

# Syntax-based Statistical Machine Translation

Yang Liu  
*Tsinghua University*



# Regularities in Natural Languages

- The way people say things has regularities

Every boy likes a car  
The girl saw a dog  
Prof. Wang gave a talk

# Regularities in Natural Languages

- The way people say things has regularities

Every boy likes a car  
The girl saw a dog  
Prof. Wang gave a talk

# Regularities in Natural Languages

- The way people say things has regularities

Every boy likes a car  
The girl saw a dog  
Prof. Wang gave a talk

How are the sentences generated?

# Context-Free Grammar

- Context-free grammar describes how natural language sentences are generated

# Context-Free Grammar

- Context-free grammar describes how natural language sentences are generated  
lexical rules

# Context-Free Grammar

- **Context-free grammar** describes how natural language sentences are generated

## lexical rules

NNP  $\rightarrow$  Bush

VBD  $\rightarrow$  held

DT  $\rightarrow$  a

NN  $\rightarrow$  talk

IN  $\rightarrow$  with

NNP  $\rightarrow$  Sharon

# Context-Free Grammar

- **Context-free grammar** describes how natural language sentences are generated

lexical rules

syntactic rules

NNP  $\rightarrow$  Bush

VBD  $\rightarrow$  held

DT  $\rightarrow$  a

NN  $\rightarrow$  talk

IN  $\rightarrow$  with

NNP  $\rightarrow$  Sharon



# Context-Free Grammar

- **Context-free grammar** describes how natural language sentences are generated

## lexical rules

NNP  $\rightarrow$  Bush

VBD  $\rightarrow$  held

DT  $\rightarrow$  a

NN  $\rightarrow$  talk

IN  $\rightarrow$  with

NNP  $\rightarrow$  Sharon

## syntactic rules

NP  $\rightarrow$  NNP

NP  $\rightarrow$  DT NN

PP  $\rightarrow$  IN NP

VP  $\rightarrow$  VBD NP PP

S  $\rightarrow$  NP VP

# Derivation

- A derivation explains how a sentence can be generated by applying CFG rules

# Derivation

- A derivation explains how a sentence can be generated by applying CFG rules

$S \rightarrow NP VP$

# Derivation

- A derivation explains how a sentence can be generated by applying CFG rules

$S \rightarrow NP VP$

$S \Rightarrow NP VP$

# Derivation

- A derivation explains how a sentence can be generated by applying CFG rules

$S \rightarrow NP VP$

$S \Rightarrow NP VP$

$NP \rightarrow NNP$

# Derivation

- A derivation explains how a sentence can be generated by applying CFG rules

$S \rightarrow NP VP$

$S \Rightarrow NP VP$

$NP \rightarrow NNP$

$\Rightarrow NNP VP$

# Derivation

- A derivation explains how a sentence can be generated by applying CFG rules

$S \rightarrow NP VP$

$S \Rightarrow NP VP$

$NP \rightarrow NNP$

$\Rightarrow NNP VP$

$NNP \rightarrow \text{Bush}$

# Derivation

- A derivation explains how a sentence can be generated by applying CFG rules

$S \rightarrow NP VP$

$S \Rightarrow NP VP$

$NP \rightarrow NNP$

$\Rightarrow NNP VP$

$NNP \rightarrow \text{Bush}$

$\Rightarrow \text{Bush VP}$



# Derivation

- A derivation explains how a sentence can be generated by applying CFG rules

$S \rightarrow NP VP$

$S \Rightarrow NP VP$

$NP \rightarrow NNP$

$\Rightarrow NNP VP$

$NNP \rightarrow \text{Bush}$

$\Rightarrow \text{Bush VP}$

$VP \rightarrow VBD NP PP$

# Derivation

- A derivation explains how a sentence can be generated by applying CFG rules

$S \rightarrow NP VP$

$S \Rightarrow NP VP$

$NP \rightarrow NNP$

$\Rightarrow NNP VP$

$NNP \rightarrow \text{Bush}$

$\Rightarrow \text{Bush VP}$

$VP \rightarrow VBD NP PP$

$\Rightarrow \text{Bush VBD NP PP}$

# Derivation

- A derivation explains how a sentence can be generated by applying CFG rules

$S \rightarrow NP VP$

$S \Rightarrow NP VP$

$NP \rightarrow NNP$

$\Rightarrow NNP VP$

$NNP \rightarrow \text{Bush}$

$\Rightarrow \text{Bush VP}$

$VP \rightarrow VBD NP PP$

$\Rightarrow \text{Bush VBD NP PP}$

$VBD \rightarrow \text{held}$

# Derivation

- A derivation explains how a sentence can be generated by applying CFG rules

$S \rightarrow NP VP$

$S \Rightarrow NP VP$

$NP \rightarrow NNP$

$\Rightarrow NNP VP$

$NNP \rightarrow \text{Bush}$

$\Rightarrow \text{Bush VP}$

$VP \rightarrow VBD NP PP$

$\Rightarrow \text{Bush VBD NP PP}$

$VBD \rightarrow \text{held}$

$\Rightarrow \text{Bush held NP PP}$

# Derivation

- A derivation explains how a sentence can be generated by applying CFG rules

$S \rightarrow NP VP$

$S \Rightarrow NP VP$

$NP \rightarrow NNP$

$\Rightarrow NNP VP$

$NNP \rightarrow \text{Bush}$

$\Rightarrow \text{Bush VP}$

$VP \rightarrow VBD NP PP$

$\Rightarrow \text{Bush VBD NP PP}$

$VBD \rightarrow \text{held}$

$\Rightarrow \text{Bush held NP PP}$

$NP \rightarrow DT NN$

# Derivation

- A derivation explains how a sentence can be generated by applying CFG rules

$S \rightarrow NP VP$

$S \Rightarrow NP VP$

$NP \rightarrow NNP$

$\Rightarrow NNP VP$

$NNP \rightarrow \text{Bush}$

$\Rightarrow \text{Bush VP}$

$VP \rightarrow VBD NP PP$

$\Rightarrow \text{Bush VBD NP PP}$

$VBD \rightarrow \text{held}$

$\Rightarrow \text{Bush held NP PP}$

$NP \rightarrow DT NN$

$\Rightarrow \text{Bush held DT NN PP}$

# Derivation

- A derivation explains how a sentence can be generated by applying CFG rules

$S \rightarrow NP VP$

$S \Rightarrow NP VP$

$NP \rightarrow NNP$

$\Rightarrow NNP VP$

$NNP \rightarrow \text{Bush}$

$\Rightarrow \text{Bush VP}$

$VP \rightarrow VBD NP PP$

$\Rightarrow \text{Bush VBD NP PP}$

$VBD \rightarrow \text{held}$

$\Rightarrow \text{Bush held NP PP}$

$NP \rightarrow DT NN$

$\Rightarrow \text{Bush held DT NN PP}$

$DT \rightarrow a$

# Derivation

- A derivation explains how a sentence can be generated by applying CFG rules

$S \rightarrow NP VP$

$S \Rightarrow NP VP$

$NP \rightarrow NNP$

$\Rightarrow NNP VP$

$NNP \rightarrow \text{Bush}$

$\Rightarrow \text{Bush VP}$

$VP \rightarrow VBD NP PP$

$\Rightarrow \text{Bush VBD NP PP}$

$VBD \rightarrow \text{held}$

$\Rightarrow \text{Bush held NP PP}$

$NP \rightarrow DT NN$

$\Rightarrow \text{Bush held DT NN PP}$

$DT \rightarrow a$

$\Rightarrow \text{Bush held a NN PP}$



# Derivation

- A derivation explains how a sentence can be generated by applying CFG rules

$S \rightarrow NP VP$

$S \Rightarrow NP VP$

$NP \rightarrow NNP$

$\Rightarrow NNP VP$

$NNP \rightarrow \text{Bush}$

$\Rightarrow \text{Bush VP}$

$VP \rightarrow VBD NP PP$

$\Rightarrow \text{Bush VBD NP PP}$

$VBD \rightarrow \text{held}$

$\Rightarrow \text{Bush held NP PP}$

$NP \rightarrow DT NN$

$\Rightarrow \text{Bush held DT NN PP}$

$DT \rightarrow a$

$\Rightarrow \text{Bush held a NN PP}$

...

# Derivation

- A derivation explains how a sentence can be generated by applying CFG rules

$S \rightarrow NP VP$

$S \Rightarrow NP VP$

$NP \rightarrow NNP$

$\Rightarrow NNP VP$

$NNP \rightarrow \text{Bush}$

$\Rightarrow \text{Bush VP}$

$VP \rightarrow VBD NP PP$

$\Rightarrow \text{Bush VBD NP PP}$

$VBD \rightarrow \text{held}$

$\Rightarrow \text{Bush held NP PP}$

$NP \rightarrow DT NN$

$\Rightarrow \text{Bush held DT NN PP}$

$DT \rightarrow a$

$\Rightarrow \text{Bush held a NN PP}$

...

...

# Derivation

- A derivation explains how a sentence can be generated by applying CFG rules

$S \rightarrow NP VP$

$S \Rightarrow NP VP$

$NP \rightarrow NNP$

$\Rightarrow NNP VP$

$NNP \rightarrow Bush$

$\Rightarrow Bush VP$

$VP \rightarrow VBD NP PP$

$\Rightarrow Bush VBD NP PP$

$VBD \rightarrow held$

$\Rightarrow Bush held NP PP$

$NP \rightarrow DT NN$

$\Rightarrow Bush held DT NN PP$

$DT \rightarrow a$

$\Rightarrow Bush held a NN PP$

...

...

$NNP \rightarrow Sharon$

# Derivation

- A derivation explains how a sentence can be generated by applying CFG rules

$S \rightarrow NP VP$

$S \Rightarrow NP VP$

$NP \rightarrow NNP$

$\Rightarrow NNP VP$

$NNP \rightarrow Bush$

$\Rightarrow Bush VP$

$VP \rightarrow VBD NP PP$

$\Rightarrow Bush VBD NP PP$

$VBD \rightarrow held$

$\Rightarrow Bush held NP PP$

$NP \rightarrow DT NN$

$\Rightarrow Bush held DT NN PP$

$DT \rightarrow a$

$\Rightarrow Bush held a NN PP$

...

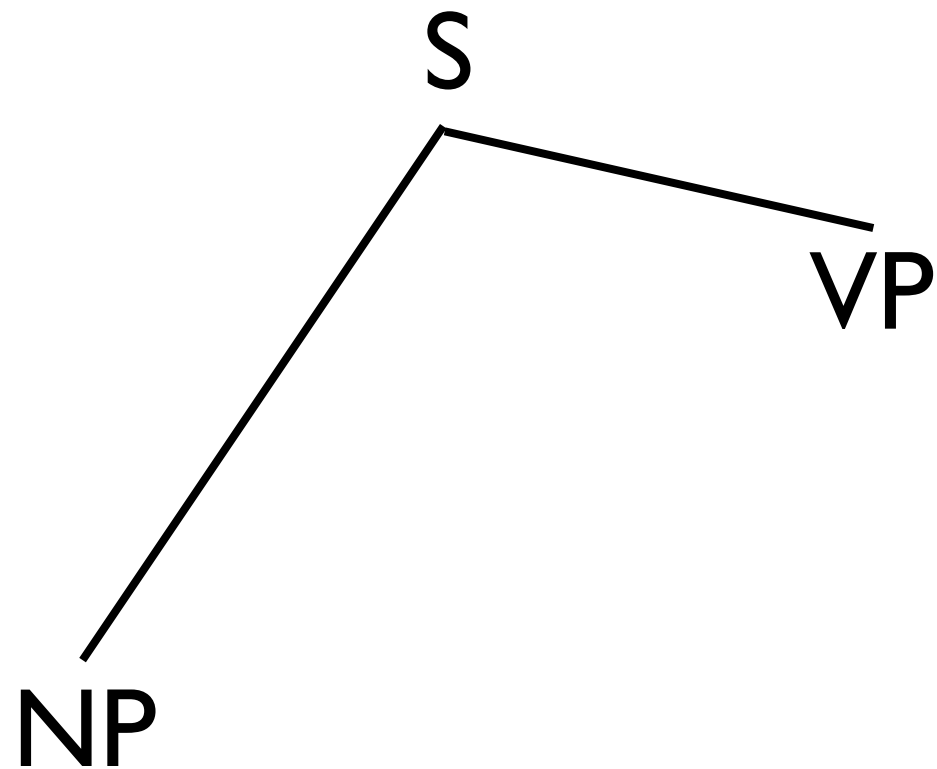
...

$NNP \rightarrow Sharon$

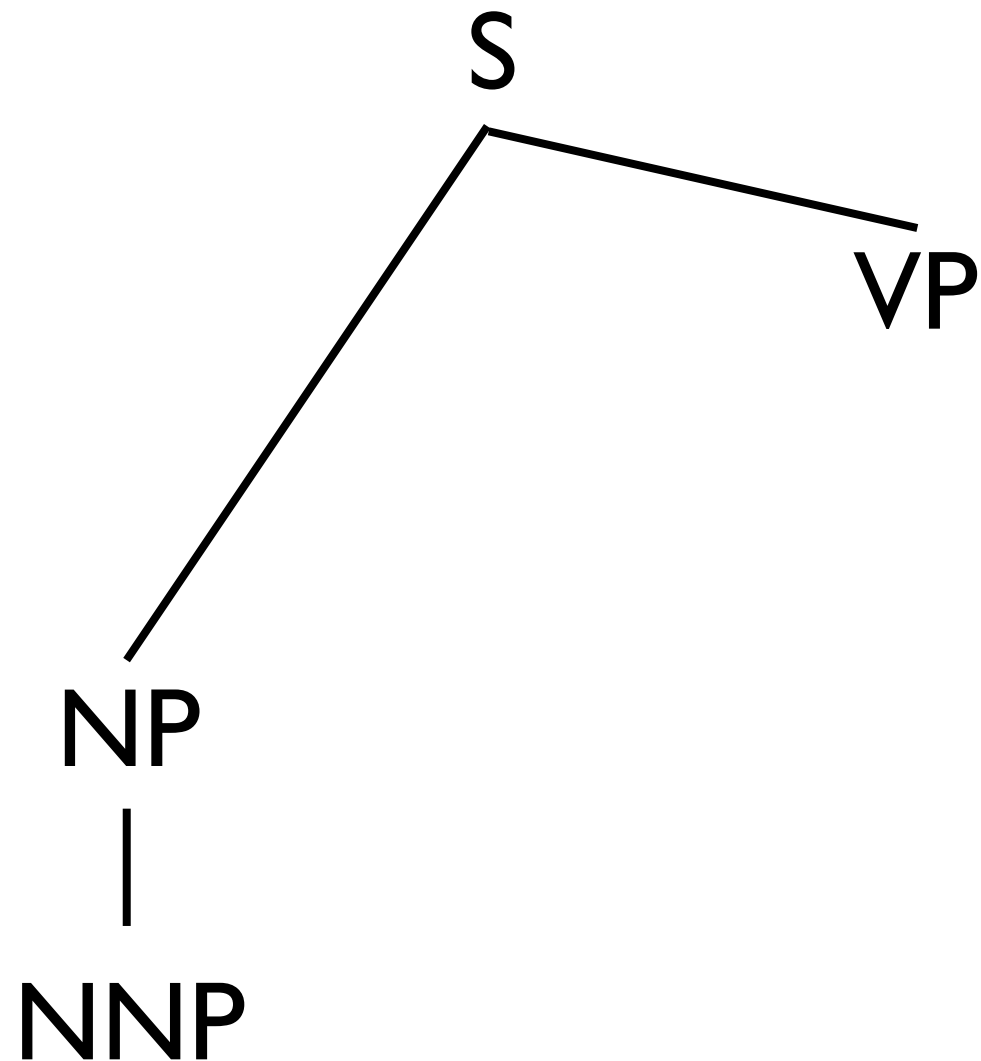
$\Rightarrow Bush held a talk with Sharon$

# Graphical Representation

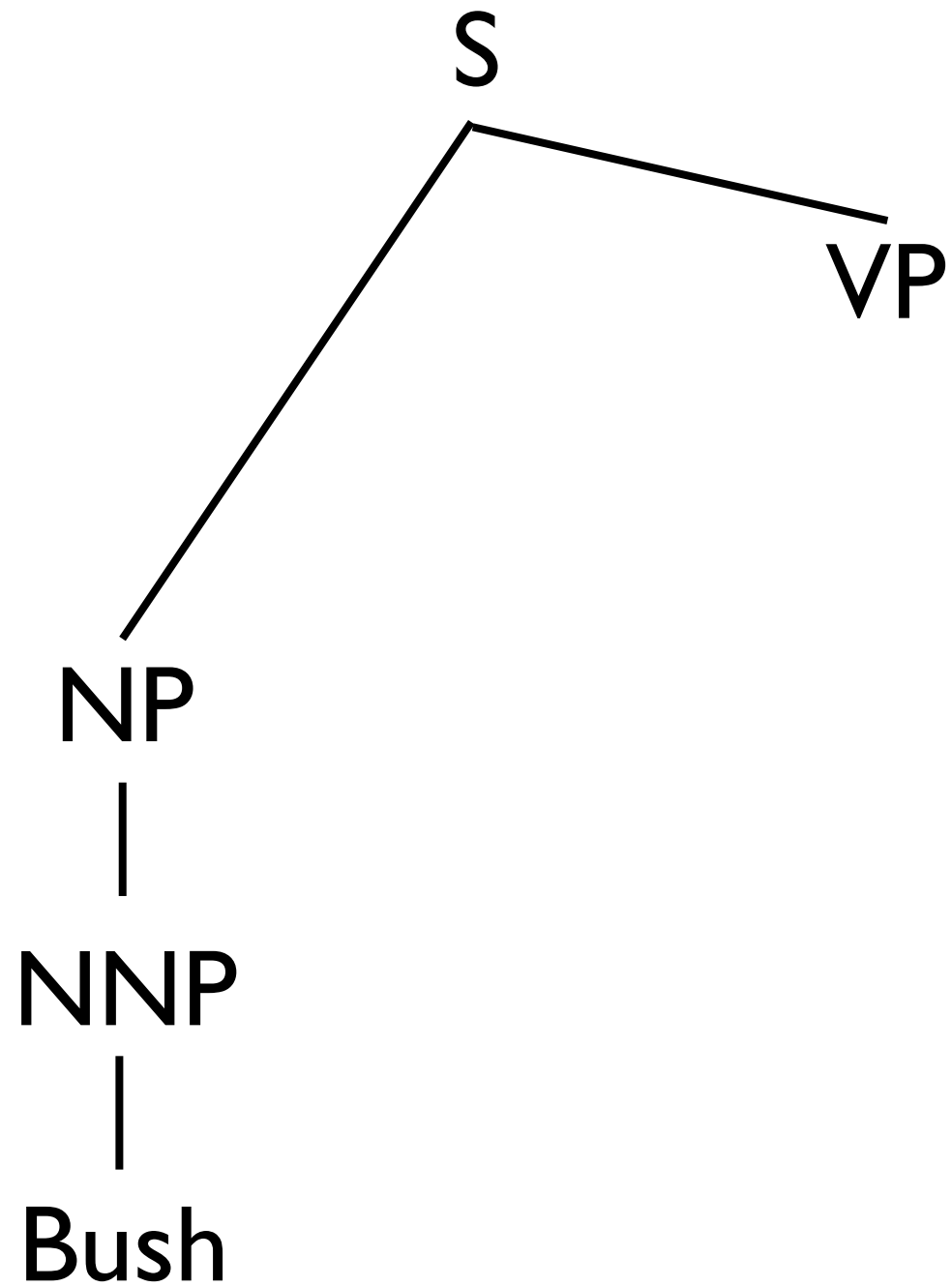
# Graphical Representation



# Graphical Representation

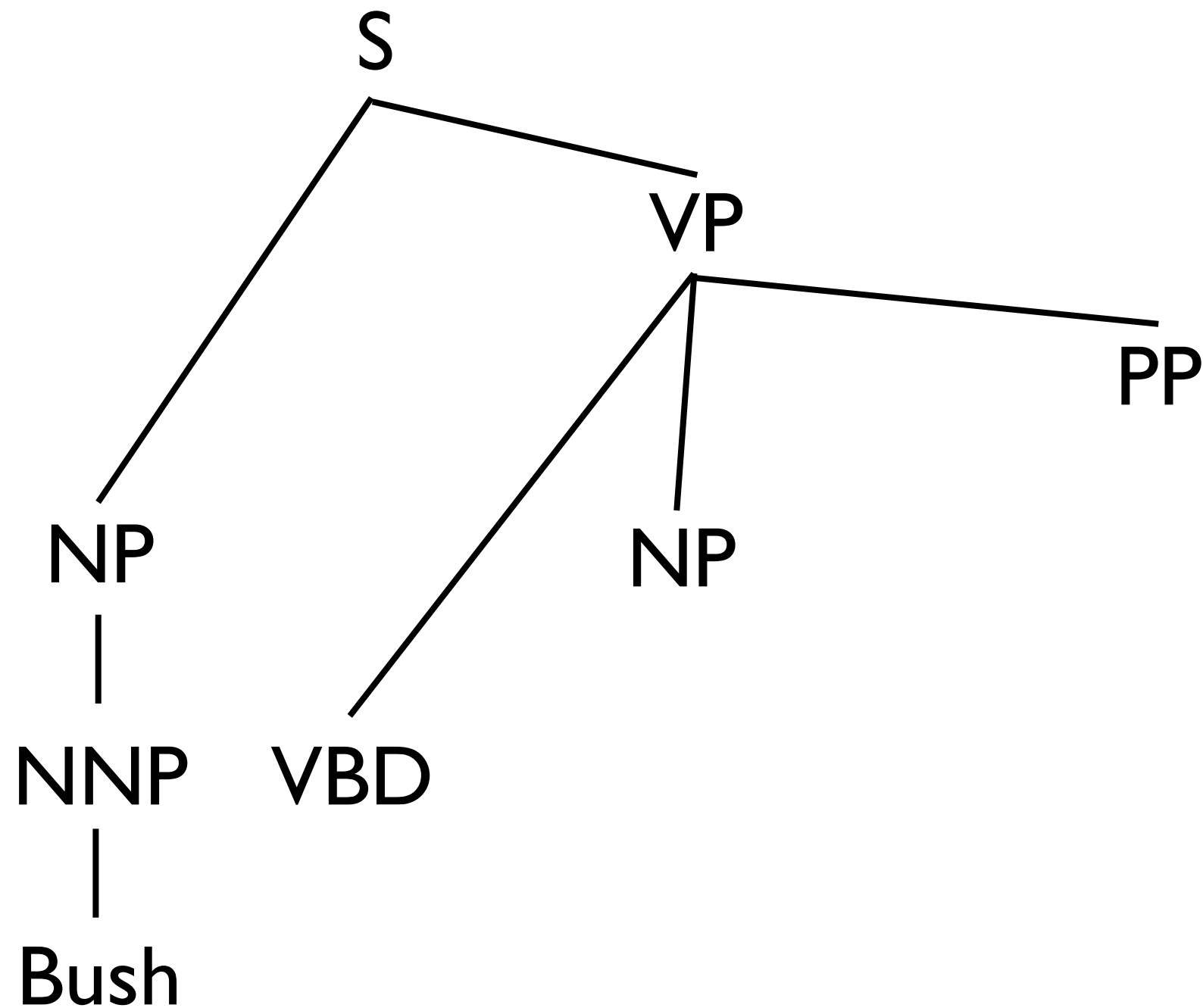


# Graphical Representation

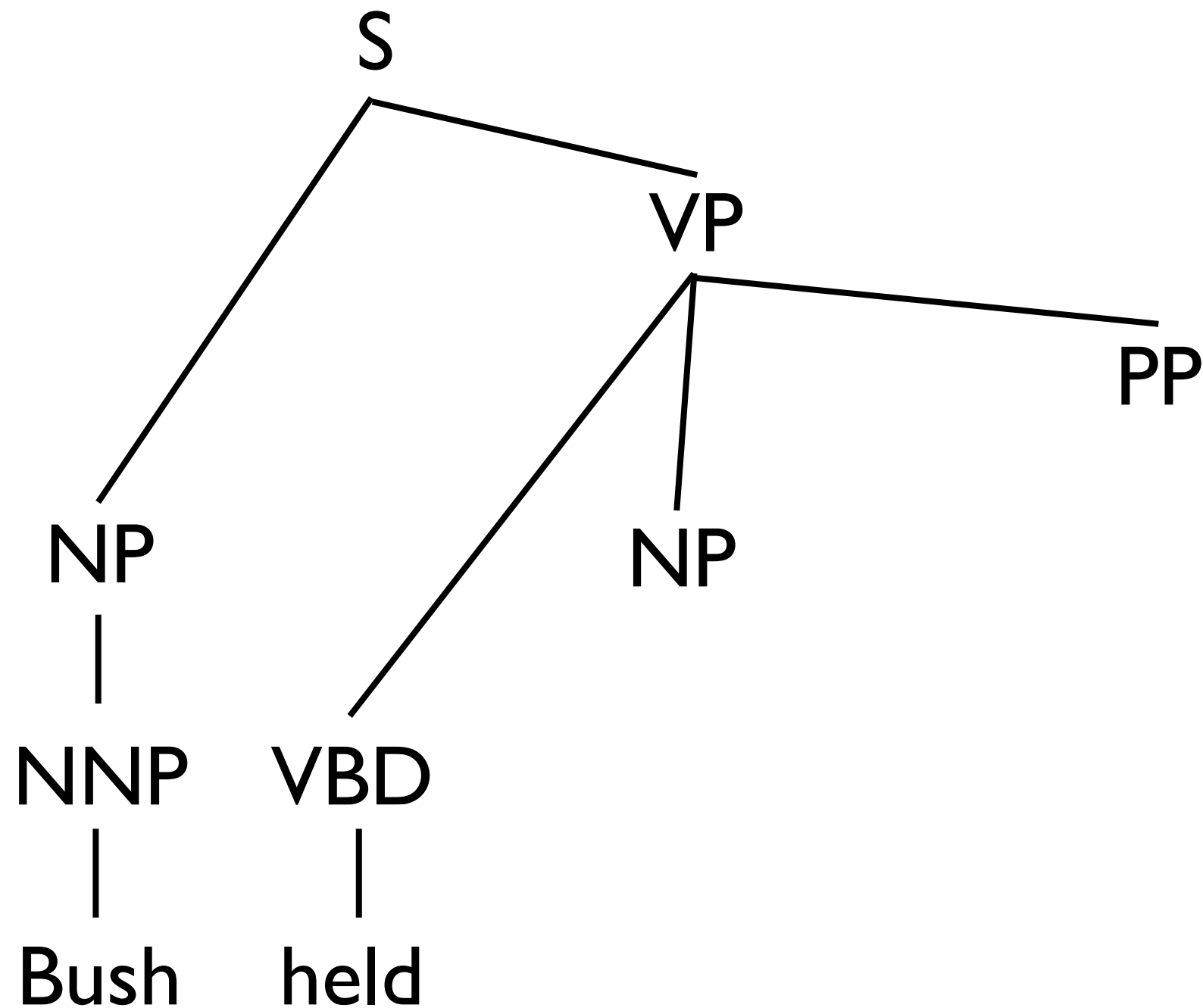




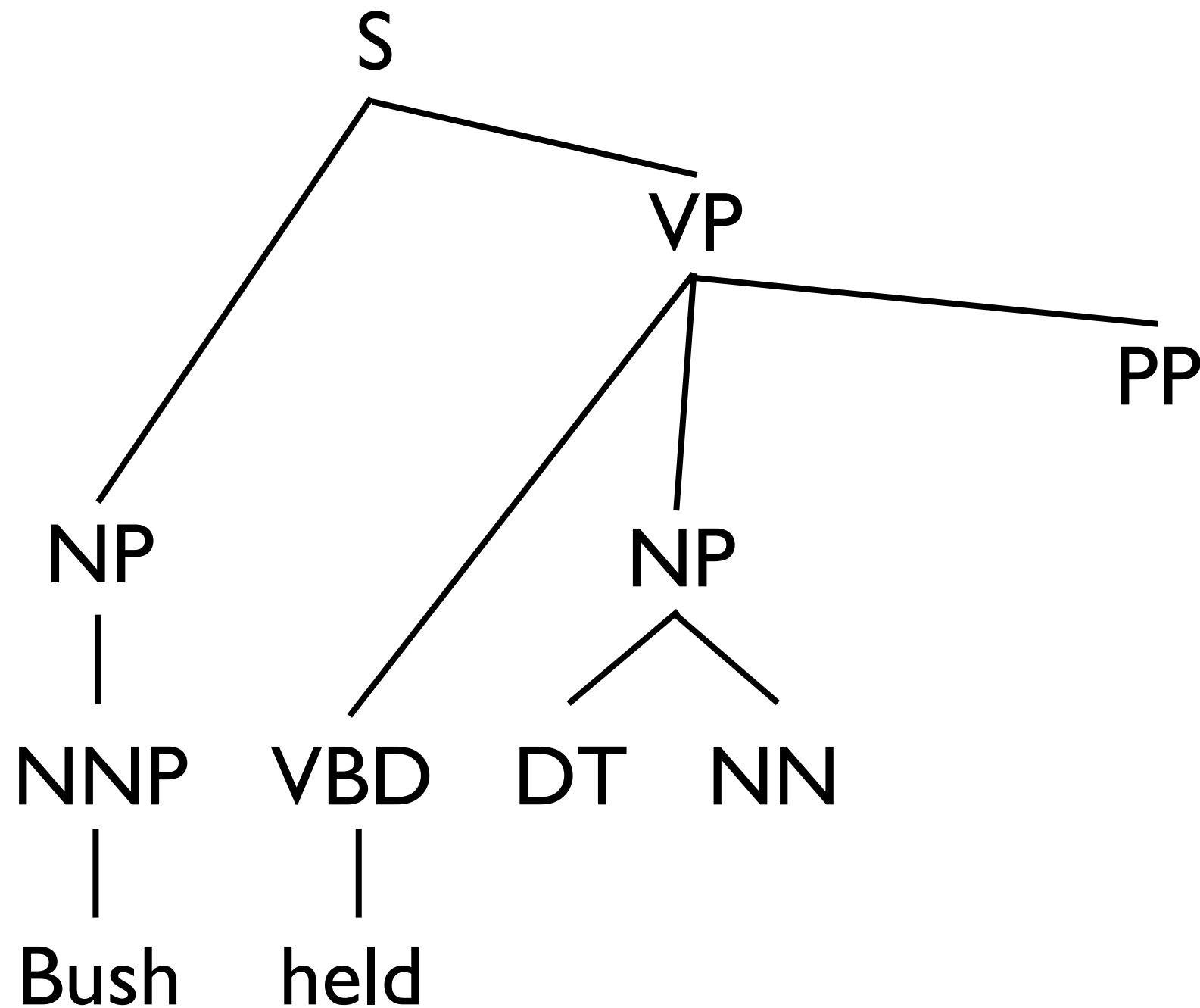
# Graphical Representation



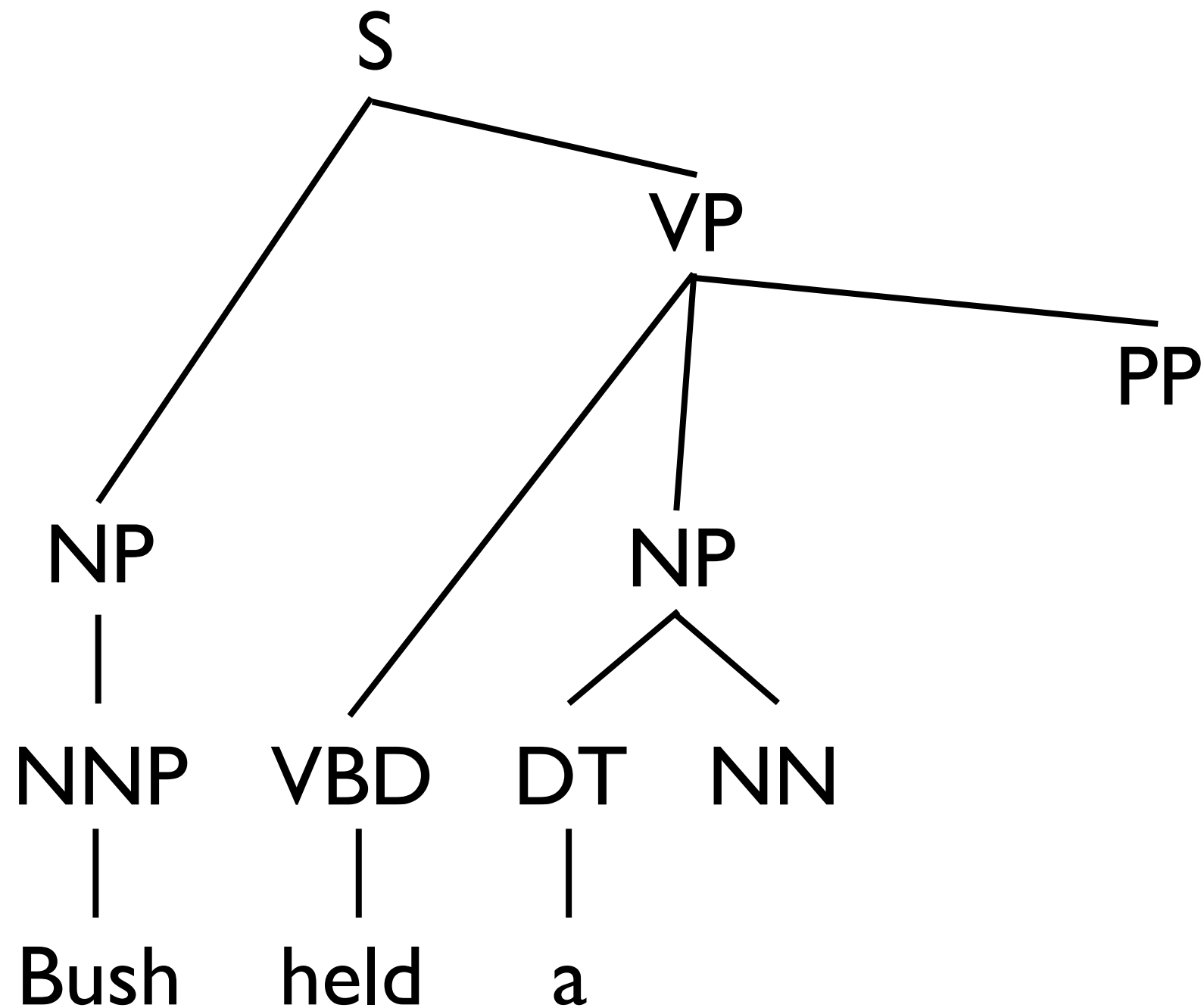
# Graphical Representation



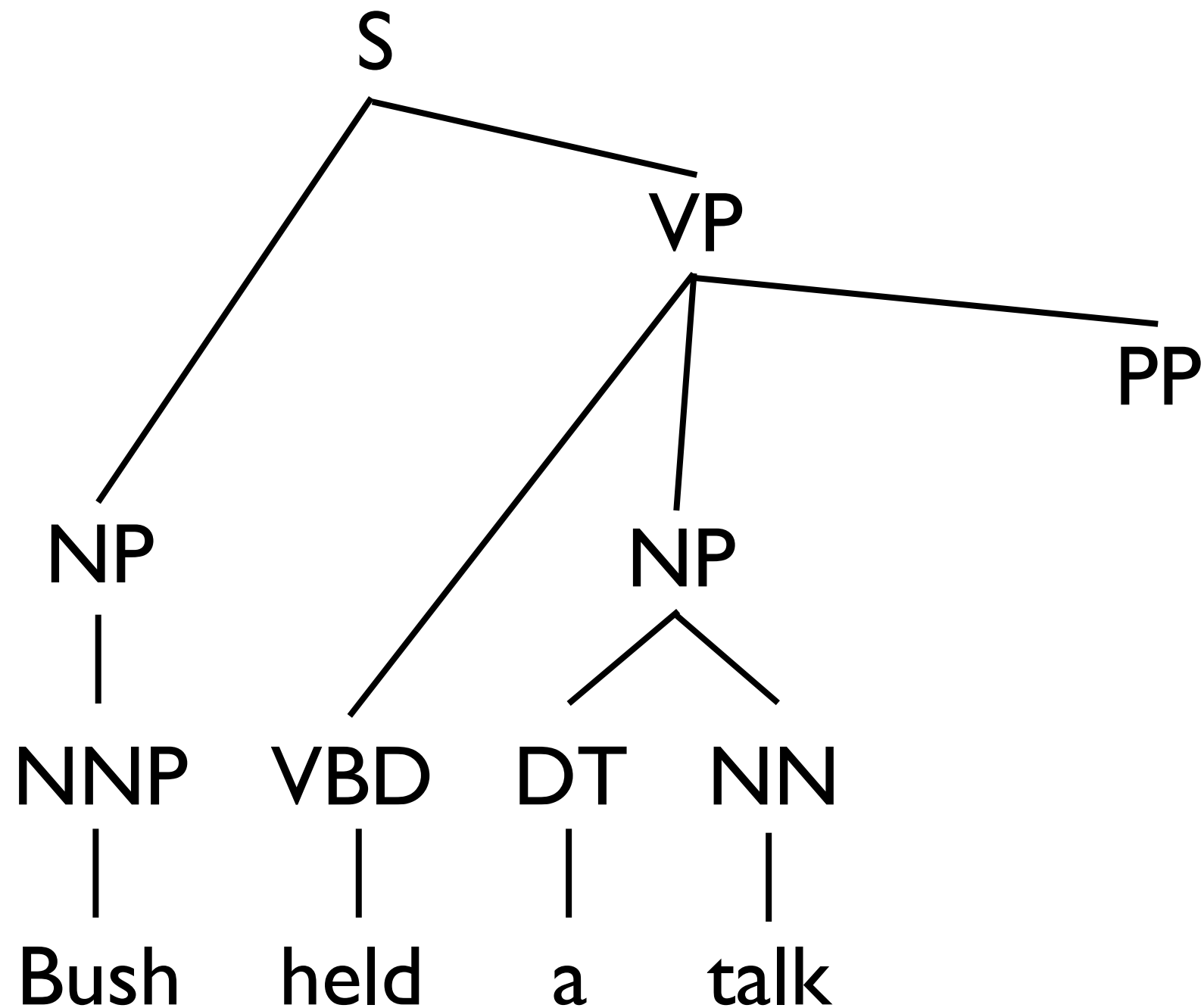
# Graphical Representation



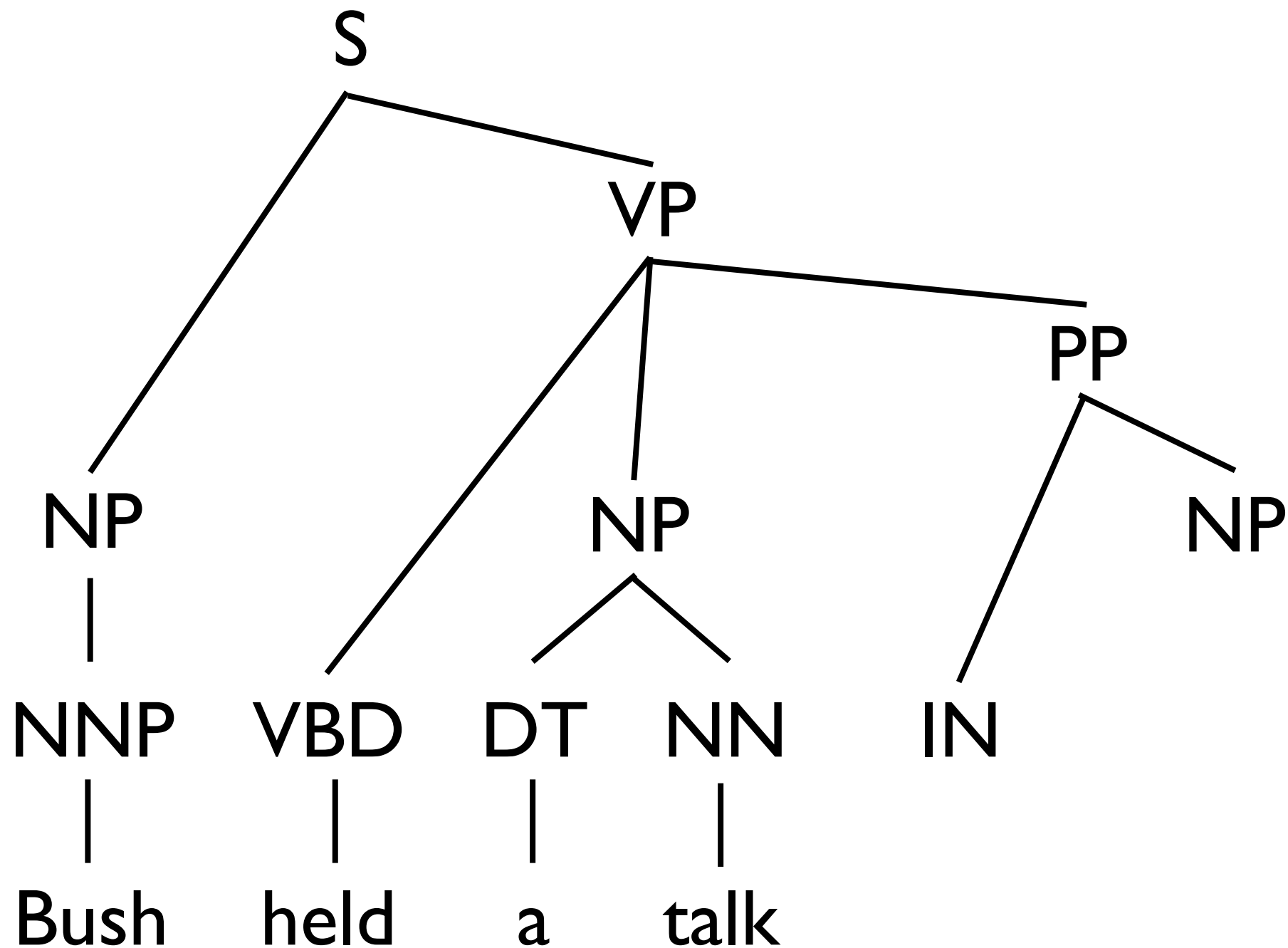
# Graphical Representation



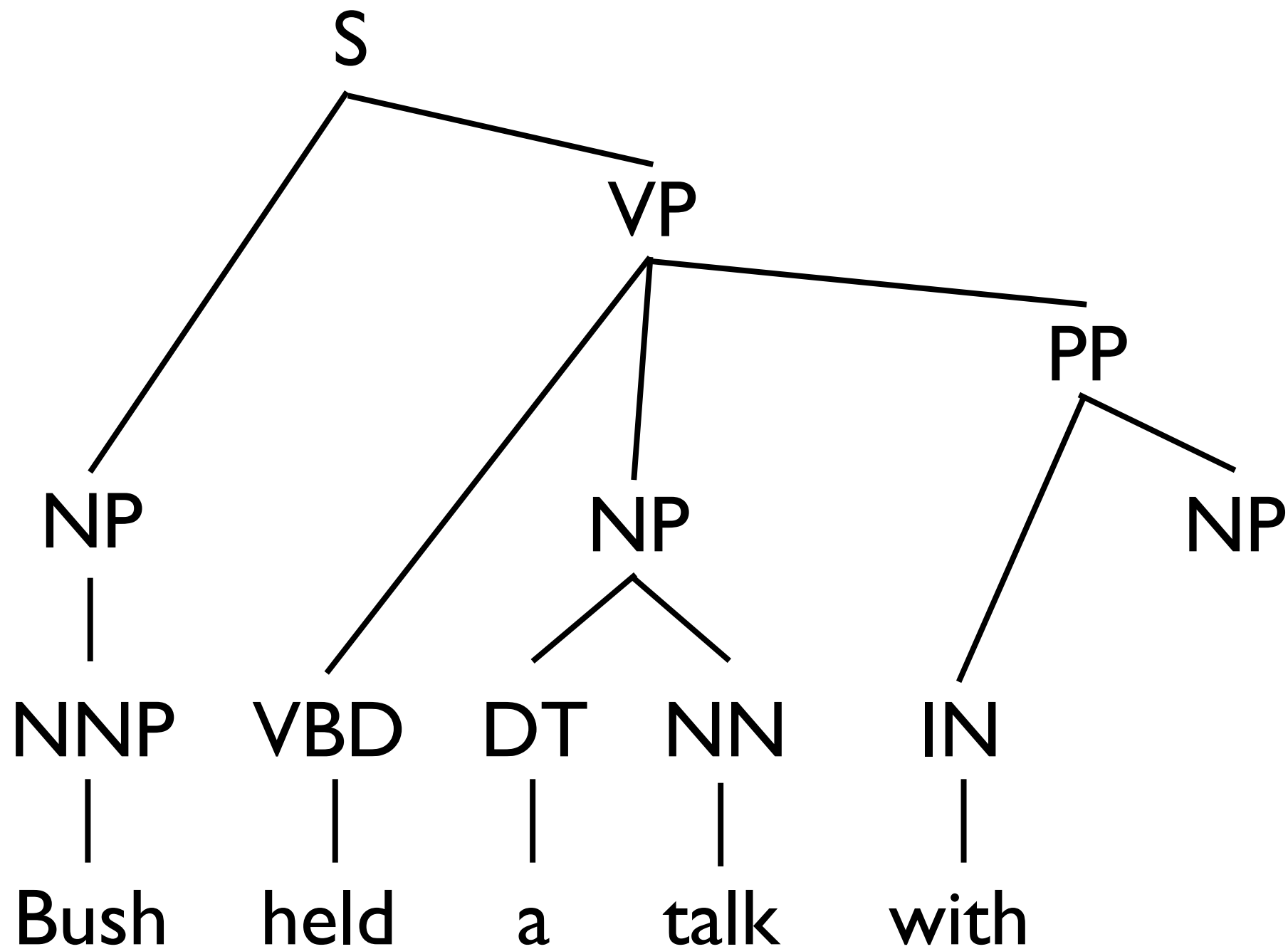
# Graphical Representation



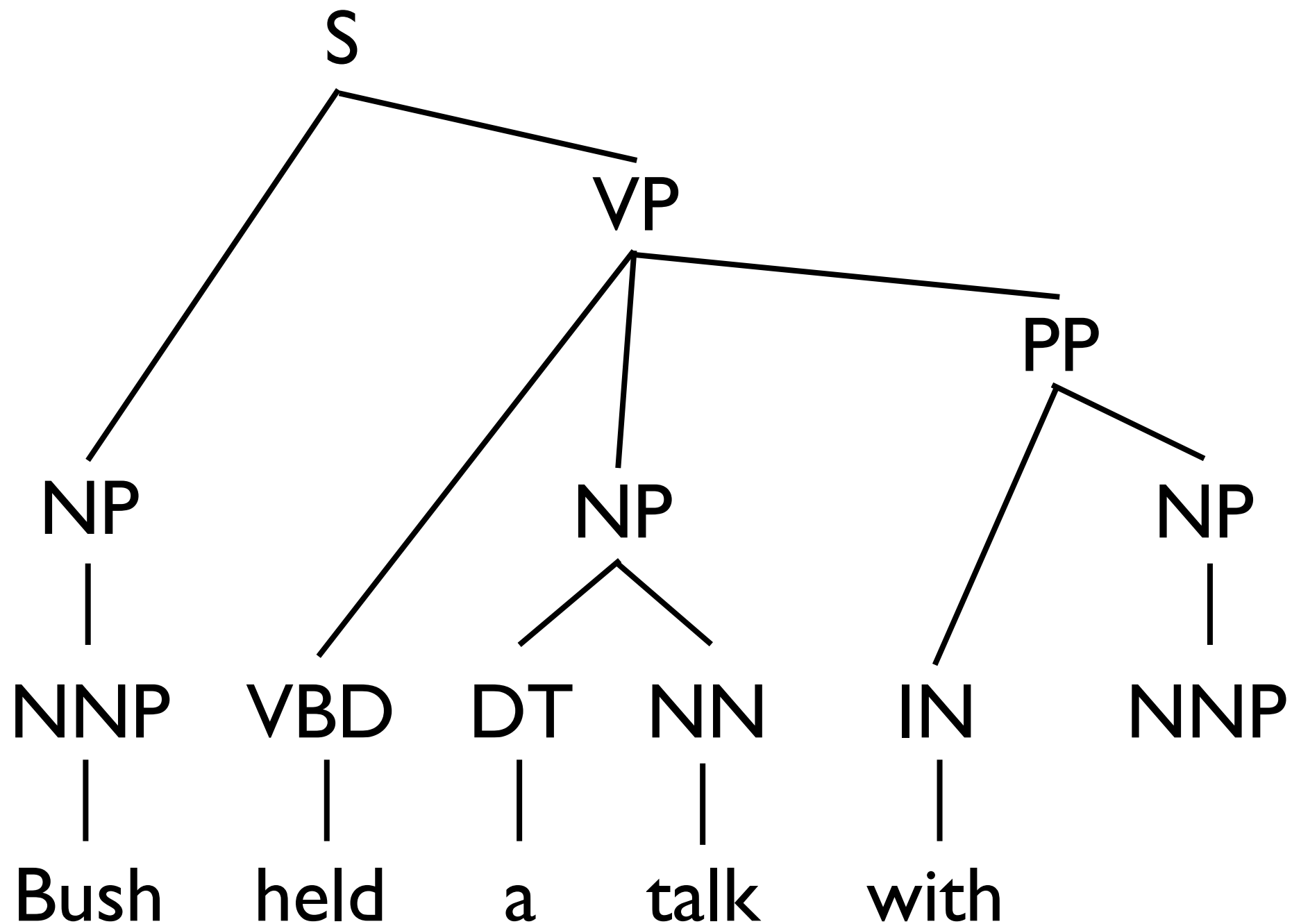
# Graphical Representation



# Graphical Representation

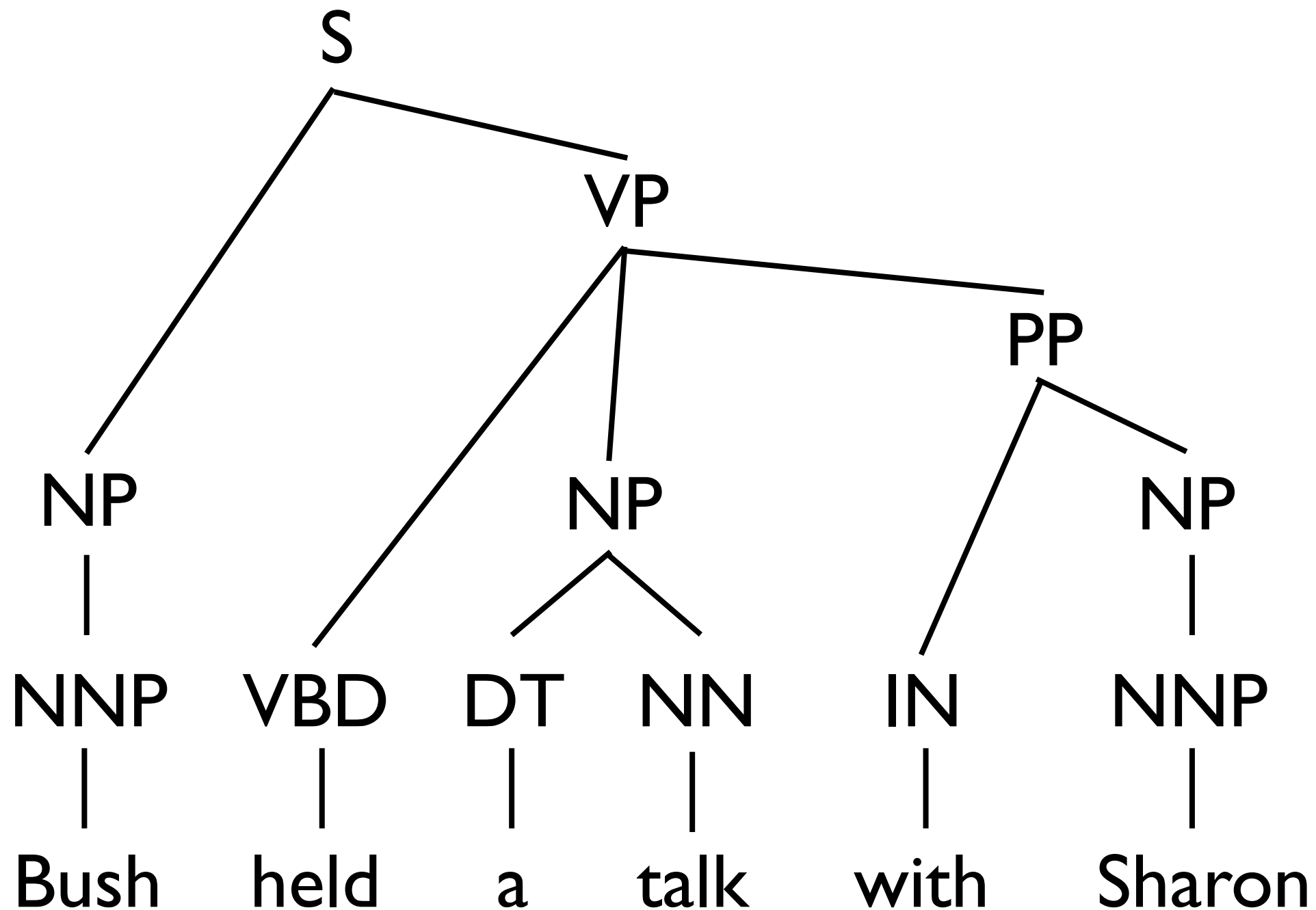


# Graphical Representation





# Graphical Representation



# Synchronous Context-Free Grammar

- Synchronous Context-free grammar describes how two natural language sentences are generated simultaneously

