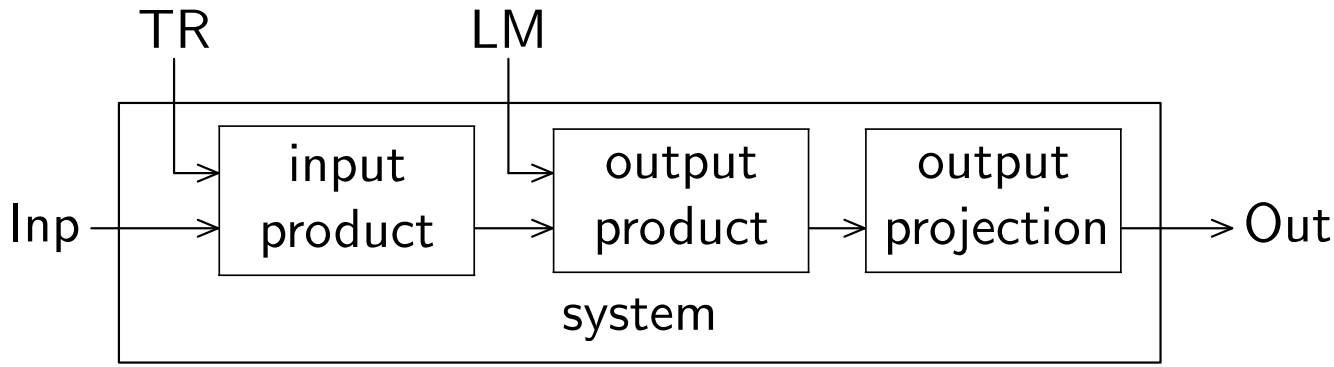
Translation System Example

$\mathrm{TR}\colon \mathrm{Tree} \times \mathrm{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ synchronous tree-adjoining grammar

$\mathrm{LM}\colon \mathrm{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ regular tree grammar

$\mathrm{Inp}\colon \mathrm{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ regular tree grammar

$\mathrm{Out}\colon \mathrm{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$???



$$\mathrm{Out} = \mathrm{Proj}((\mathrm{Inp} \lhd \mathrm{TR}) \rhd \mathrm{LM})$$

Finitely representable?

$\rightsquigarrow$ Tree Parsing

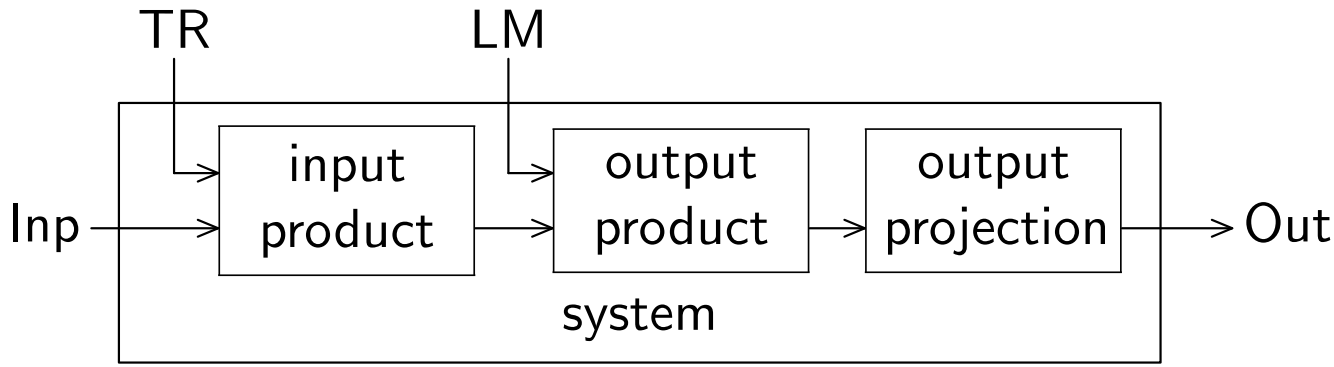
Translation System Example

$\mathrm{TR}\colon \mathrm{Tree} \times \mathrm{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ synchronous tree-adjoining grammar

$\mathrm{LM}\colon \mathrm{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ regular tree grammar

$\mathrm{Inp}\colon \mathrm{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ regular tree grammar

$\mathrm{Out}\colon \mathrm{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$???



$$\mathrm{Out} = \mathrm{Proj}((\mathrm{Inp} \lhd \mathrm{TR}) \rhd \mathrm{LM})$$

Finitely representable?
$\rightsquigarrow$ Tree Parsing

Small selection of grammar types?