# Synchronous Context-Free Grammar

- **Synchronous Context-free grammar** describes how two natural language sentences are generated simultaneously

**lexical rules**       $NN \rightarrow \langle \text{bushi}, \text{Bush} \rangle$

# Synchronous Context-Free Grammar

- **Synchronous Context-free grammar** describes how two natural language sentences are generated simultaneously

**lexical rules**       $NN \rightarrow \langle \text{bushi}, \text{Bush} \rangle$

**syntactic rules**       $S \rightarrow \langle NP_1 \ VP_2, NP_1 \ VP_2 \rangle$

# Synchronous Context-Free Grammar

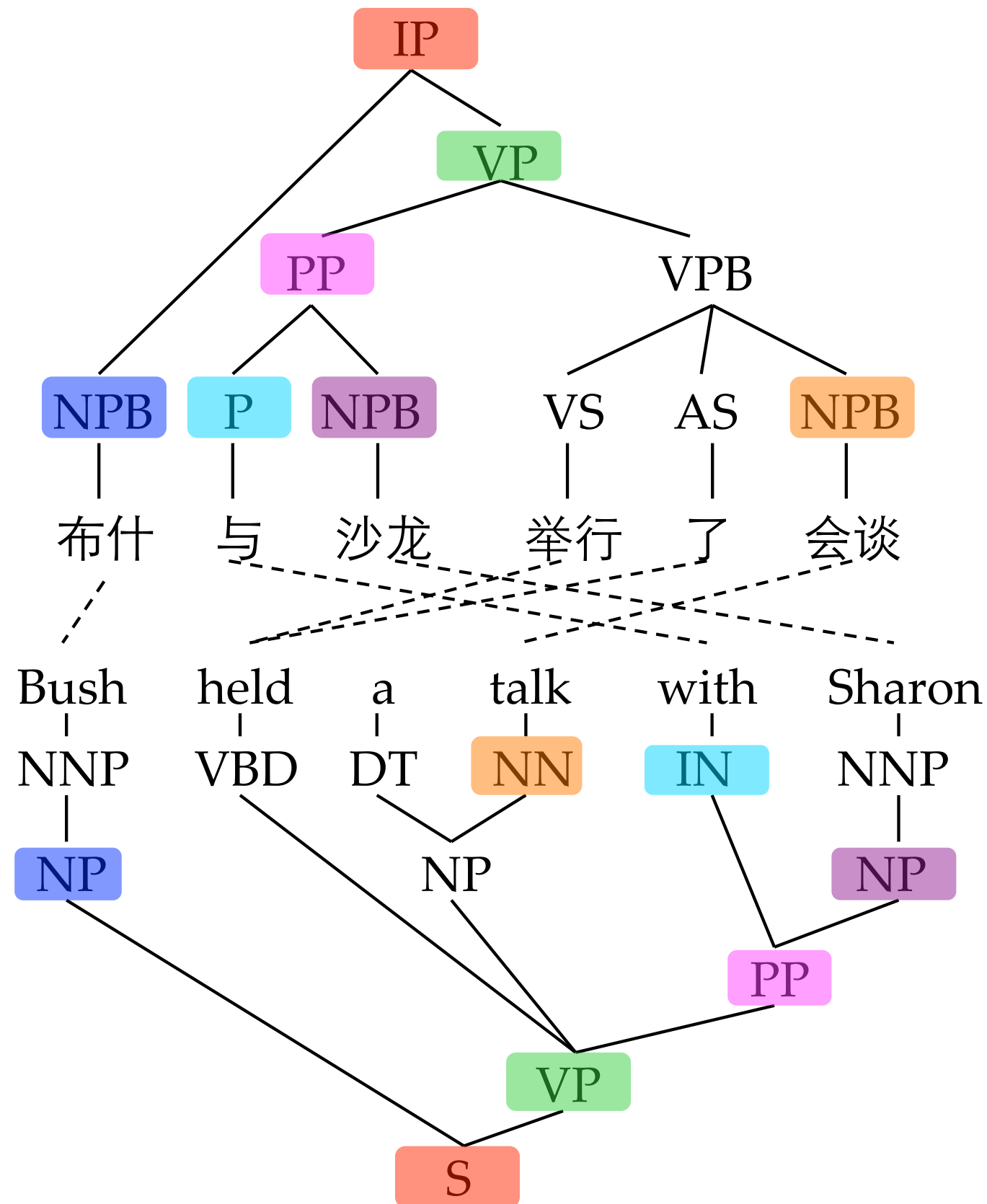
- **Synchronous Context-free grammar** describes how two natural language sentences are generated simultaneously

**lexical rules**       $NN \rightarrow \langle \text{bushi}, \text{Bush} \rangle$

**syntactic rules**       $S \rightarrow \langle NP_1 \ VP_2, NP_1 \ VP_2 \rangle$

Unfortunately, SCFG suffers from the **non-isomorphism** problem

# Non-Isomorphism



# Syntax-based MT

SCFGs without linguistic syntax

*inverted transduction grammar*

*hierarchical phrase-based model*

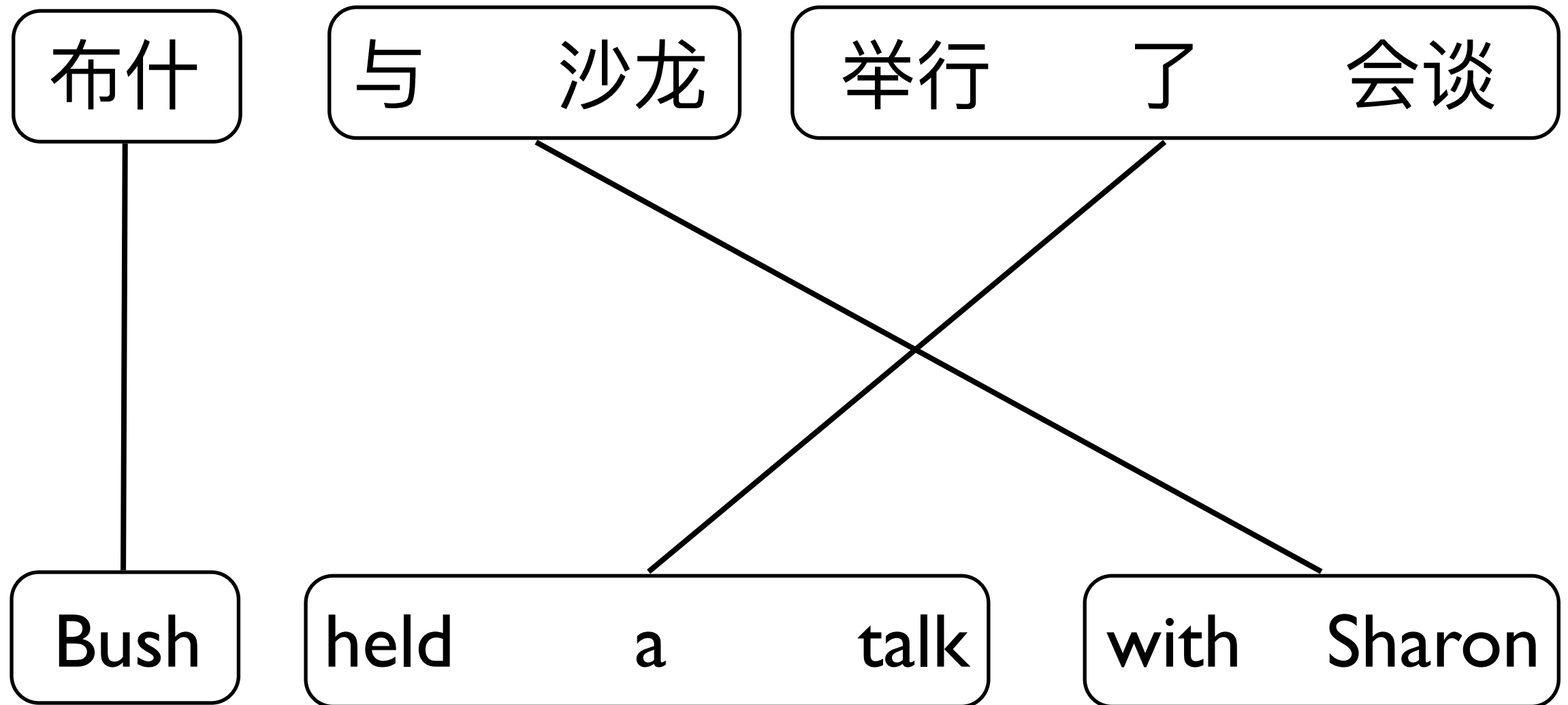
STSGs with linguistic syntax

*string-to-tree*

*tree-to-string*

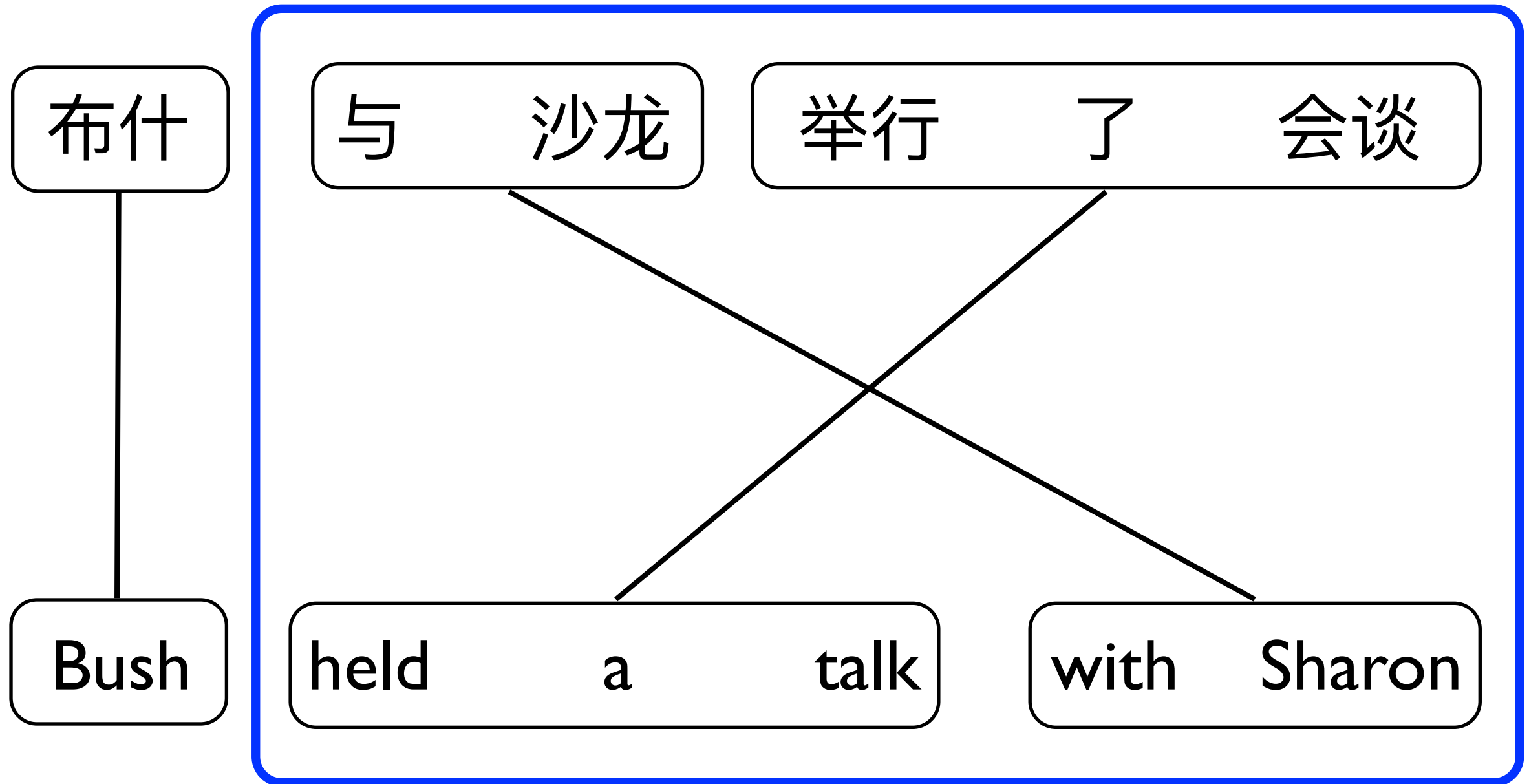
*tree-to-tree*

# Block Merging

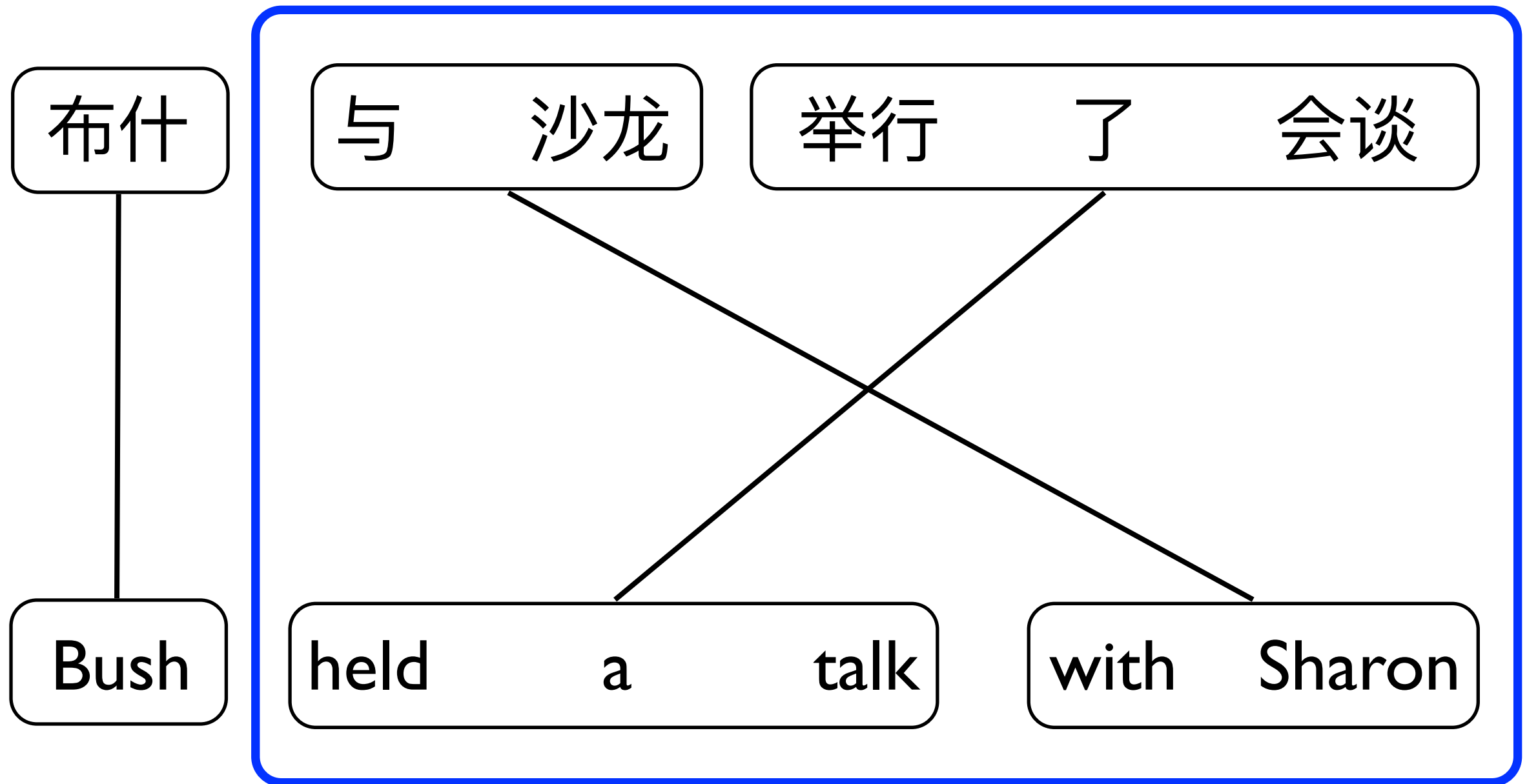




# Block Merging



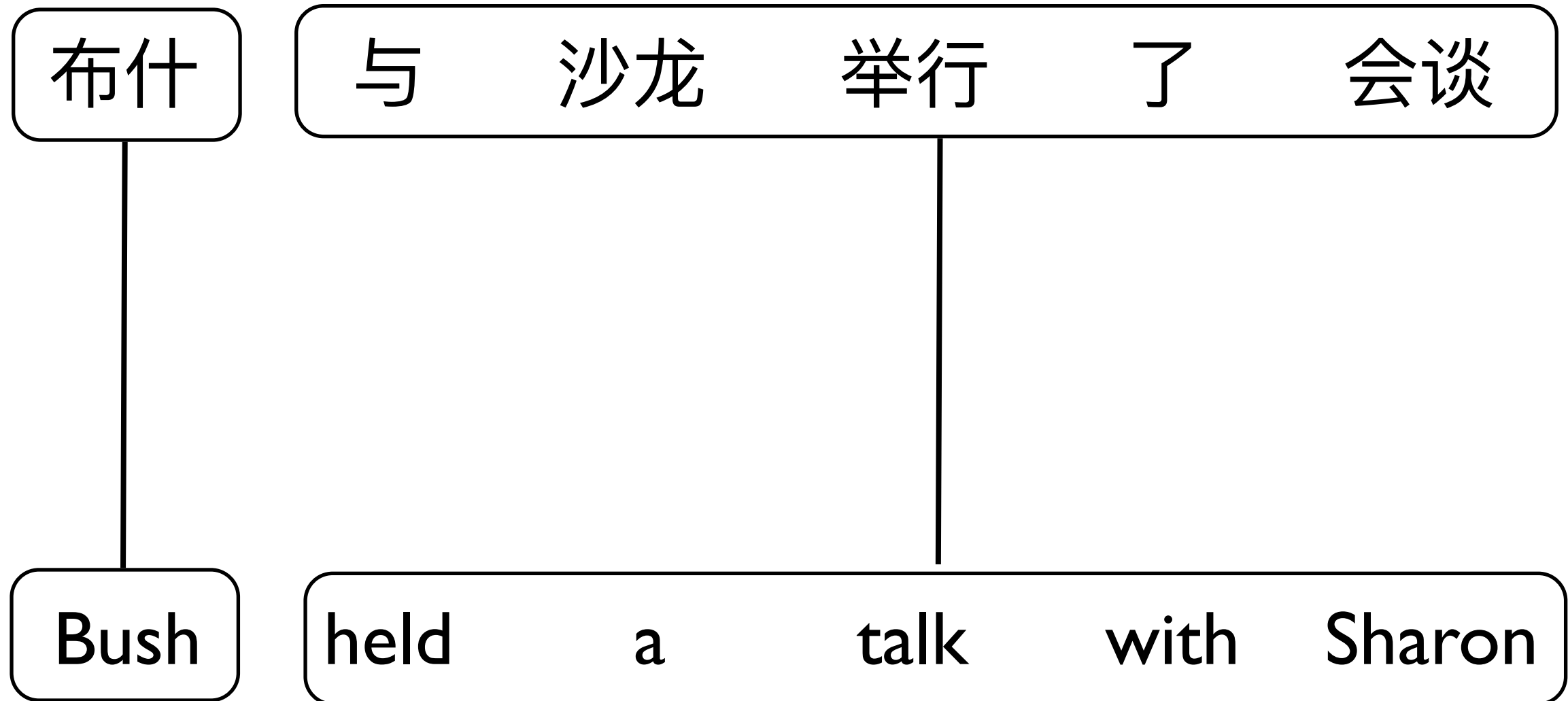
# Block Merging



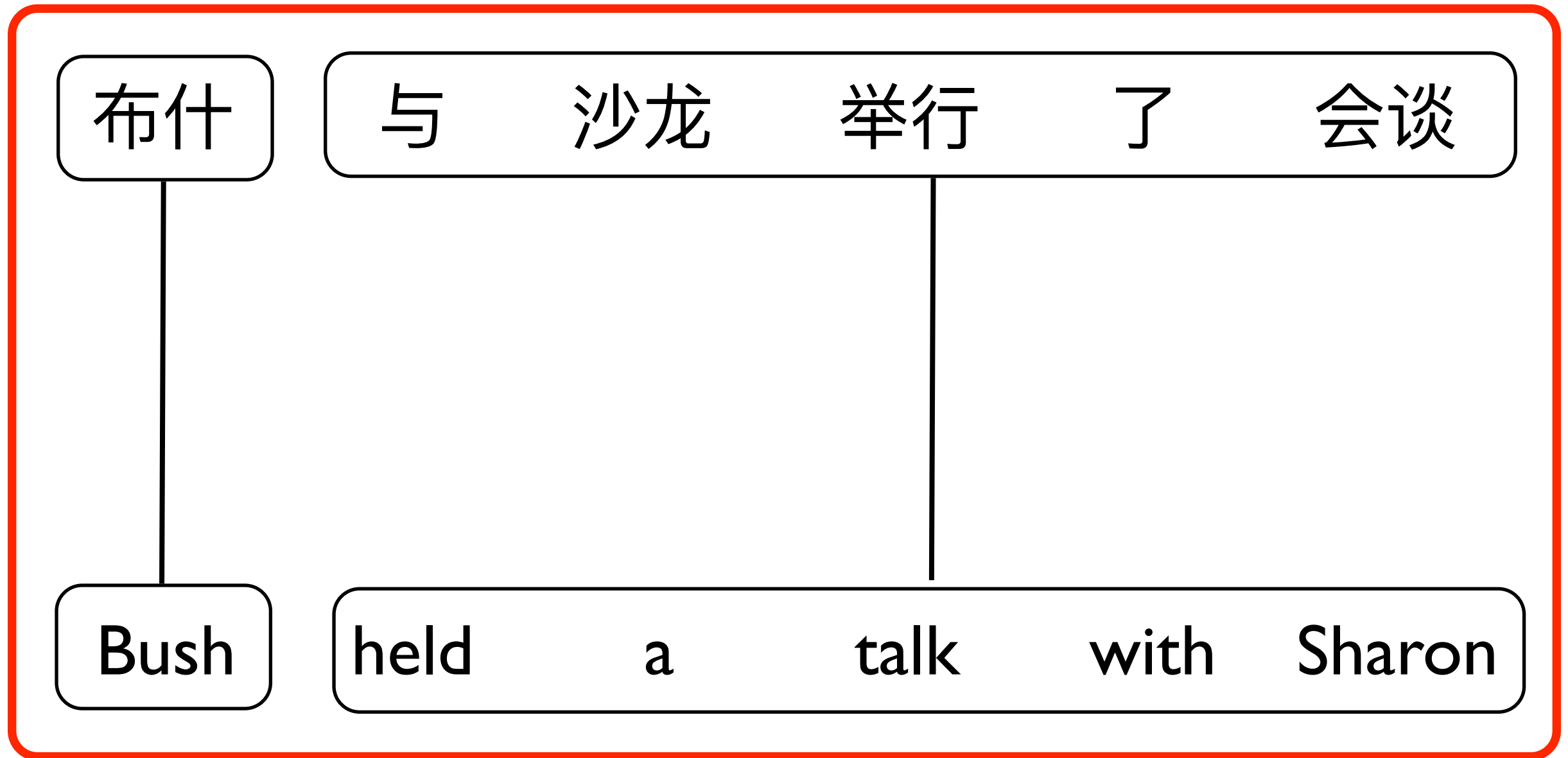
inverted

(Wu, 1997; Xiong et al., 2006)

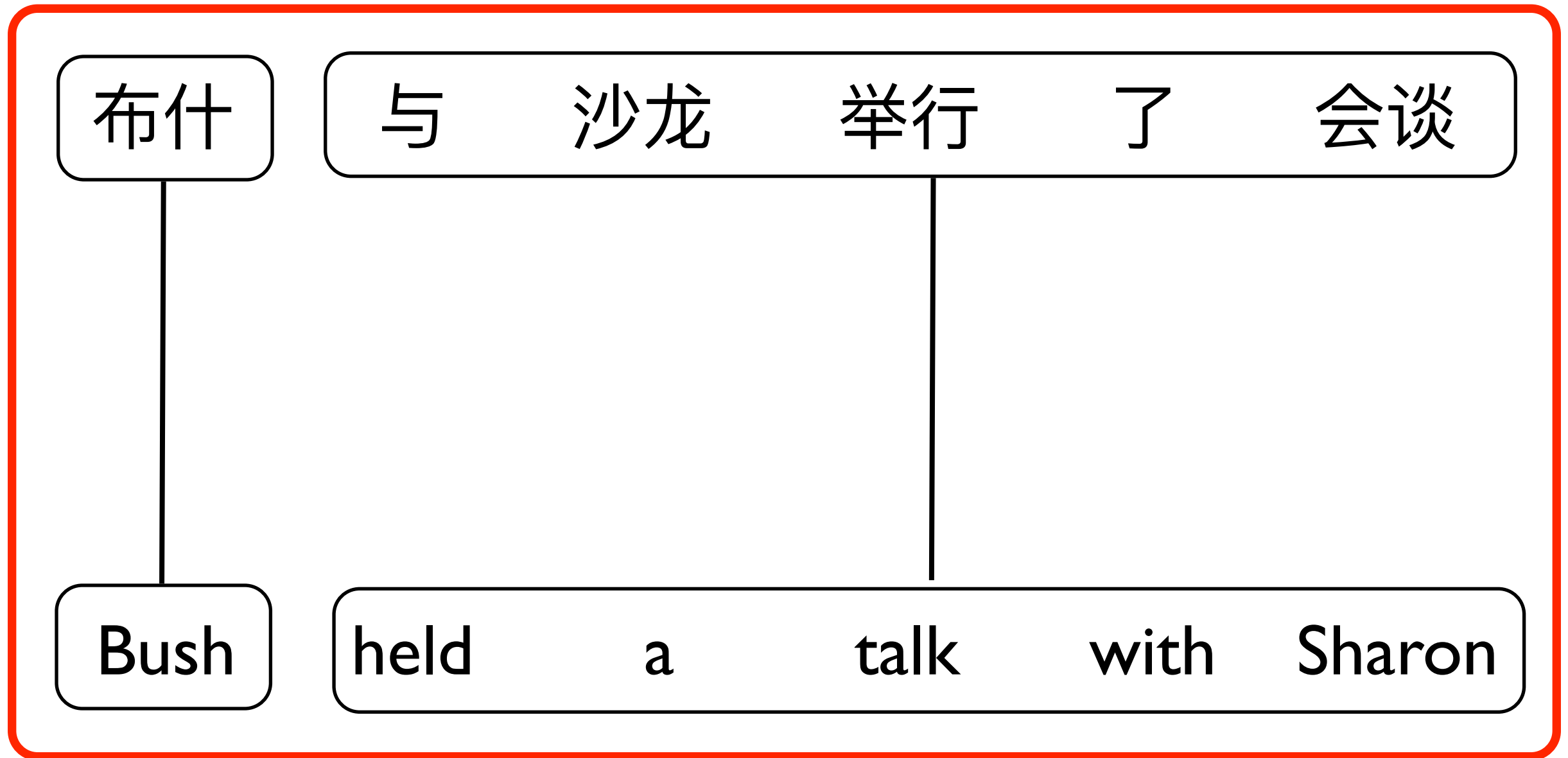
# Block Merging



# Block Merging



# Block Merging



straight

(Wu, 1997; Xiong et al., 2006)

# Block Merging

布什 与 沙龙 举行 了 会谈

Bush held a talk with Sharon

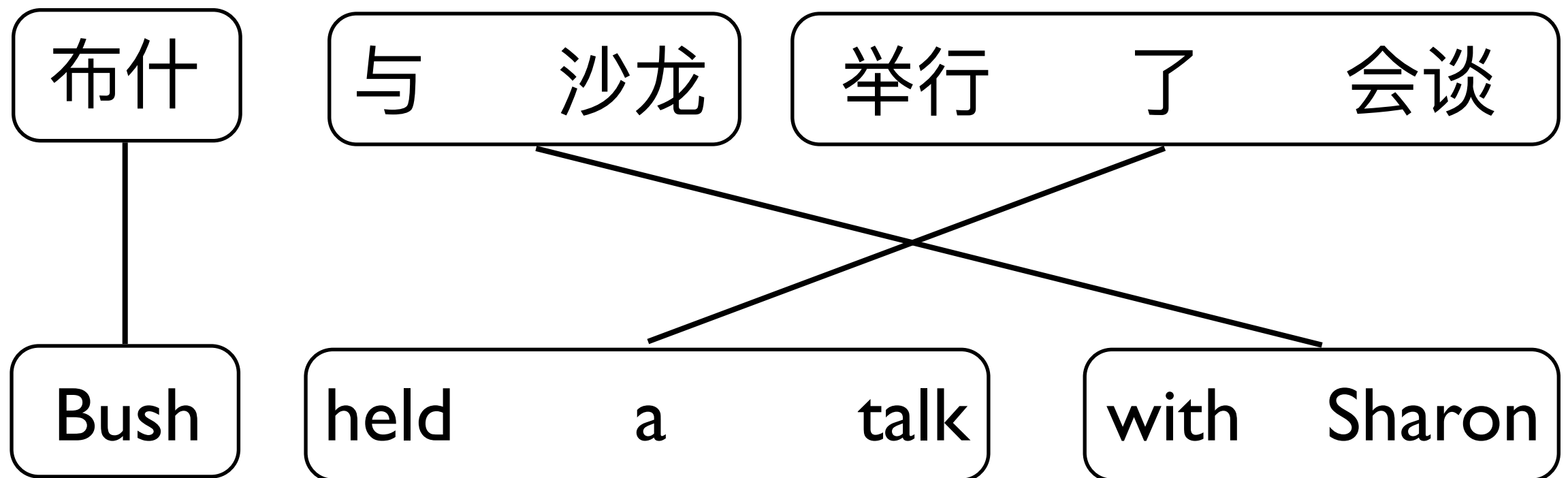
# Block Merging

布什 与 沙龙 举行 了 会谈

Bush held a talk with Sharon

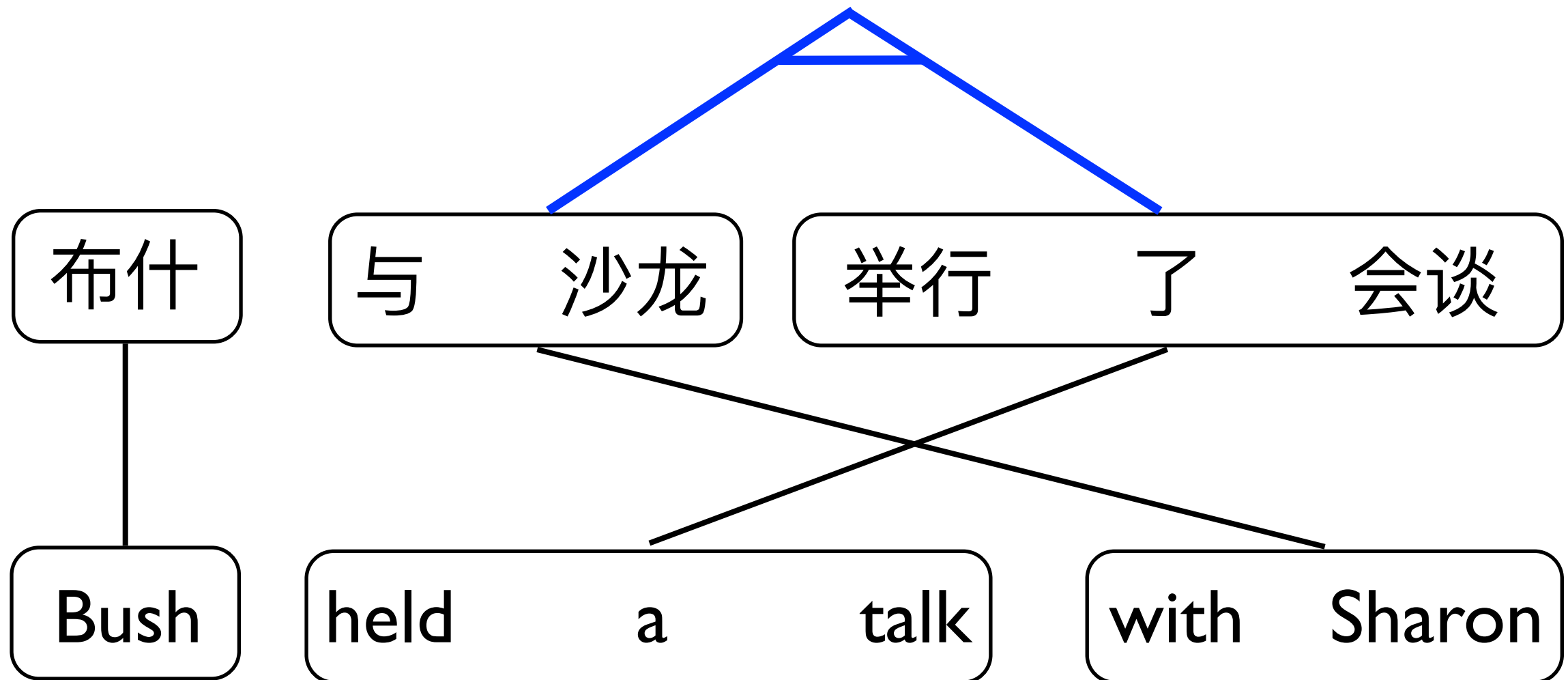
describe reordering using two operators

# Block Merging

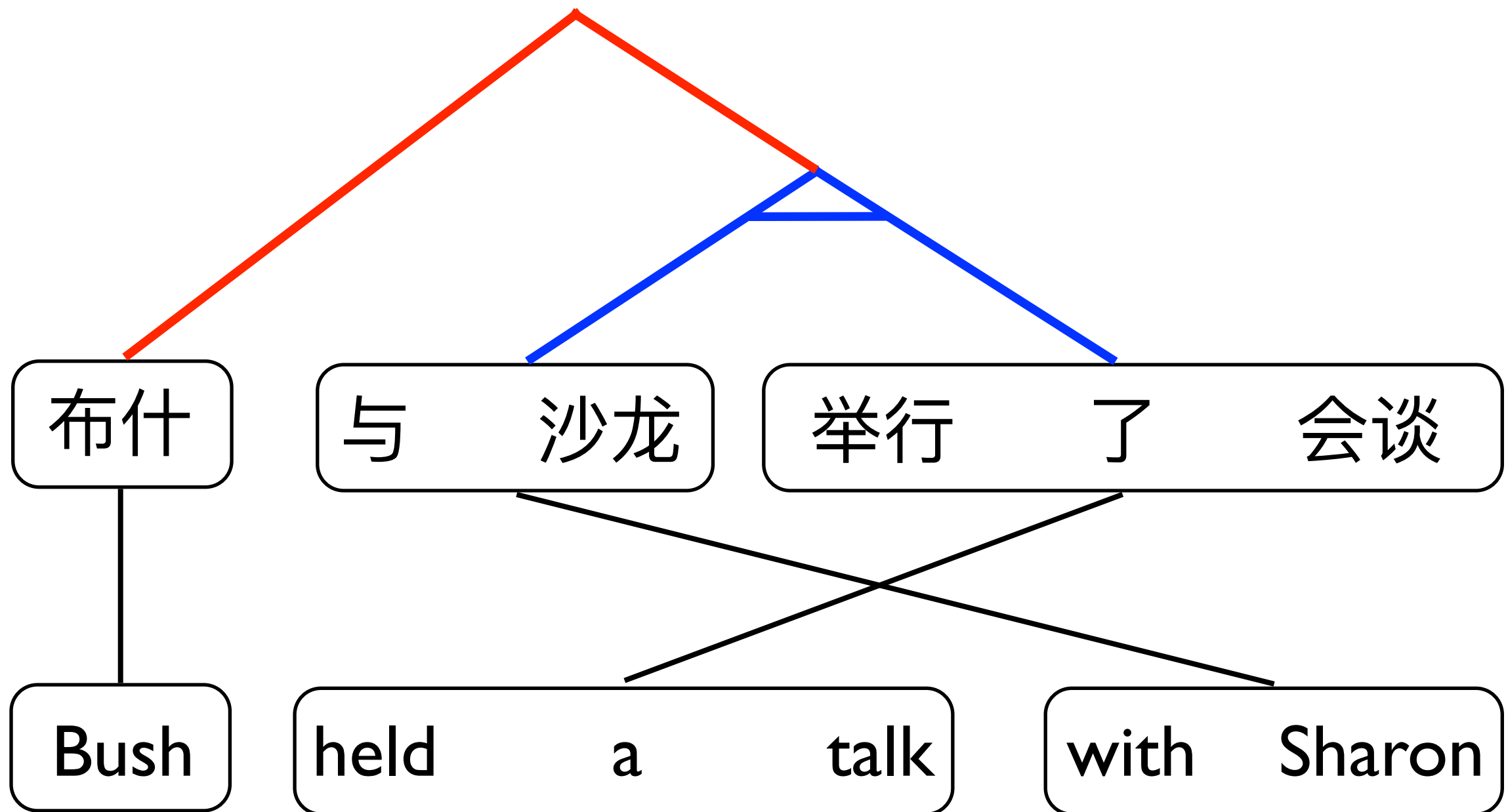




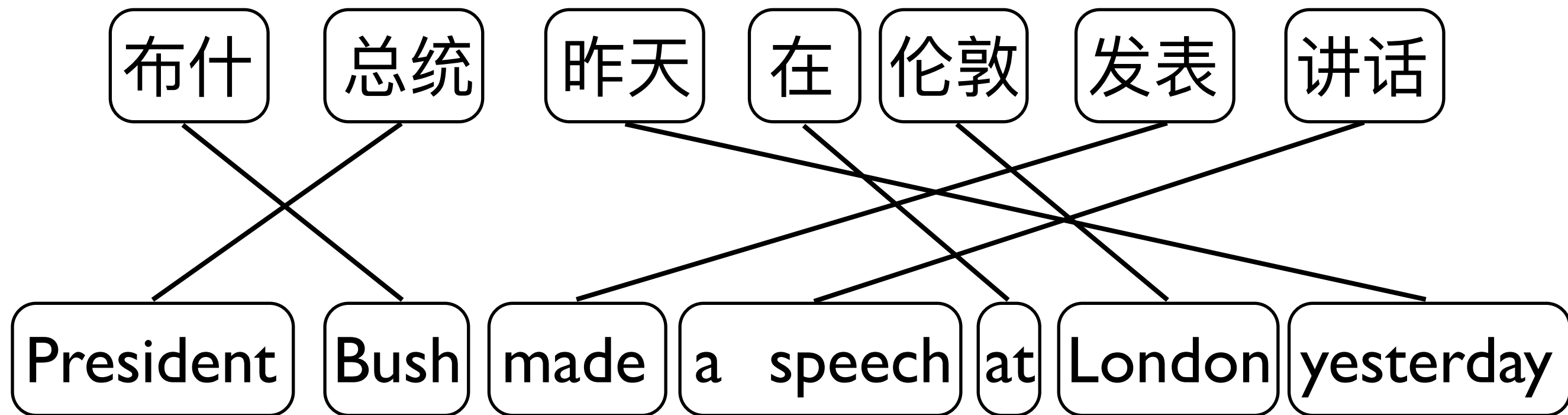
# Block Merging



# Block Merging

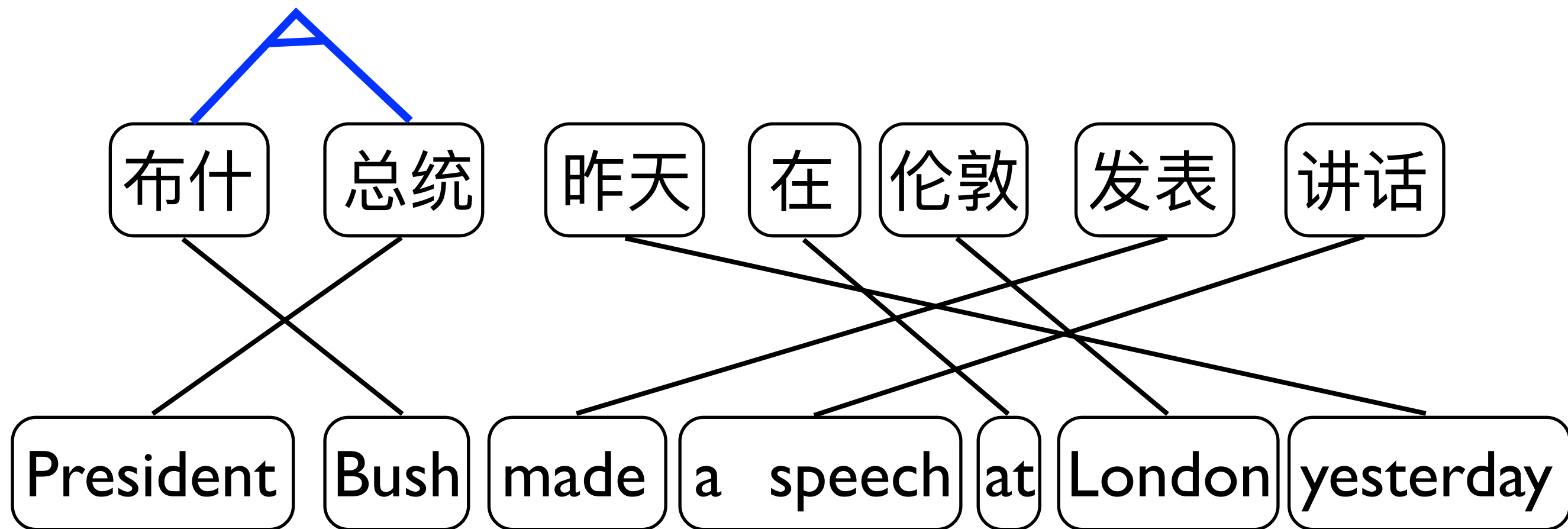


# Block Merging



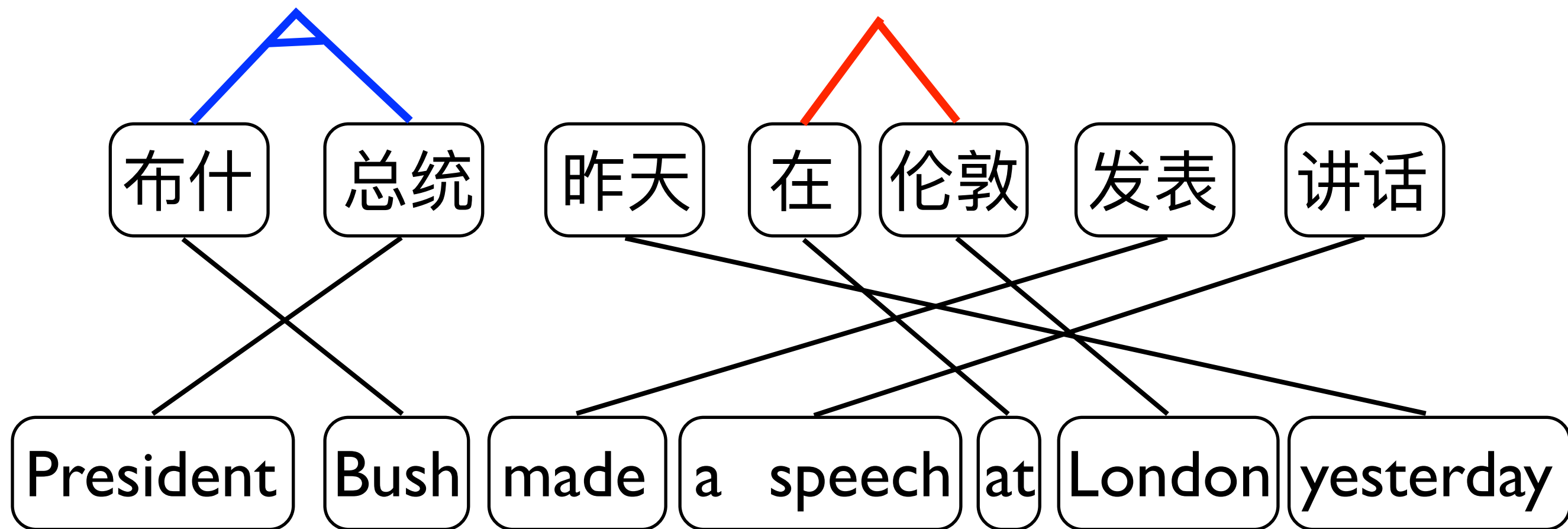
(Wu, 1997; Xiong et al., 2006)

# Block Merging



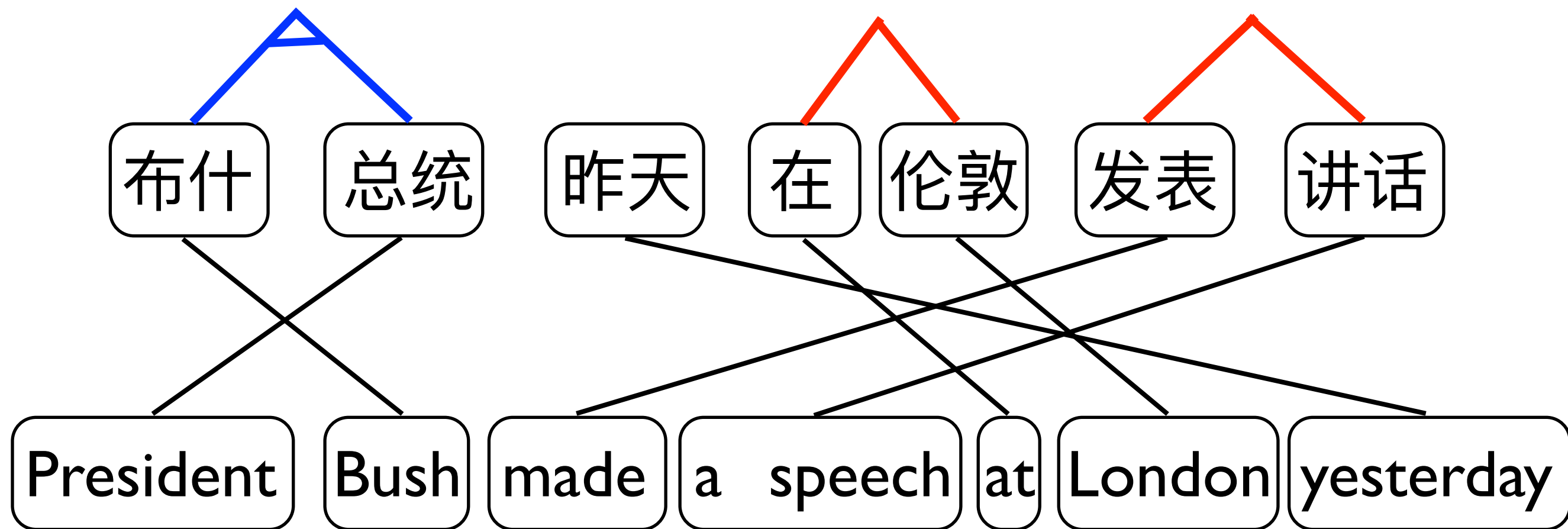
(Wu, 1997; Xiong et al., 2006)

# Block Merging



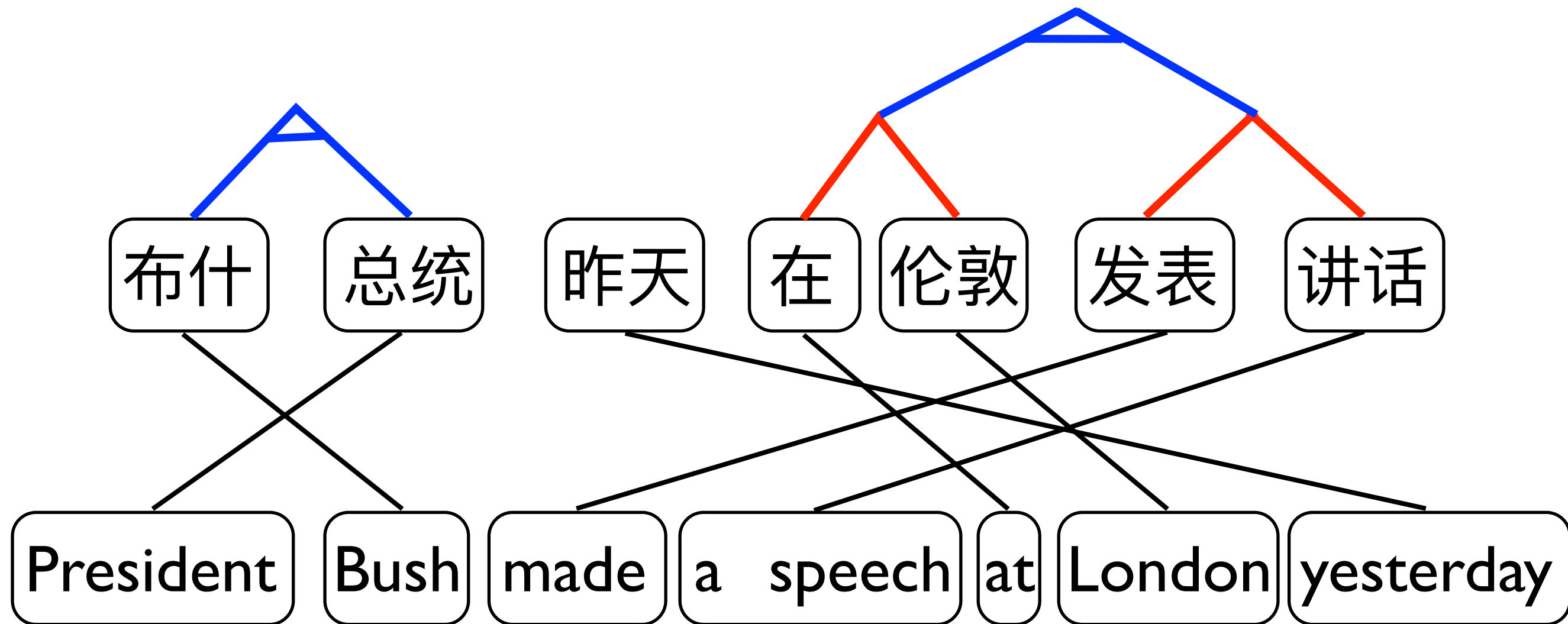
(Wu, 1997; Xiong et al., 2006)

# Block Merging



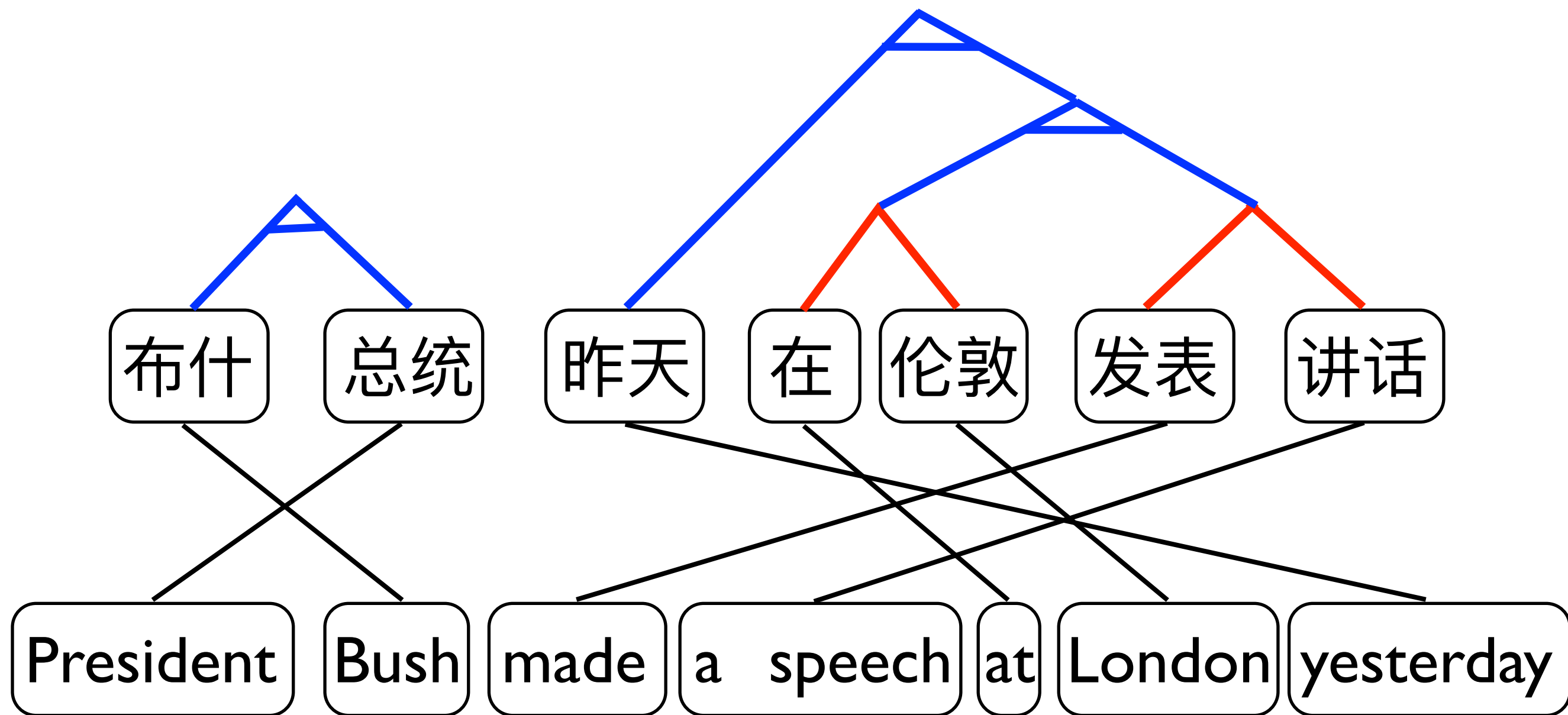
(Wu, 1997; Xiong et al., 2006)

# Block Merging



(Wu, 1997; Xiong et al., 2006)

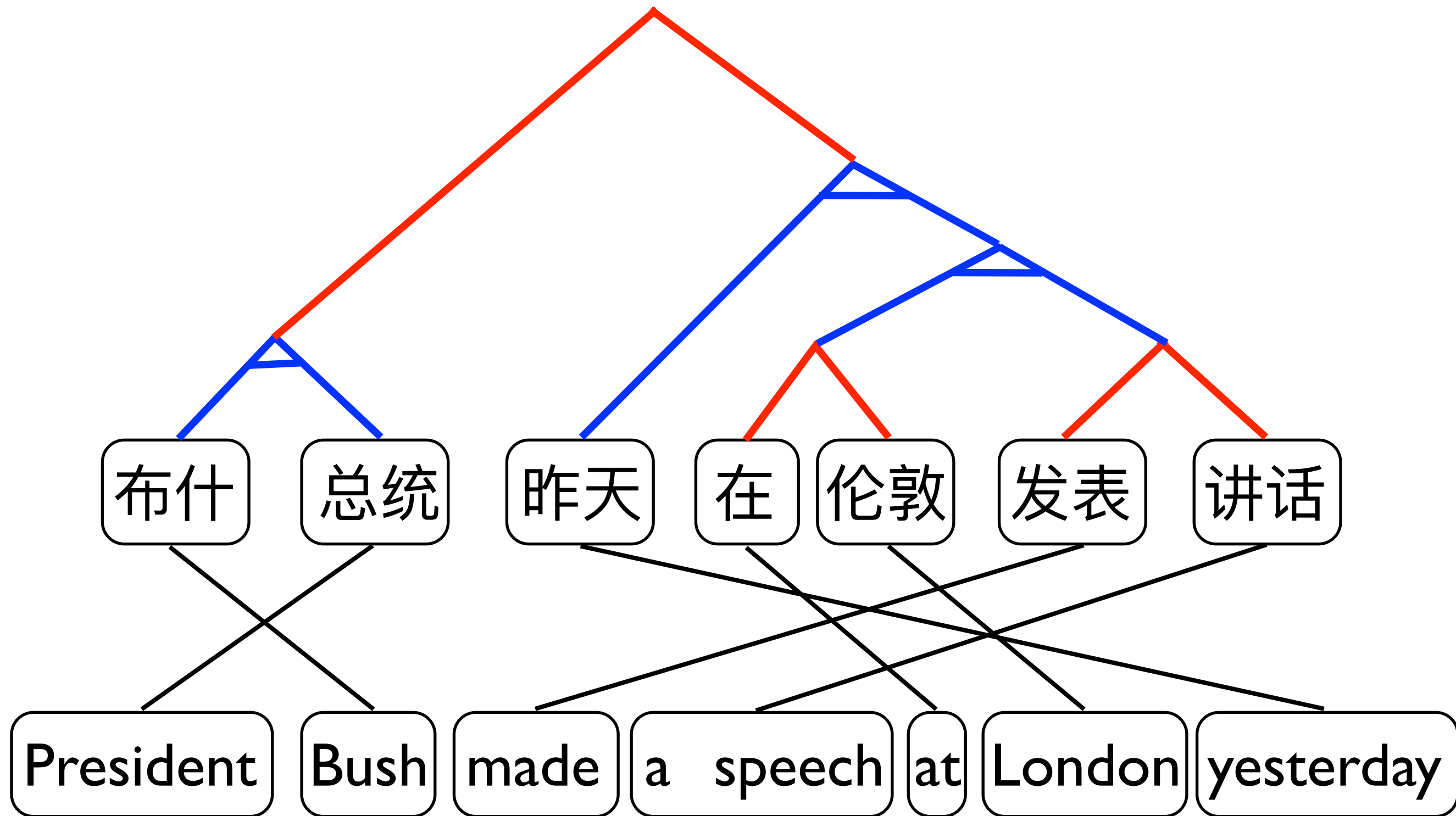
# Block Merging



(Wu, 1997; Xiong et al., 2006)



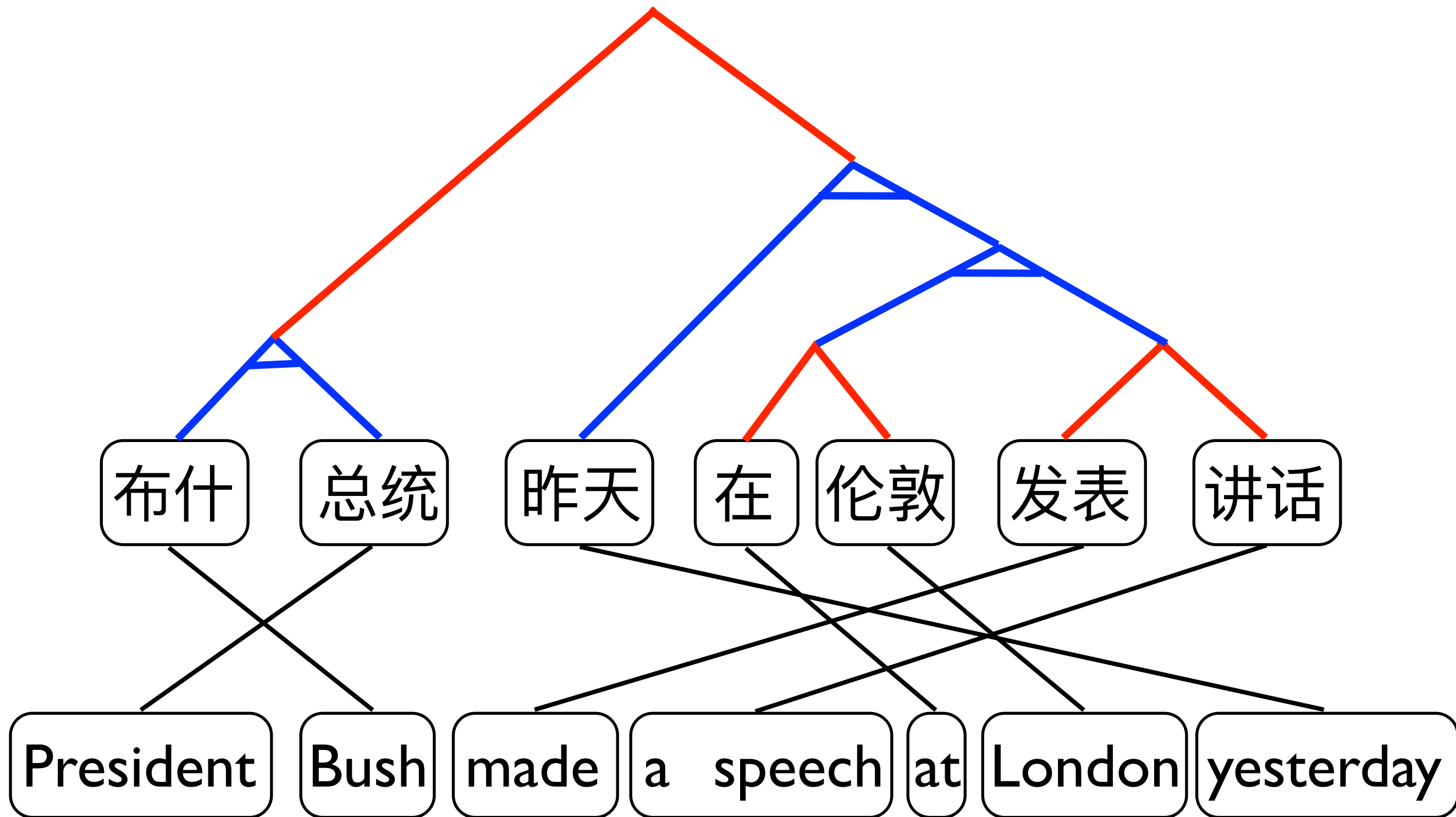
# Block Merging



(Wu, 1997; Xiong et al., 2006)

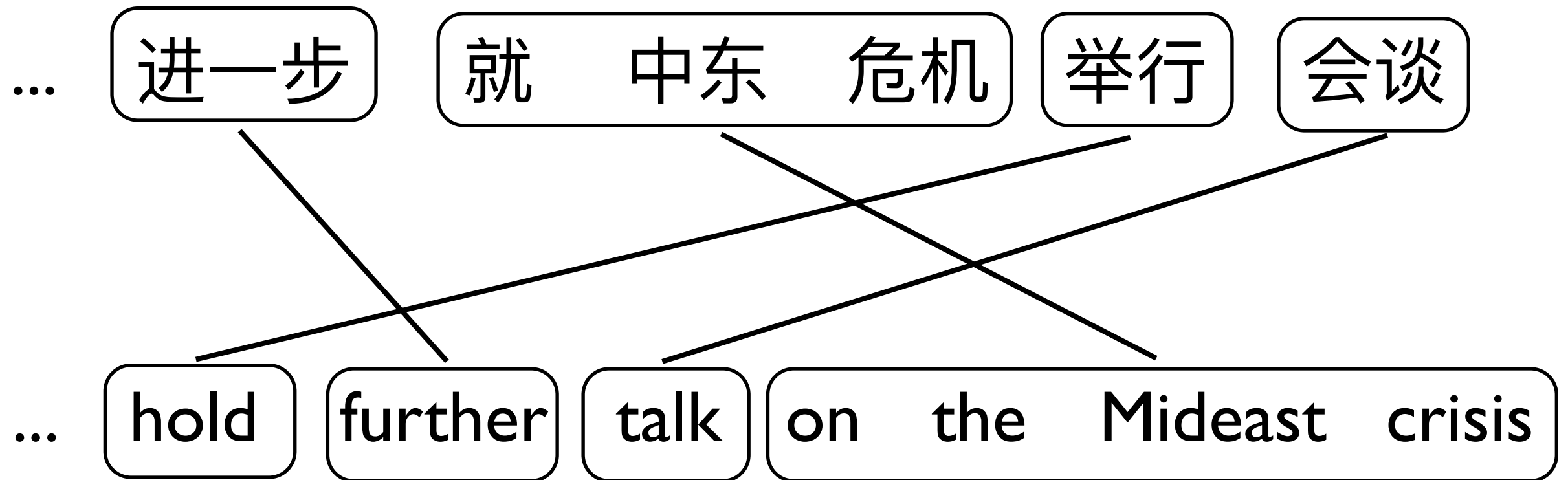
# Block Merging

Q: can you find a counter example?



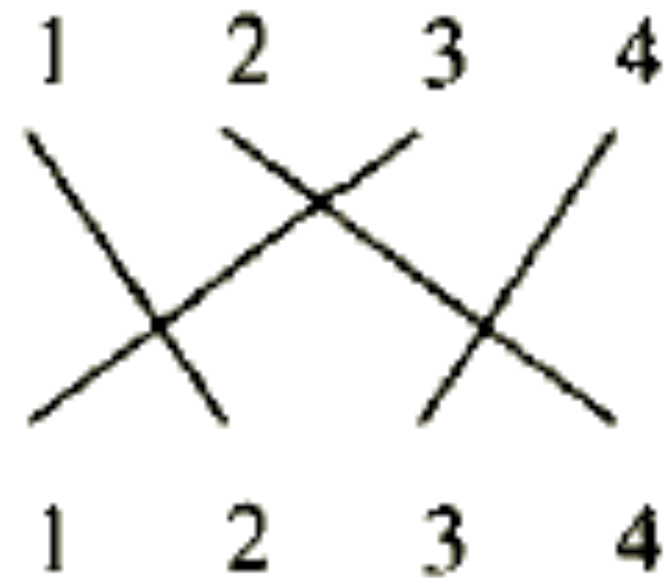
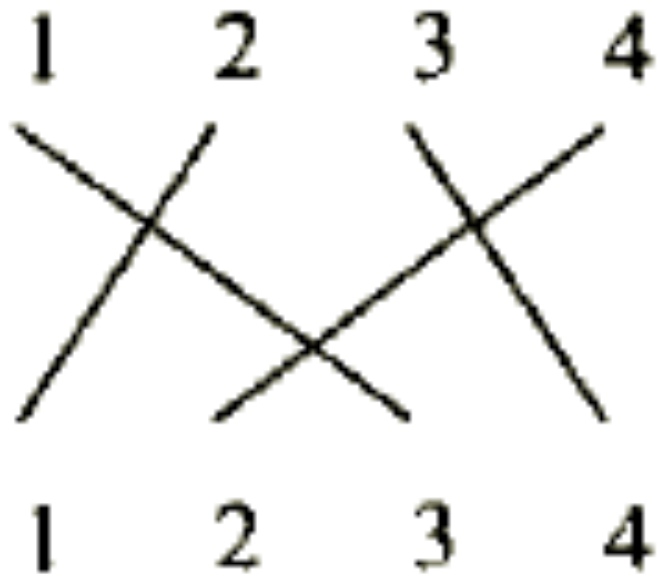
(Wu, 1997; Xiong et al., 2006)

# Counter Example



# Expressiveness of ITG

“inside-out”



(Wu, 1997)

# Inverted Transduction Grammar

- Inverted Transduction Grammar explains how two natural language sentences are generated synchronously using two block-merging operators

# Inverted Transduction Grammar

- Inverted Transduction Grammar explains how two natural language sentences are generated synchronously using two block-merging operators

$$X \rightarrow f/e$$

# Inverted Transduction Grammar

- Inverted Transduction Grammar explains how two natural language sentences are generated synchronously using two block-merging operators

lexical rules

$$X \rightarrow f/e$$

# Inverted Transduction Grammar

- Inverted Transduction Grammar explains how two natural language sentences are generated synchronously using two block-merging operators

lexical rules

$$X \rightarrow f/e$$

syntactic rules

$$X \rightarrow [X^1, X^2]$$

$$X \rightarrow \langle X^1, X^2 \rangle$$



# Inverted Transduction Grammar

- Inverted Transduction Grammar explains how two natural language sentences are generated synchronously using two block-merging operators

lexical rules

$$X \rightarrow f/e$$

syntactic rules

$$X \rightarrow [X^1, X^2] \quad \text{straight}$$

$$X \rightarrow \langle X^1, X^2 \rangle$$

# Inverted Transduction Grammar

- Inverted Transduction Grammar explains how two natural language sentences are generated synchronously using two block-merging operators

lexical rules



$$X \rightarrow f/e$$

syntactic rules

$$X \rightarrow [X^1, X^2] \quad \text{straight}$$

$$X \rightarrow \langle X^1, X^2 \rangle \quad \text{inverted}$$

# CKY Parsing

布什 与 沙龙 举行 了 会谈

# CKY Parsing

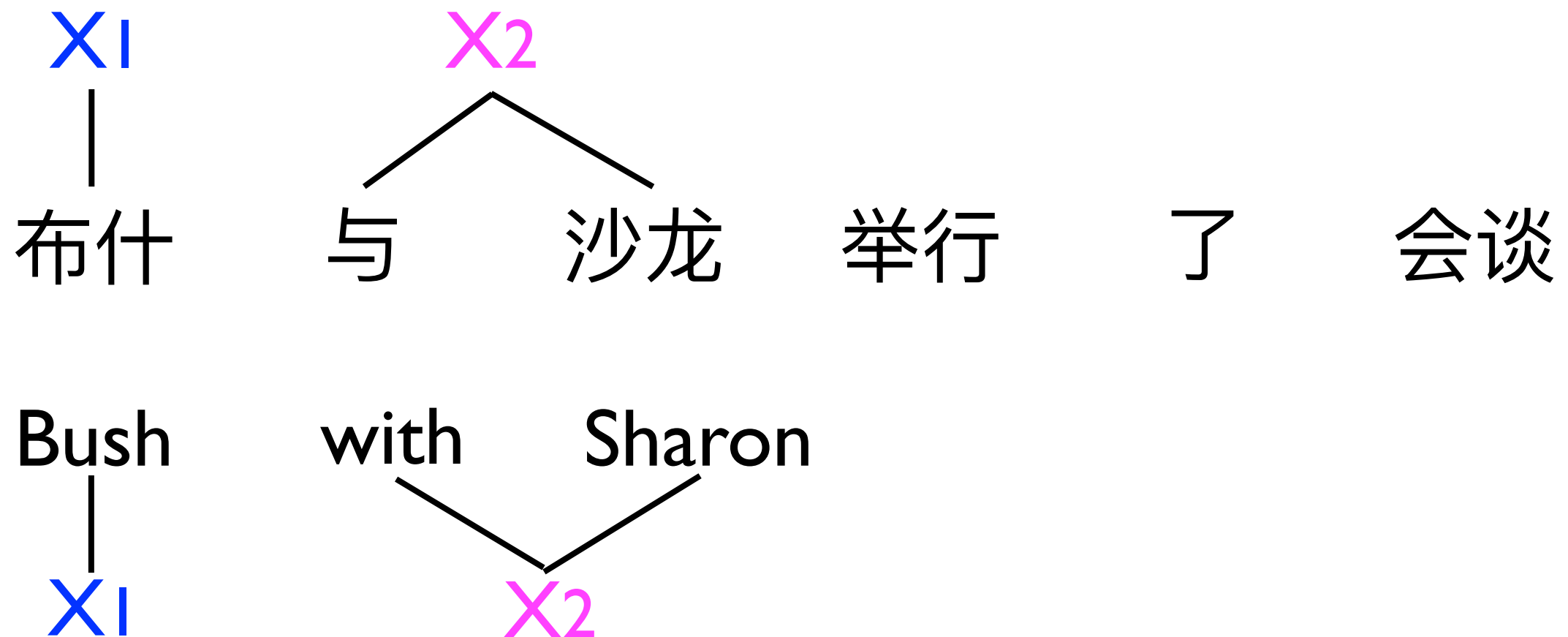
布什 与 沙龙 举行 了 会谈

Bush

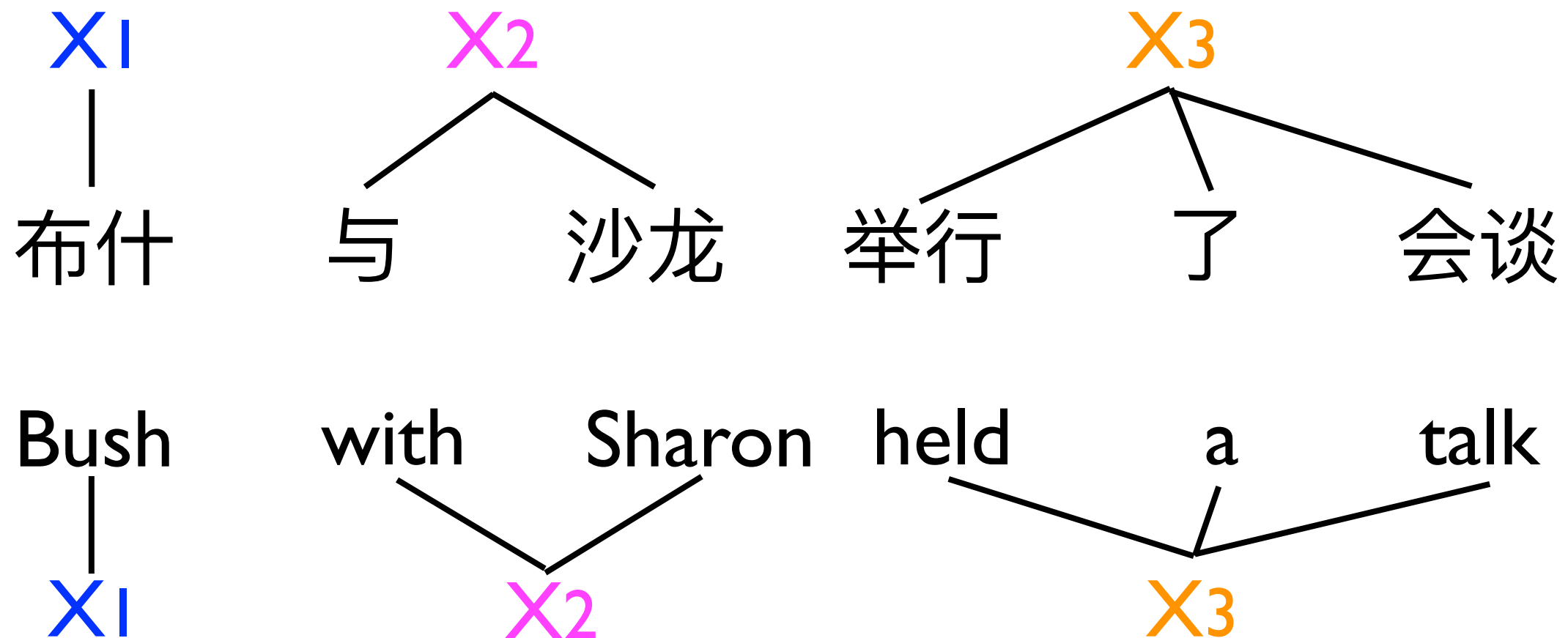
# CKY Parsing

Xi  
|  
布什 与 沙龙 举行 了 会谈  
Bush  
|  
Xi

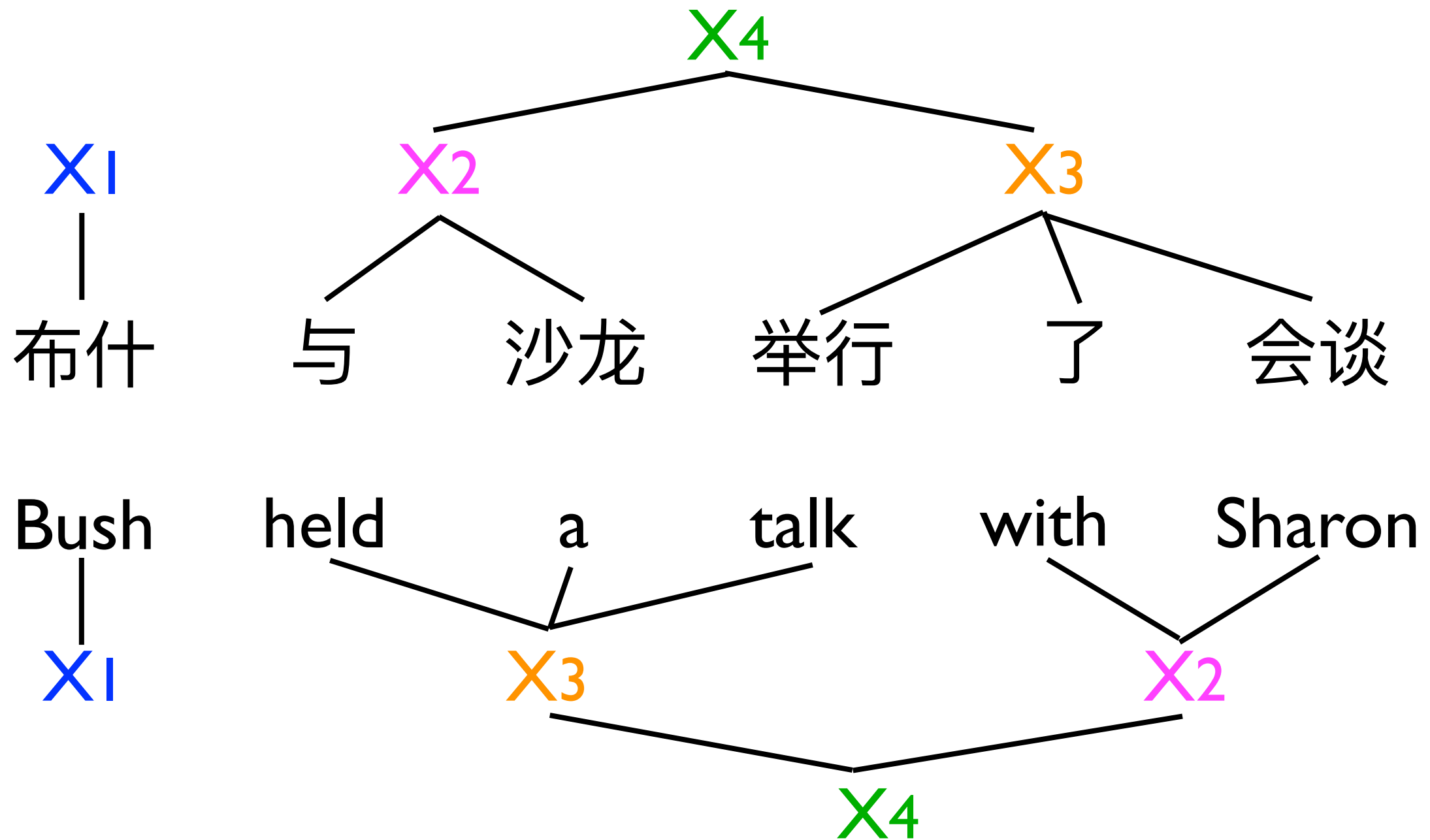
# CKY Parsing



# CKY Parsing

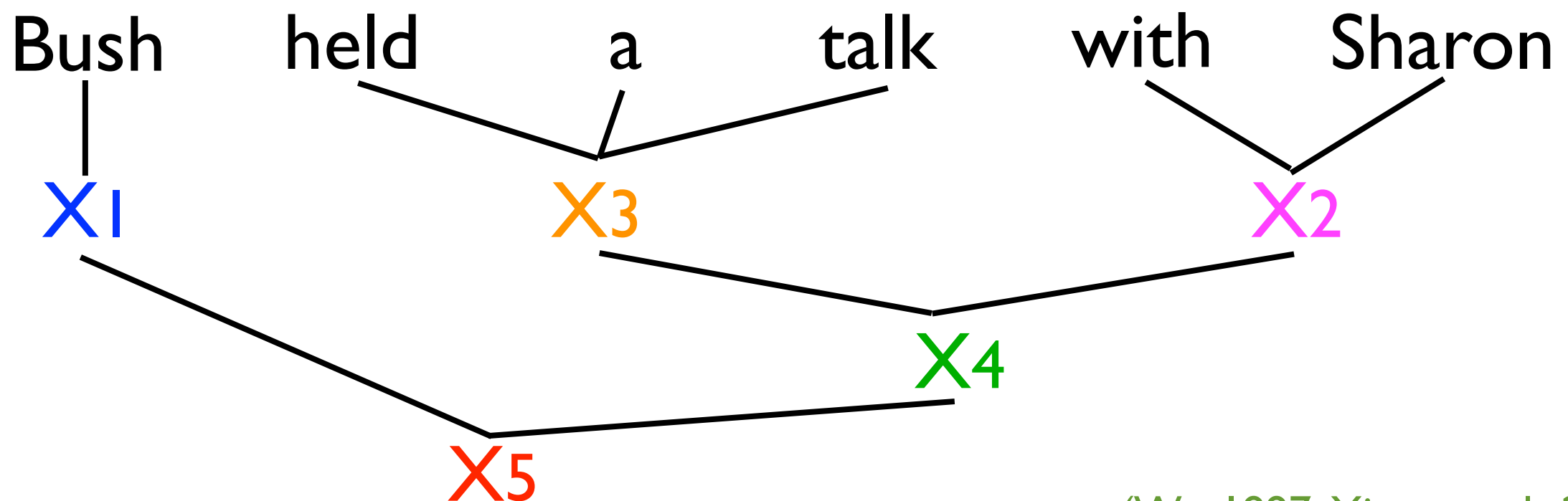
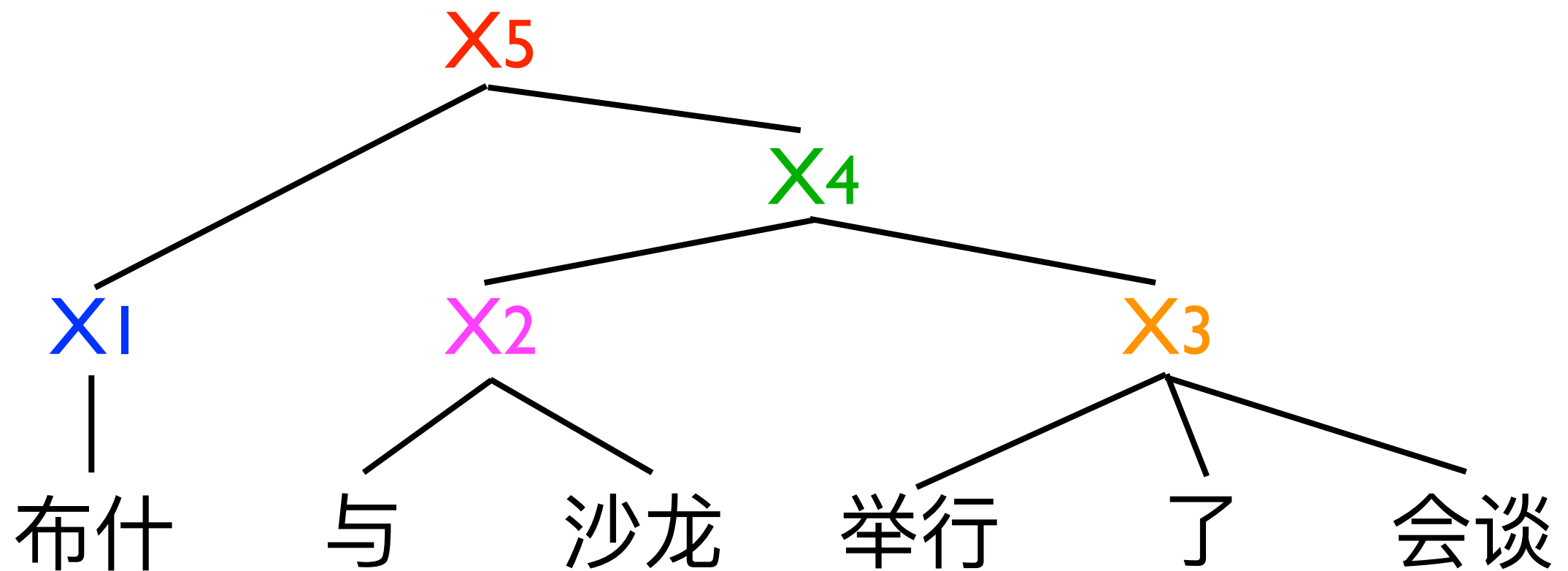


# CKY Parsing





# CKY Parsing



(Wu, 1997; Xiong et al., 2006)

# Chart

Bush ... Sharon					
	held ... Sharon				
			held ... talk		
	with Sharon		held		
Bush	with	Sharon			a talk

布什      与      沙龙      举行      了      会谈

# Syntax-based MT

SCFGs without linguistic syntax

*inverted transduction grammar*

*hierarchical phrase-based model*

STSGs with linguistic syntax

*string-to-tree*

*tree-to-string*

*tree-to-tree*

# Translation Templates

从北京到上海

from Beijing to Shanghai

从武汉到天津

from Wuhan to Tianjin

从广州到重庆

from Guangzhou to Chongqing

# Translation Templates

从北京到上海

from Beijing to Shanghai

从武汉到天津

from Wuhan to Tianjin

从广州到重庆

from Guangzhou to Chongqing

# Translation Templates

从北京到上海

from Beijing to Shanghai

从武汉到天津

from Wuhan to Tianjin

从广州到重庆

from Guangzhou to Chongqing

(从  $X_1$  到  $X_2$ , from  $X_1$  to  $X_2$ )

# Translation Templates

从北京到上海

from Beijing to Shanghai

从武汉到天津

from Wuhan to Tianjin

从广州到重庆

from Guangzhou to Chongqing

(从  $X_1$  到  $X_2$ , from  $X_1$  to  $X_2$ )

(北京, Beijing) (上海, Shanghai) (武汉, Wuhan)

(天津, Tianjin) (广州, Guangzhou) (重庆, Chongqing)

(Chiang, 2005)

# Translation Templates

从北京到上海

from Beijing to Shanghai

从武汉到天津

from Wuhan to Tianjin

从广州到重庆

from Guangzhou to Chongqing

(从  $X_1$  到  $X_2$ , from  $X_1$  to  $X_2$ ) *hierarchical phrase pair*

(北京, Beijing) (上海, Shanghai) (武汉, Wuhan)

(天津, Tianjin) (广州, Guangzhou) (重庆, Chongqing)

(Chiang, 2005)

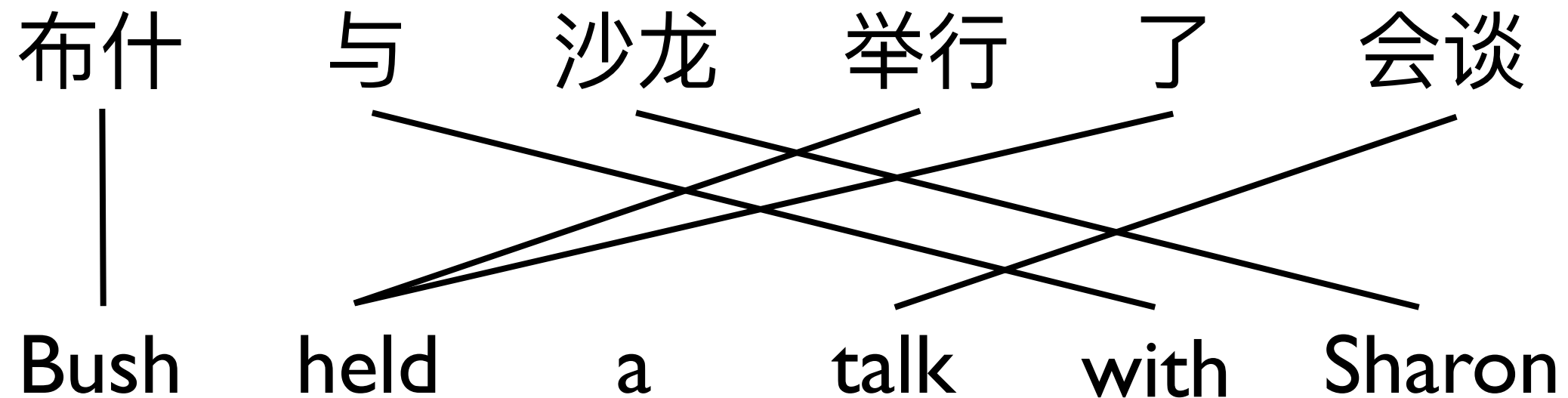


# Hierarchical Phrase Extraction

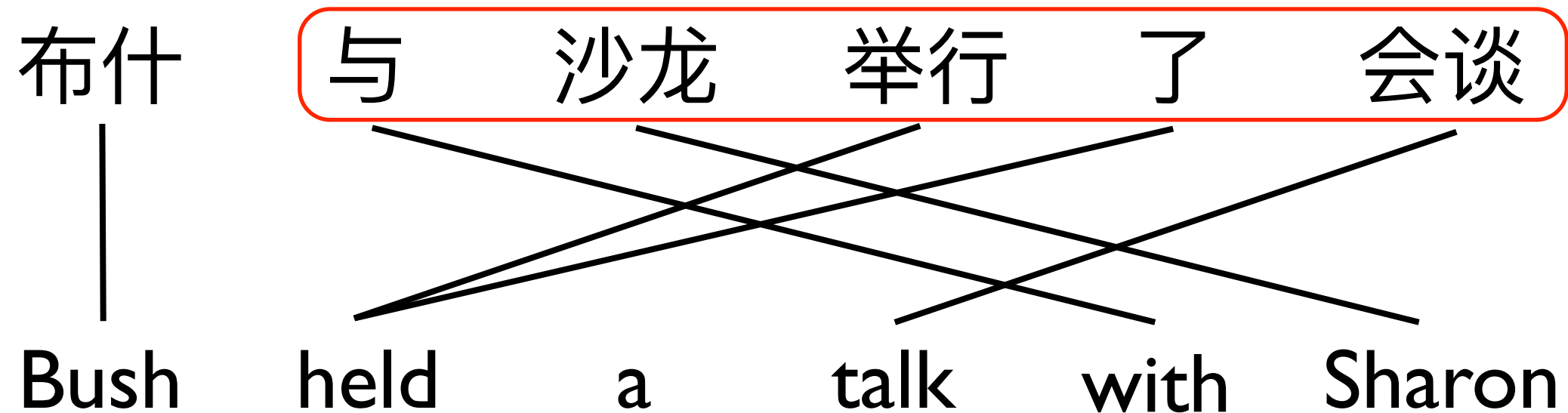
布什      与      沙龙      举行      了      会谈