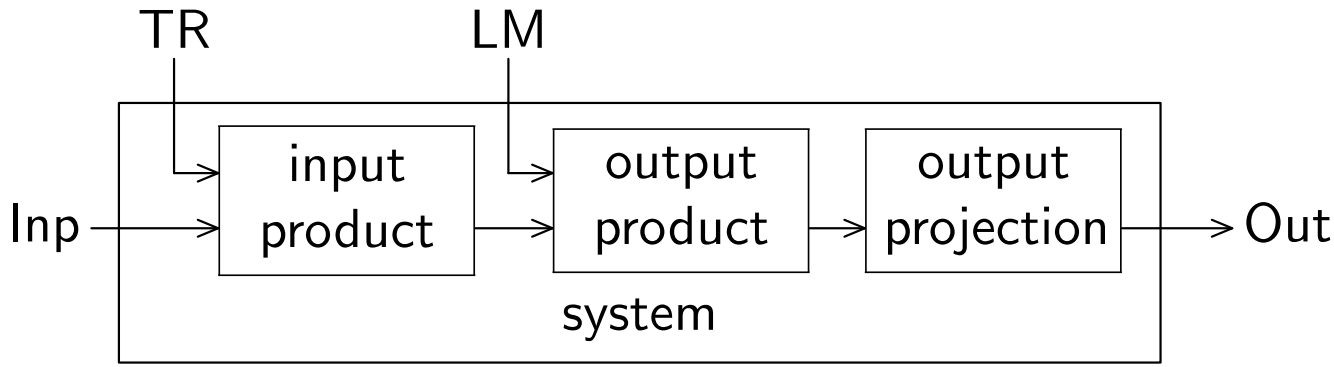
Translation System Example

$\mathrm{TR}\colon \mathrm{Tree} \times \mathrm{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ synchronous tree-adjoining grammar

$\mathrm{LM}\colon \mathrm{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ regular tree grammar

$\mathrm{Inp}\colon \mathrm{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$ regular tree grammar

$\mathrm{Out}\colon \mathrm{Tree} \to \mathbb{R}_{\infty}^{\geq 0}$???



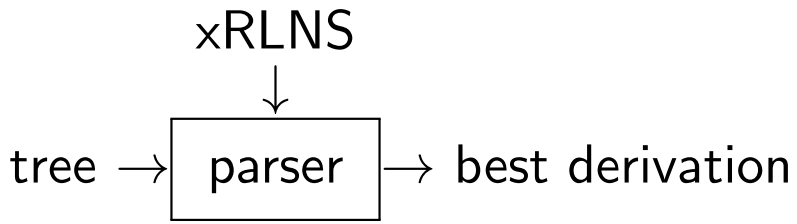
$$\mathrm{Out} = \mathrm{Proj}((\mathrm{Inp} \lhd \mathrm{TR}) \rhd \mathrm{LM})$$

Finitely representable?
$\rightsquigarrow$ Tree Parsing

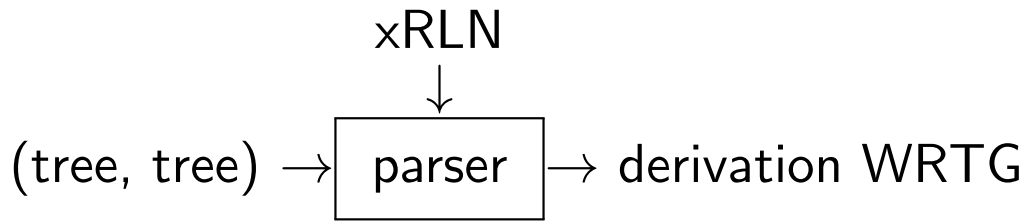
Small selection of grammar types?
$\rightsquigarrow$ Toolbox

Tree Parsing Results

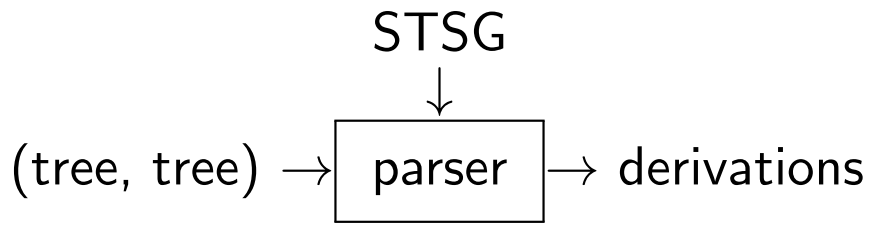
Huang et al. (2006)



Graehl et al. (2008)

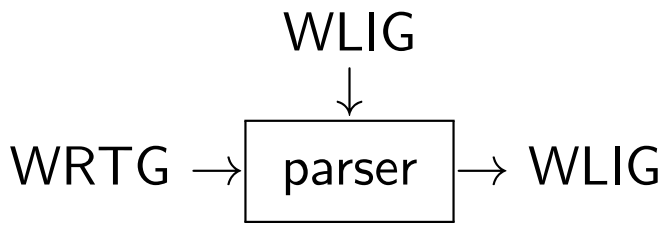


Eisner (2003)

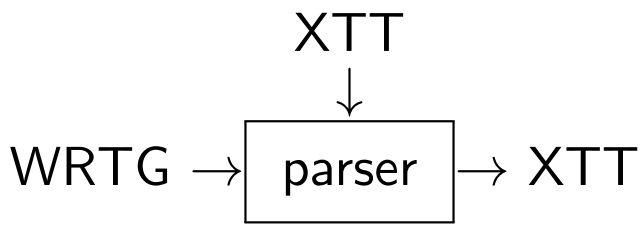


Tree Parsing Results

Nederhof (2009)

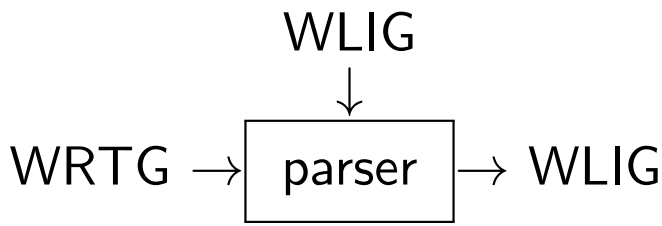


Maletti (2010)