Bush      held      a      talk      with      Sharon

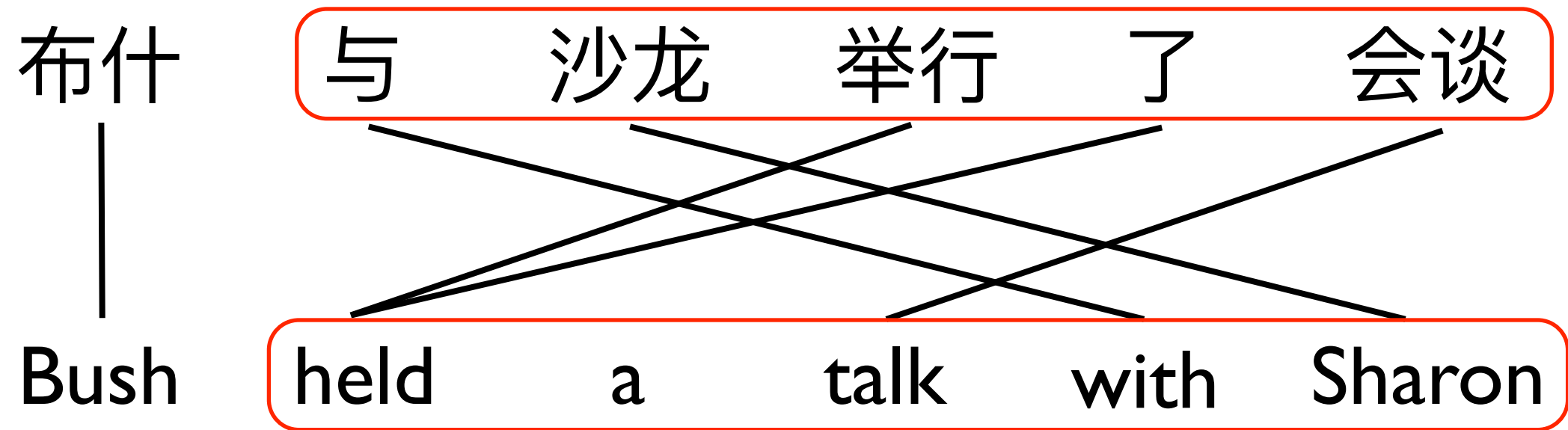
# Hierarchical Phrase Extraction



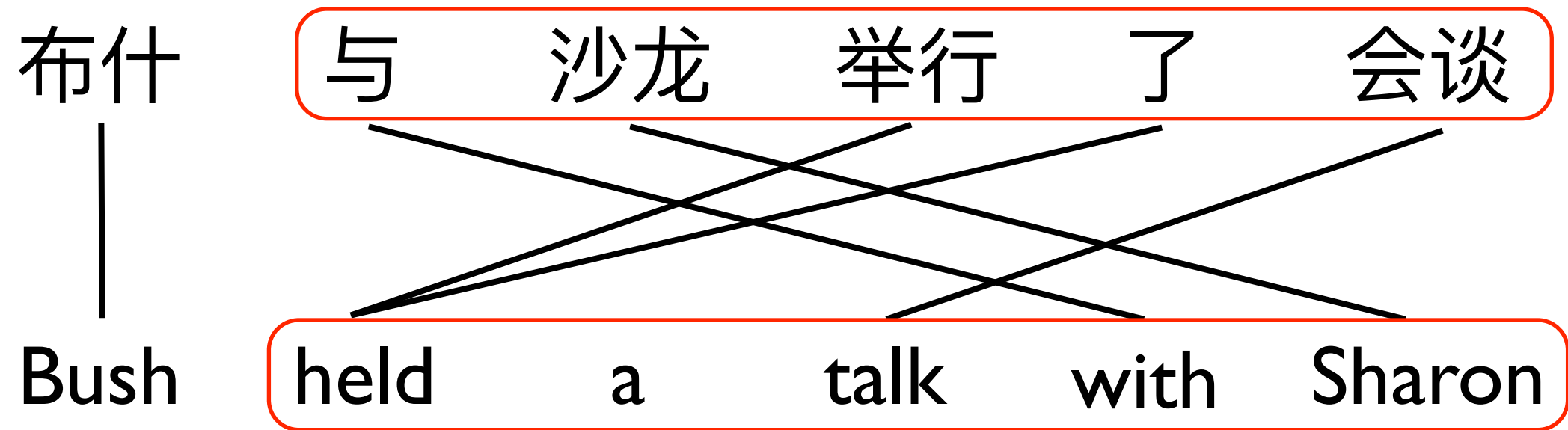
# Hierarchical Phrase Extraction



# Hierarchical Phrase Extraction

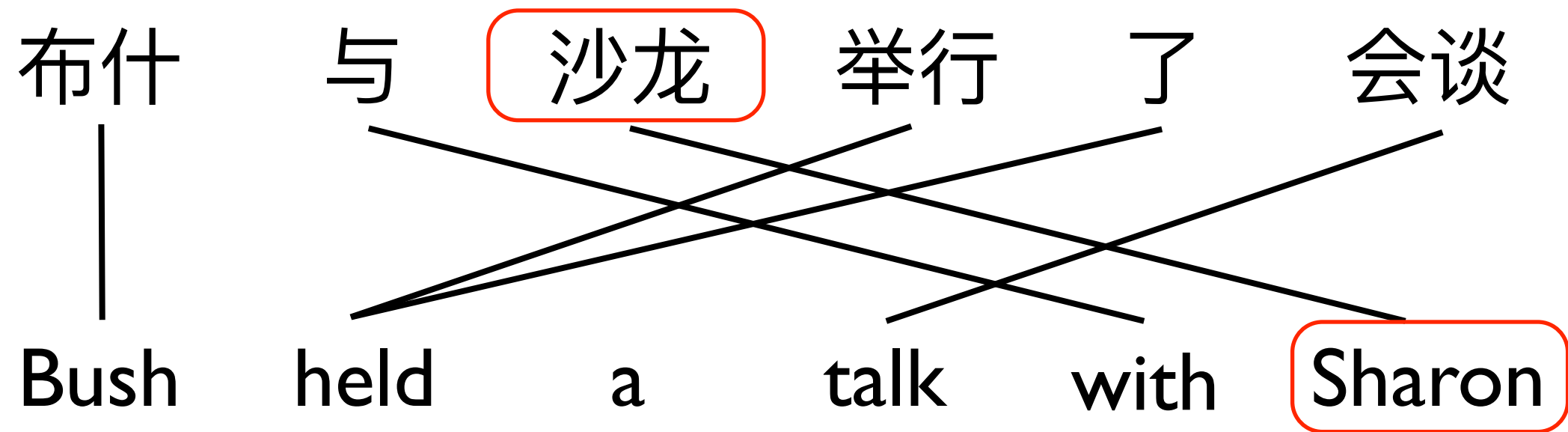


# Hierarchical Phrase Extraction



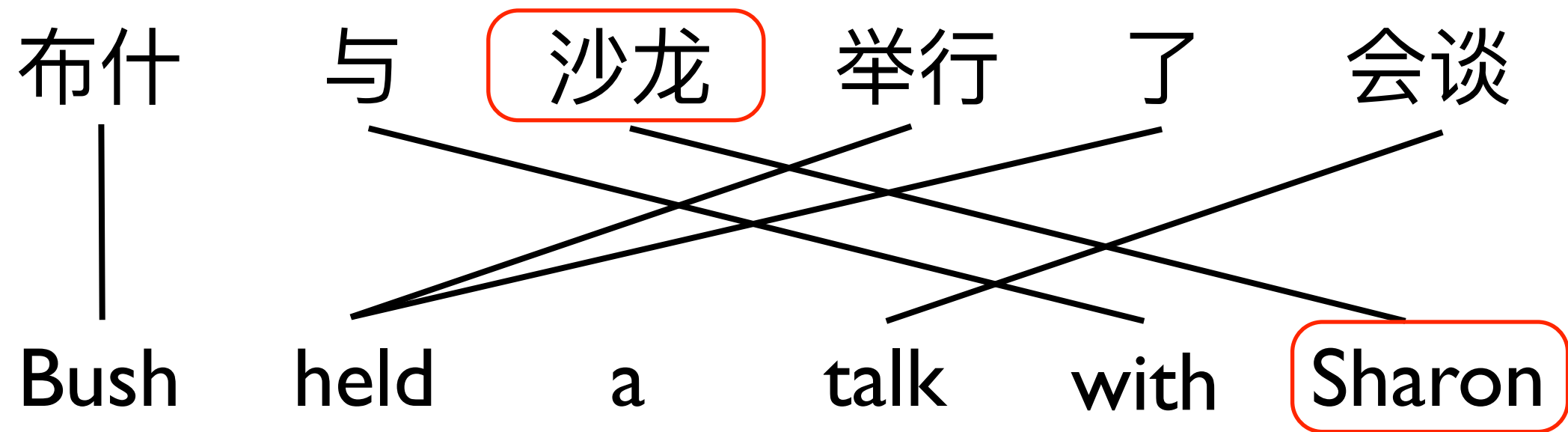
(与 沙龙 举行 了 会谈, held a talk with Sharon)

# Hierarchical Phrase Extraction



(与 沙龙 举行 了 会谈, held a talk with Sharon)

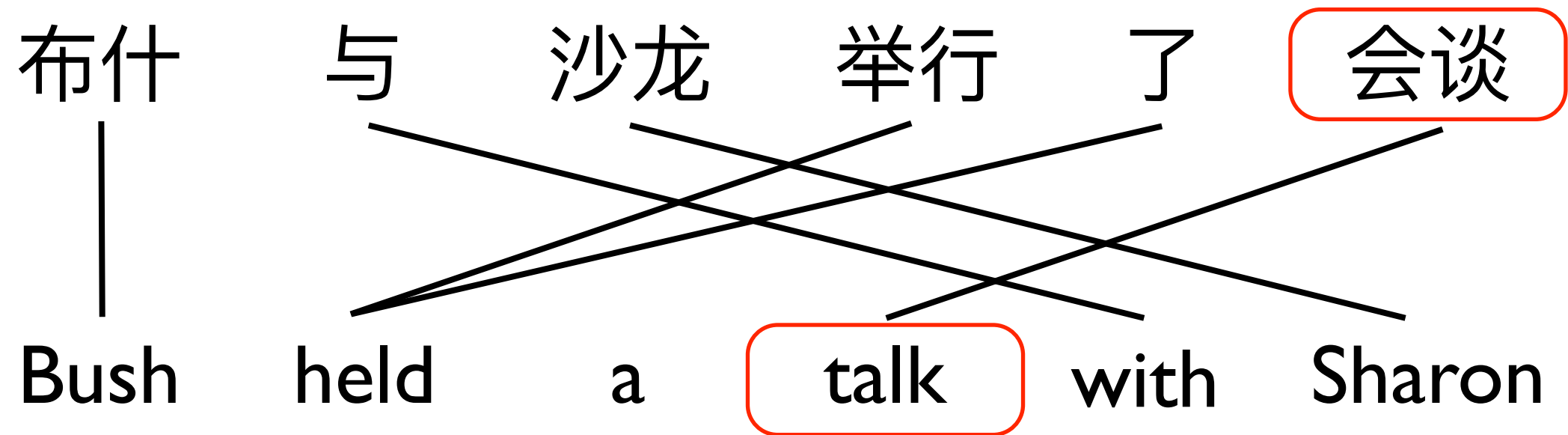
# Hierarchical Phrase Extraction



(与 沙龙 举行 了 会谈, held a talk with Sharon)

(沙龙, Sharon)

# Hierarchical Phrase Extraction

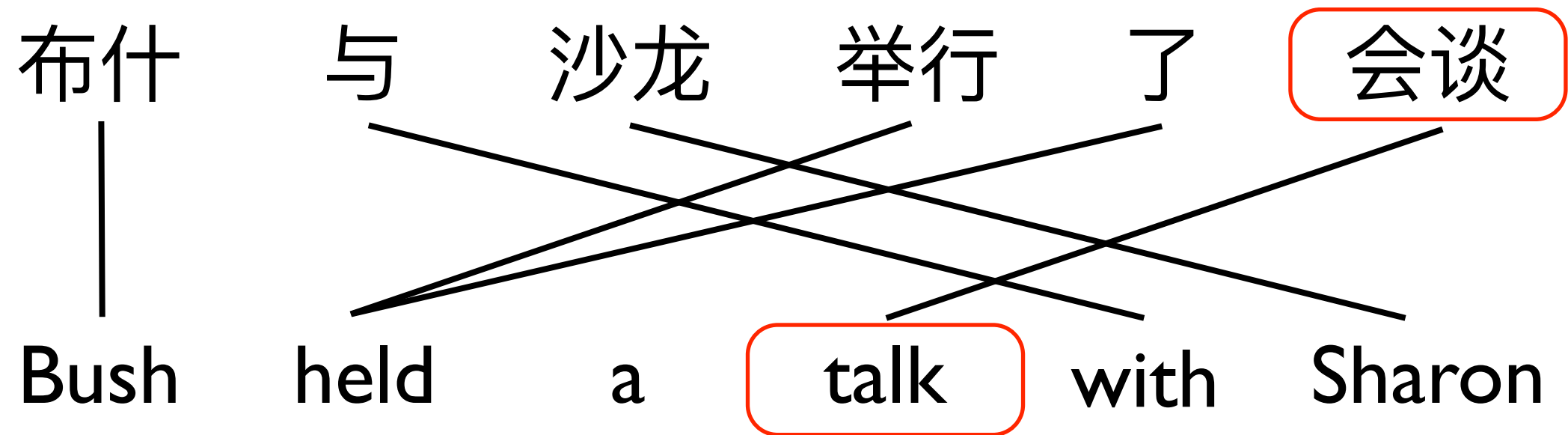


(与 沙龙 举行 了 会谈, held a talk with Sharon)

(沙龙, Sharon)



# Hierarchical Phrase Extraction

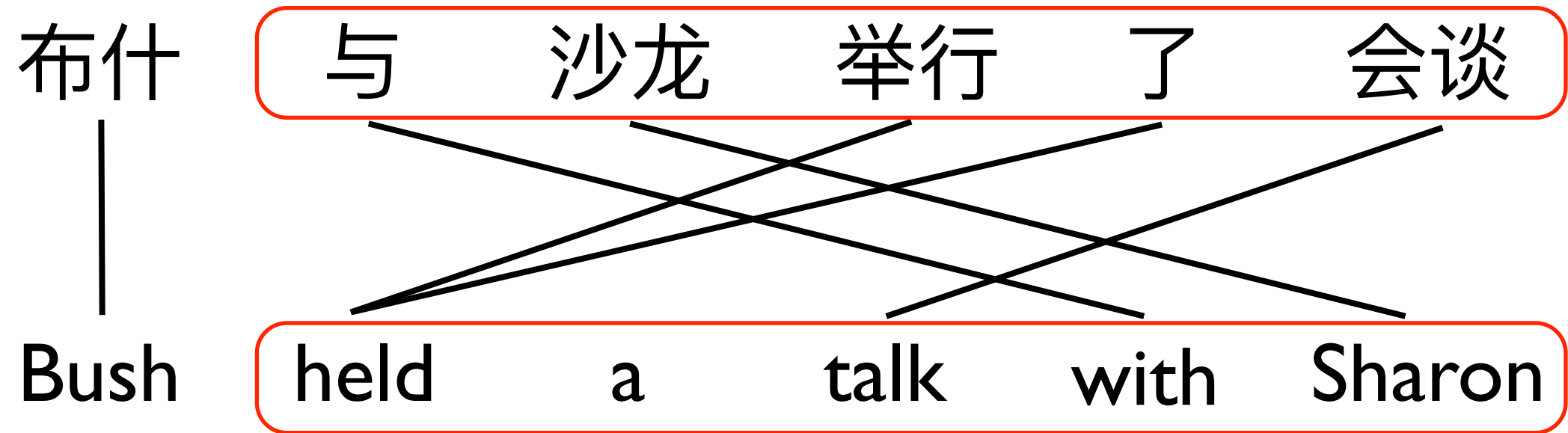


(与 沙龙 举行 了 会谈, held a talk with Sharon)

(沙龙, Sharon)

(会谈, talk)

# Hierarchical Phrase Extraction

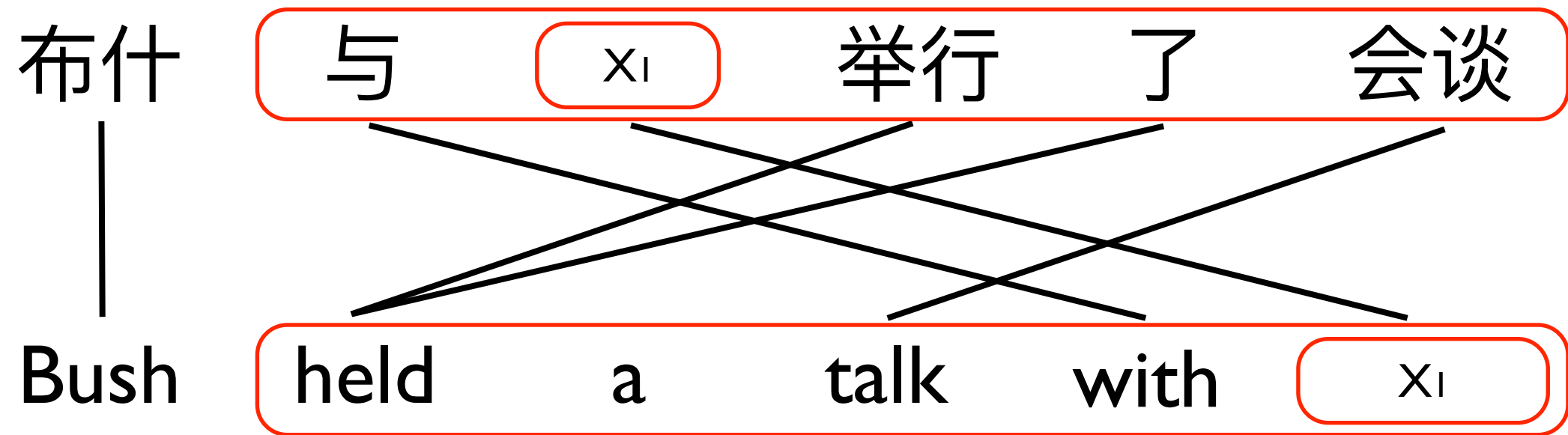


(与 沙龙 举行 了 会谈, held a talk with Sharon)

(沙龙, Sharon)

(会谈, talk)

# Hierarchical Phrase Extraction

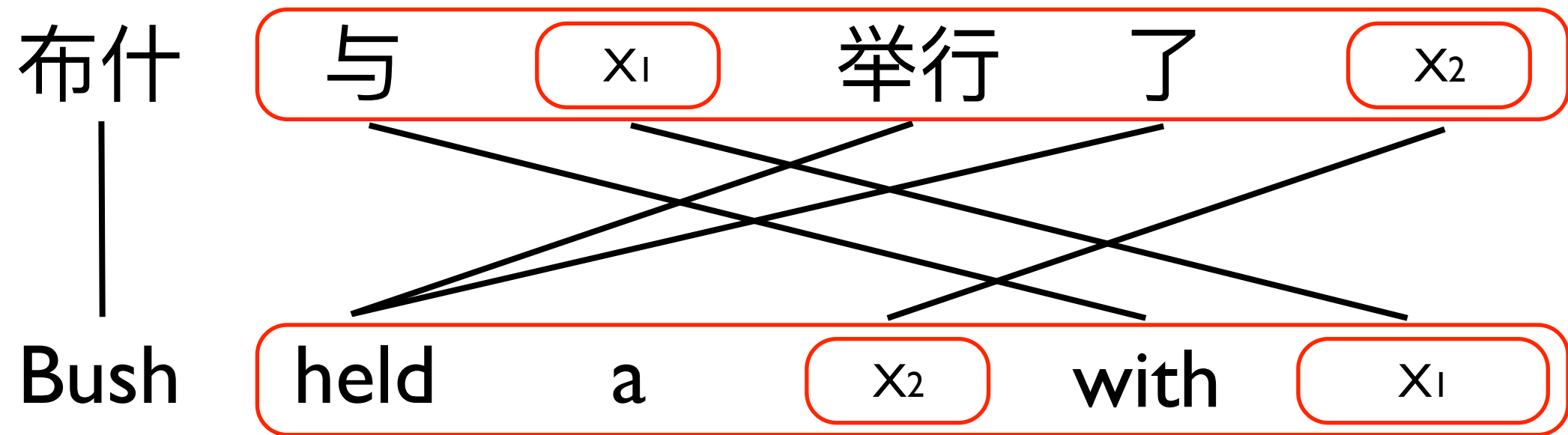


(与 沙龙 举行 了 会谈, held a talk with Sharon)

(沙龙, Sharon)

(会谈, talk)

# Hierarchical Phrase Extraction

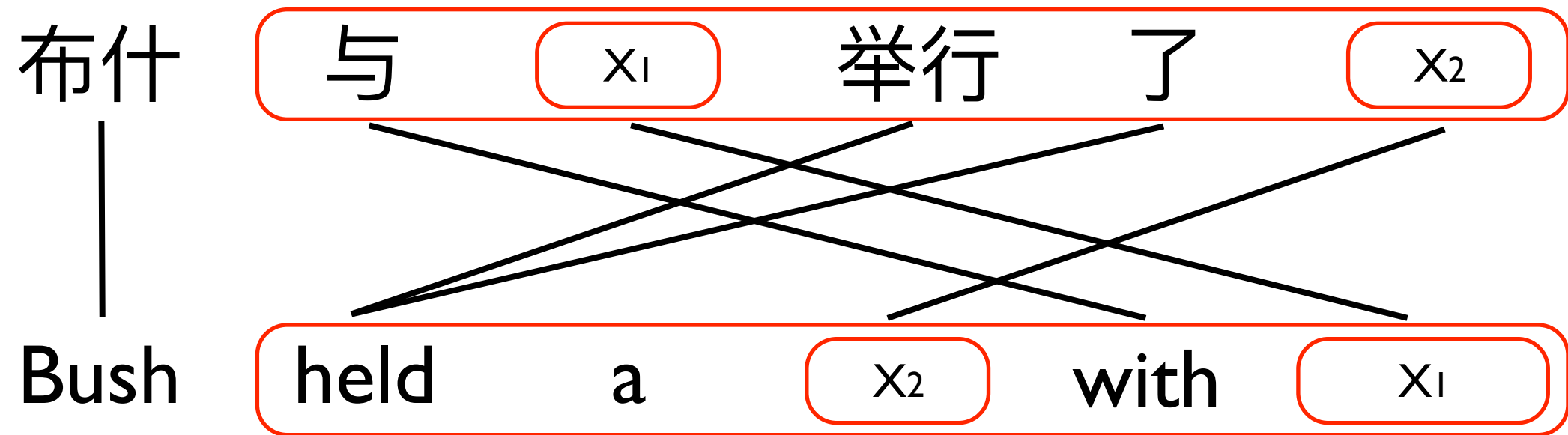


(与 沙龙 举行 了 会谈, held a talk with Sharon)

(沙龙, Sharon)

(会谈, talk)

# Hierarchical Phrase Extraction



(与 沙龙 举行了 会谈, held a talk with Sharon)

(沙龙, Sharon)

(会谈, talk)

(与  $X_1$  举行了  $X_2$ , held a  $X_2$  with  $X_1$ )

# Hierarchical Phrase-based Translation

- Hierarchical phrase-based model is based on SCFG without linguistic syntax

# Hierarchical Phrase-based Translation

- Hierarchical phrase-based model is based on SCFG without linguistic syntax

lexical rules

$X \rightarrow (\text{沙龙}, \text{Sharon})$

$X \rightarrow (\text{会谈}, \text{talk})$

# Hierarchical Phrase-based Translation

- Hierarchical phrase-based model is based on SCFG without linguistic syntax

lexical rules

$X \rightarrow (\text{沙龙}, \text{Sharon})$

$X \rightarrow (\text{会谈}, \text{talk})$

syntactic rules



# Hierarchical Phrase-based Translation

- Hierarchical phrase-based model is based on SCFG without linguistic syntax

## lexical rules

$X \rightarrow (\text{沙龙}, \text{Sharon})$

$X \rightarrow (\text{会谈}, \text{talk})$

## syntactic rules

$X \rightarrow (\text{与 } X_1 \text{ 举行了 } X_2, \text{held a } X_2 \text{ with } X_1)$

# Hierarchical Phrase-based Translation

- Hierarchical phrase-based model is based on SCFG without linguistic syntax

## lexical rules

$X \rightarrow (\text{沙龙}, \text{Sharon})$

$X \rightarrow (\text{会谈}, \text{talk})$

## syntactic rules

$X \rightarrow (\text{与 } X_1 \text{ 举行了 } X_2, \text{held a } X_2 \text{ with } X_1)$

$X \rightarrow (\text{布什 } X_1, \text{Bush } X_1)$

# Hierarchical Phrase-based Translation

- Hierarchical phrase-based model is based on SCFG without linguistic syntax

## lexical rules

$X \rightarrow (\text{沙龙}, \text{Sharon})$

$X \rightarrow (\text{会谈}, \text{talk})$

## syntactic rules

$X \rightarrow (\text{与 } X_1 \text{ 举行了 } X_2, \text{held a } X_2 \text{ with } X_1)$

$X \rightarrow (\text{布什 } X_1, \text{Bush } X_1)$

**ITG is a special case of SCFG**

(Chiang, 2005)

# CKY Parsing

布什 与 沙龙 举行 了 会谈

# CKY Parsing

$X \rightarrow (\text{沙龙}, \text{Sharon})$

布什      与      沙龙      举行      了      会谈

# CKY Parsing

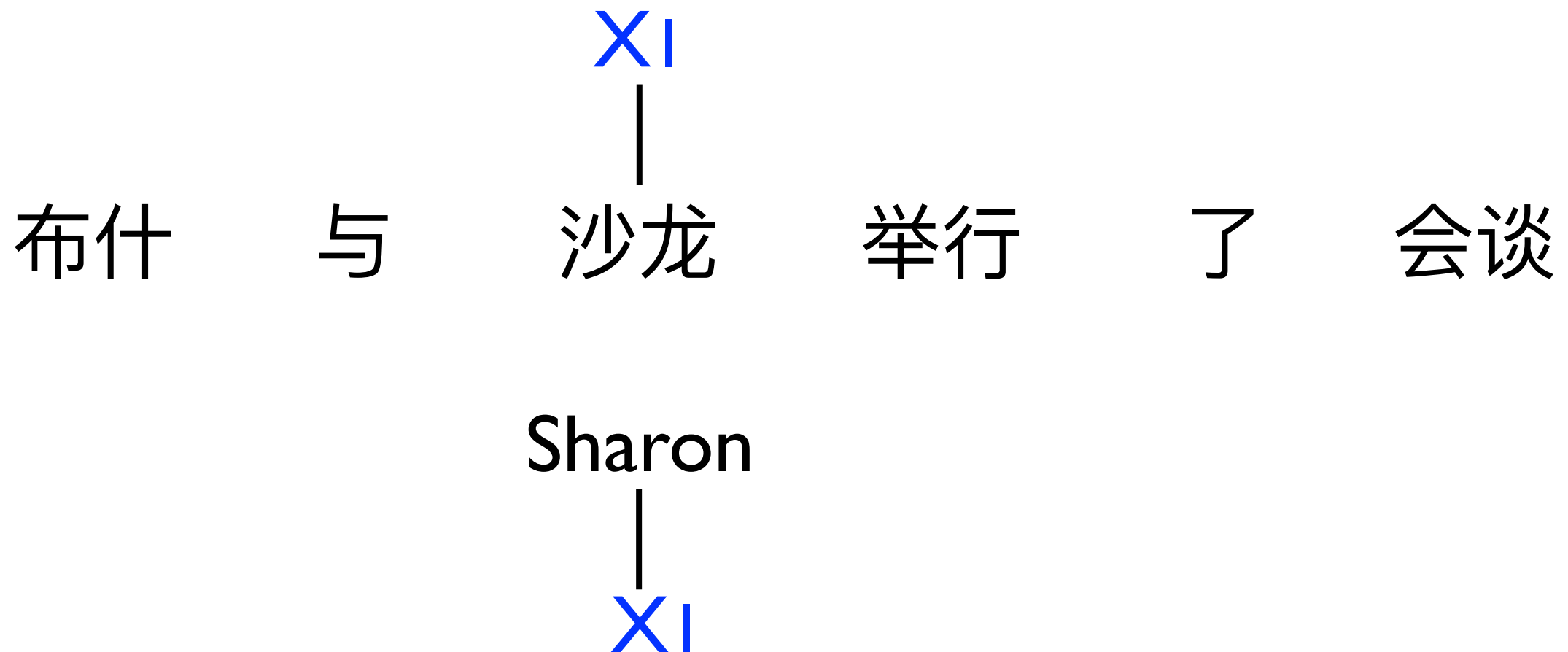
$X \rightarrow (\text{沙龙}, \text{Sharon})$

布什      与      沙龙      举行      了      会谈

Sharon

# CKY Parsing

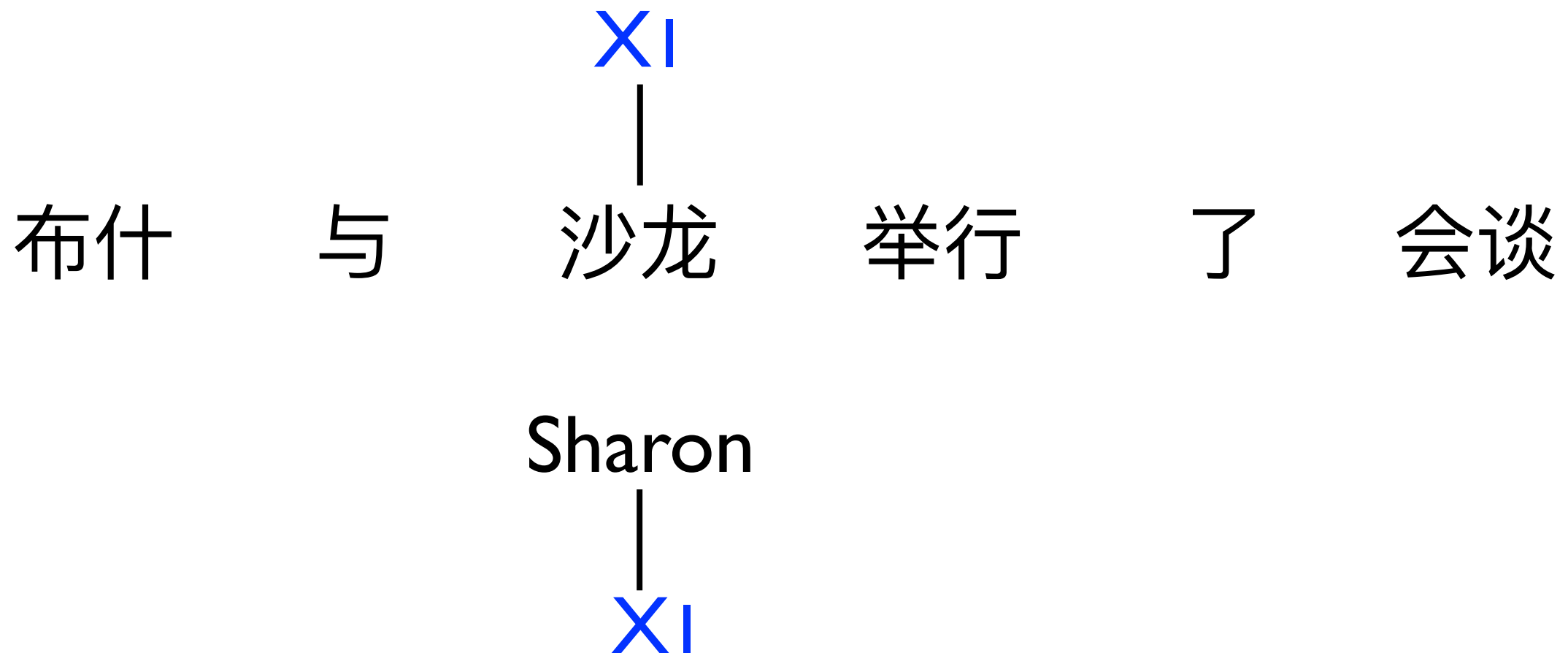
$X \rightarrow (\text{沙龙}, \text{Sharon})$



# CKY Parsing

$X \rightarrow (\text{沙龙}, \text{Sharon})$

$X \rightarrow (\text{会谈}, \text{talk})$

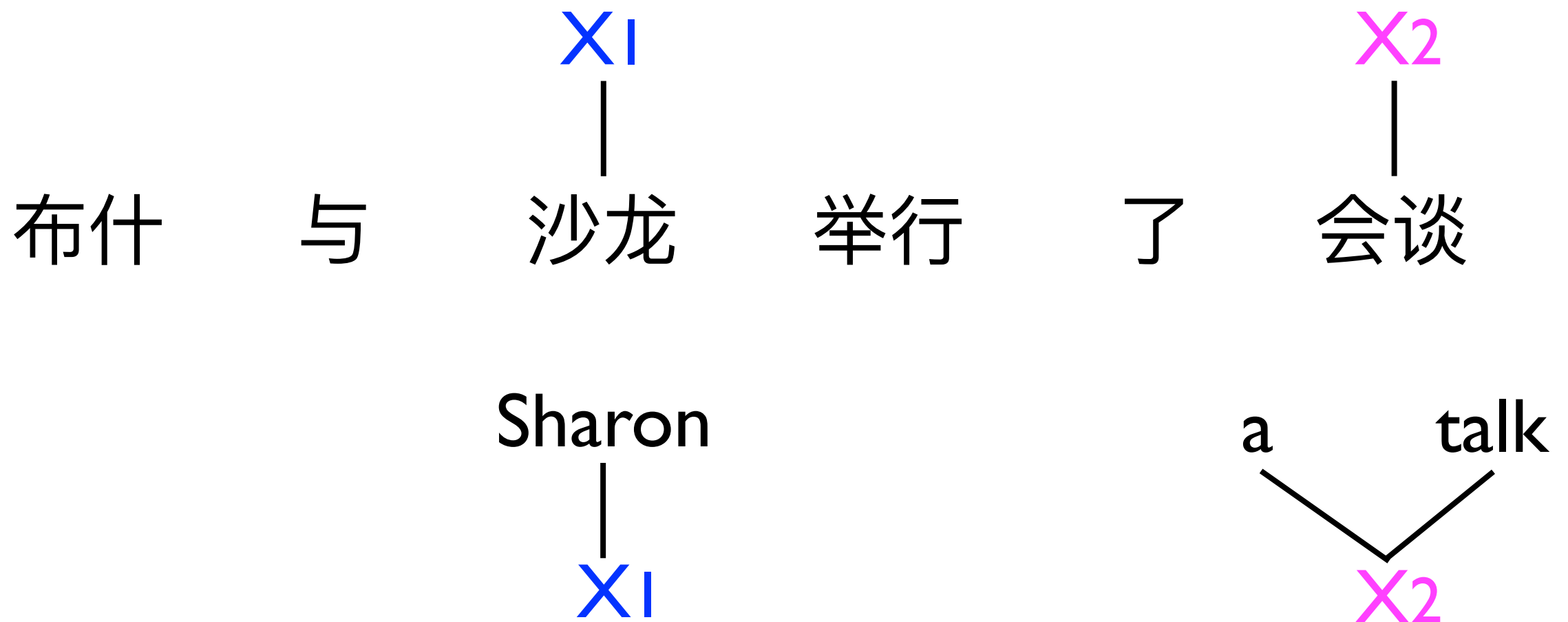




# CKY Parsing

$X \rightarrow (\text{沙龙}, \text{Sharon})$

$X \rightarrow (\text{会谈}, \text{talk})$



# CKY Parsing

$X \rightarrow (\text{与 } X_1 \text{ 举行 了 } X_2, \text{held a } X_2 \text{ with } X_1)$

布什      与       $X_1$   
                 |  
                沙龙      举行      了       $X_2$   
   |  
                                        会谈

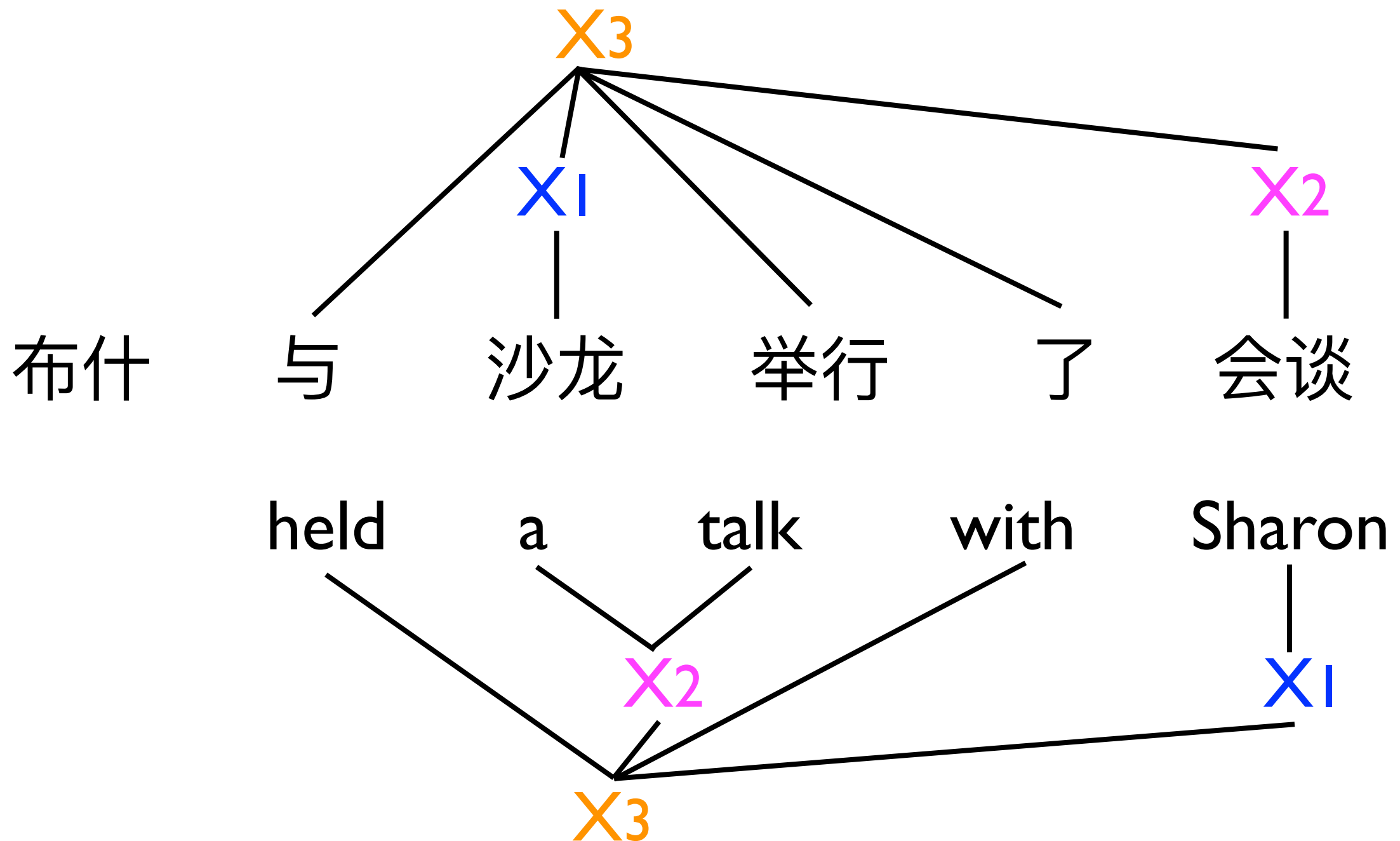
Sharon  
|  
 $X_1$

a      talk  
      \    /  
        $X_2$

(Chiang, 2005)

# CKY Parsing

$X \rightarrow (\text{与 } X_1 \text{ 举行了 } X_2, \text{held a } X_2 \text{ with } X_1)$

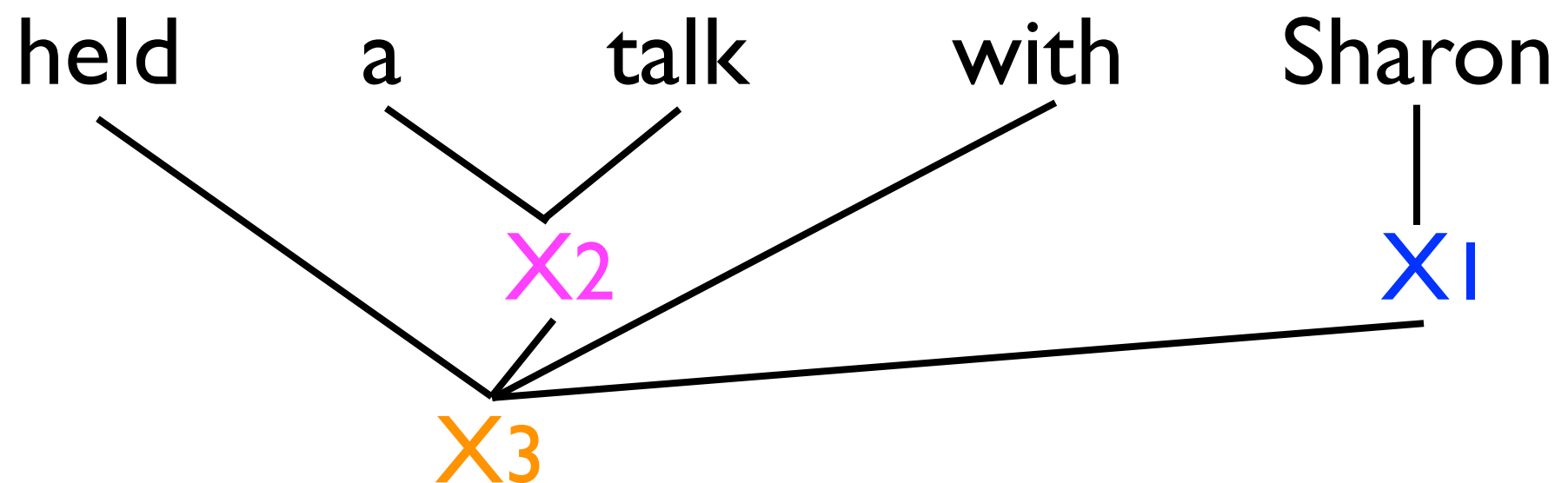
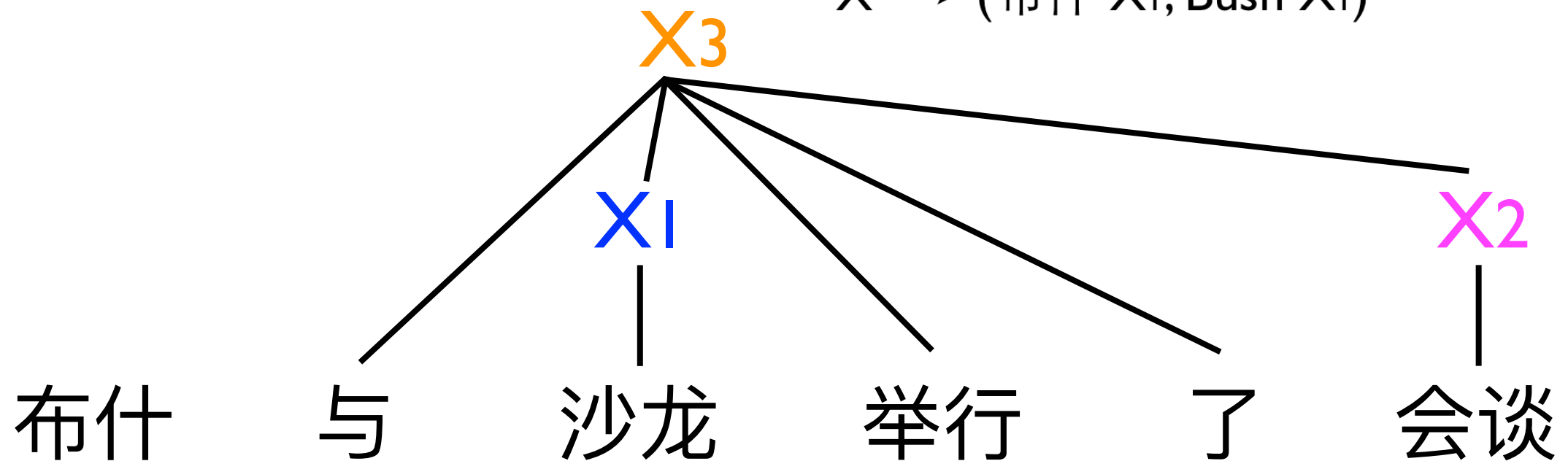


(Chiang, 2005)

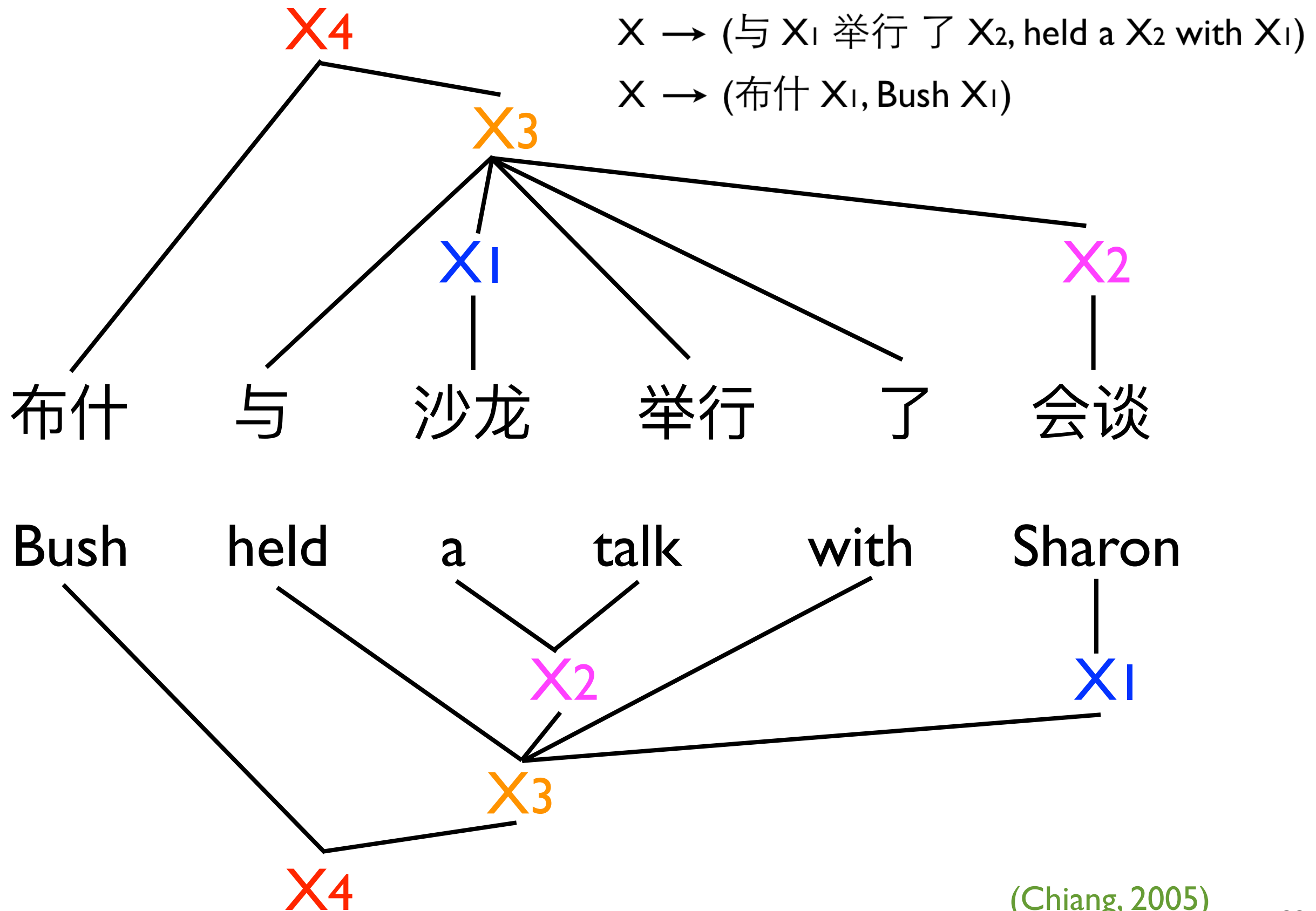
# CKY Parsing

$X \rightarrow (\text{与 } X_1 \text{ 举行了 } X_2, \text{held a } X_2 \text{ with } X_1)$

$X \rightarrow (\text{布什 } X_1, \text{Bush } X_1)$



# CKY Parsing



(Chiang, 2005)

# Chart

$X \rightarrow$  (沙龙, Sharon)

$X \rightarrow$  (会谈, talk)

$X \rightarrow$  (与  $X_1$  举行了  $X_2$ , held a  $X_2$  with  $X_1$ )

$X \rightarrow$  (布什  $X_1$ , Bush  $X_1$ )