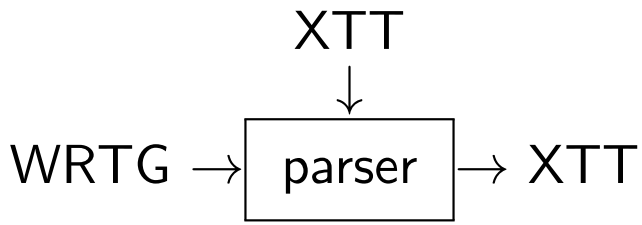


Tree Parsing Results

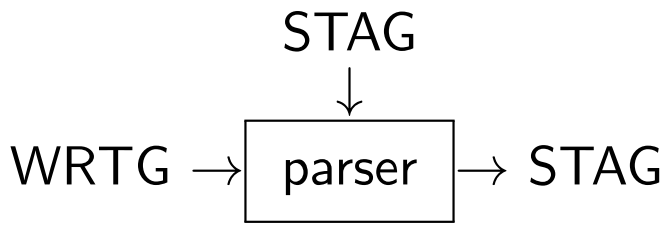
Nederhof (2009)



Maletti (2010)



this talk



Outline

Outline

Weighted Regular Tree Grammar (WRTG)

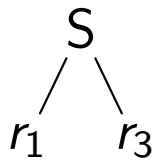
$$\rho_1: \quad r_0 \to \overset{\mathsf{S}}{\underset{r_1 \quad r_3}{/ \ \backslash}} \qquad \# \ p(\rho_1)$$

$$\rho_2: \quad r_1 \to \overset{\mathsf{Adv}}{\underset{r_2}{|}} \qquad \# \ p(\rho_2)$$

$$\rho_3: \quad r_2 \to \mathsf{yesterday} \qquad \# \ p(\rho_3)$$

$$\vdots$$

Weighted Regular Tree Grammar (WRTG)

ρ_1 : $r_0 \to$ 

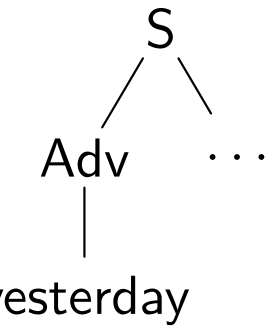
$\#\ p(\rho_1)$

ρ_2 : $r_1 \to$ 

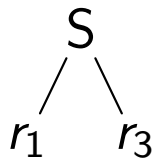
$\#\ p(\rho_2)$

ρ_3 : $r_2 \to$ yesterday $\#\ p(\rho_3)$

$\vdots$

tree t 

Weighted Regular Tree Grammar (WRTG)

ρ_1 : $r_0 \to$ 

$\#\ p(\rho_1)$

ρ_2 : $r_1 \to$ 

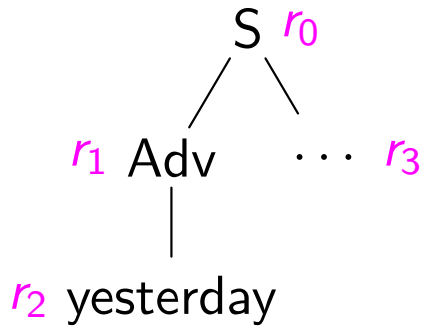
$\#\ p(\rho_2)$

ρ_3 : $r_2 \to$ yesterday $\#\ p(\rho_3)$

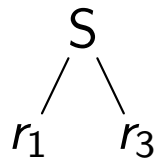
$\vdots$

tree t

run κ



Weighted Regular Tree Grammar (WRTG)

ρ_1 : $r_0 \to$ 

$\#\ p(\rho_1)$

ρ_2 : $r_1 \to$ 

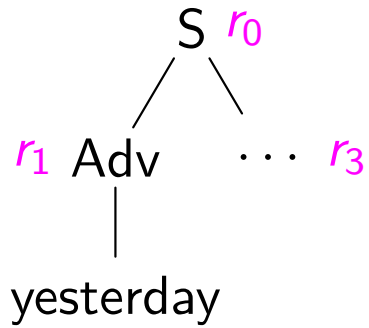
$\#\ p(\rho_2)$

ρ_3 : $r_2 \to$ yesterday $\#\ p(\rho_3)$

$\vdots$

tree t

run κ



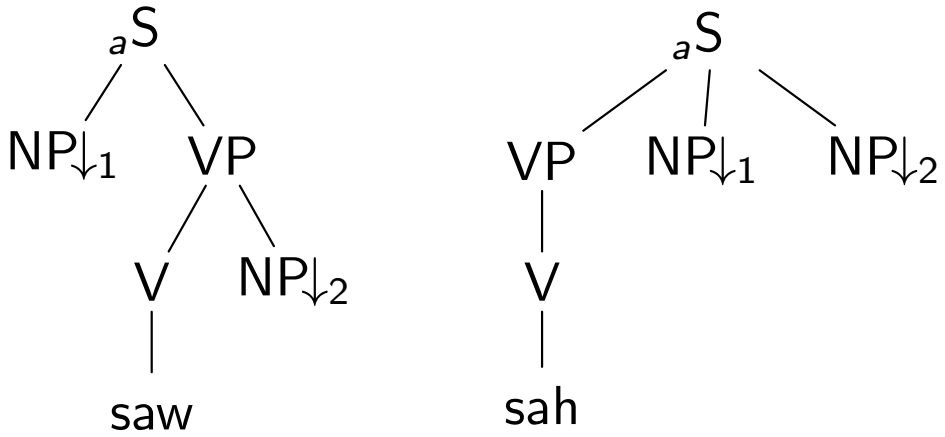
weight of κ $\quad p(\rho_1) \cdot p(\rho_2) \cdot p(\rho_3) \cdot \ldots$

weight of t $\quad \sum_\kappa p(\kappa)$

Synchronous Tree-Adjoining Grammar

following DeNeefe and Knight (2009)

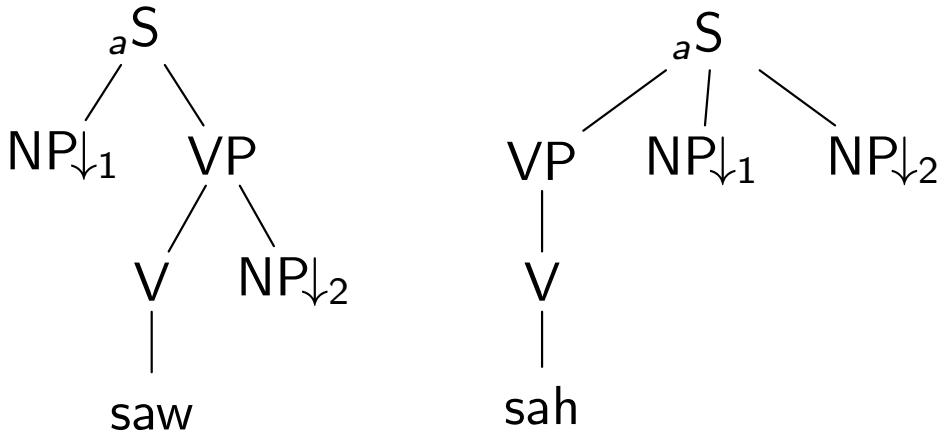
ρ_1:



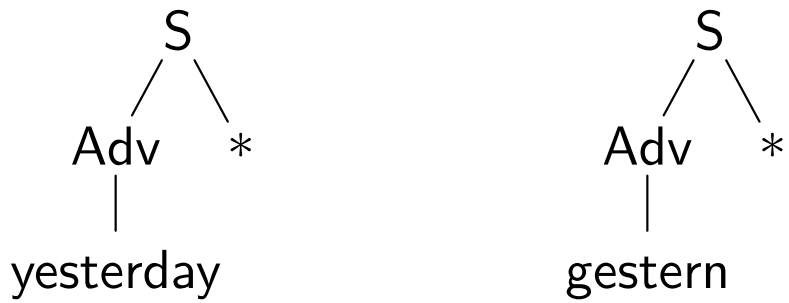
Synchronous Tree-Adjoining Grammar

following DeNeefe and Knight (2009)

ρ_1:



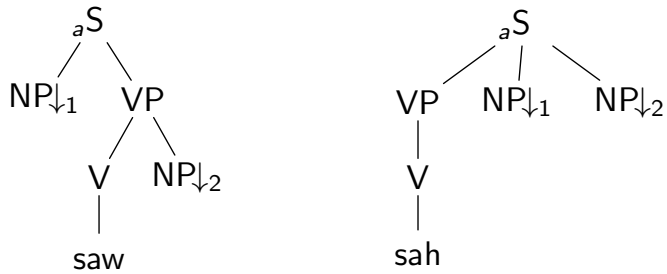
ρ_5:



Synchronous Tree-Adjoining Grammar

following DeNeefe and Knight (2009)

ρ_1 :



ρ_2 :