Bush ... Sharon					
	held ... Sharon				
		Sharon			talk

布什      与      沙龙      举行      了      会谈

# Syntax-based MT

## SCFGs without linguistic syntax

*inverted transduction grammar*

*hierarchical phrase-based model*

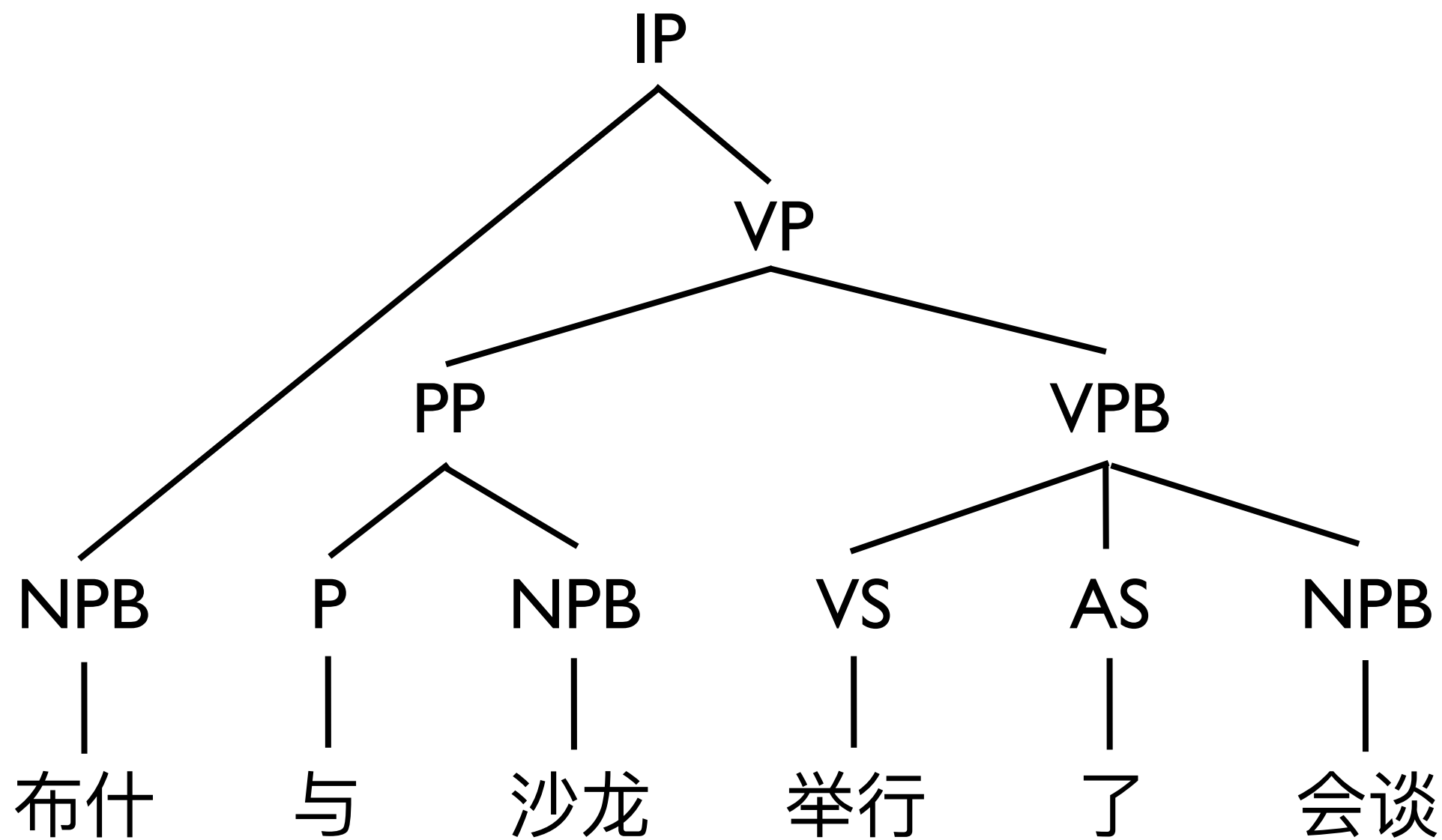
## STSGs with linguistic syntax

*string-to-tree*

*tree-to-string*

*tree-to-tree*

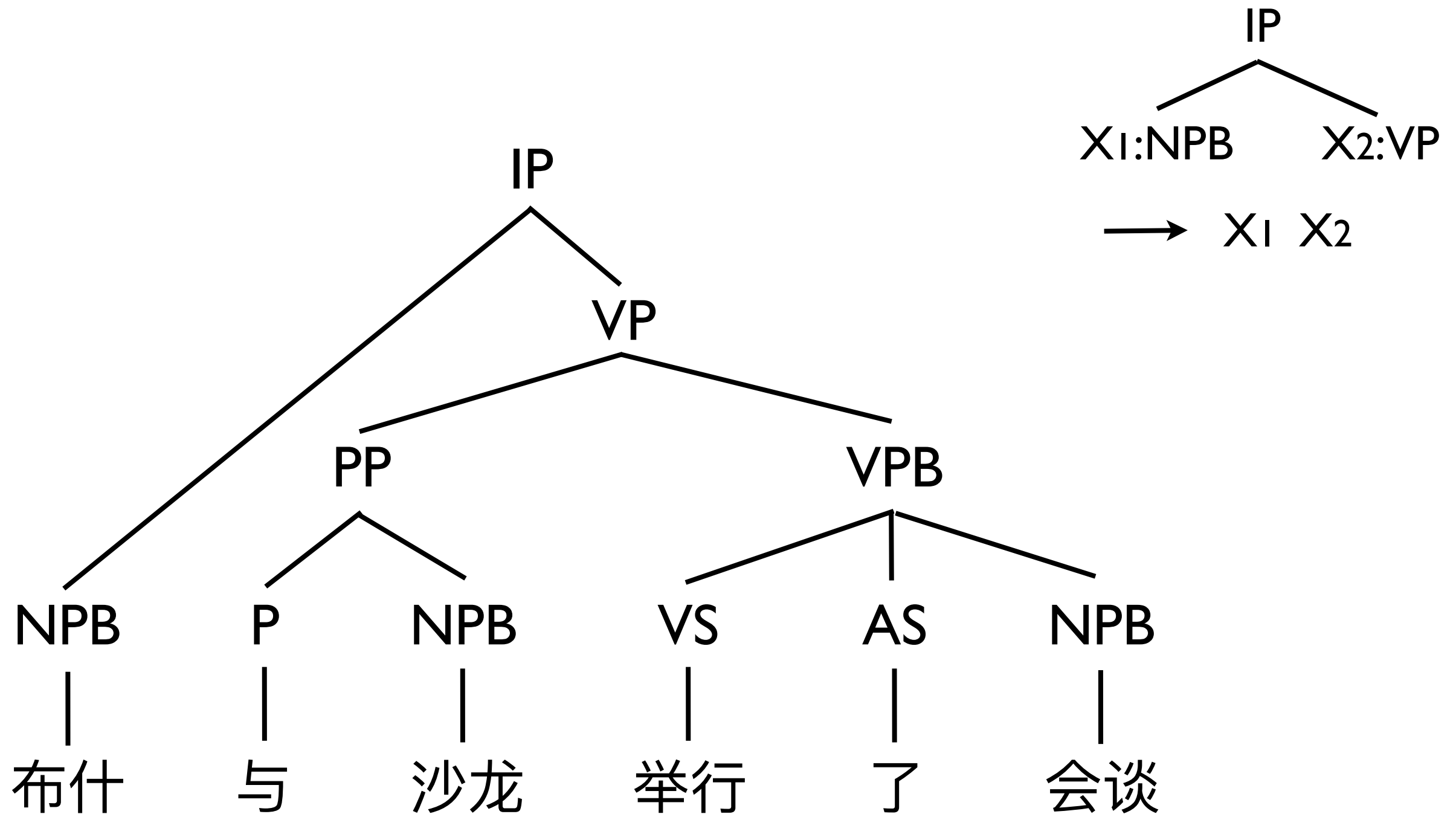
# Tree-to-String Translation



(Liu et al., 2006; Huang et al., 2006)

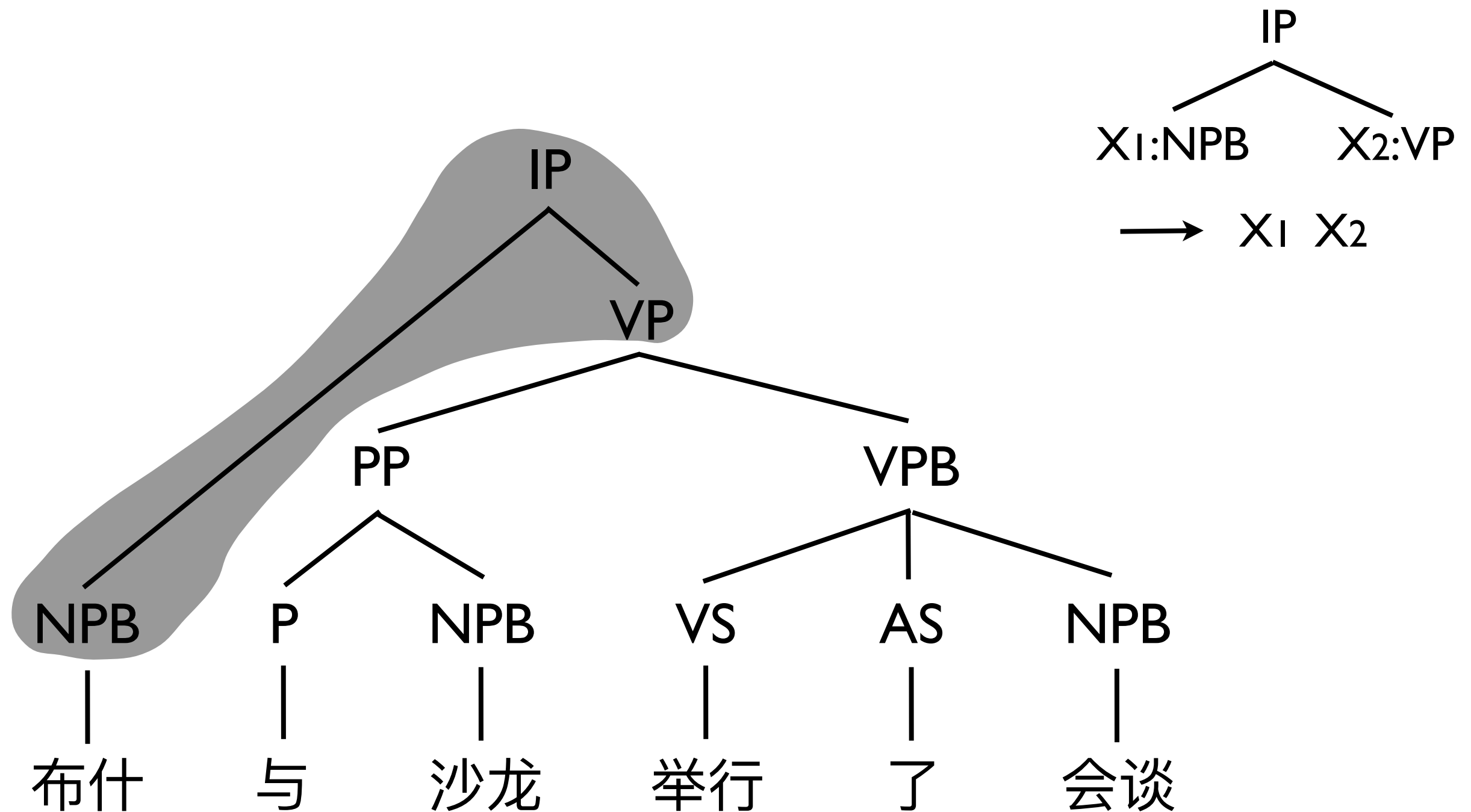


# Tree-to-String Translation



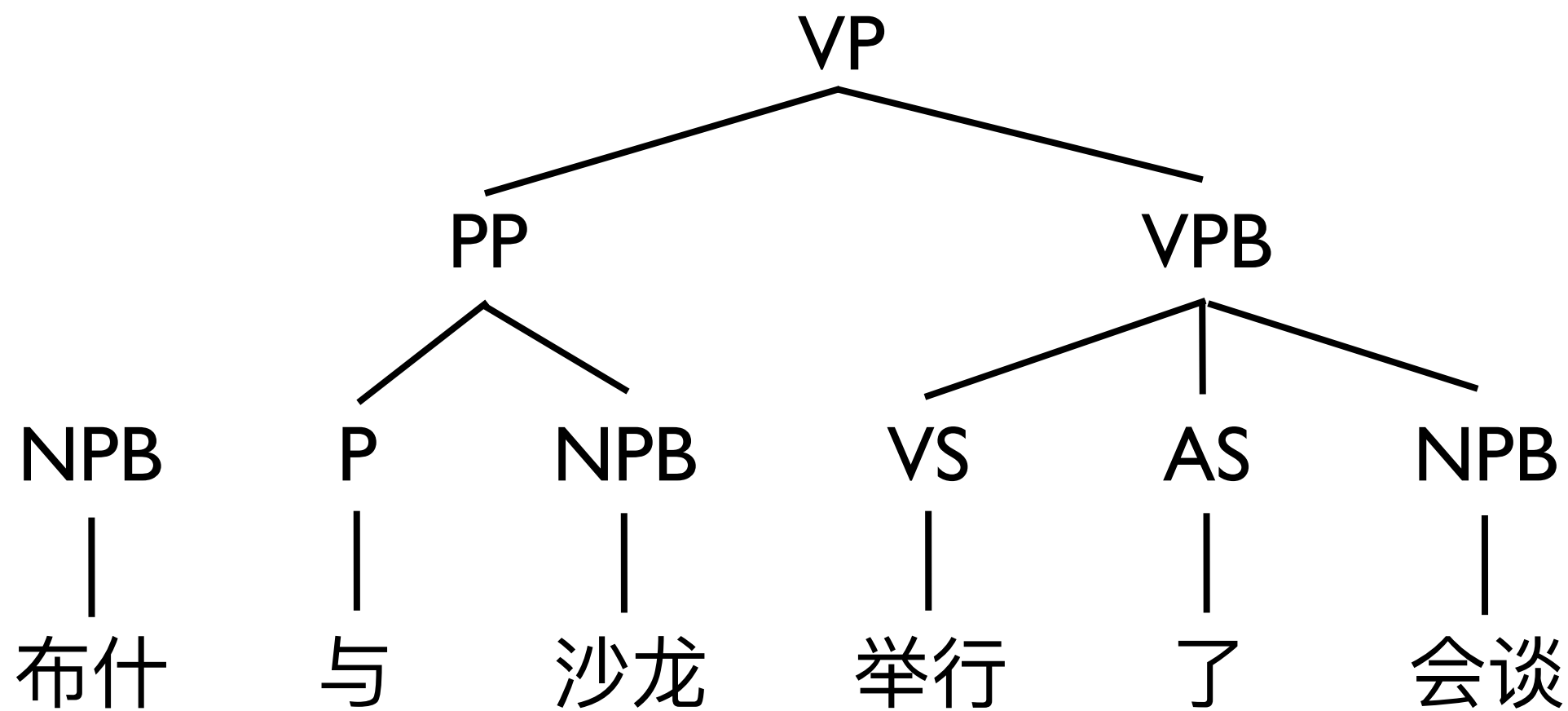
(Liu et al., 2006; Huang et al., 2006)

# Tree-to-String Translation



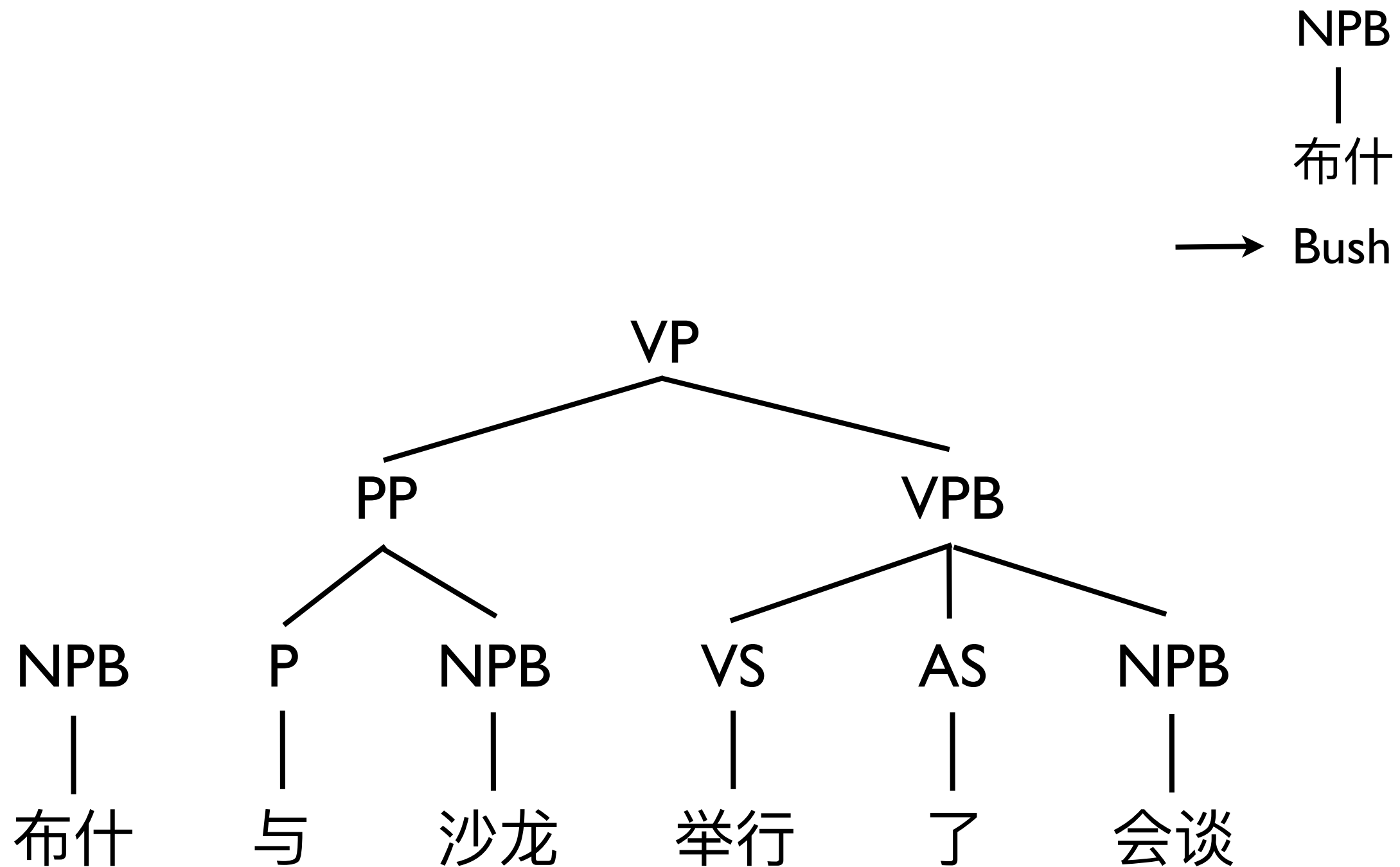
(Liu et al., 2006; Huang et al., 2006)

# Tree-to-String Translation



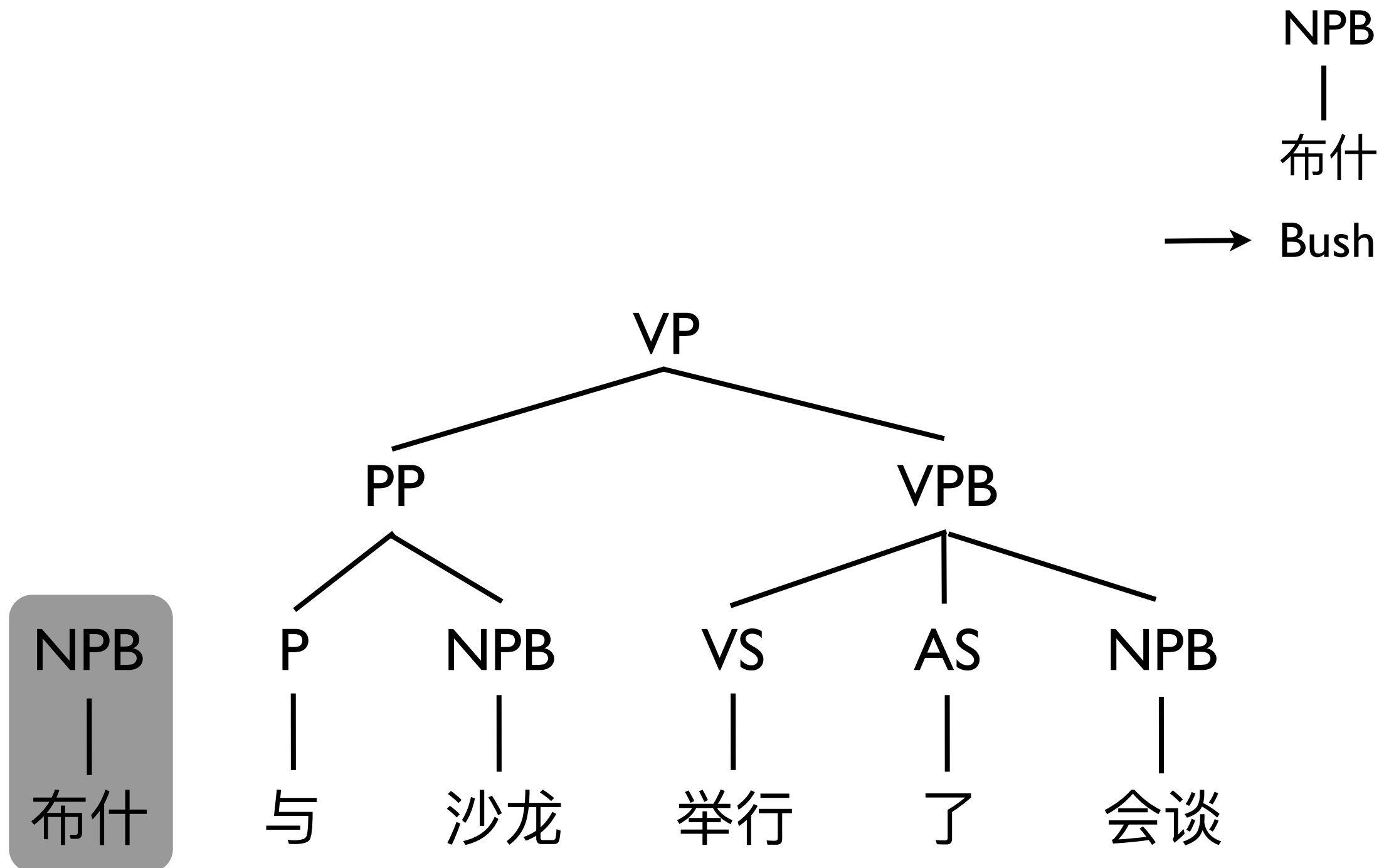
(Liu et al., 2006; Huang et al., 2006)

# Tree-to-String Translation



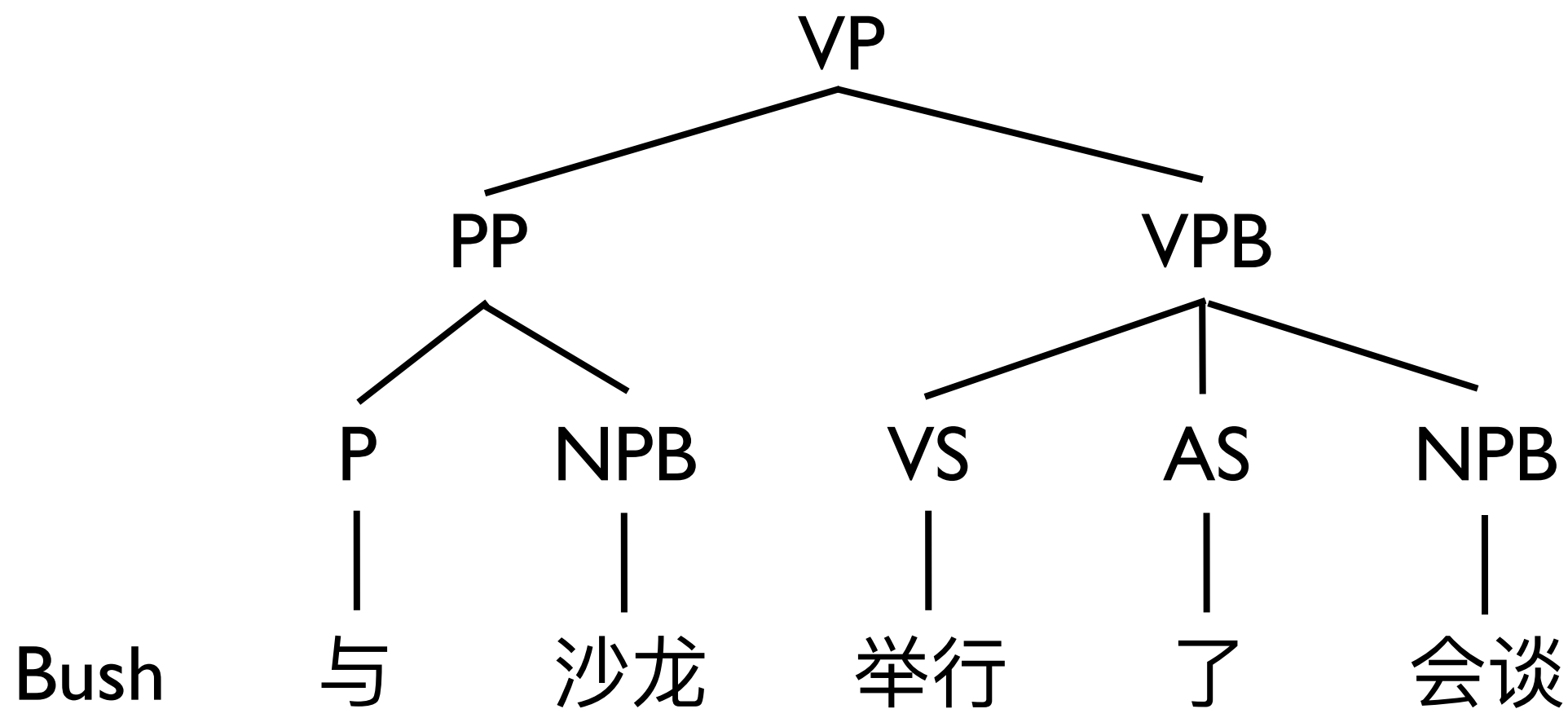
(Liu et al., 2006; Huang et al., 2006)

# Tree-to-String Translation



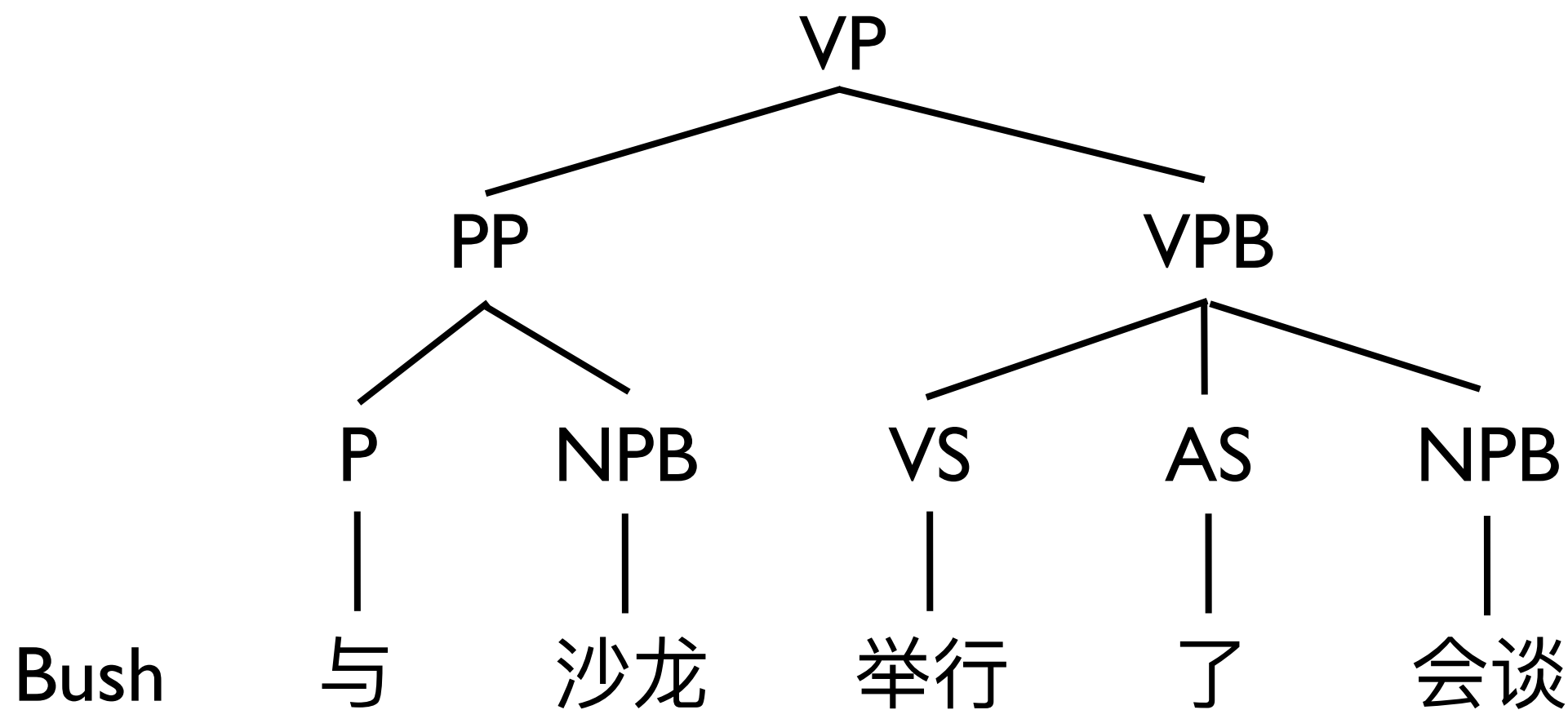
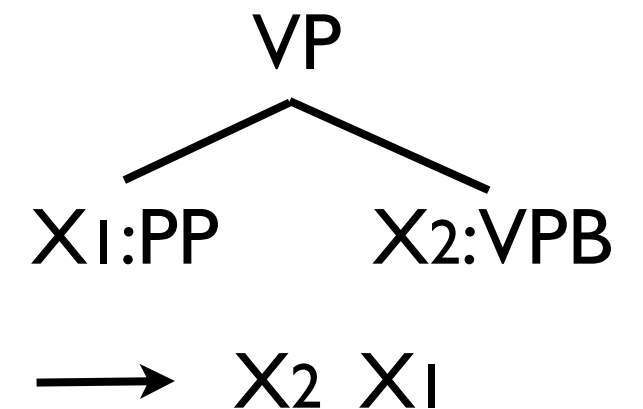
(Liu et al., 2006; Huang et al., 2006)

# Tree-to-String Translation



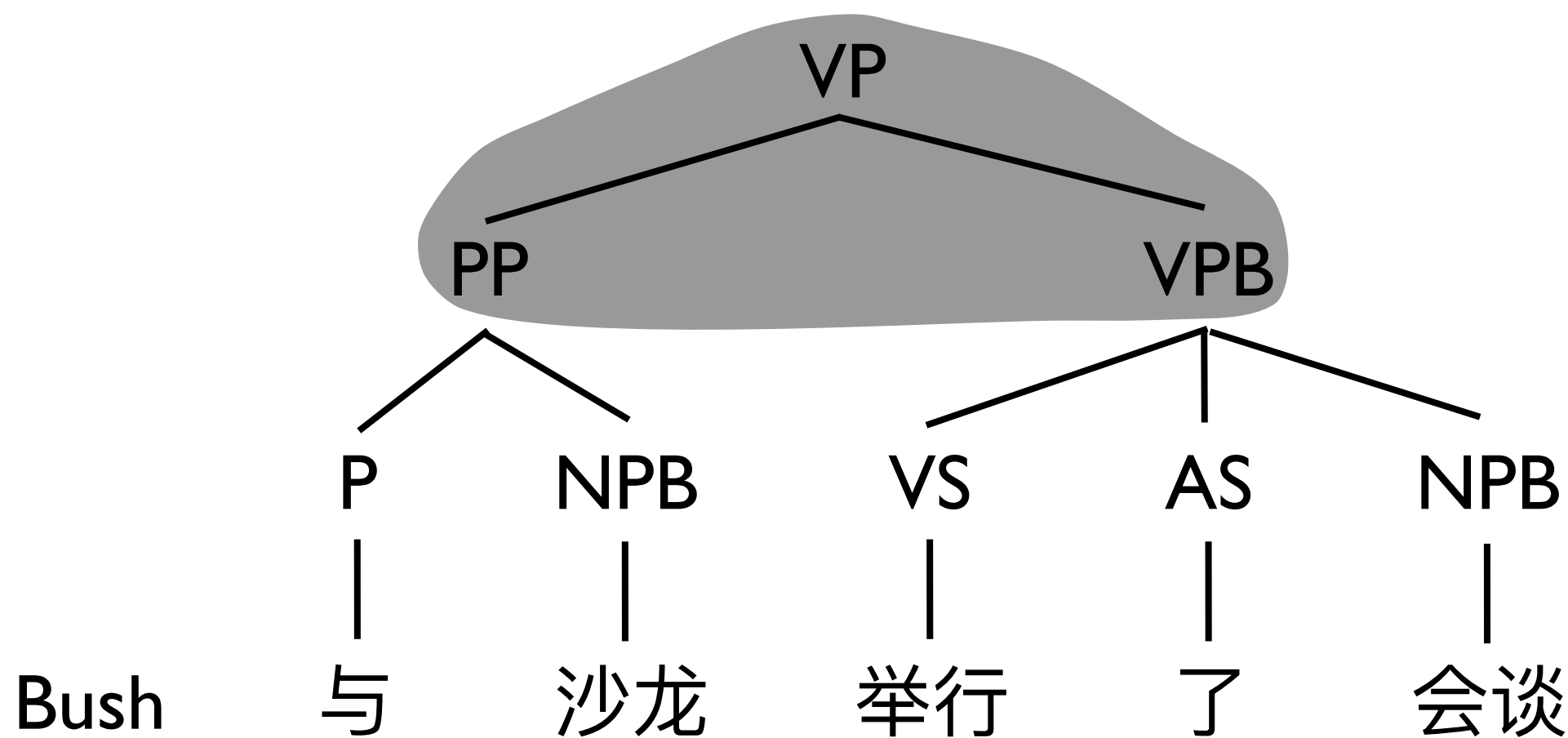
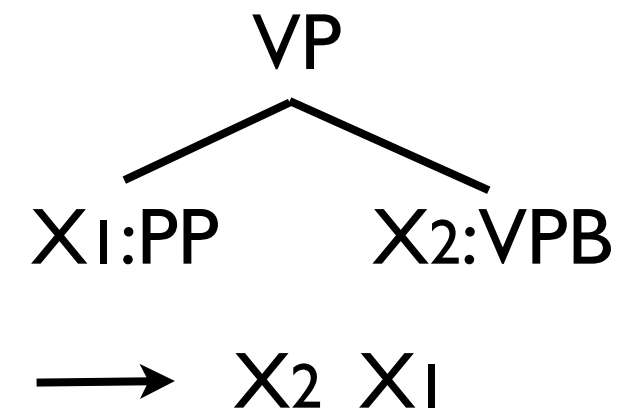
(Liu et al., 2006; Huang et al., 2006)

# Tree-to-String Translation



(Liu et al., 2006; Huang et al., 2006)

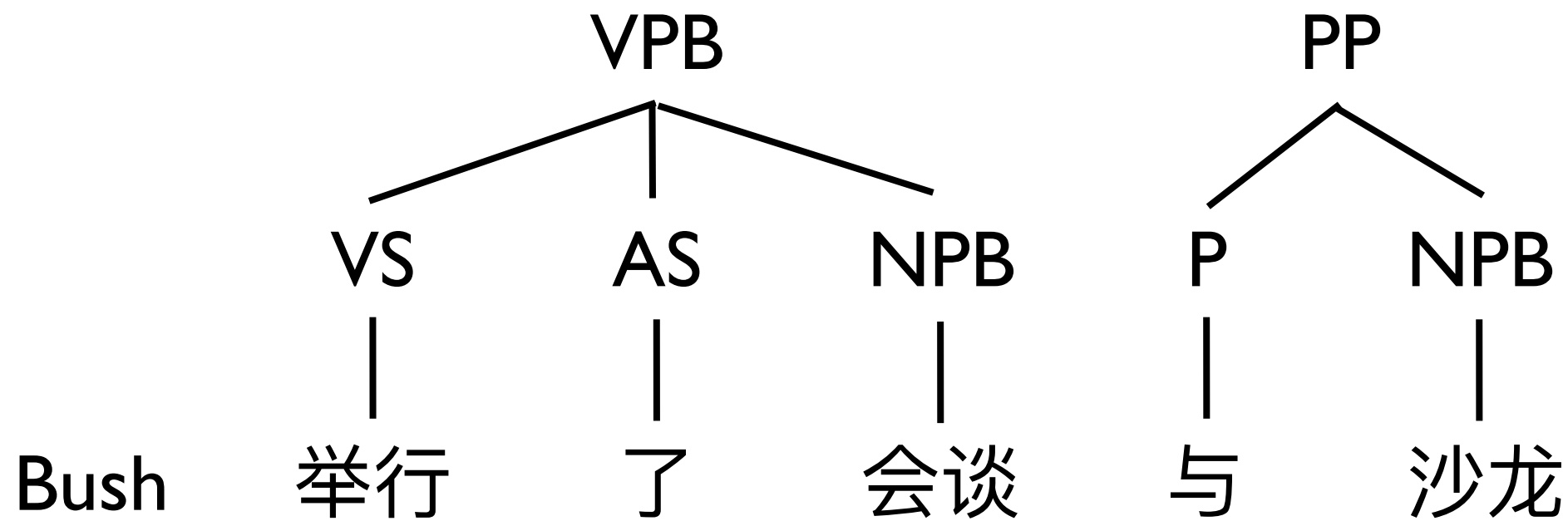
# Tree-to-String Translation



(Liu et al., 2006; Huang et al., 2006)

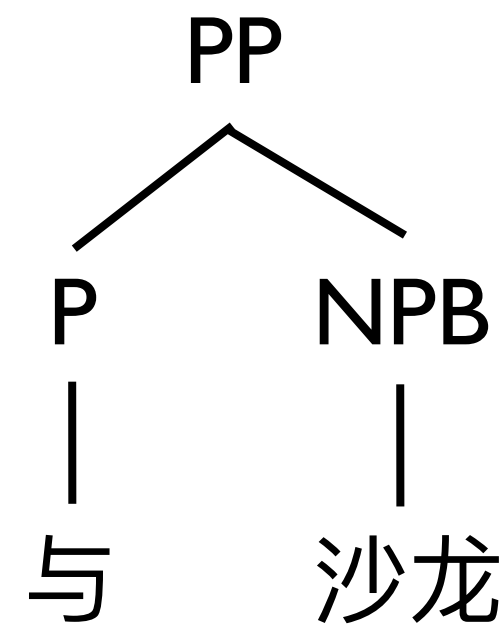
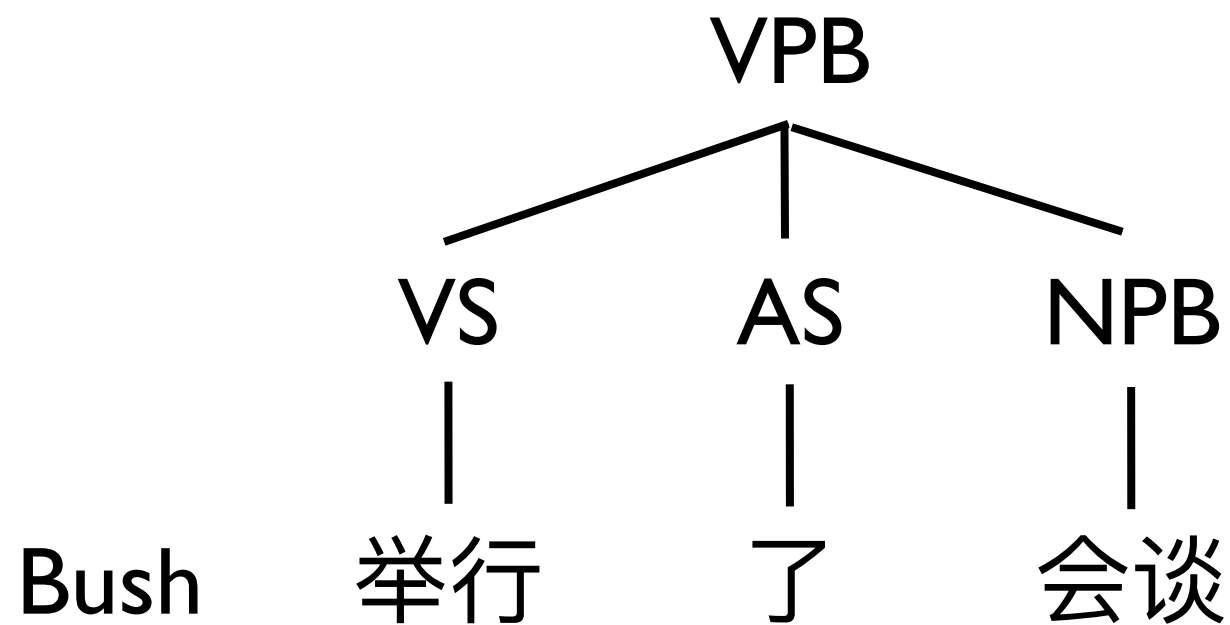
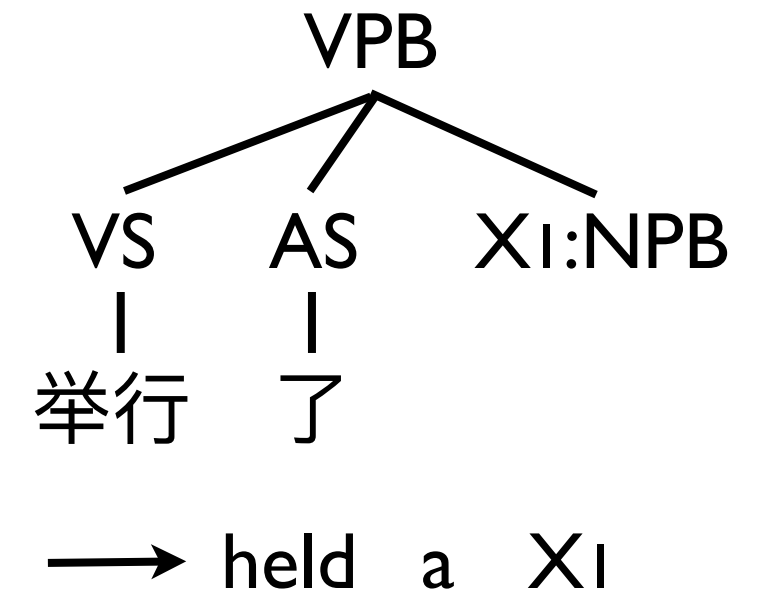


# Tree-to-String Translation



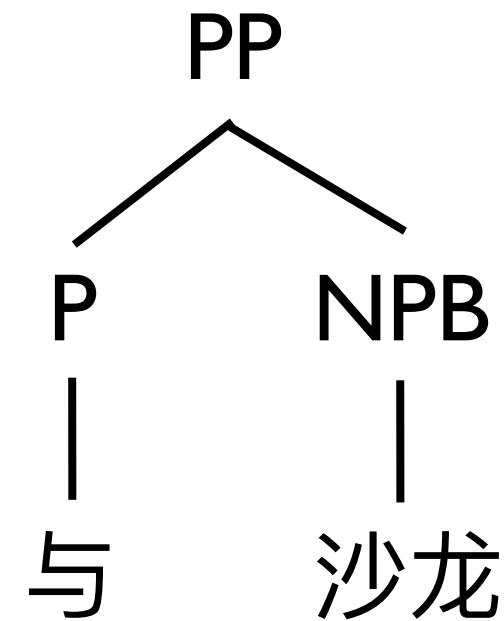
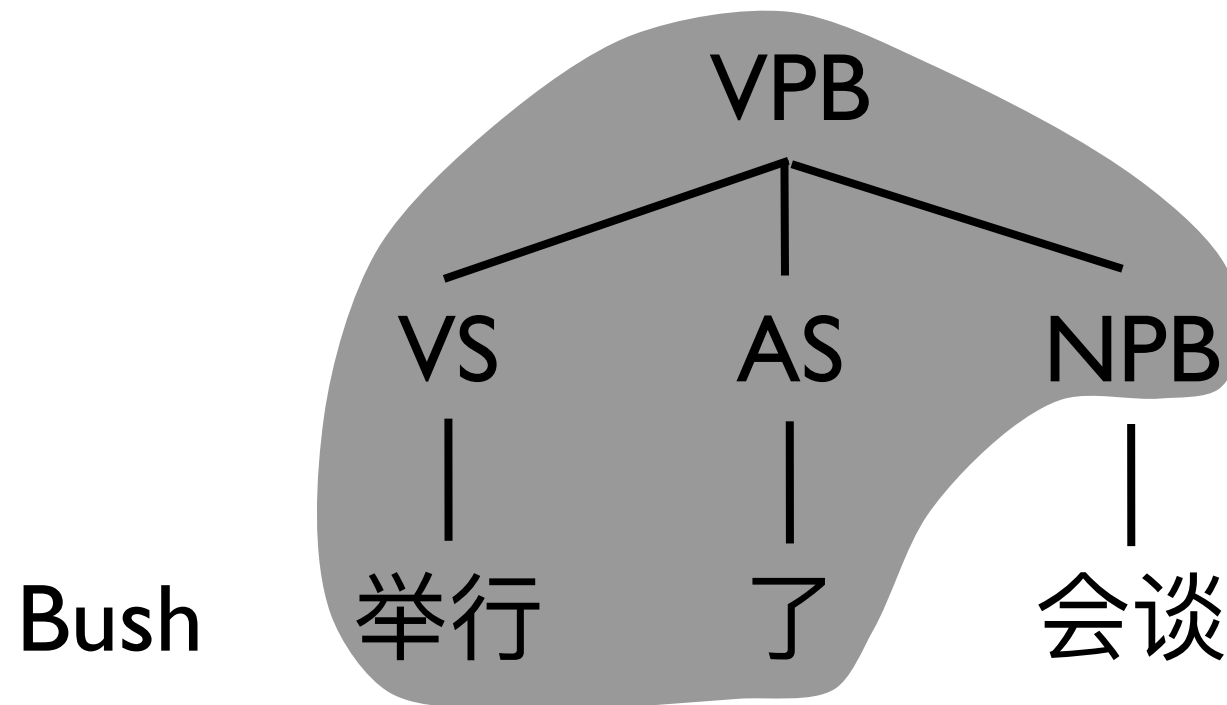
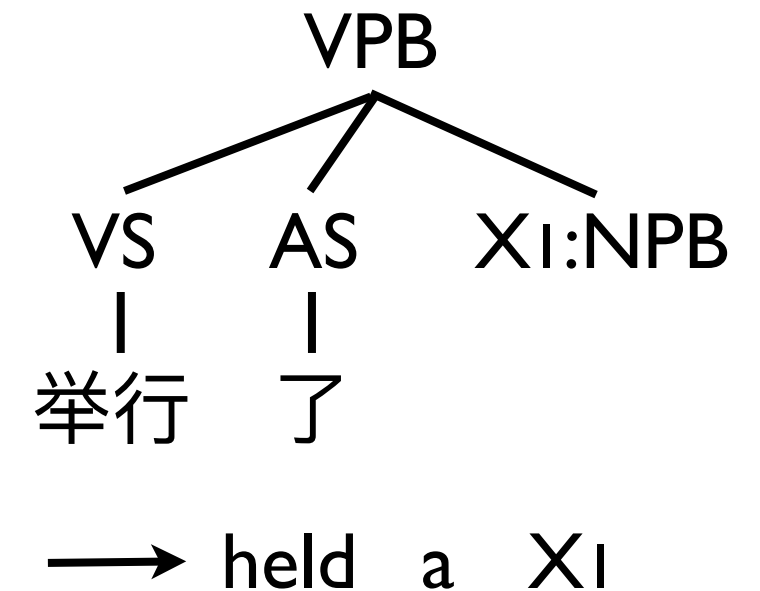
(Liu et al., 2006; Huang et al., 2006)

# Tree-to-String Translation



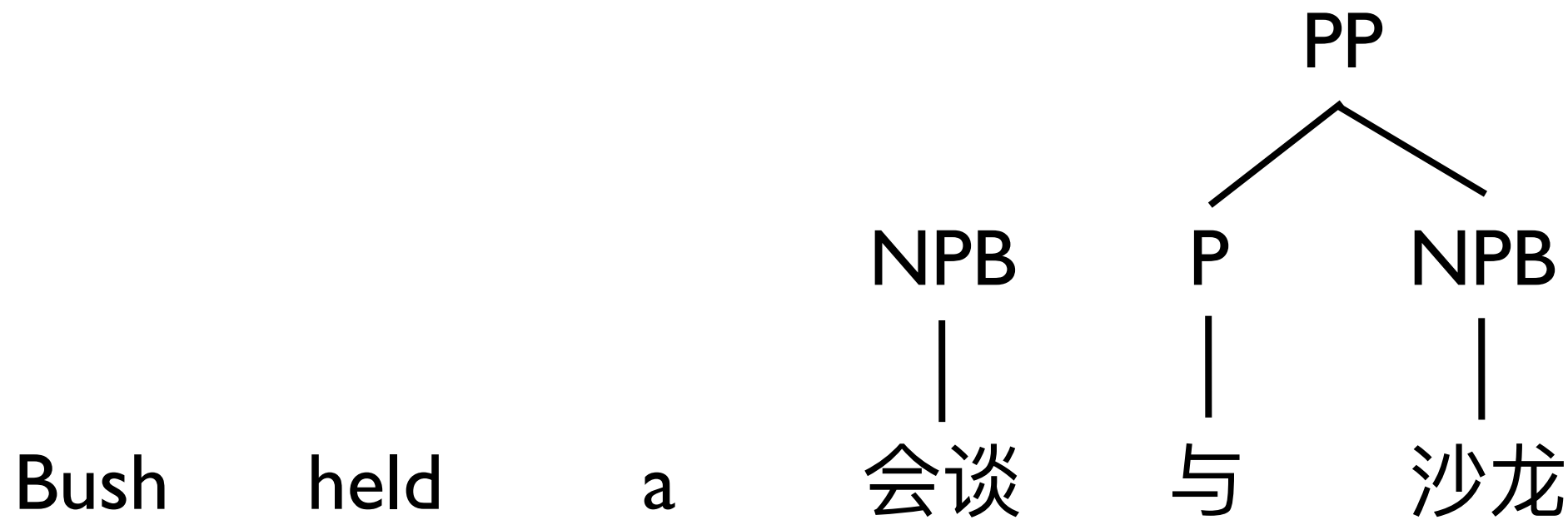
(Liu et al., 2006; Huang et al., 2006)