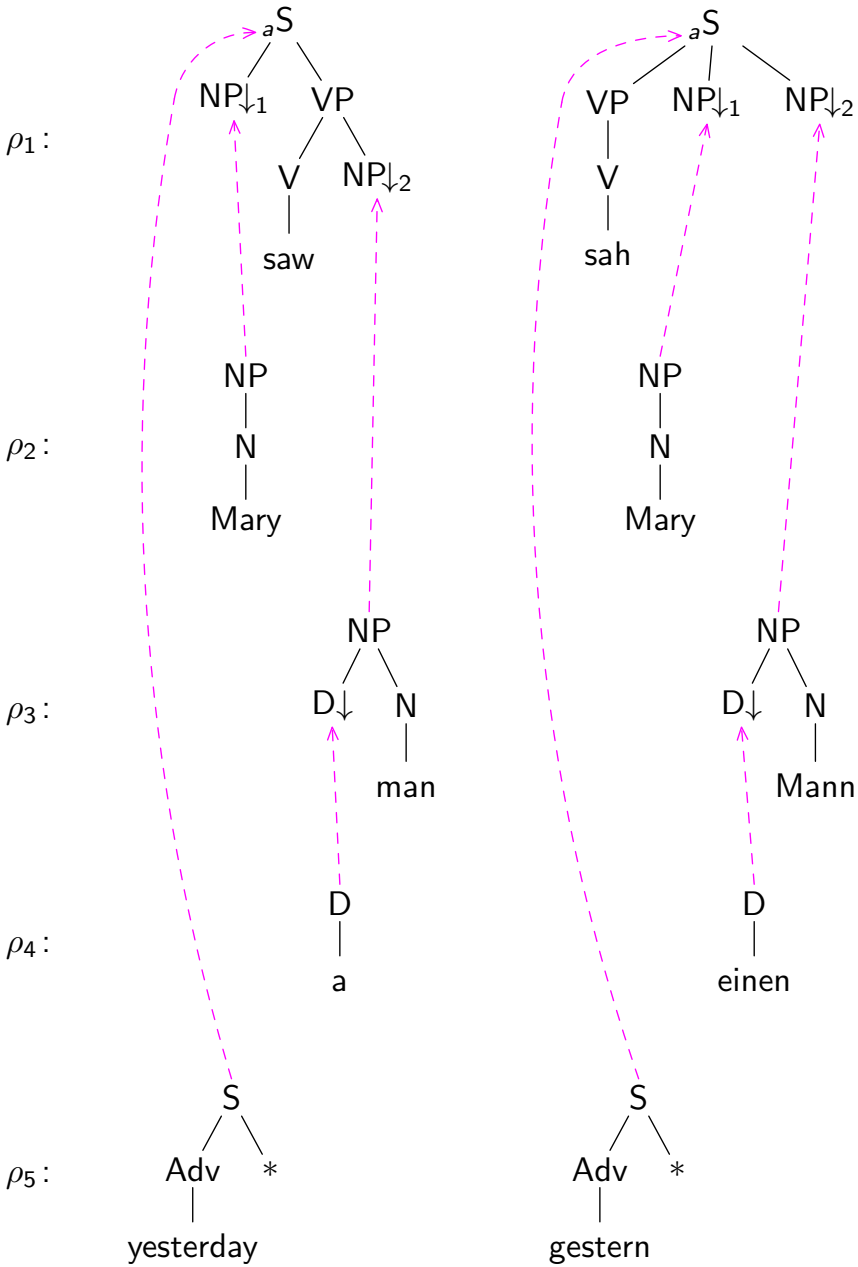
ρ_3 :

ρ_4 :

ρ_5 :

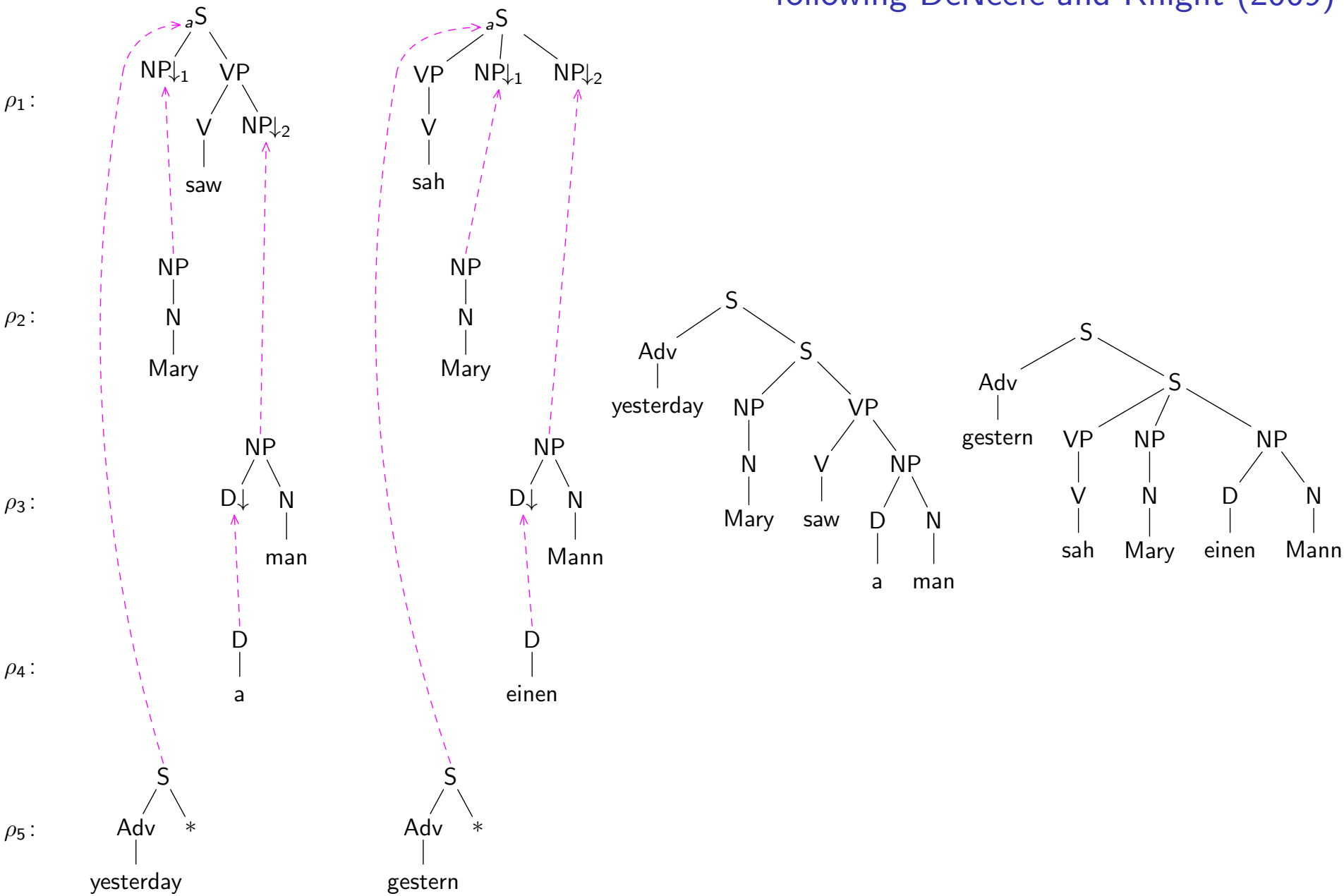
Synchronous Tree-Adjoining Grammar

following DeNeefe and Knight (2009)