# Tree-to-String Translation



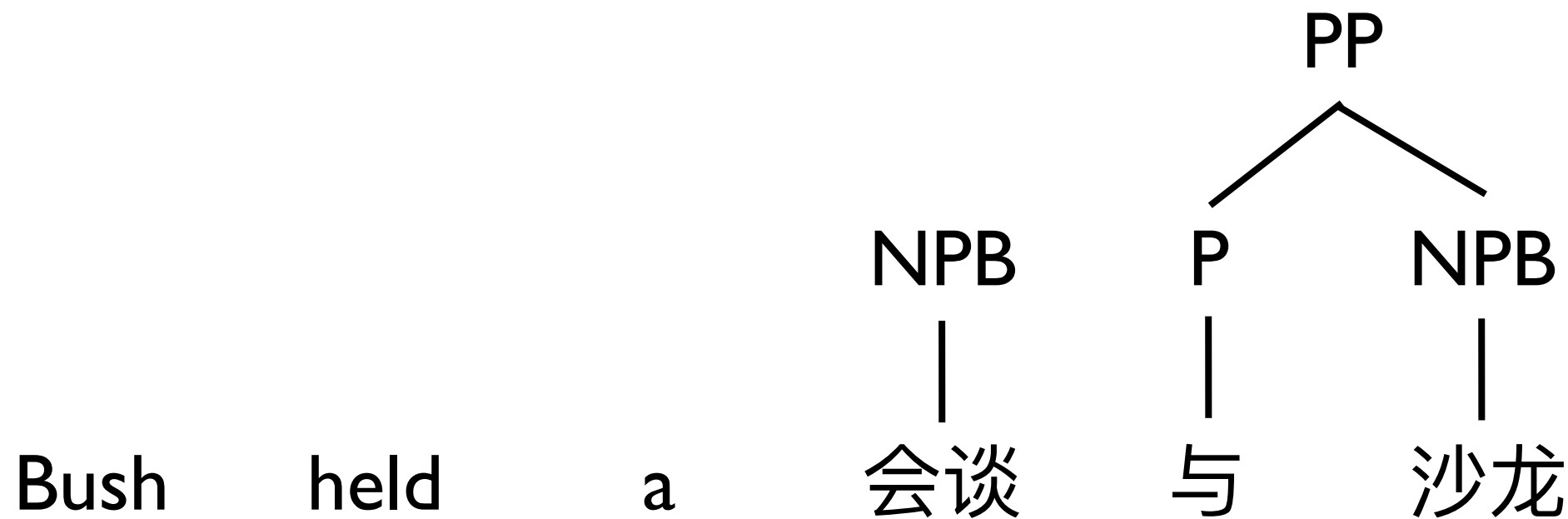
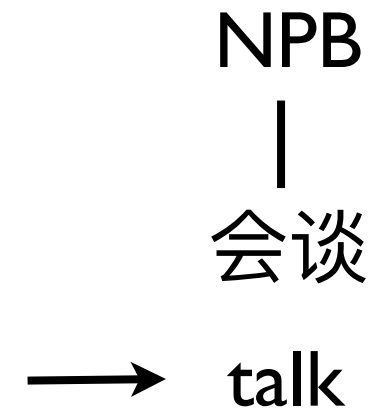
(Liu et al., 2006; Huang et al., 2006)

# Tree-to-String Translation



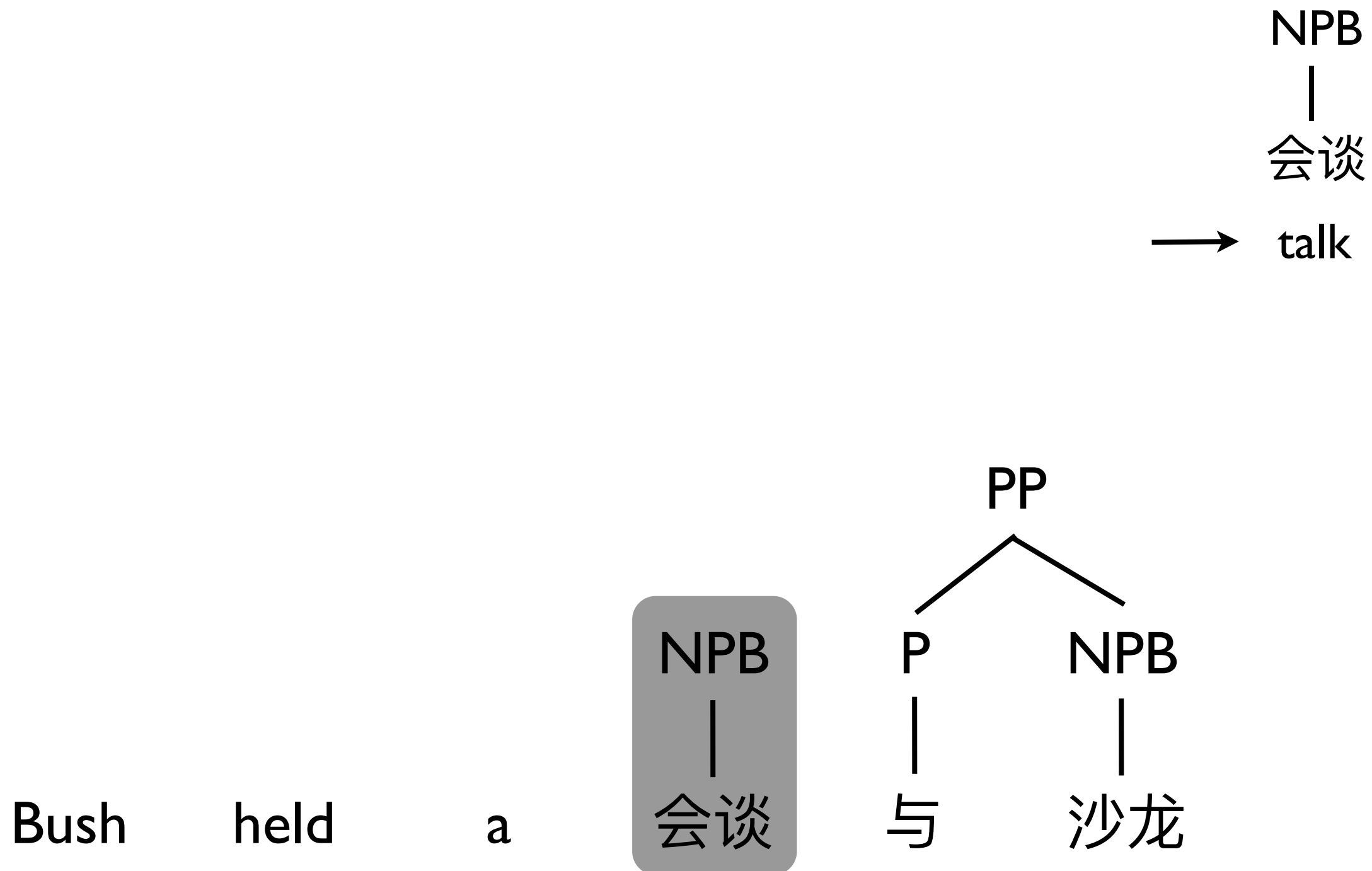
(Liu et al., 2006; Huang et al., 2006)

# Tree-to-String Translation



(Liu et al., 2006; Huang et al., 2006)

# Tree-to-String Translation



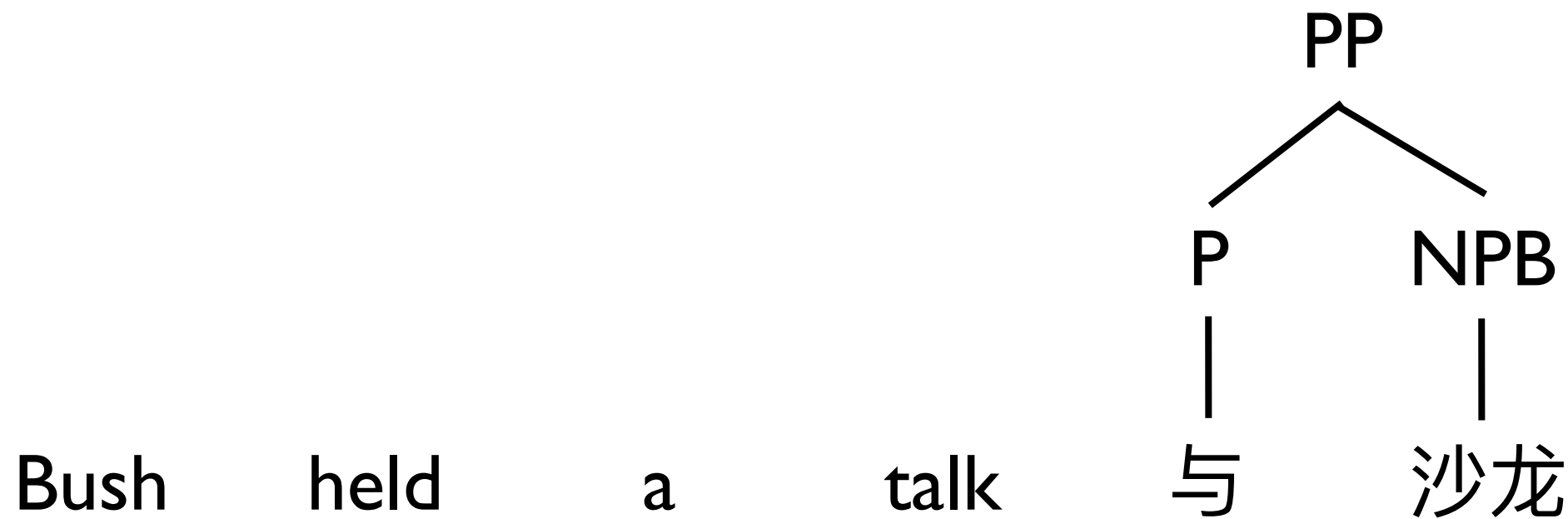
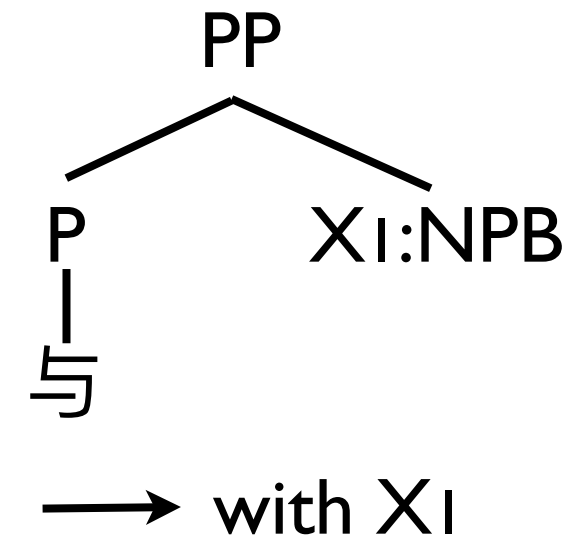
(Liu et al., 2006; Huang et al., 2006)

# Tree-to-String Translation



(Liu et al., 2006; Huang et al., 2006)

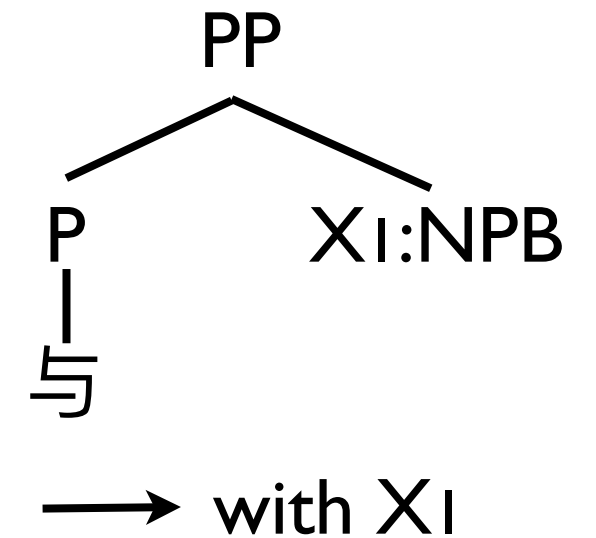
# Tree-to-String Translation



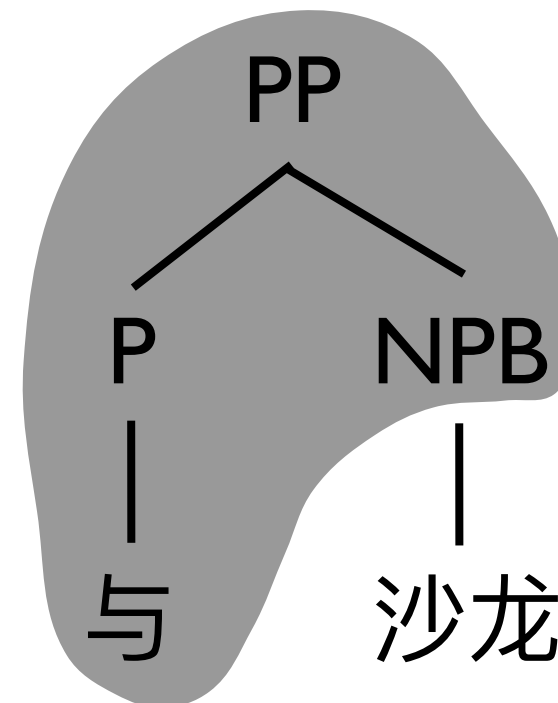
(Liu et al., 2006; Huang et al., 2006)



# Tree-to-String Translation



Bush held a talk



(Liu et al., 2006; Huang et al., 2006)

# Tree-to-String Translation

Bush held a talk with NPB  
沙龙

(Liu et al., 2006; Huang et al., 2006)

# Tree-to-String Translation

**NPB**

1

# 沙龙

→ Sharon

# Bush

# held

a

# talk

with

# NPB

1

沙龙

(Liu et al., 2006; Huang et al., 2006)

# Tree-to-String Translation

NPB  
|  
沙龙  
→ Sharon

Bush held a talk with

NPB  
|  
沙龙

(Liu et al., 2006; Huang et al., 2006)

# Tree-to-String Translation

Bush held a talk with Sharon

(Liu et al., 2006; Huang et al., 2006)

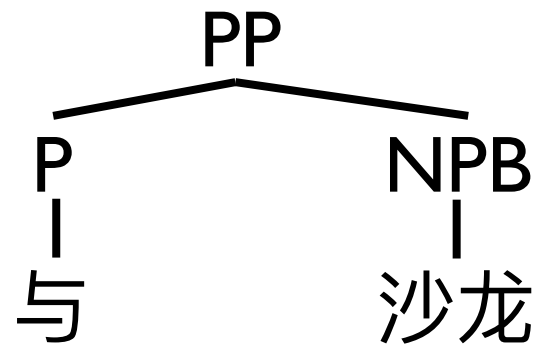
# Expressive Power

# Expressive Power

phrase translation

# Expressive Power

phrase translation



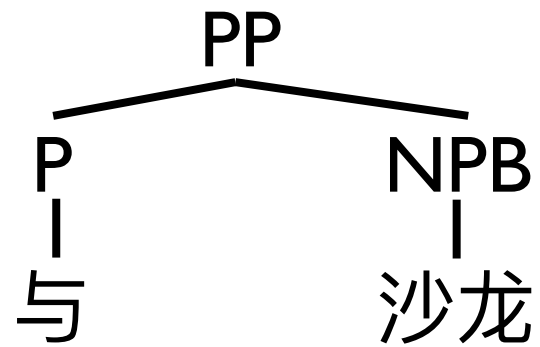
→ with Sharon



# Expressive Power

phrase translation

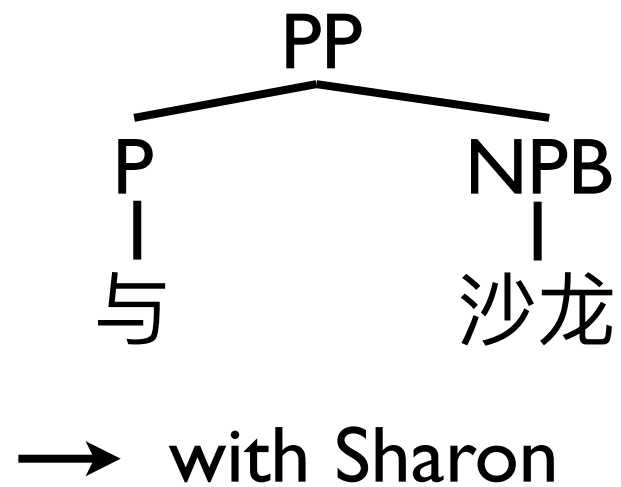
non-constituent



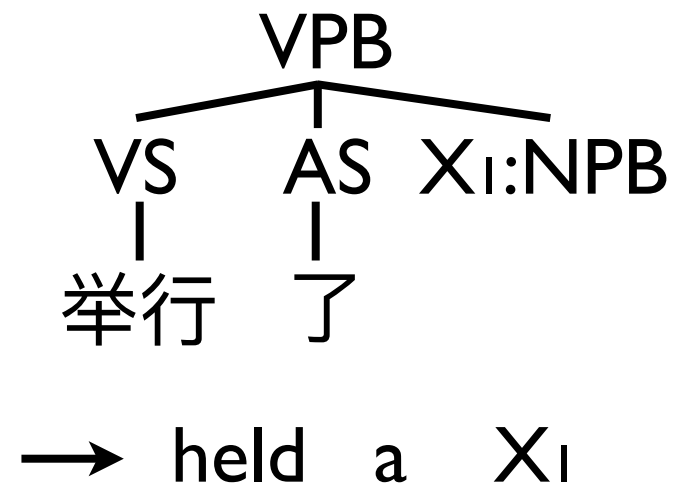
→ with Sharon

# Expressive Power

phrase translation

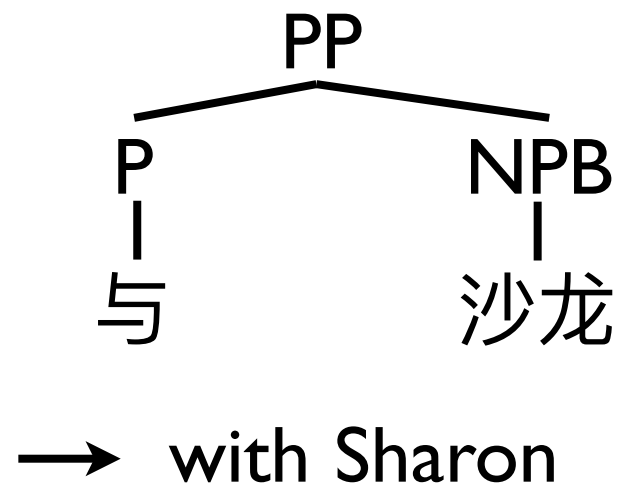


non-constituent

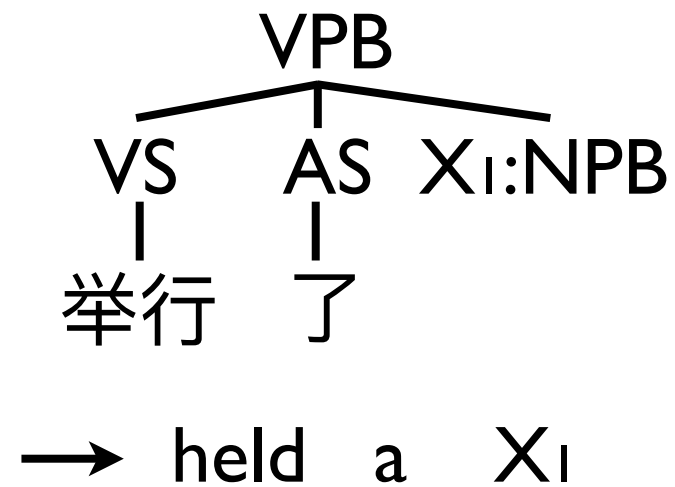


# Expressive Power

phrase translation



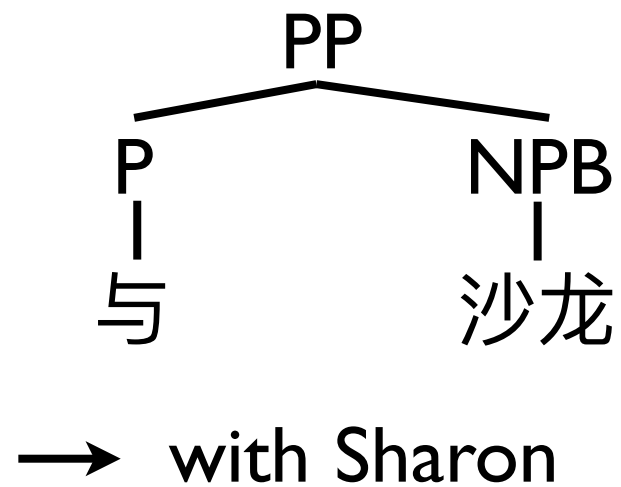
non-constituent



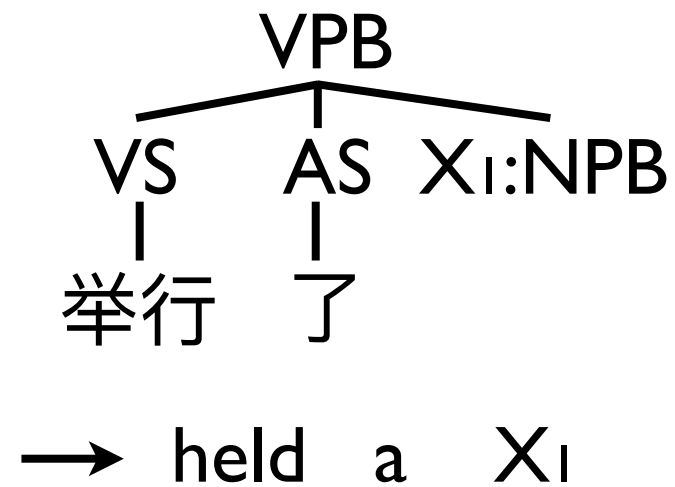
discontinuous

# Expressive Power

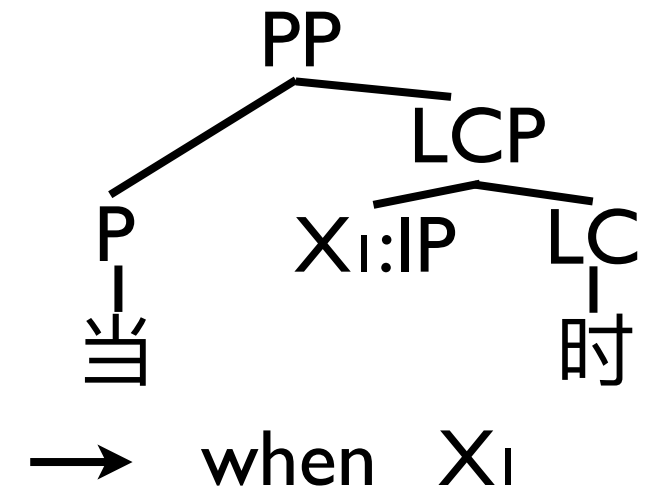
phrase translation



non-constituent

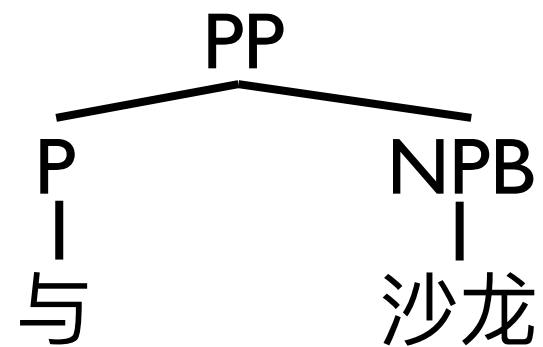


discontinuous



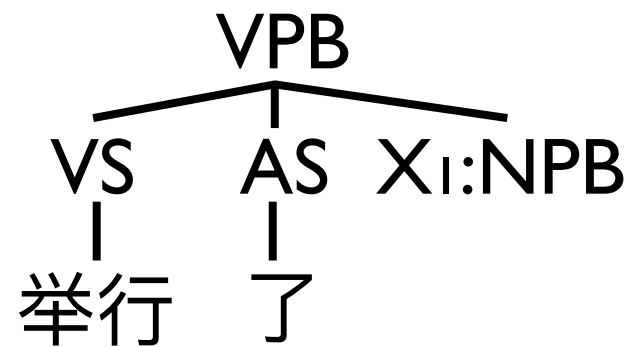
# Expressive Power

phrase translation



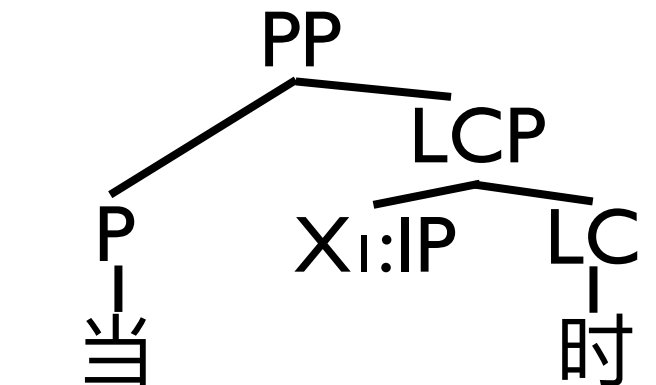
→ with Sharon

non-constituent



→ held a Xi

discontinuous

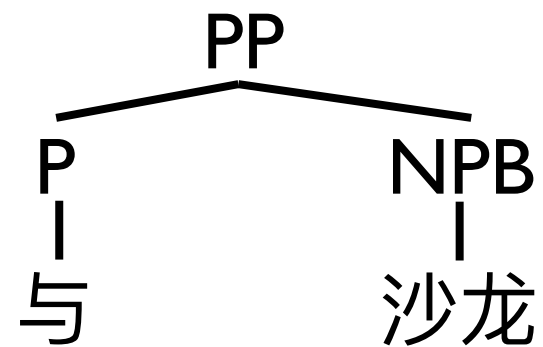


→ when Xi

word deletion

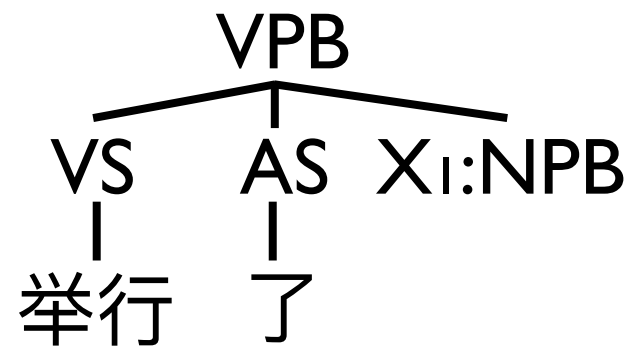
# Expressive Power

phrase translation



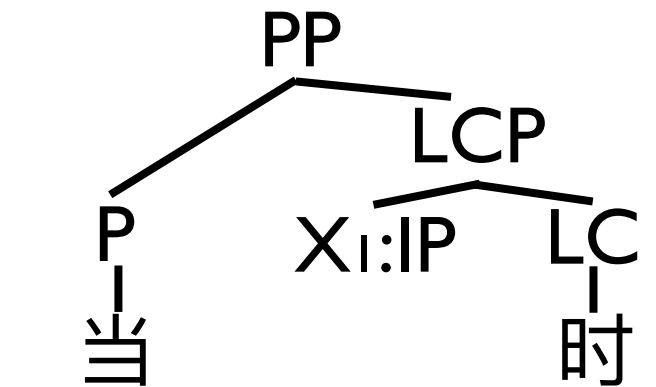
→ with Sharon

non-constituent



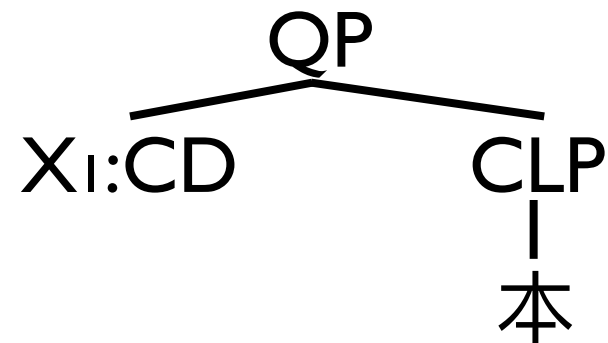
→ held a Xi

discontinuous



→ when Xi

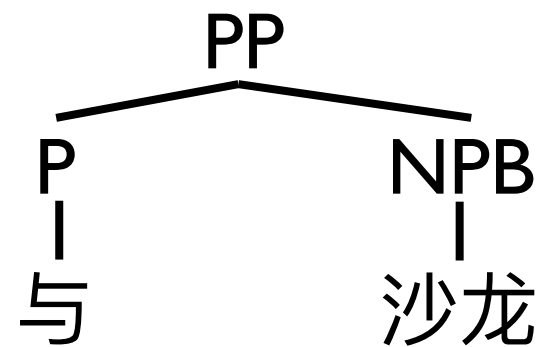
word deletion



→ Xi

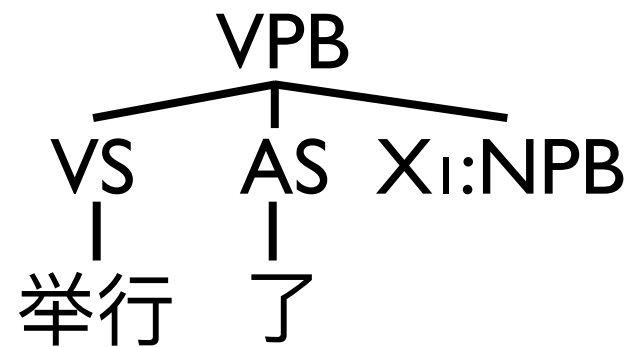
# Expressive Power

phrase translation



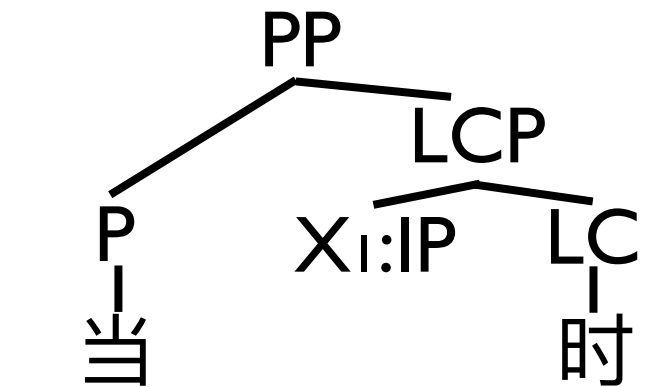
→ with Sharon

non-constituent



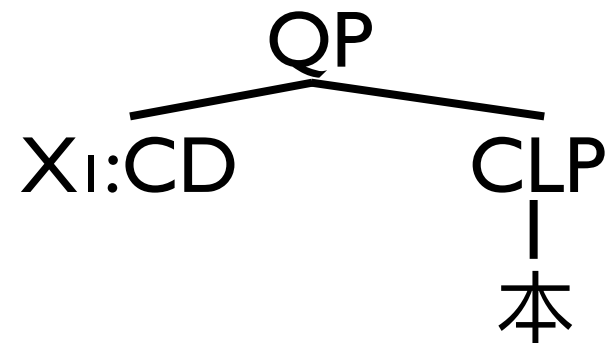
→ held a Xi

discontinuous



→ when Xi

word deletion

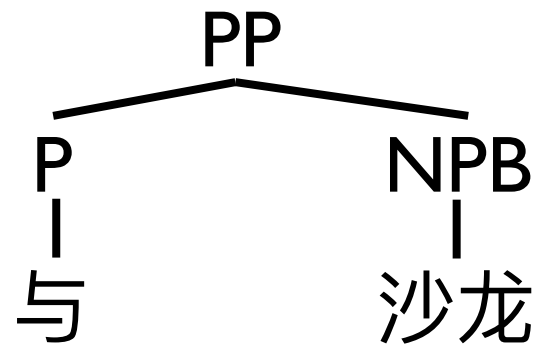


→ Xi

multi-level reordering

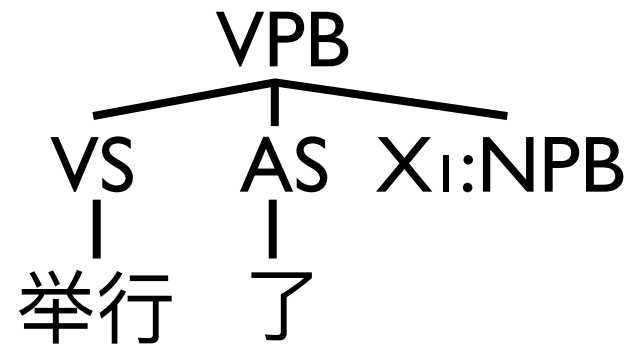
# Expressive Power

phrase translation



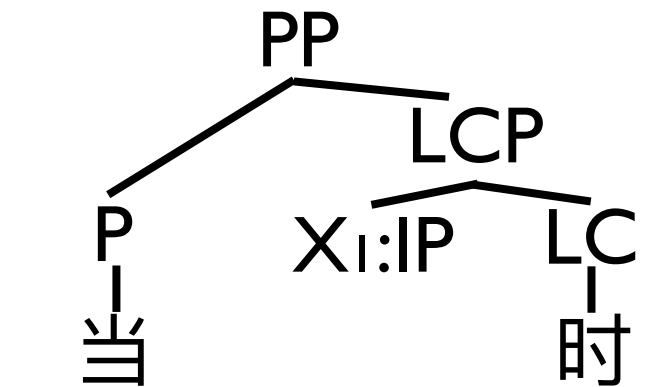
→ with Sharon

non-constituent



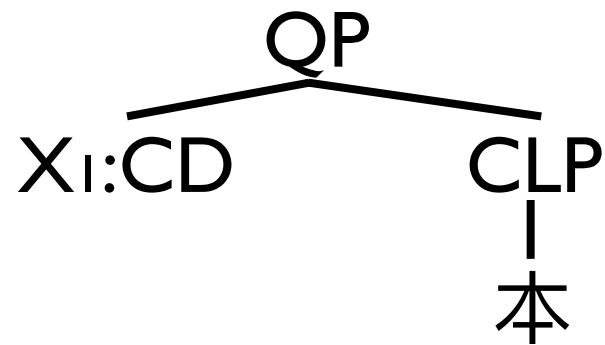
→ held a Xi

discontinuous



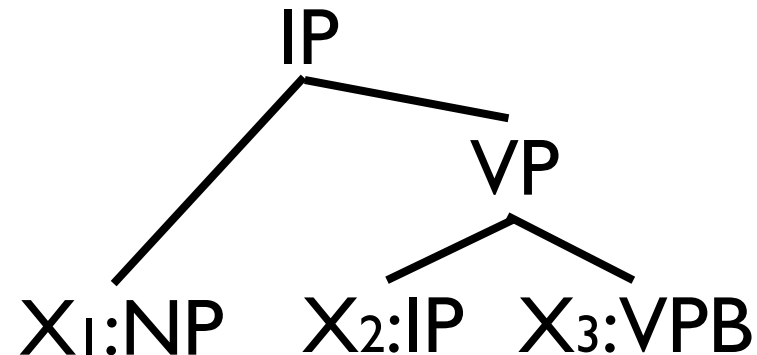
→ when Xi

word deletion



→ Xi

multi-level reordering



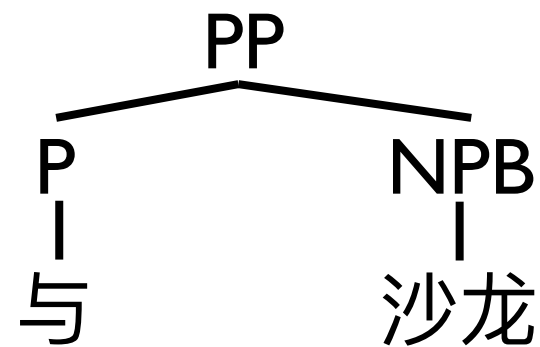
→ Xi X3 X2

(Knight and Graehl, 2005)



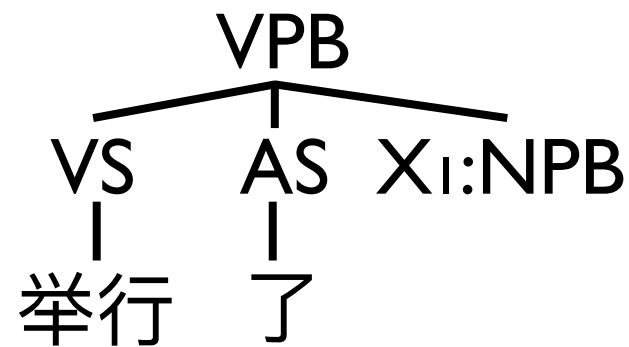
# Expressive Power

phrase translation



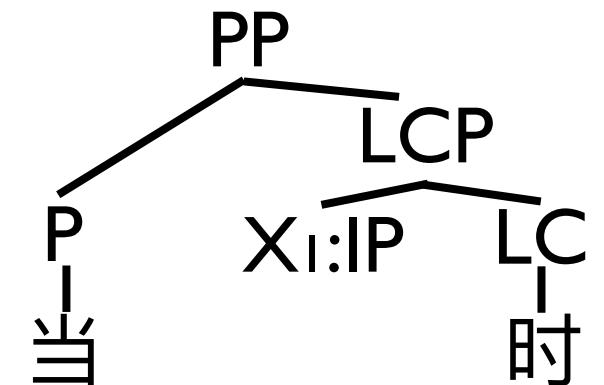
→ with Sharon

non-constituent



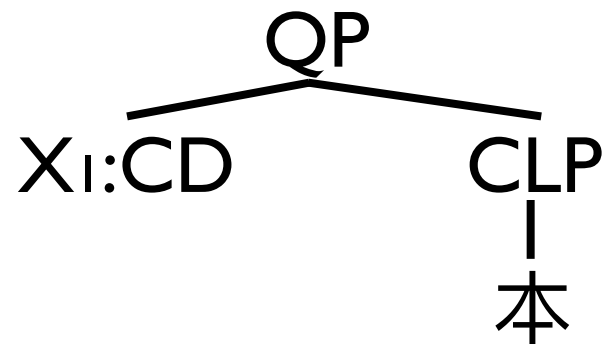
→ held a  $X_1$

discontinuous



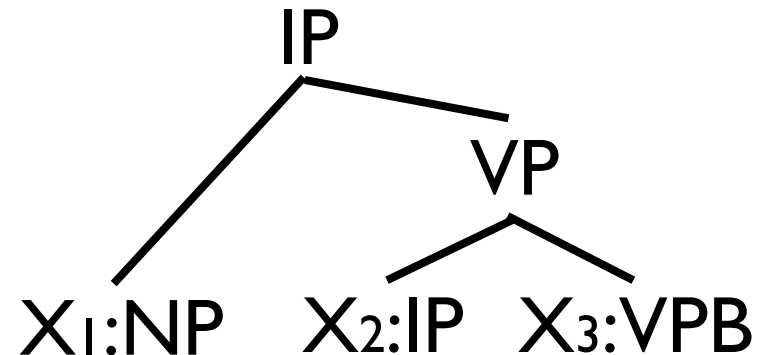
→ when  $X_1$

word deletion



→  $X_1$

multi-level reordering



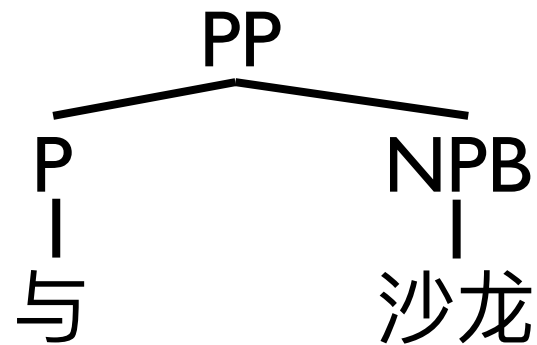
→  $X_1$   $X_3$   $X_2$

lexicalized reordering

(Knight and Graehl, 2005)

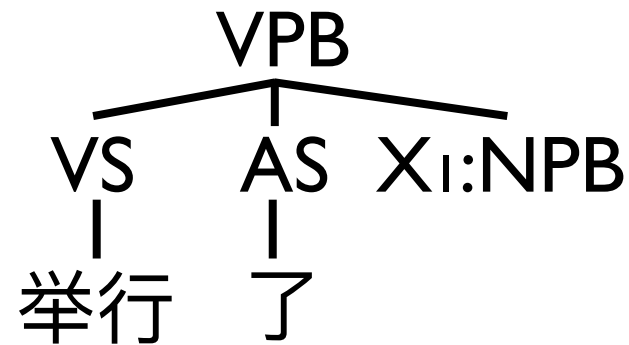
# Expressive Power

phrase translation



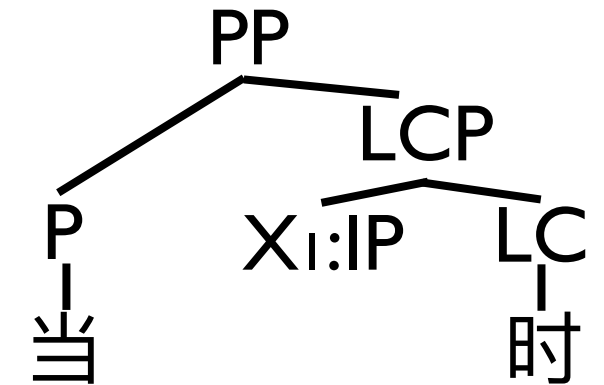
→ with Sharon

non-constituent



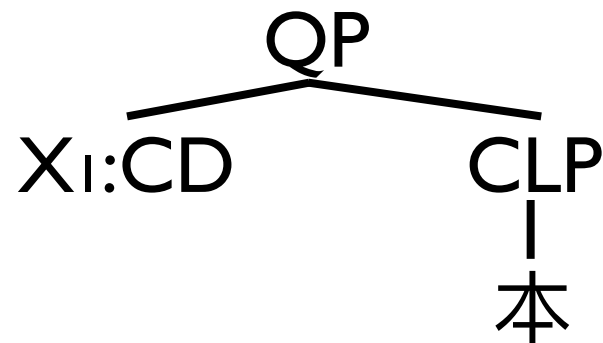
→ held a Xi

discontinuous



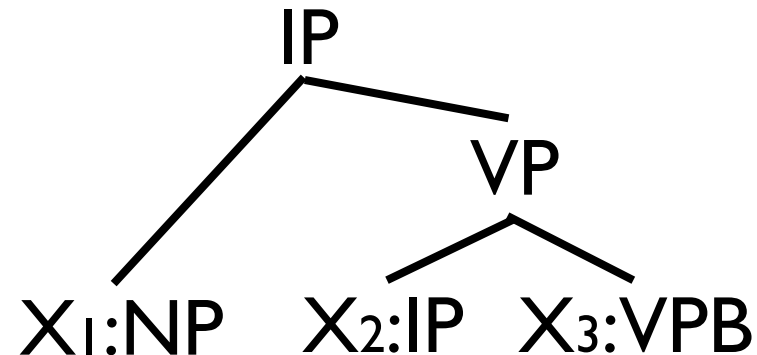
→ when Xi

word deletion



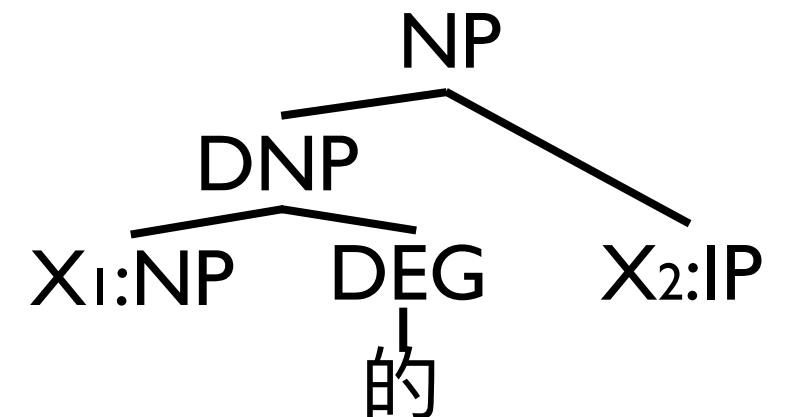
→ Xi

multi-level reordering



→ Xi X3 X2

lexicalized reordering



→ X2 of Xi

(Knight and Graehl, 2005)

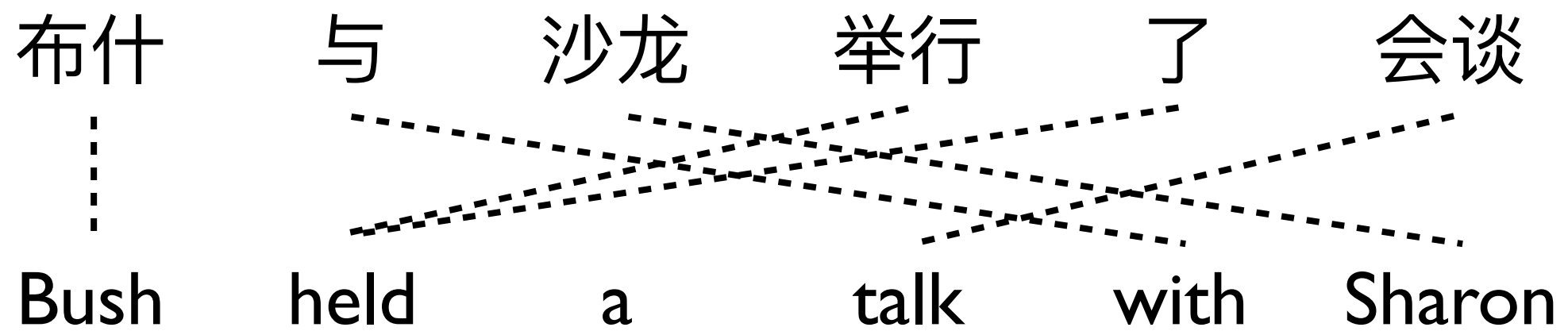
# Rule Extraction

布什 与 沙龙 举行 了 会谈

Bush held a talk with Sharon

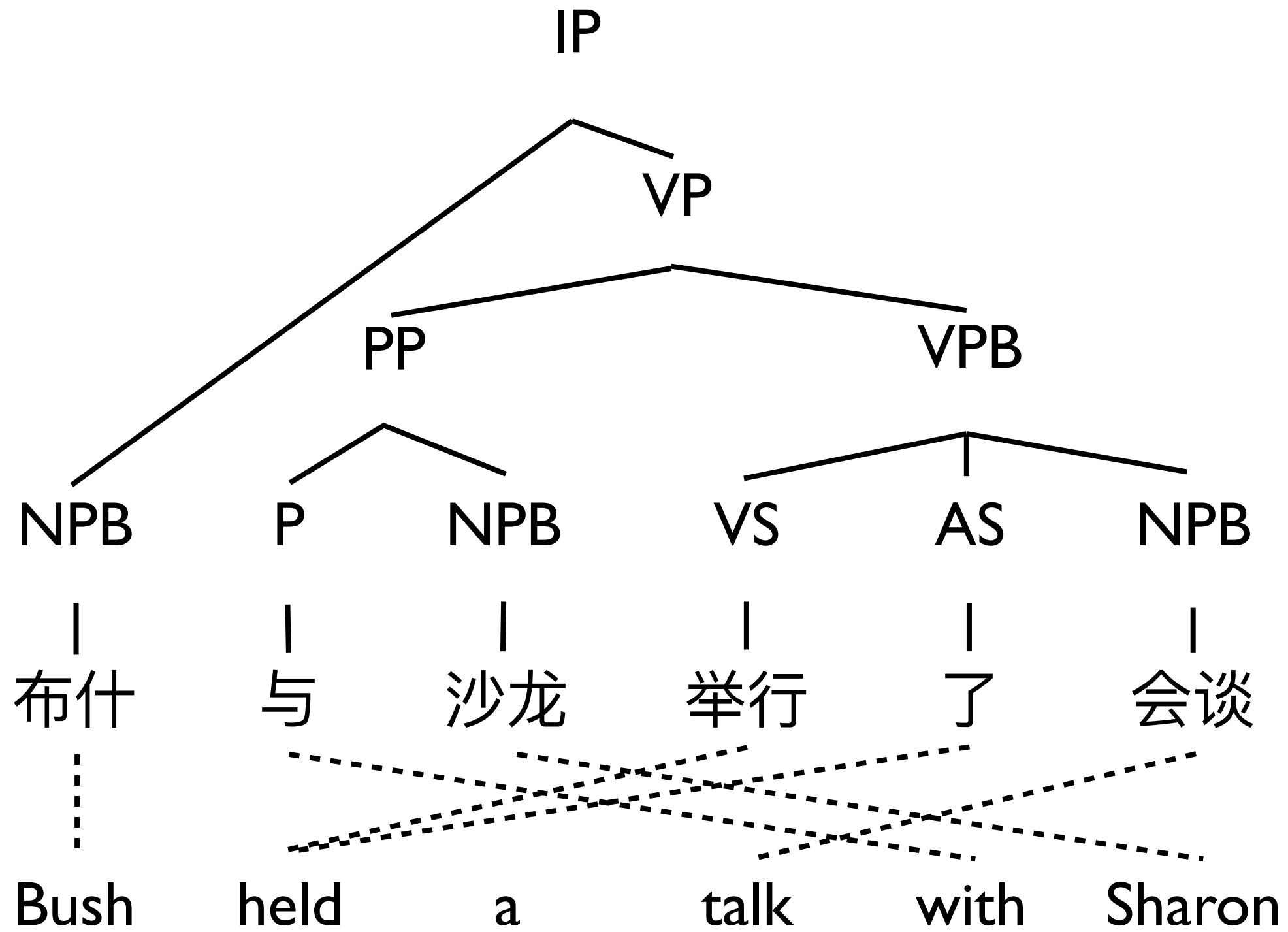
(Galley et al., 2004)

# Rule Extraction



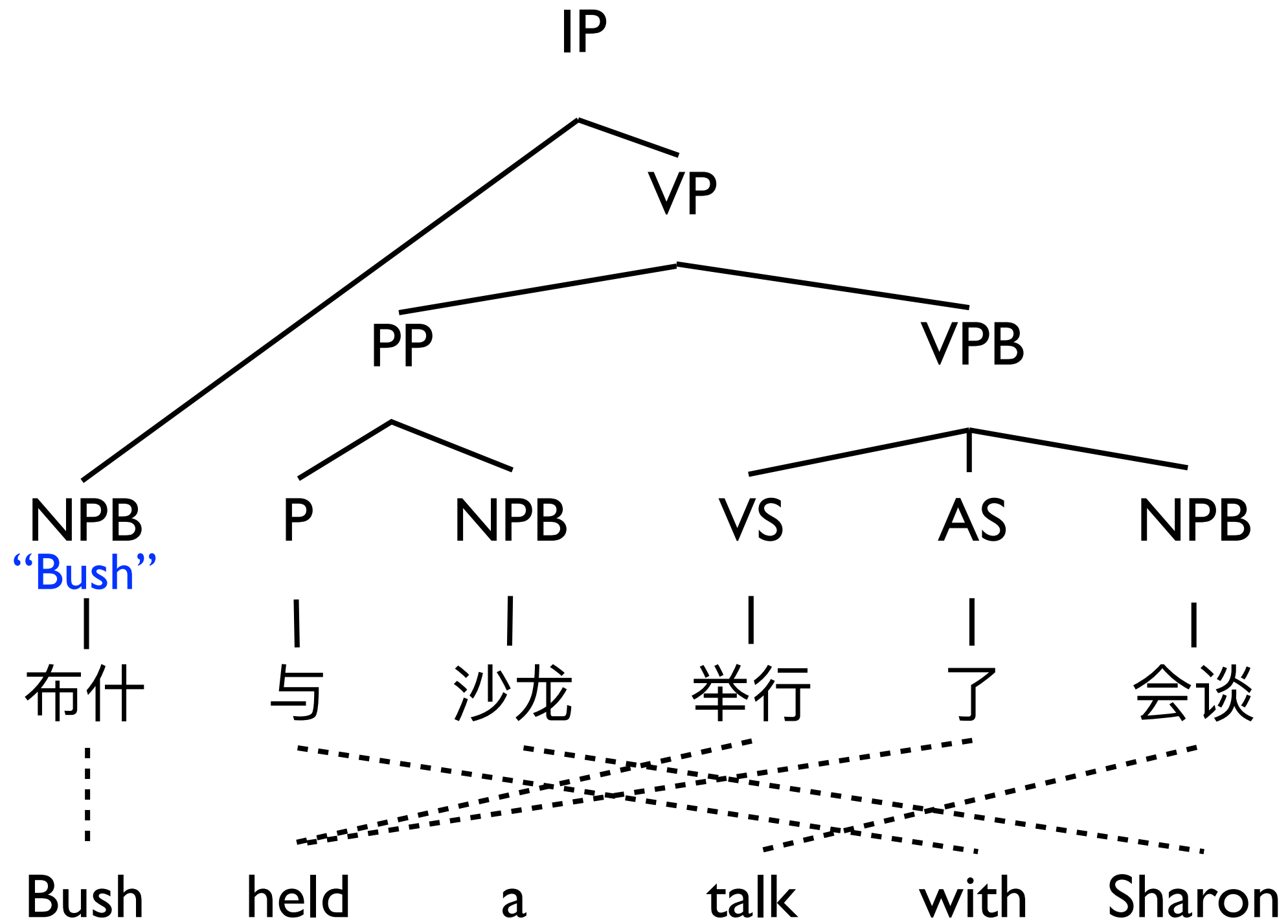
(Galley et al., 2004)

# Rule Extraction



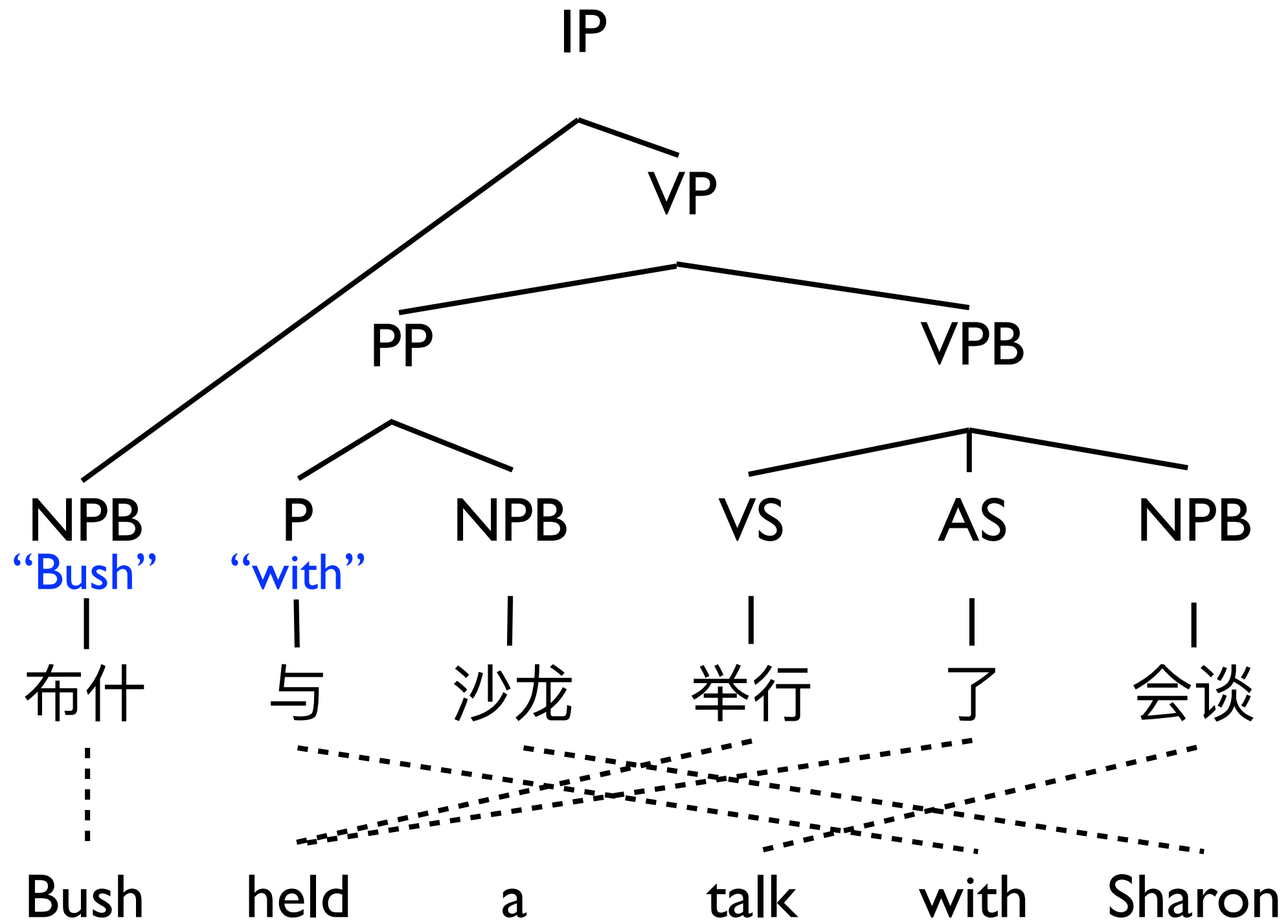
(Galley et al., 2004)

# Rule Extraction



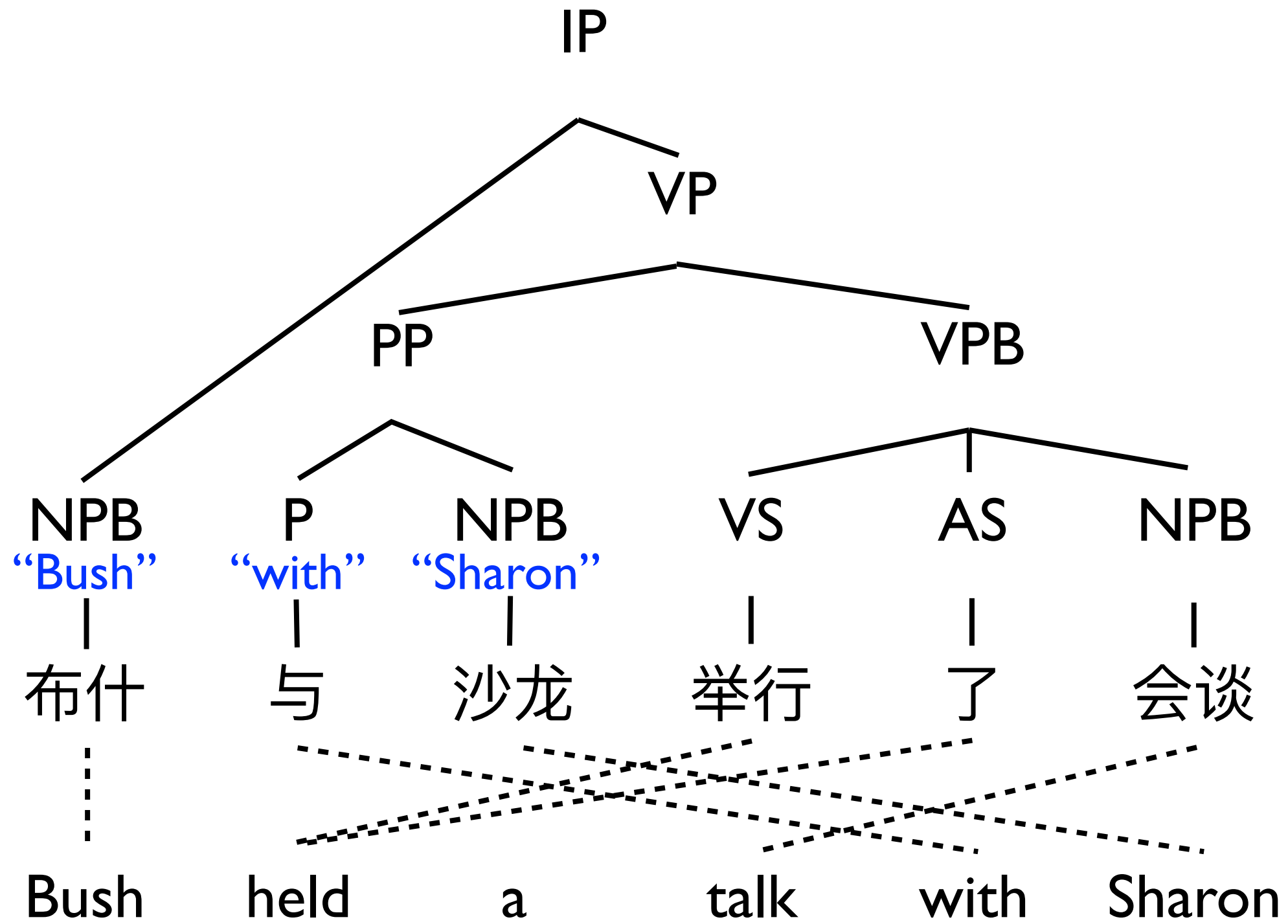
(Galley et al., 2004)

# Rule Extraction



(Galley et al., 2004)

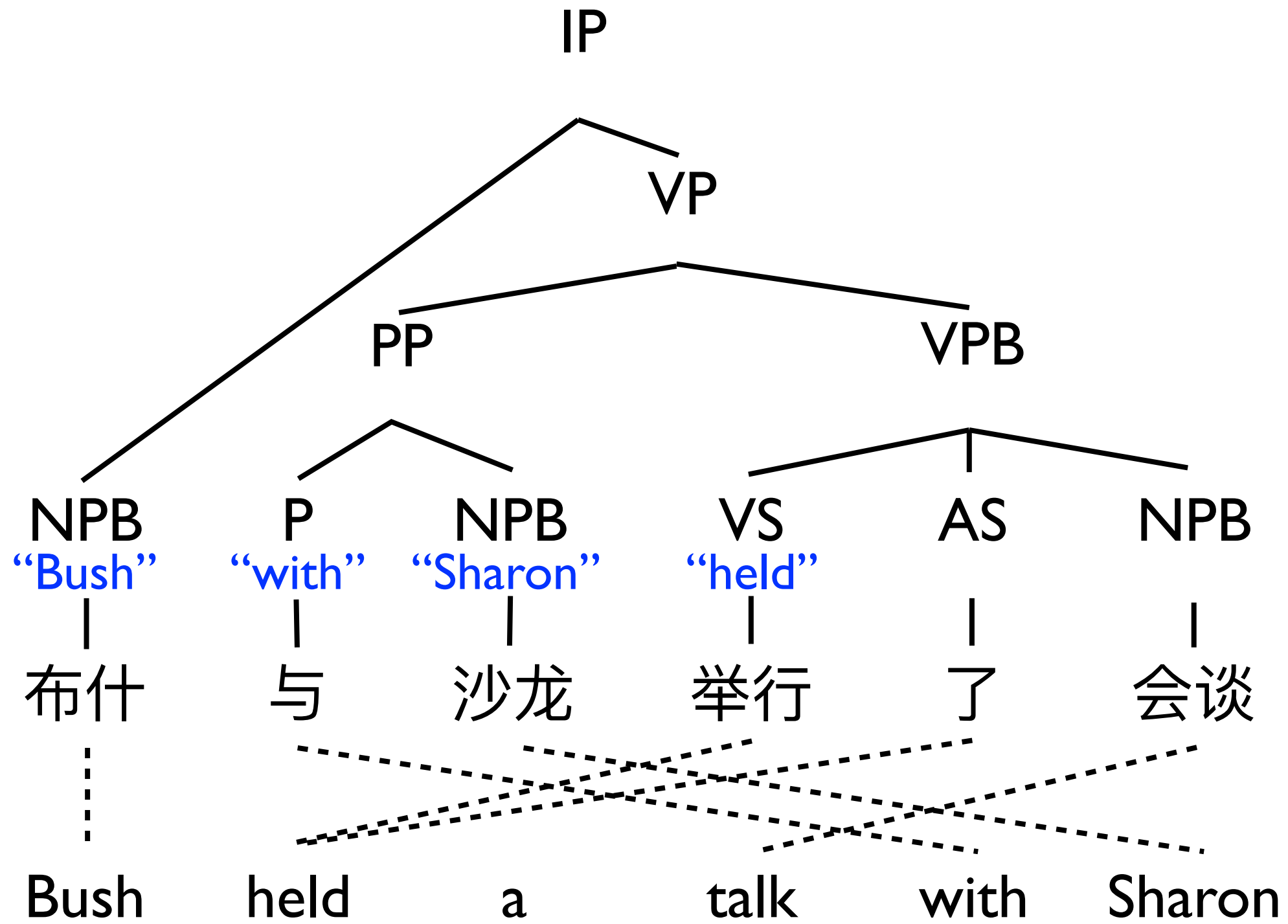
# Rule Extraction



(Galley et al., 2004)

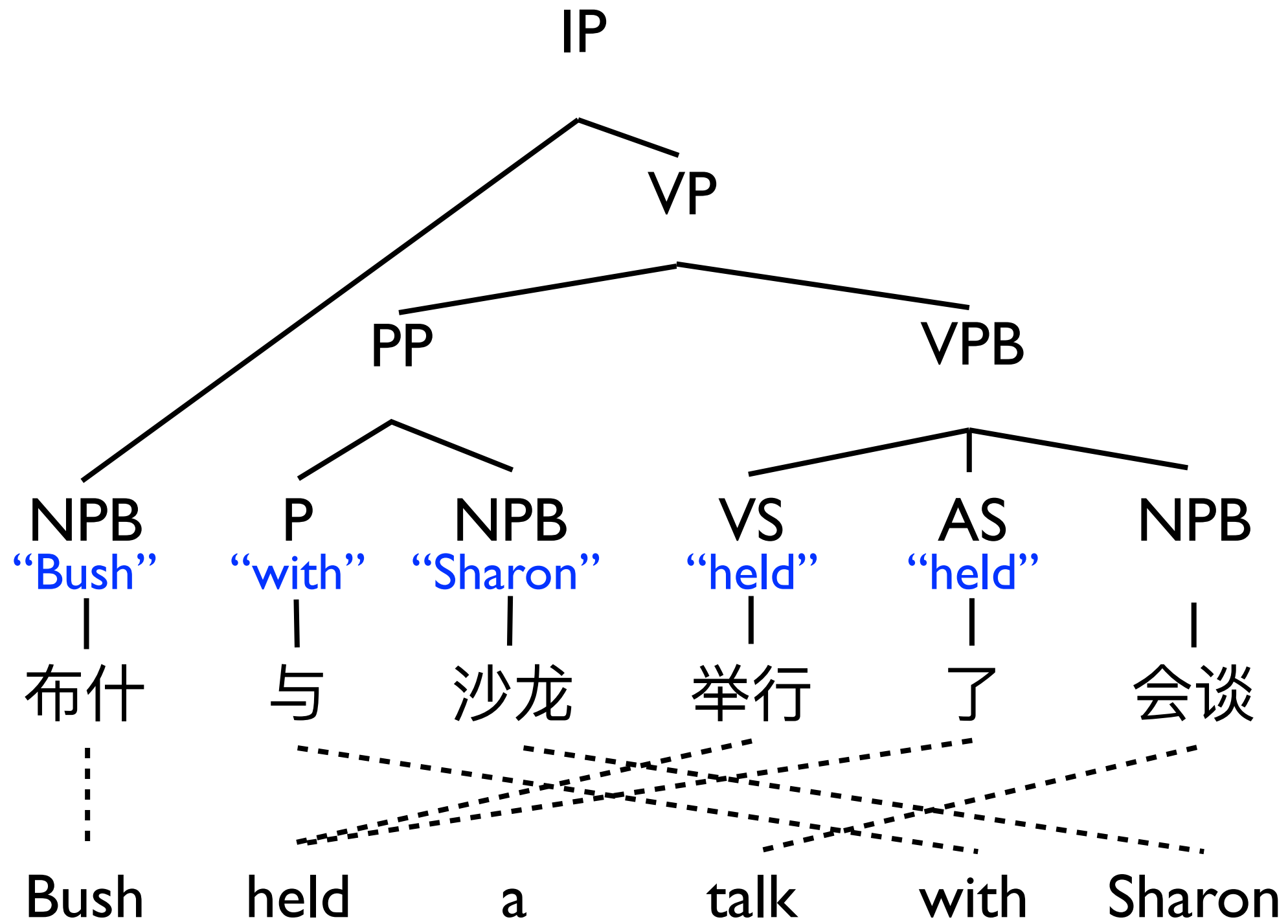


# Rule Extraction



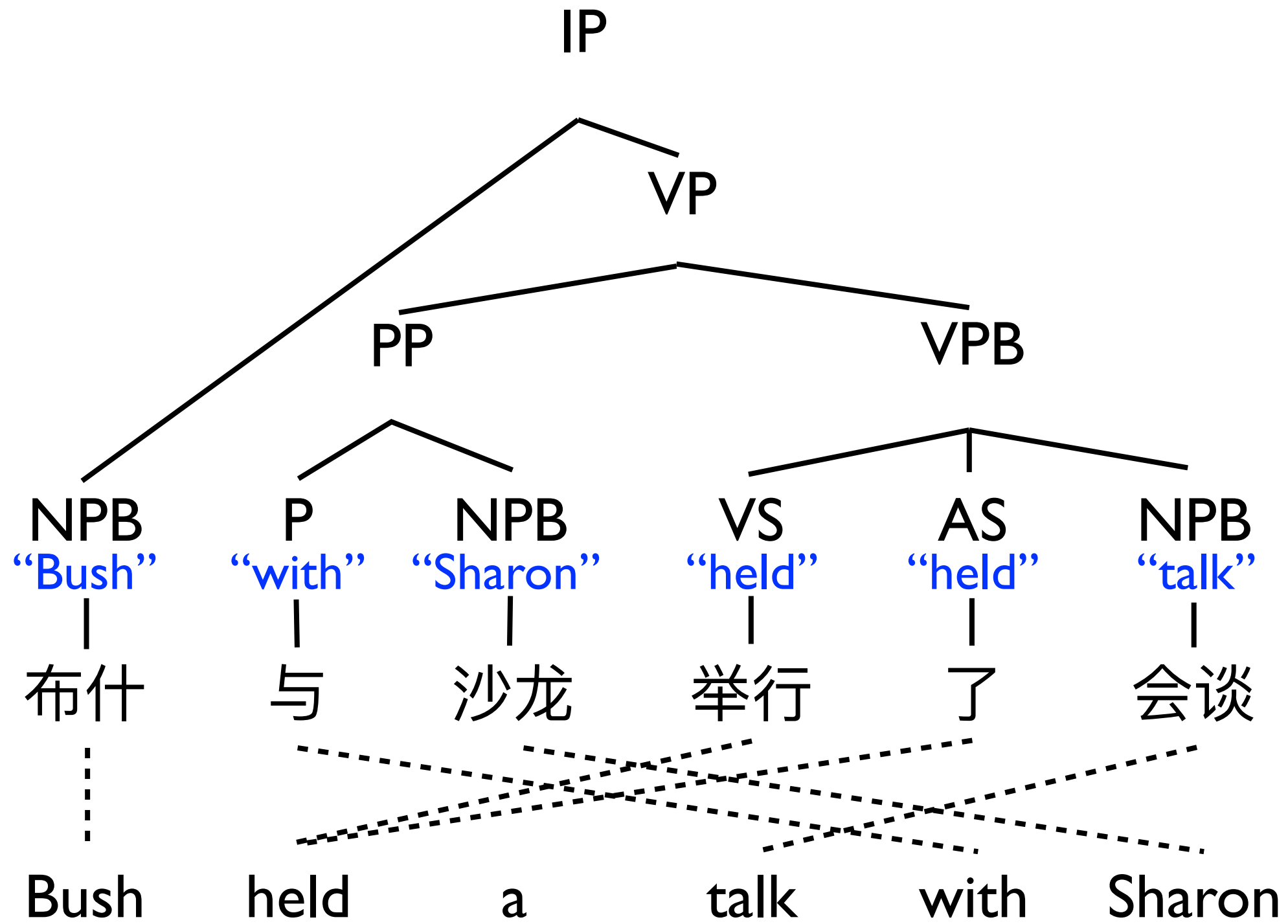
(Galley et al., 2004)

# Rule Extraction



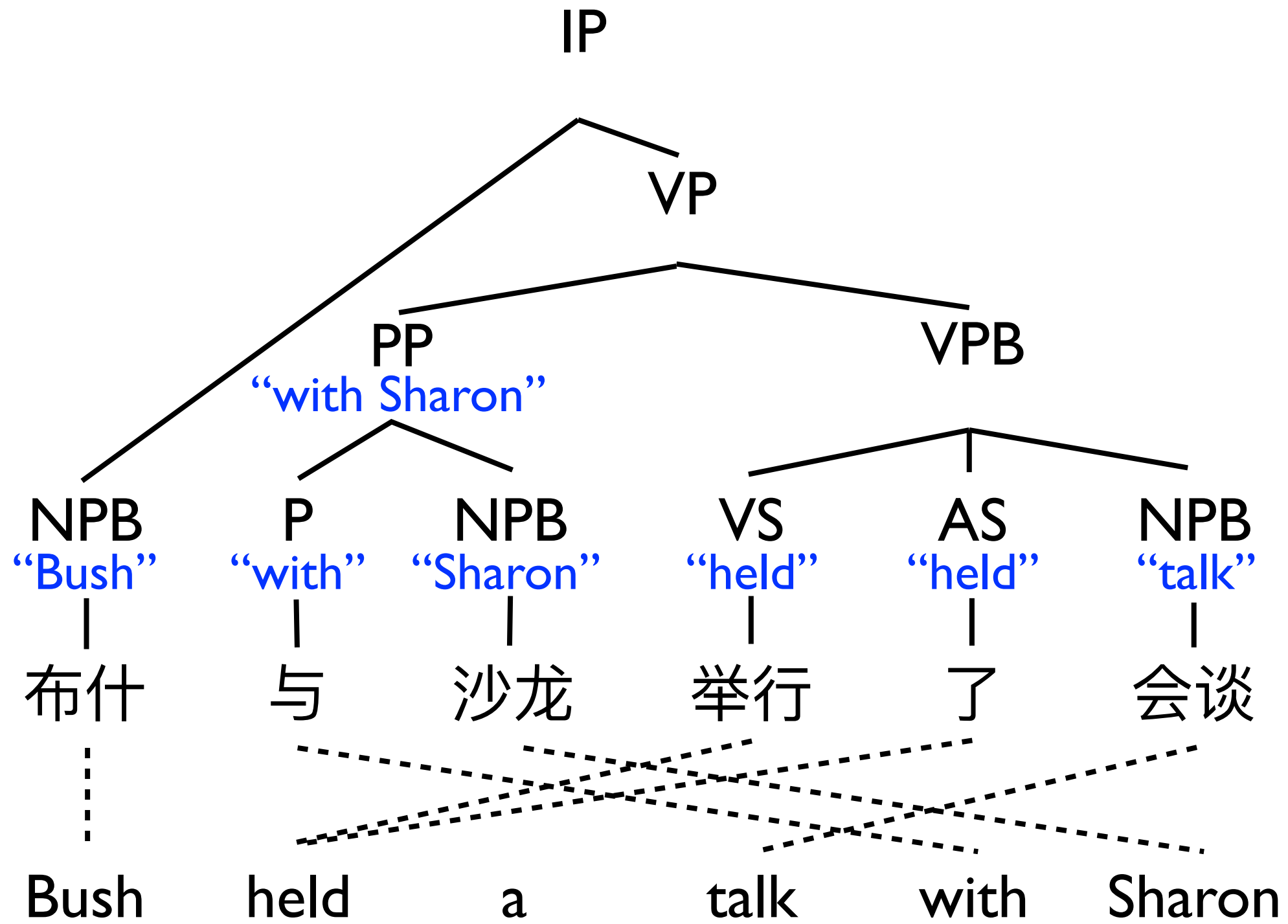
(Galley et al., 2004)

# Rule Extraction



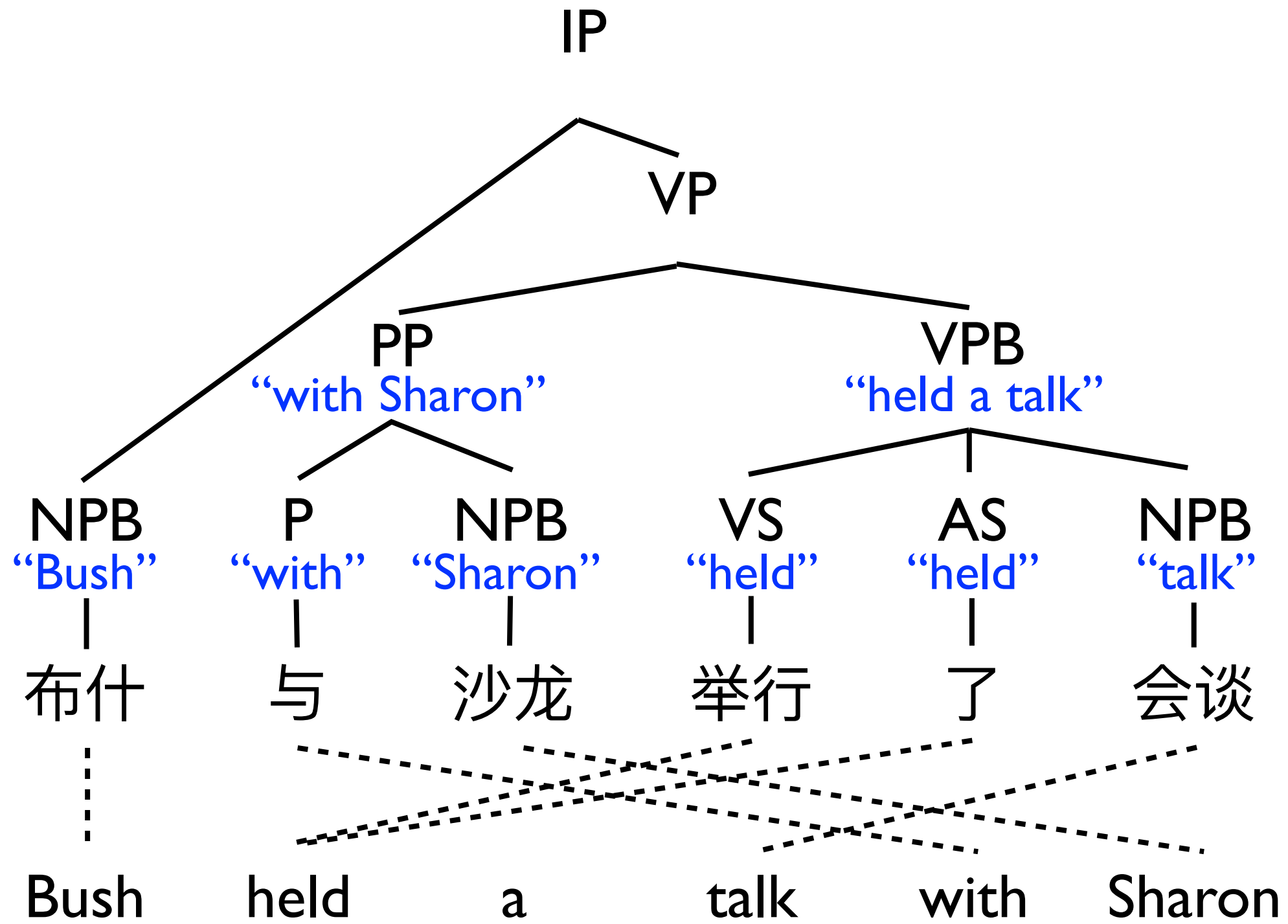
(Galley et al., 2004)

# Rule Extraction



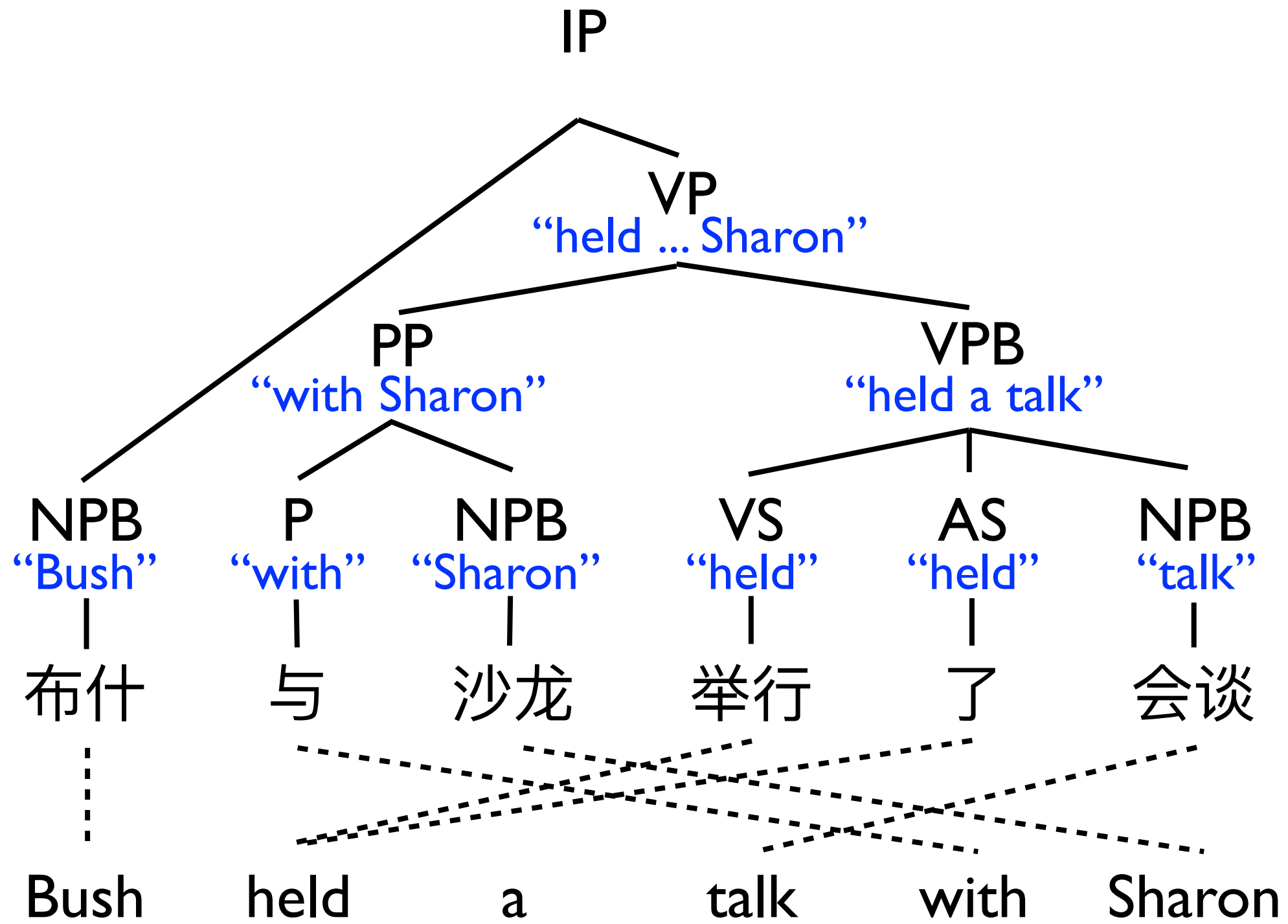
(Galley et al., 2004)

# Rule Extraction



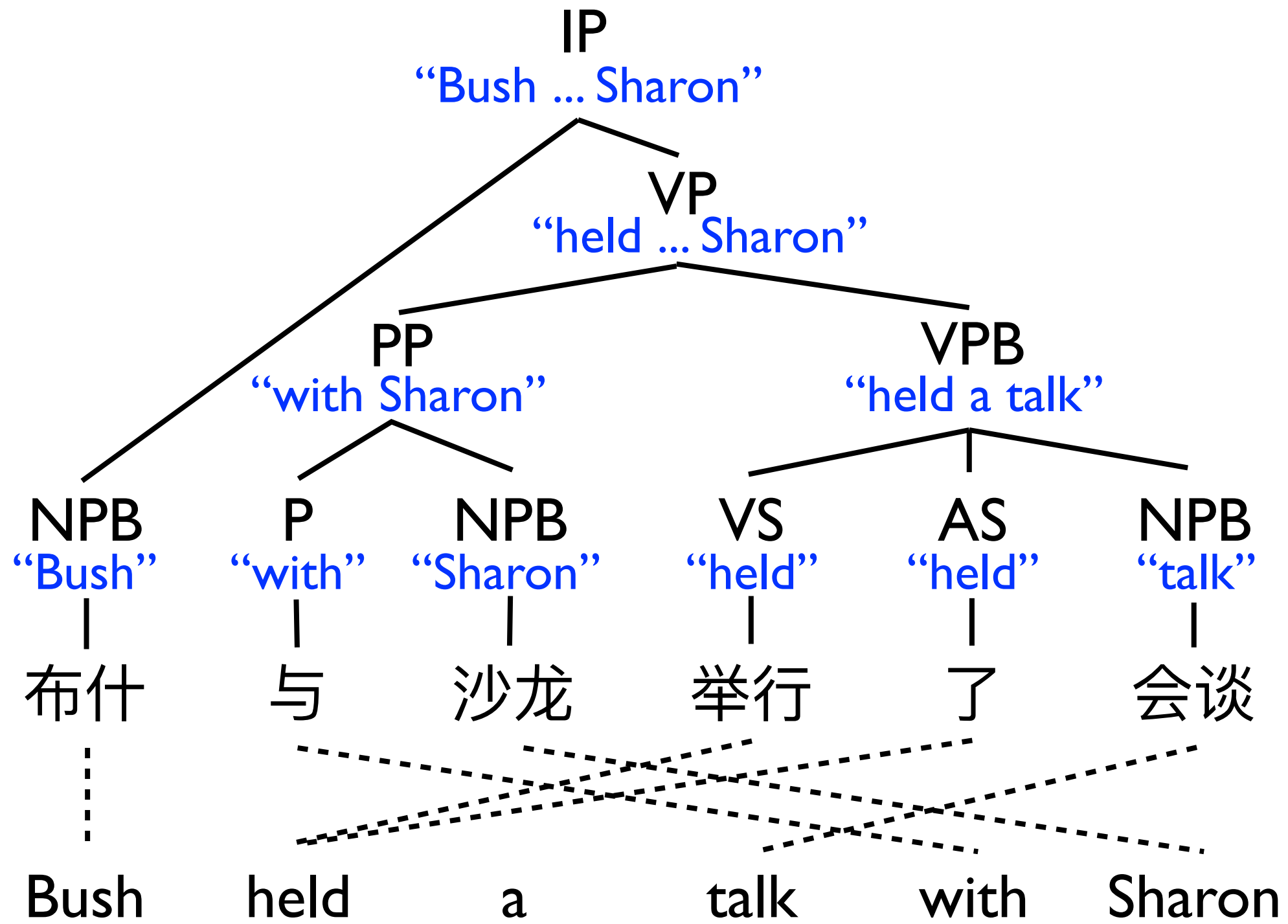
(Galley et al., 2004)

# Rule Extraction



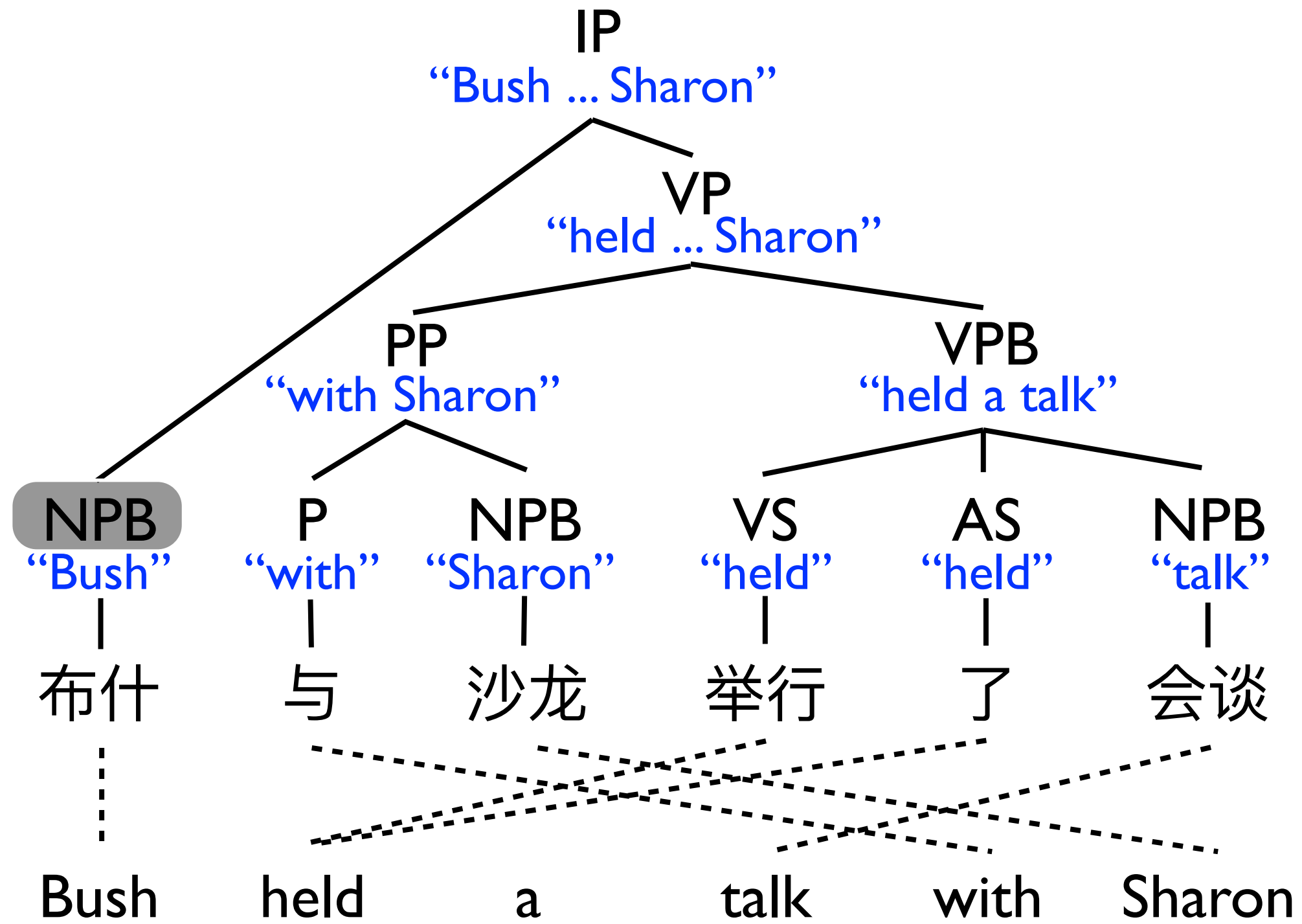
(Galley et al., 2004)

# Rule Extraction



(Galley et al., 2004)

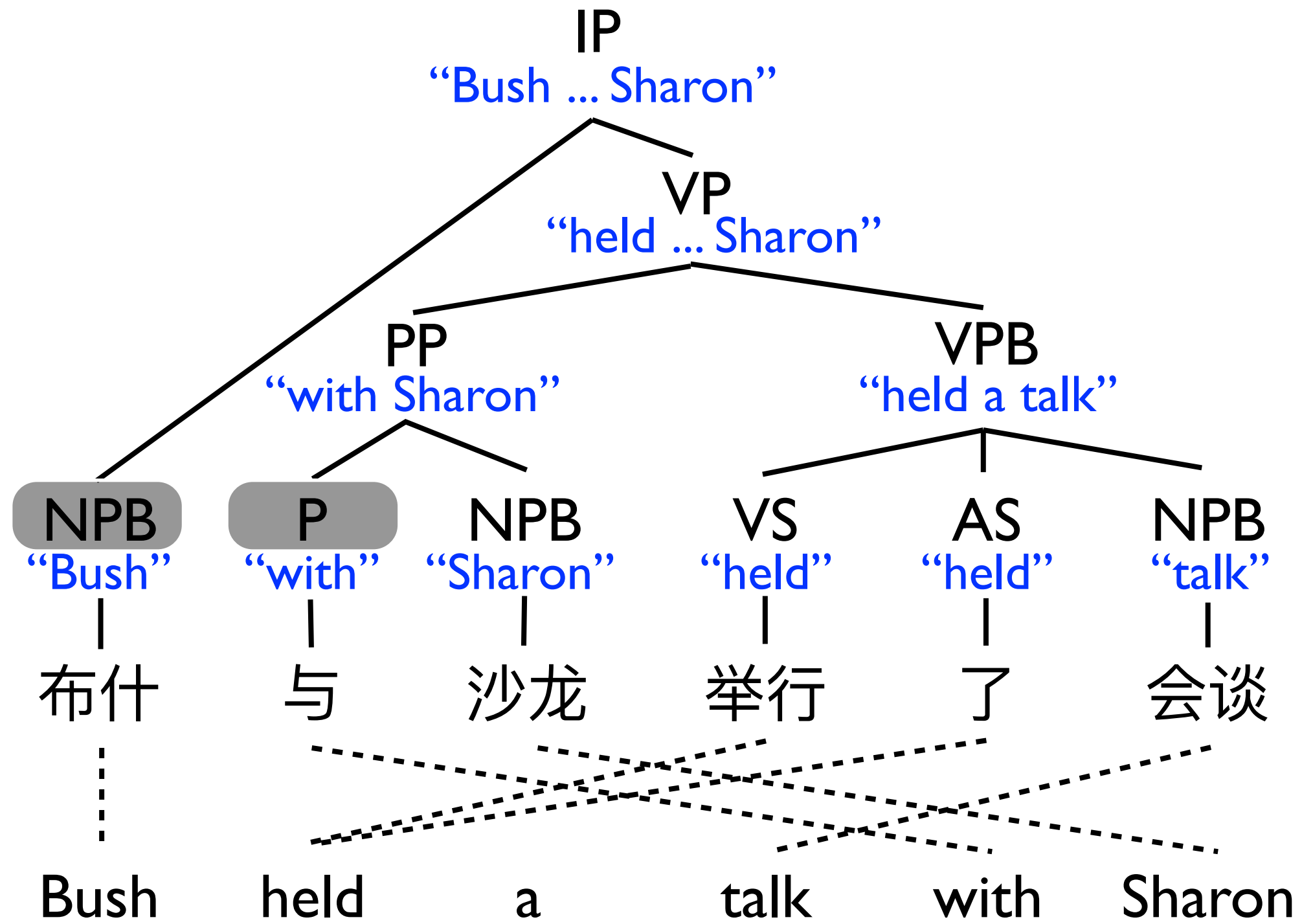
# Rule Extraction



(Galley et al., 2004)

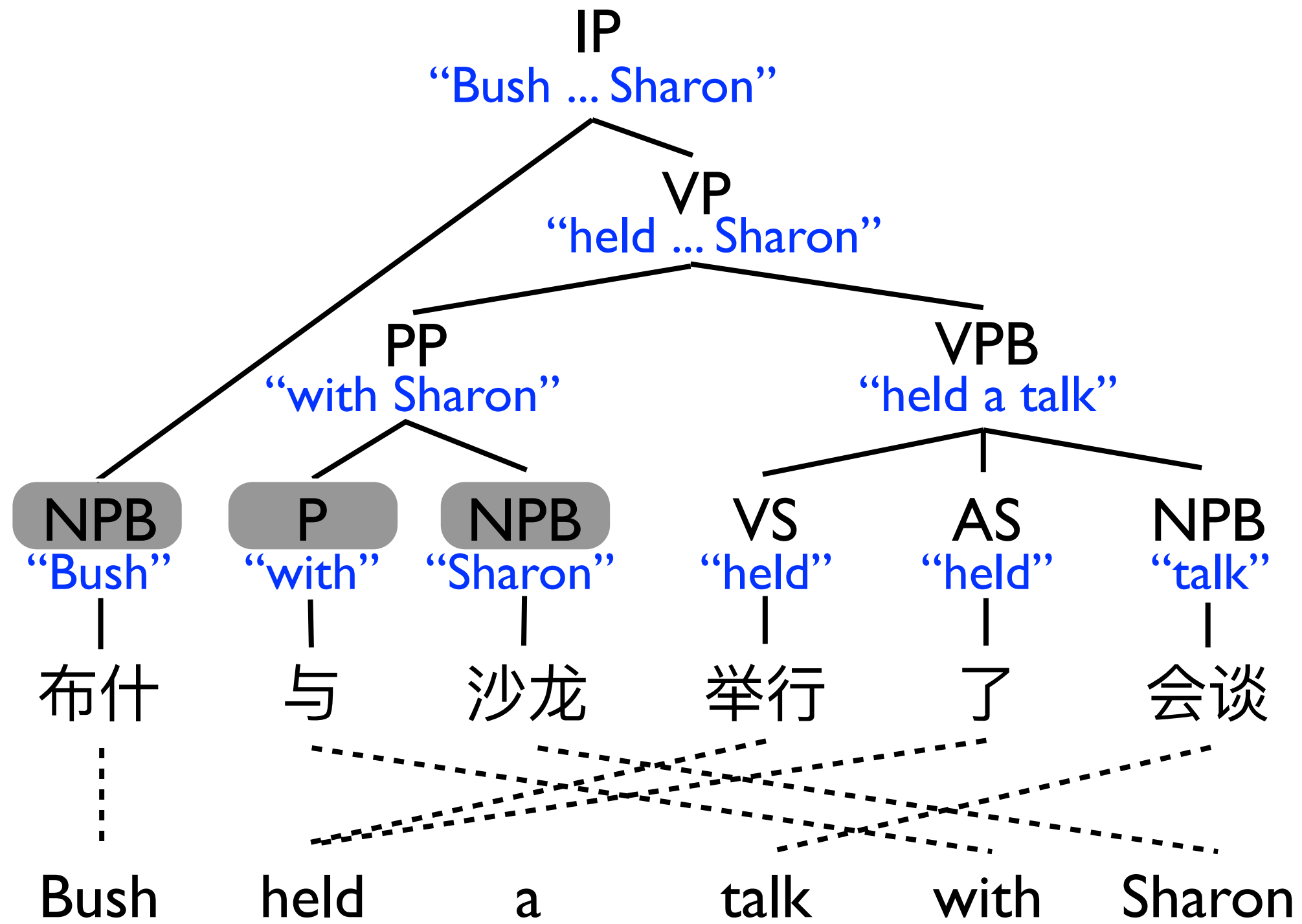


# Rule Extraction



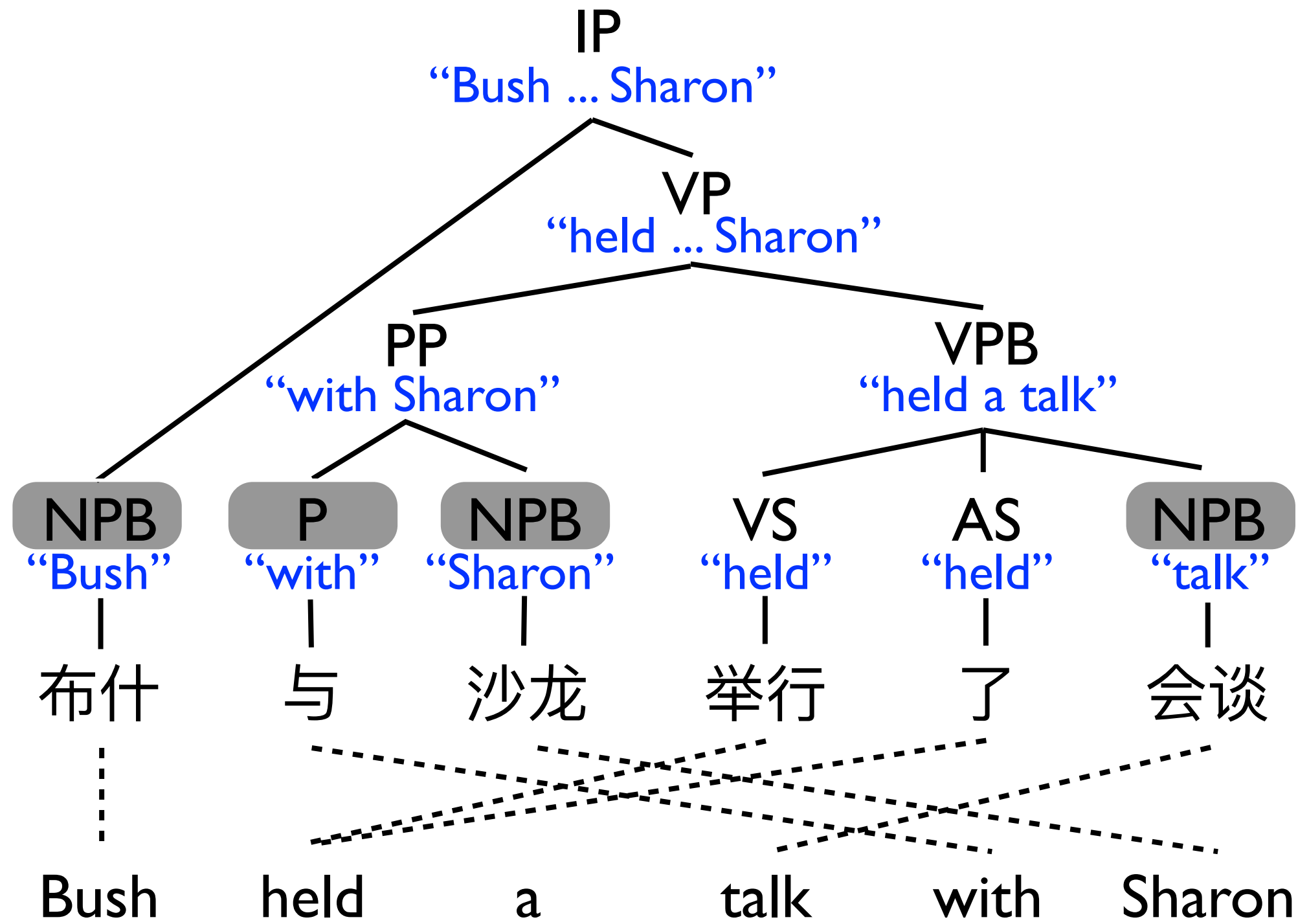
(Galley et al., 2004)