Synchronous Tree-Adjoining Grammar

following DeNeefe and Knight (2009)



Outline

Example Revisited

ρ_1: $q_S \to \left\langle \begin{array}{c} z_1 \\ S \\ x_1 \quad VP \\ \quad V \quad x_2 \\ \quad saw \end{array} \quad \begin{array}{c} z_1 \\ S \\ VP \quad x_1 \quad x_2 \\ V \\ sah \end{array} , q_{NP}\, q_{NP}, f_S \right\rangle \quad \# \quad 1.0$

ρ_2: $q_{NP} \to \left\langle \begin{array}{c} NP \\ N \\ Mary \end{array} \quad \begin{array}{c} NP \\ N \\ Mary \end{array} , \varepsilon, \varepsilon \right\rangle \quad \# \quad 0.6$

ρ_3: $q_{NP} \to \left\langle \begin{array}{c} NP \\ x_1 \quad N \\ \quad man \end{array} \quad \begin{array}{c} NP \\ x_1 \quad N \\ \quad Mann \end{array} , q_D, \varepsilon \right\rangle \quad \# \quad 0.4$

ρ_4: $q_D \to \left\langle \begin{array}{c} D \\ a \end{array} \quad \begin{array}{c} D \\ einen \end{array} , \varepsilon, \varepsilon \right\rangle \quad \# \quad 1.0$

ρ_5: $f_S(*) \to \left\langle \begin{array}{c} S \\ Adv \quad * \\ yesterday \end{array} \quad \begin{array}{c} S \\ Adv \quad * \\ gestern \end{array} , \varepsilon, \varepsilon \right\rangle \quad \# \quad 1.0$

Example Revisited

ρ_1: $q_S \to \left\langle \begin{array}{c} z_1 \\ S \\ x_1 \quad VP \\ \quad V \quad x_2 \\ \quad saw \end{array} \quad \begin{array}{c} z_1 \\ S \\ VP \quad x_1 \quad x_2 \\ V \\ sah \end{array} , q_{NP}\ q_{NP}, f_S \right\rangle$ # 1.0

x_1 x_2 z_1

ρ_2: $q_{NP} \to \left\langle \begin{array}{c} NP \\ N \\ Mary \end{array} \quad \begin{array}{c} NP \\ N \\ Mary \end{array} , \varepsilon, \varepsilon \right\rangle$ # 0.6

ρ_3: $q_{NP} \to \left\langle \begin{array}{c} NP \\ x_1 \quad N \\ \quad man \end{array} \quad \begin{array}{c} NP \\ x_1 \quad N \\ \quad Mann \end{array} , q_D, \varepsilon \right\rangle$ # 0.4

ρ_4: $q_D \to \left\langle \begin{array}{c} D \\ a \end{array} \quad \begin{array}{c} D \\ einen \end{array} , \varepsilon, \varepsilon \right\rangle$ # 1.0

ρ_5: $f_S(*) \to \left\langle \begin{array}{c} S \\ Adv \quad * \\ yesterday \end{array} \quad \begin{array}{c} S \\ Adv \quad * \\ gestern \end{array} , \varepsilon, \varepsilon \right\rangle$ # 1.0

Example Revisited

ρ_1: $q_S \to \left\langle \begin{array}{c} z_1 \\ S \\ x_1 \quad VP \\ \quad V \quad x_2 \\ \quad saw \end{array} \quad \begin{array}{c} z_1 \\ S \\ VP \quad x_1 \quad x_2 \\ V \\ sah \end{array} , q_{NP}\ q_{NP}, f_S \right\rangle$ # 1.0

x_1 x_2 z_1

ρ_2: $q_{NP} \to \left\langle \begin{array}{c} NP \\ N \\ Mary \end{array} \quad \begin{array}{c} NP \\ N \\ Mary \end{array} , \varepsilon, \varepsilon \right\rangle$ # 0.6

ρ_3: $q_{NP} \to \left\langle \begin{array}{c} NP \\ x_1 \quad N \\ \quad man \end{array} \quad \begin{array}{c} NP \\ x_1 \quad N \\ \quad Mann \end{array} , q_D, \varepsilon \right\rangle$ # 0.4

ρ_4: $q_D \to \left\langle \begin{array}{c} D \\ a \end{array} \quad \begin{array}{c} D \\ einen \end{array} , \varepsilon, \varepsilon \right\rangle$ # 1.0

ρ_5: $f_S(*) \to \left\langle \begin{array}{c} S \\ Adv \quad * \\ yesterday \end{array} \quad \begin{array}{c} S \\ Adv \quad * \\ gestern \end{array} , \varepsilon, \varepsilon \right\rangle$ # 1.0

derivation tree WRTG:

$D(\rho_1)$:	$q_S \to \rho_1(q_{NP}, q_{NP}, f_S)$	# 1.0
$D(\rho_2)$:	$q_{NP} \to \rho_2()$	# 0.6
$D(\rho_3)$:	$q_{NP} \to \rho_3(q_D)$	# 0.4
$D(\rho_4)$:	$q_D \to \rho_4()$	# 1.0
$D(\rho_5)$:	$f_S \to \rho_5()$	# 1.0

Example Revisited

ρ_1: $q_S \to \left\langle \begin{array}{c} z_1 \\ S \\ x_1 \quad VP \\ \quad V \quad x_2 \\ \quad saw \end{array} \quad \begin{array}{c} z_1 \\ S \\ VP \quad x_1 \quad x_2 \\ V \\ sah \end{array} , q_{NP}\ q_{NP}, f_S \right\rangle$ # 1.0

(x_1 x_2 z_1)

ρ_2: $q_{NP} \to \left\langle \begin{array}{c} NP \\ N \\ Mary \end{array} \quad \begin{array}{c} NP \\ N \\ Mary \end{array} , \varepsilon, \varepsilon \right\rangle$ # 0.6

ρ_3: $q_{NP} \to \left\langle \begin{array}{c} NP \\ x_1 \quad N \\ \quad man \end{array} \quad \begin{array}{c} NP \\ x_1 \quad N \\ \quad Mann \end{array} , q_D, \varepsilon \right\rangle$ # 0.4

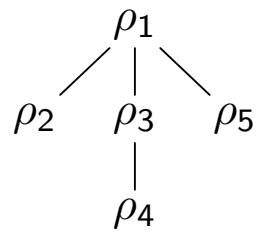
ρ_4: $q_D \to \left\langle \begin{array}{c} D \\ a \end{array} \quad \begin{array}{c} D \\ einen \end{array} , \varepsilon, \varepsilon \right\rangle$ # 1.0

ρ_5: $f_S(*) \to \left\langle \begin{array}{c} S \\ Adv \quad * \\ yesterday \end{array} \quad \begin{array}{c} S \\ Adv \quad * \\ gestern \end{array} , \varepsilon, \varepsilon \right\rangle$ # 1.0

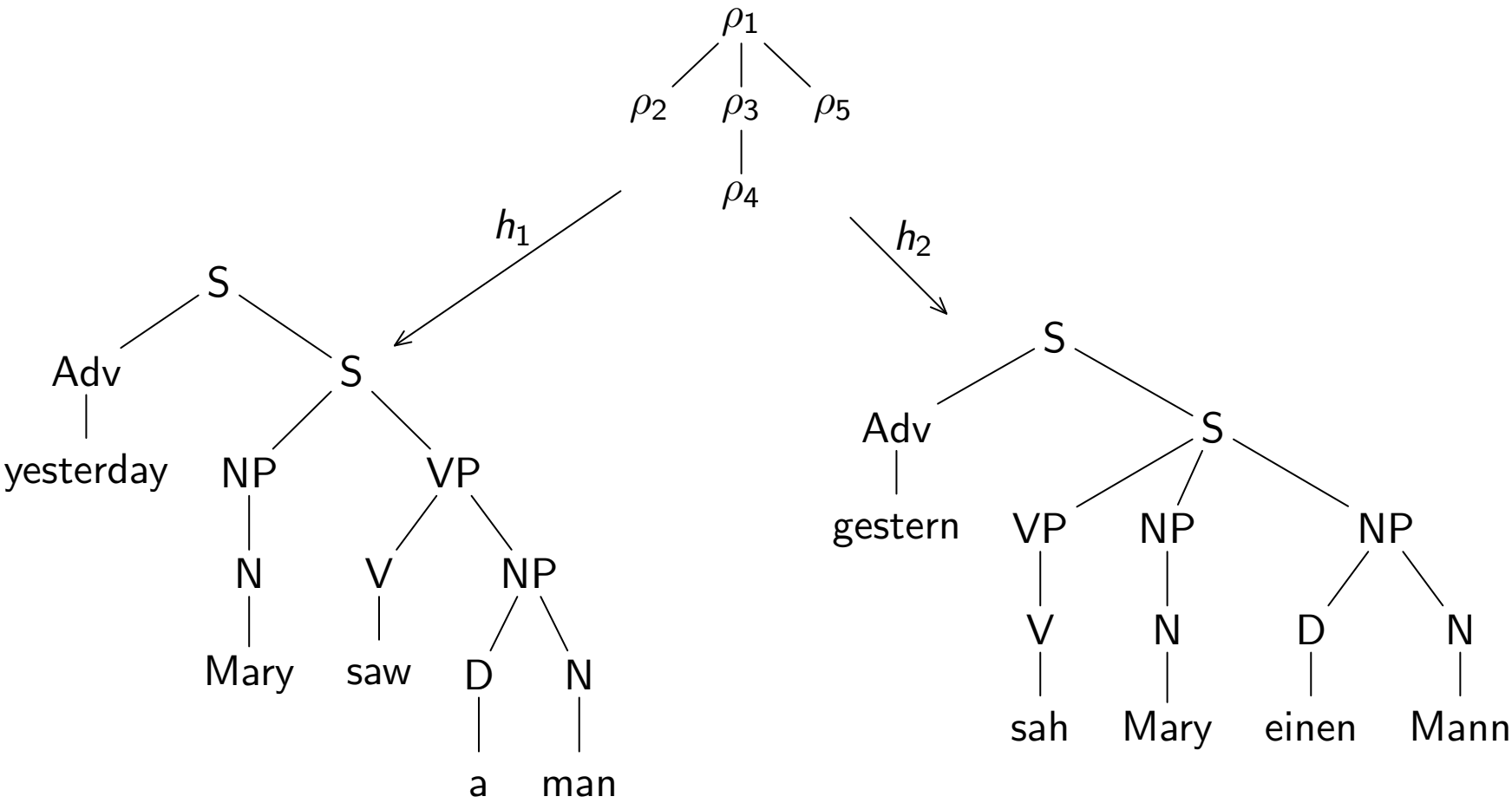
derivation tree WRTG:

$D(\rho_1)$:	$q_S \to \rho_1(q_{NP}, q_{NP}, f_S)$	# 1.0
$D(\rho_2)$:	$q_{NP} \to \rho_2()$	# 0.6
$D(\rho_3)$:	$q_{NP} \to \rho_3(q_D)$	# 0.4
$D(\rho_4)$:	$q_D \to \rho_4()$	# 1.0
$D(\rho_5)$:	$f_S \to \rho_5()$	# 1.0

derivation tree:



Bimorphism Semantics



Input Product Construction

WSTAG	WRTG	product WSTAG
$q \to \left\langle \begin{array}{c} \mathsf{A} \\ \mid \\ z_1 \\ \mid \\ \mathsf{A} \\ \mid \\ x_1 \end{array}, q, f \right\rangle \quad \# \quad 0.2$		
$q \to \langle \mathsf{a}, \varepsilon, \varepsilon \rangle \quad \# \quad 0.8$		
$f(*) \to \left\langle \begin{array}{ccc} & \mathsf{B} & \\ / & & \backslash \\ \mathsf{b} & & * \end{array}, \varepsilon, \varepsilon \right\rangle \quad \# \quad 0.1$		
$f(*) \to \langle *, \varepsilon, \varepsilon \rangle \quad \# \quad 0.9$		

Input Product Construction

WSTAG	WRTG	product WSTAG
$q \to \left\langle \begin{array}{c} \mathsf{A} \\ \mid \\ z_1 \\ \mid \\ \mathsf{A} \\ \mid \\ x_1 \end{array}, q, f \right\rangle$ # 0.2	$r_0 \to \begin{array}{c} \mathsf{A} \\ \mid \\ r_0 \end{array}$ # 0.1	
$q \to \langle \mathsf{a}, \varepsilon, \varepsilon \rangle$ # 0.8	$r_0 \to \begin{array}{ccc} & \mathsf{B} & \\ / & & \backslash \\ r_0 & & r_1 \end{array}$ # 0.5	
$f(*) \to \left\langle \begin{array}{ccc} & \mathsf{B} & \\ / & & \backslash \\ \mathsf{b} & & * \end{array}, \varepsilon, \varepsilon \right\rangle$ # 0.1	$r_0 \to \mathsf{b}$ # 0.4	
$f(*) \to \langle *, \varepsilon, \varepsilon \rangle$ # 0.9	$r_1 \to \begin{array}{c} \mathsf{A} \\ \mid \\ r_1 \end{array}$ # 0.5	
	$r_1 \to \mathsf{a}$ # 0.5	

Input Product Construction

WSTAG

$q \to \left\langle \begin{array}{c} \mathrm{A} \\ | \\ z_1 \\ | \\ \mathrm{A} \\ | \\ x_1 \end{array}, q, f \right\rangle \quad \# \quad 0.2$

$q \to \langle \mathrm{a}, \varepsilon, \varepsilon \rangle \quad \# \quad 0.8$

$f(*) \to \left\langle \begin{array}{c} \mathrm{B} \\ / \quad \backslash \\ \mathrm{b} \quad * \end{array}, \varepsilon, \varepsilon \right\rangle \quad \# \quad 0.1$

$f(*) \to \langle *, \varepsilon, \varepsilon \rangle \quad \# \quad 0.9$

WRTG

$r_0 \to \begin{array}{c} \mathrm{A} \\ | \\ r_0 \end{array} \quad \# \quad 0.1$

$r_0 \to \begin{array}{c} \mathrm{B} \\ / \quad \backslash \\ r_0 \quad r_1 \end{array} \quad \# \quad 0.5$

$r_0 \to \mathrm{b} \quad \# \quad 0.4$

$r_1 \to \begin{array}{c} \mathrm{A} \\ | \\ r_1 \end{array} \quad \# \quad 0.5$

$r_1 \to \mathrm{a} \quad \# \quad 0.5$

product WSTAG

$qr_0 \to \left\langle \begin{array}{c} \mathrm{A} \\ | \\ z_1 \\ | \\ \mathrm{A} \\ | \\ x_1 \end{array}, qr_0, fr_0 r_0 \right\rangle \quad \# \quad 0.002$

$fr_0 r_0(*) \to \langle *, \varepsilon, \varepsilon \rangle \quad \# \quad 0.9$

$qr_0 \to \left\langle \begin{array}{c} \mathrm{A} \\ | \\ z_1 \\ | \\ \mathrm{A} \\ | \\ x_1 \end{array}, qr_1, fr_0 r_1 \right\rangle \quad \# \quad 0.01$

$fr_0 r_1(*) \to \left\langle \begin{array}{c} \mathrm{B} \\ / \quad \backslash \\ \mathrm{b} \quad * \end{array}, \varepsilon, \varepsilon \right\rangle \quad \# \quad 0.002$

$qr_1 \to \langle \mathrm{a}, \varepsilon, \varepsilon \rangle \quad \# \quad 0.4$

Outline