# Rule Extraction



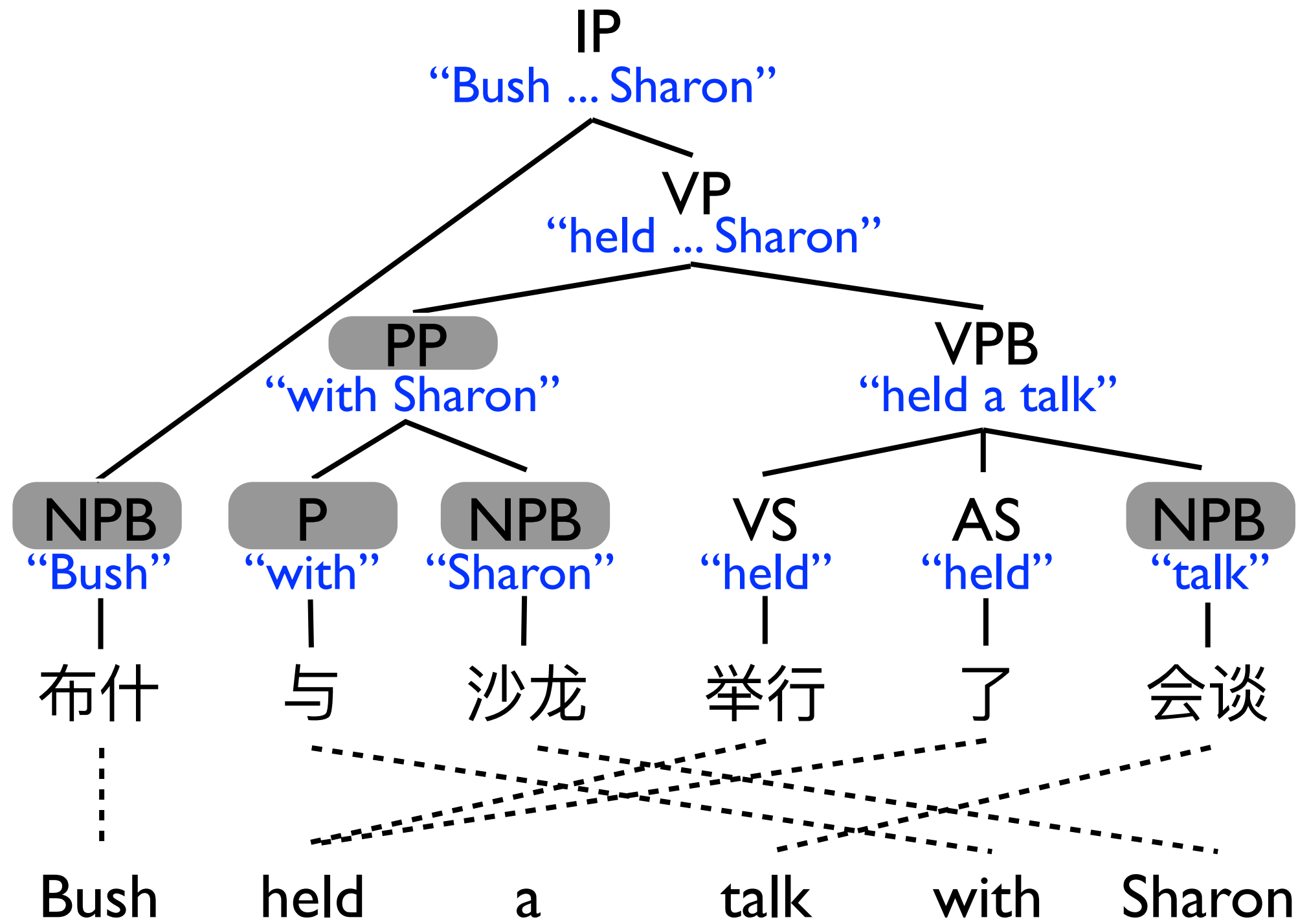
(Galley et al., 2004)

# Rule Extraction



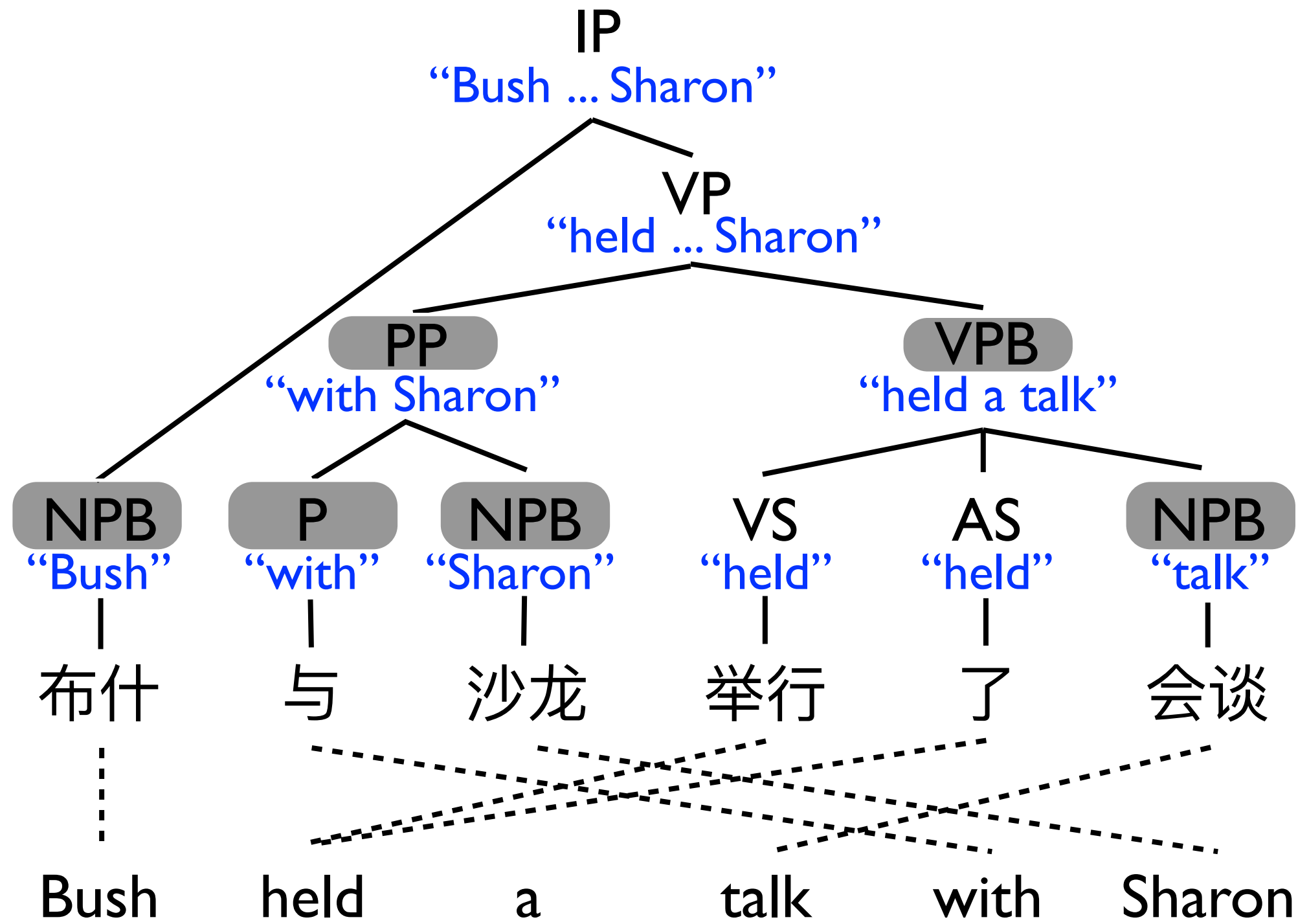
(Galley et al., 2004)

# Rule Extraction



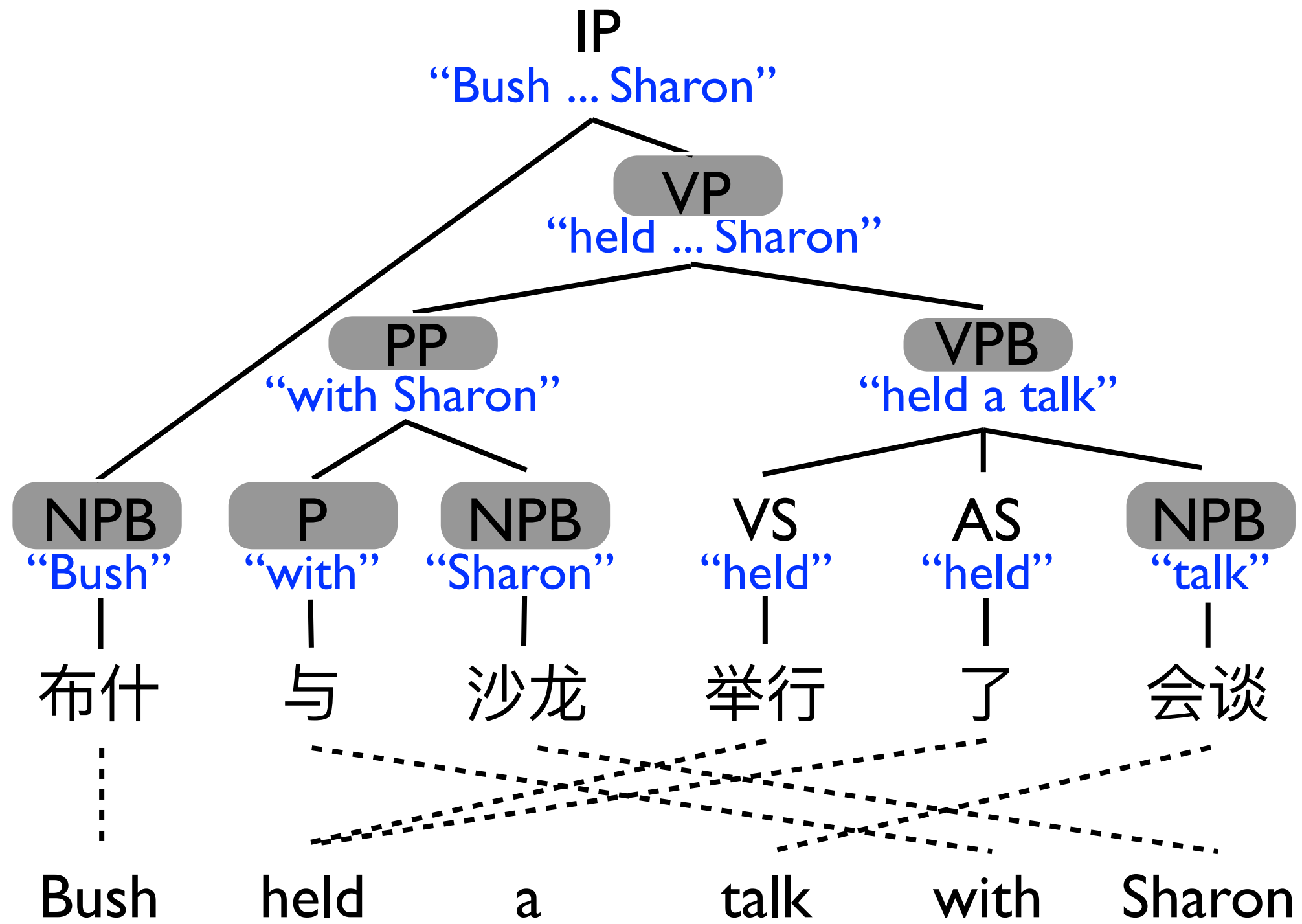
(Galley et al., 2004)

# Rule Extraction



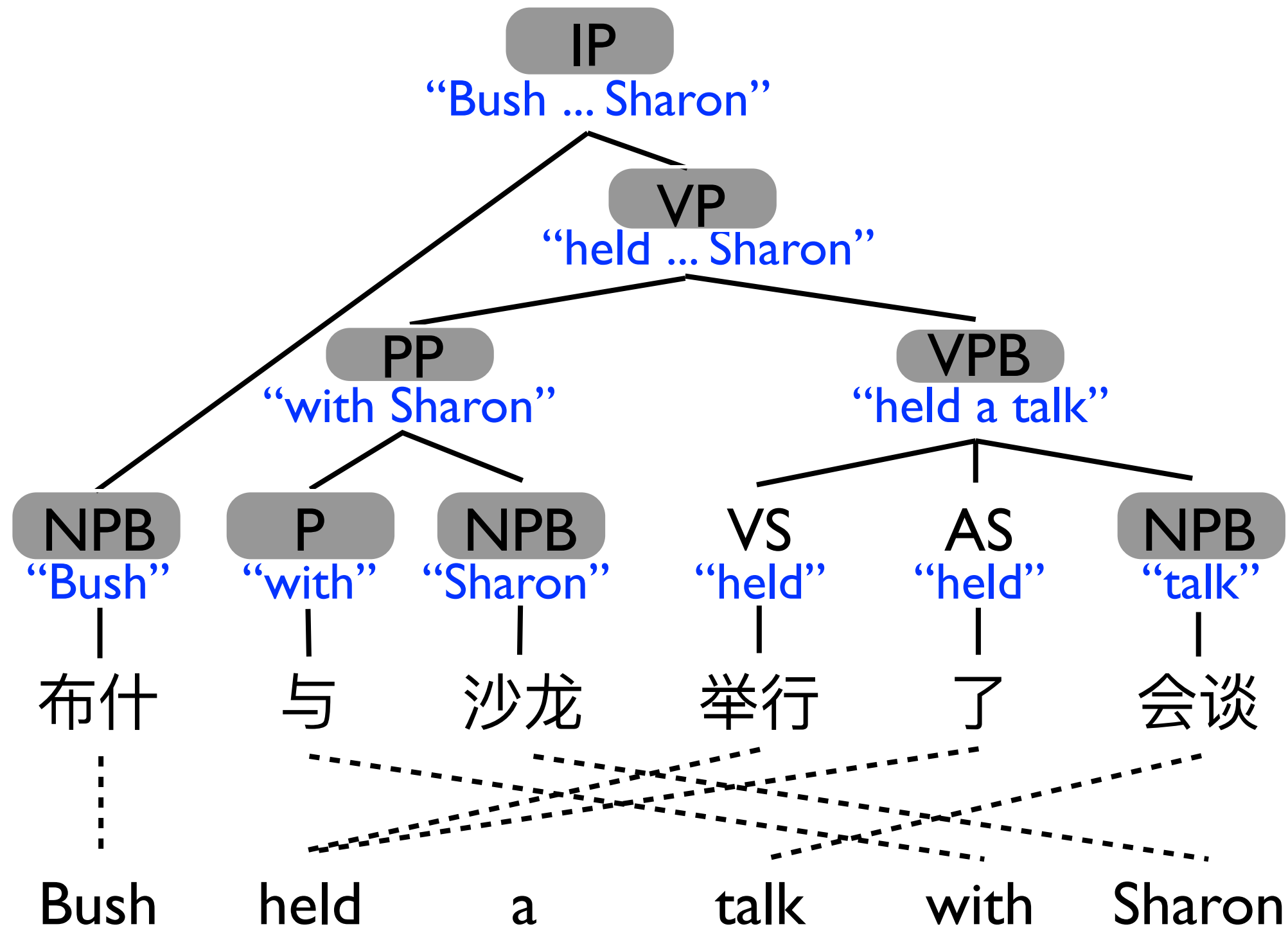
(Galley et al., 2004)

# Rule Extraction



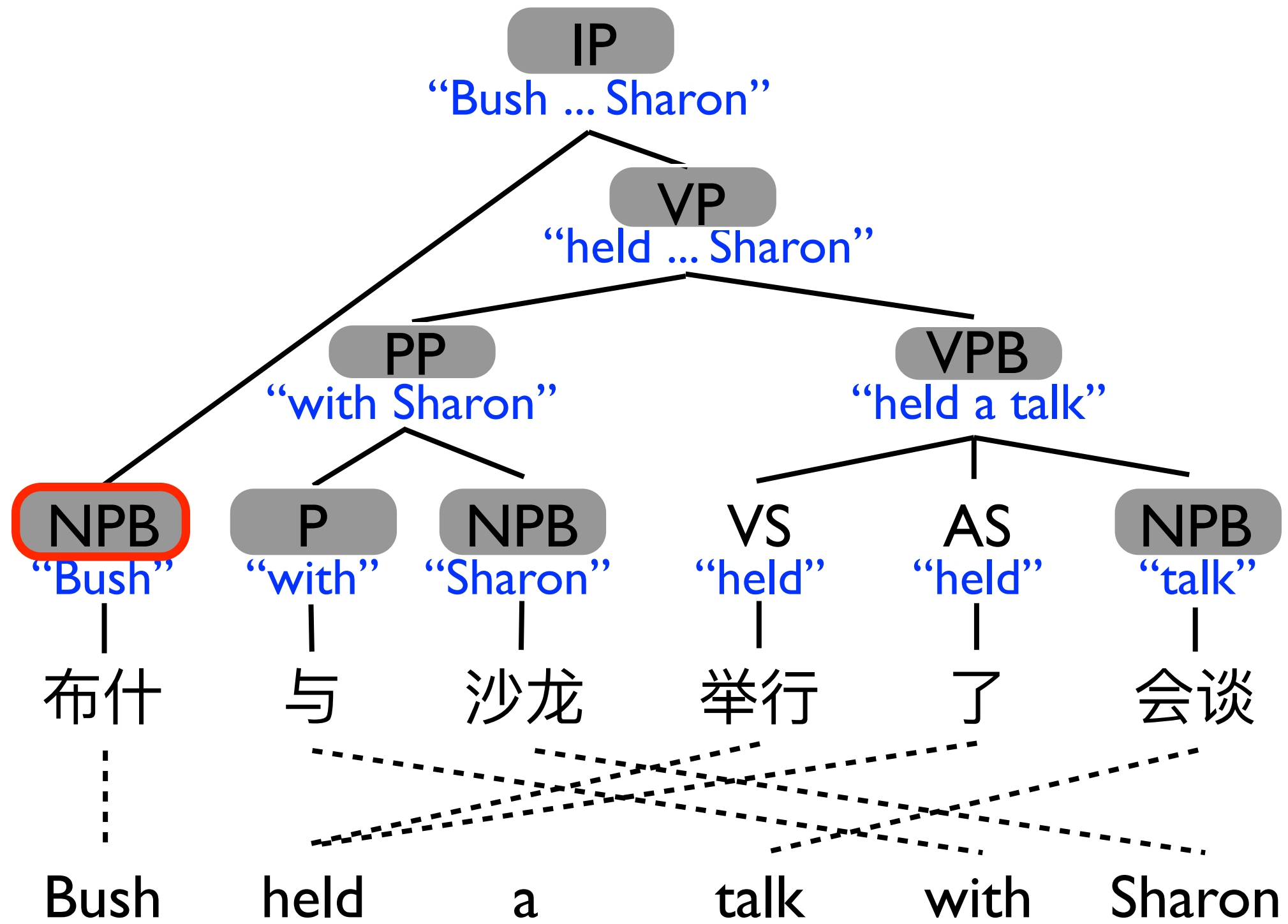
(Galley et al., 2004)

# Rule Extraction



(Galley et al., 2004)

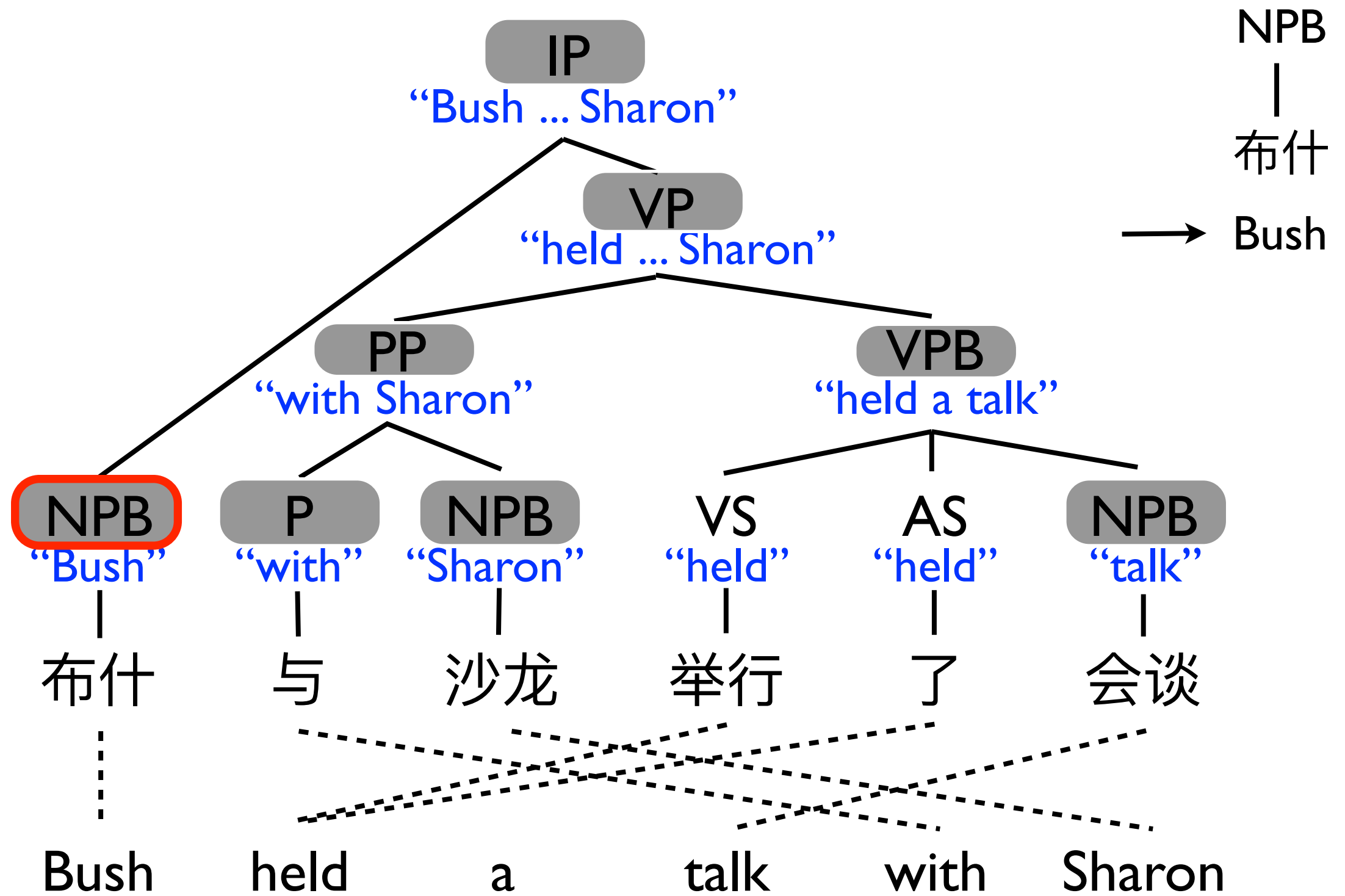
# Rule Extraction



(Galley et al., 2004)

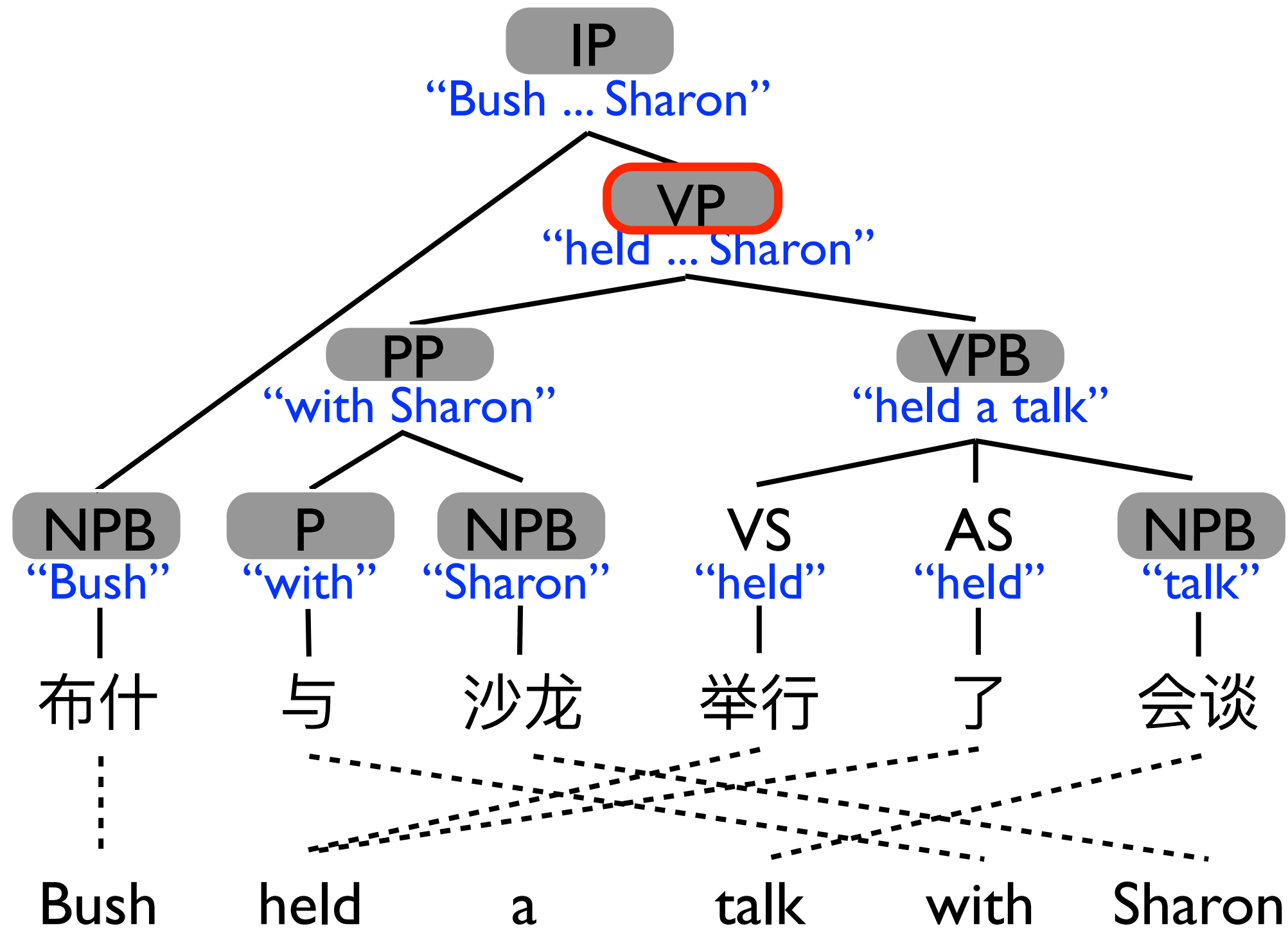


# Rule Extraction



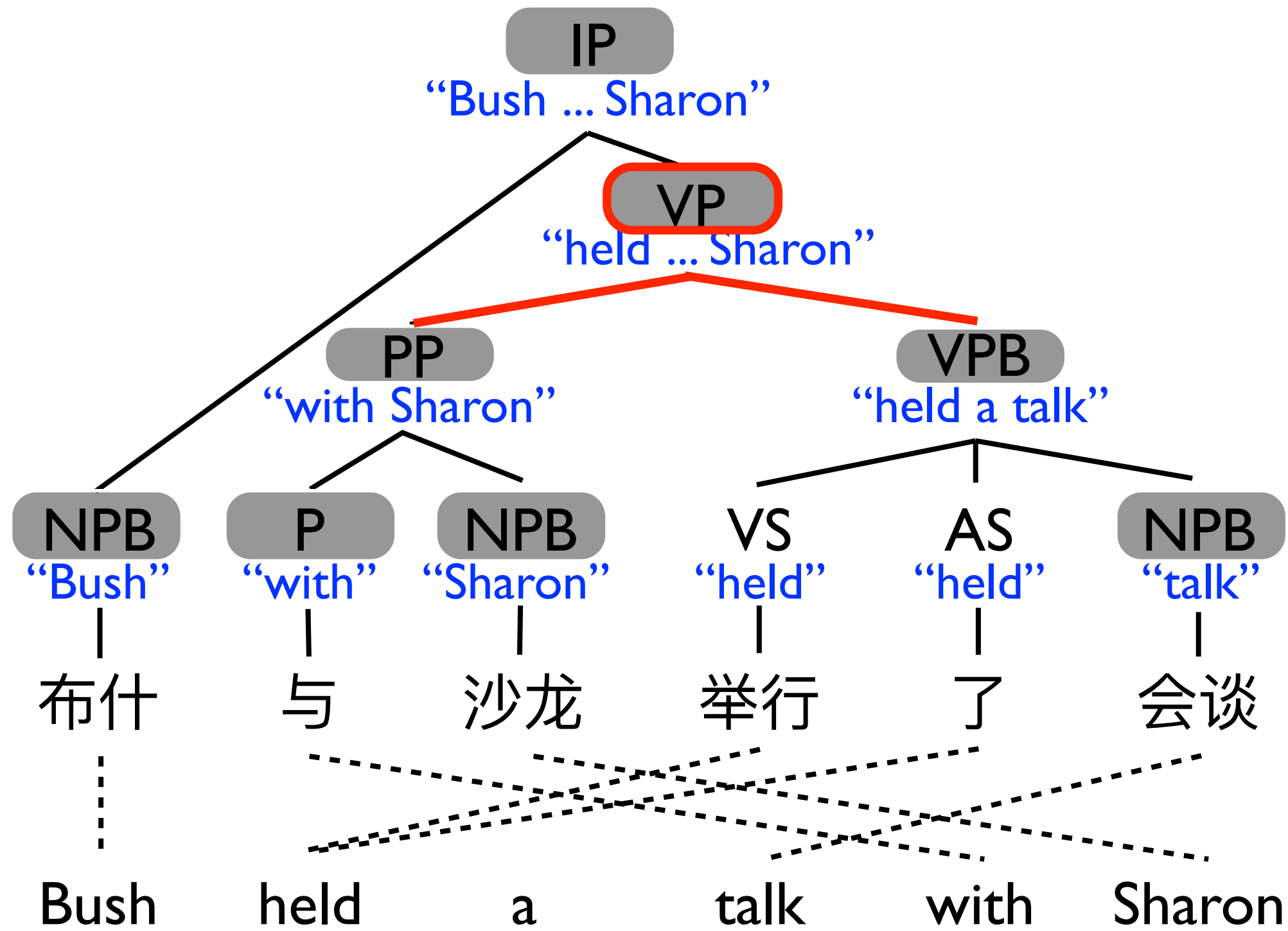
(Galley et al., 2004)

# Rule Extraction



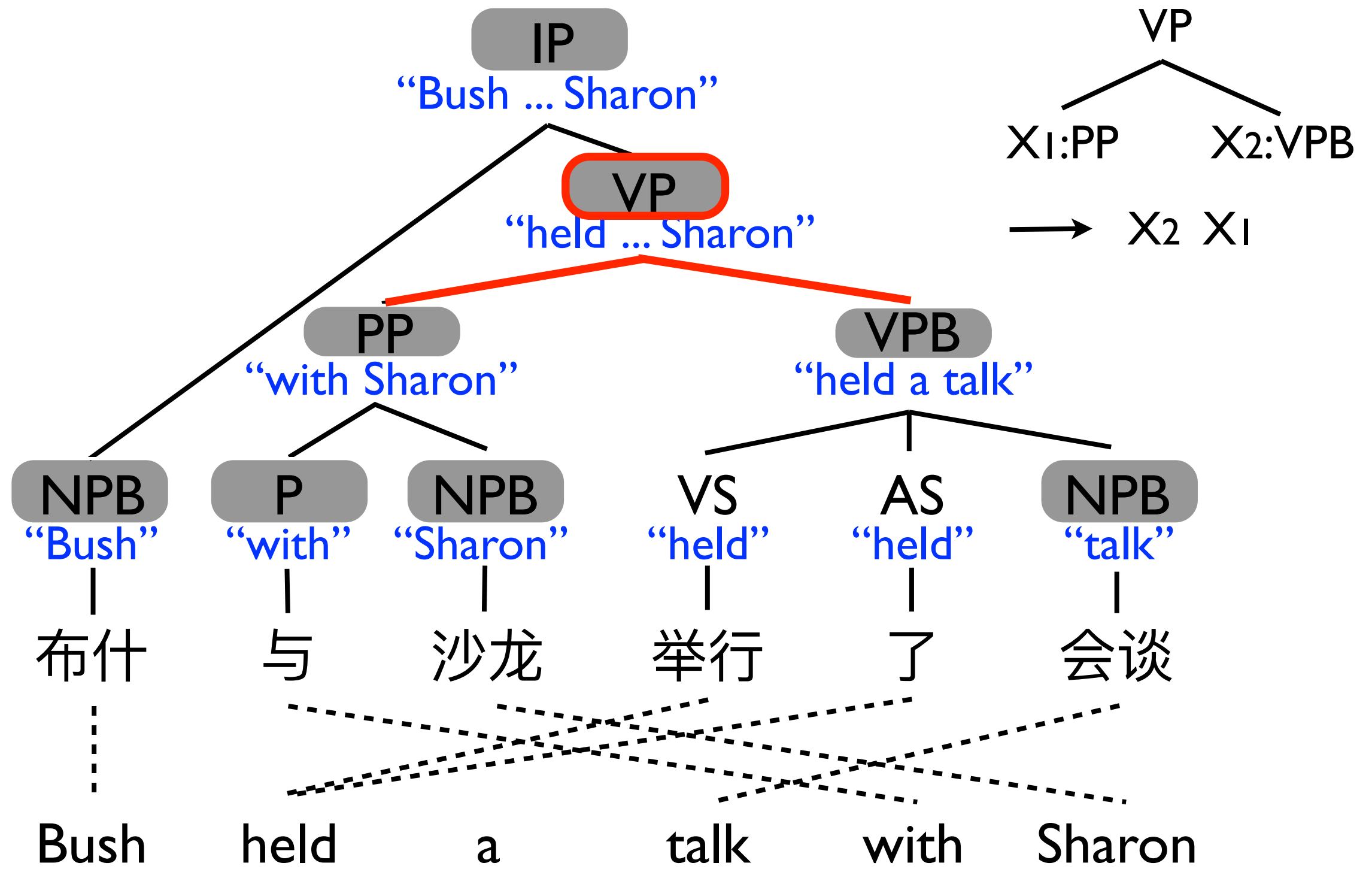
(Galley et al., 2004)

# Rule Extraction



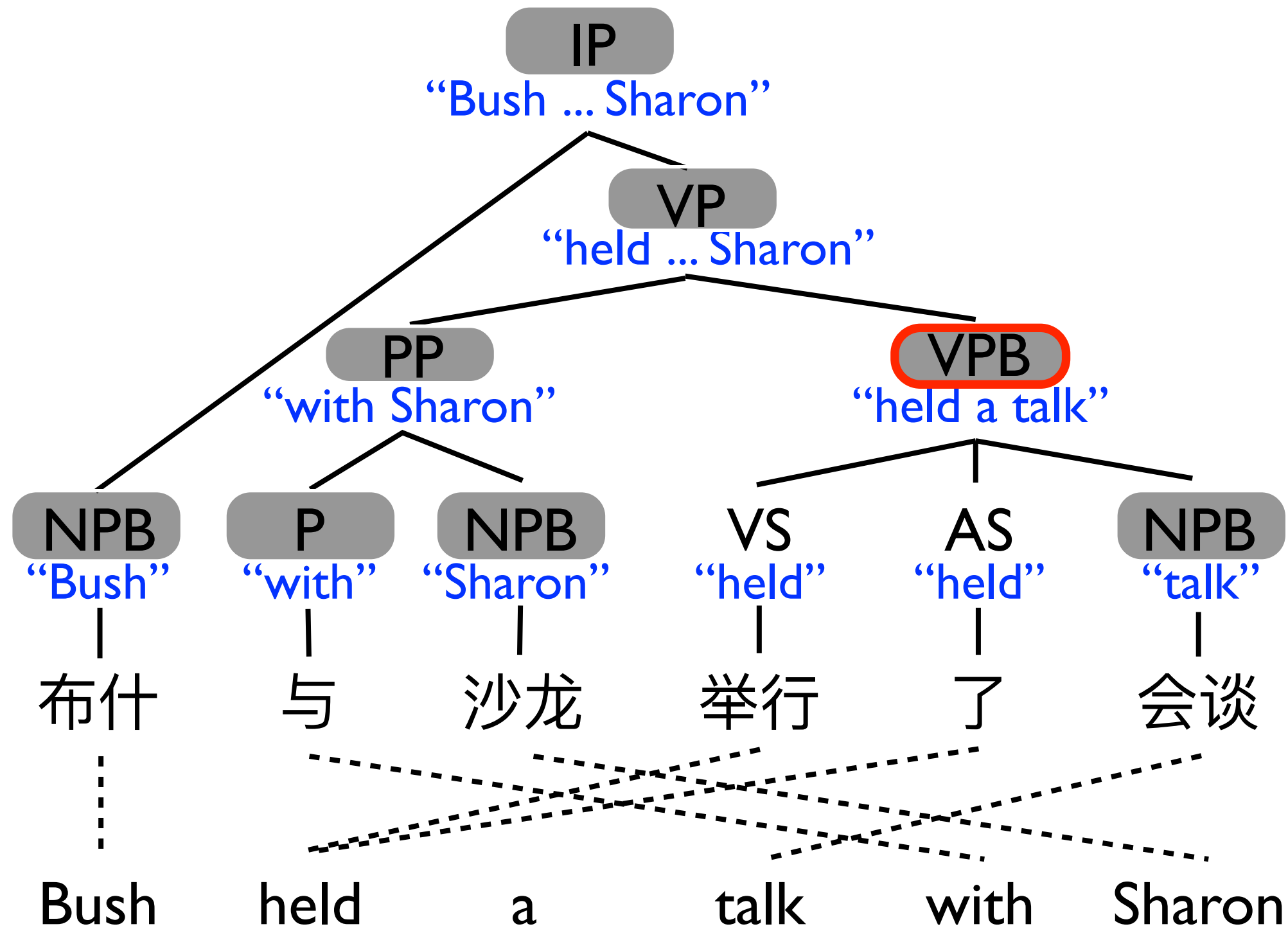
(Galley et al., 2004)

# Rule Extraction



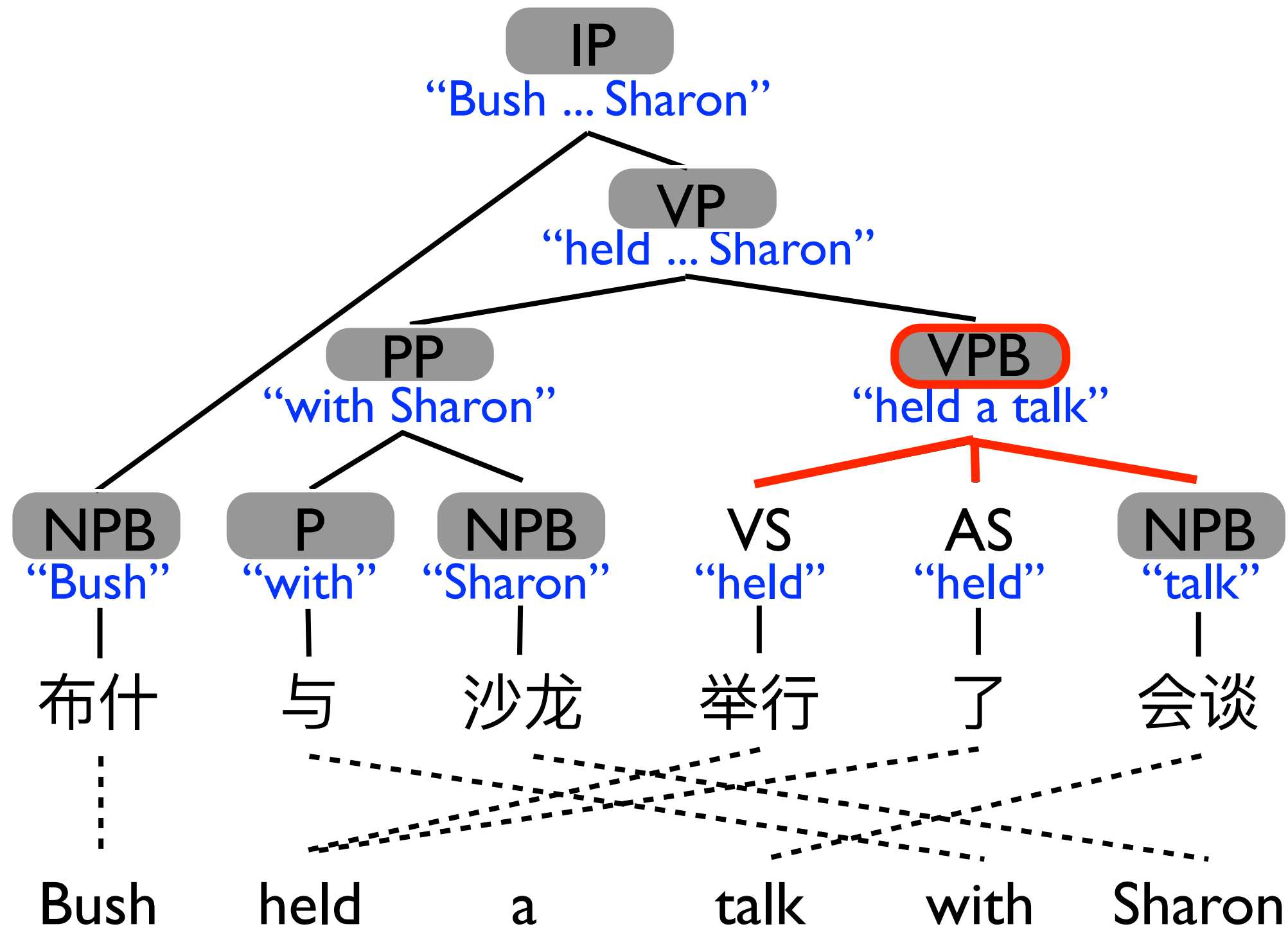
(Galley et al., 2004)

# Rule Extraction



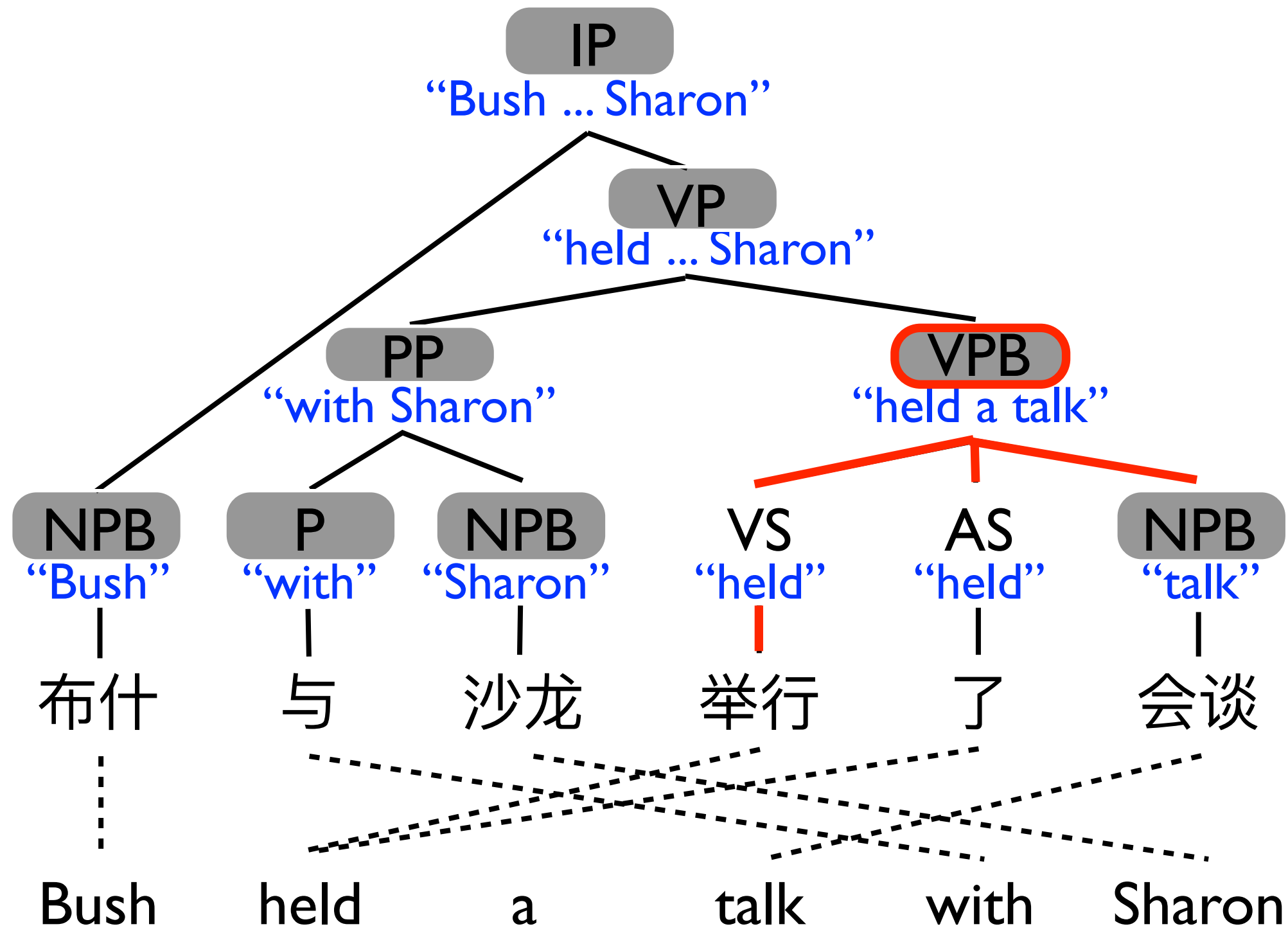
(Galley et al., 2004)

# Rule Extraction



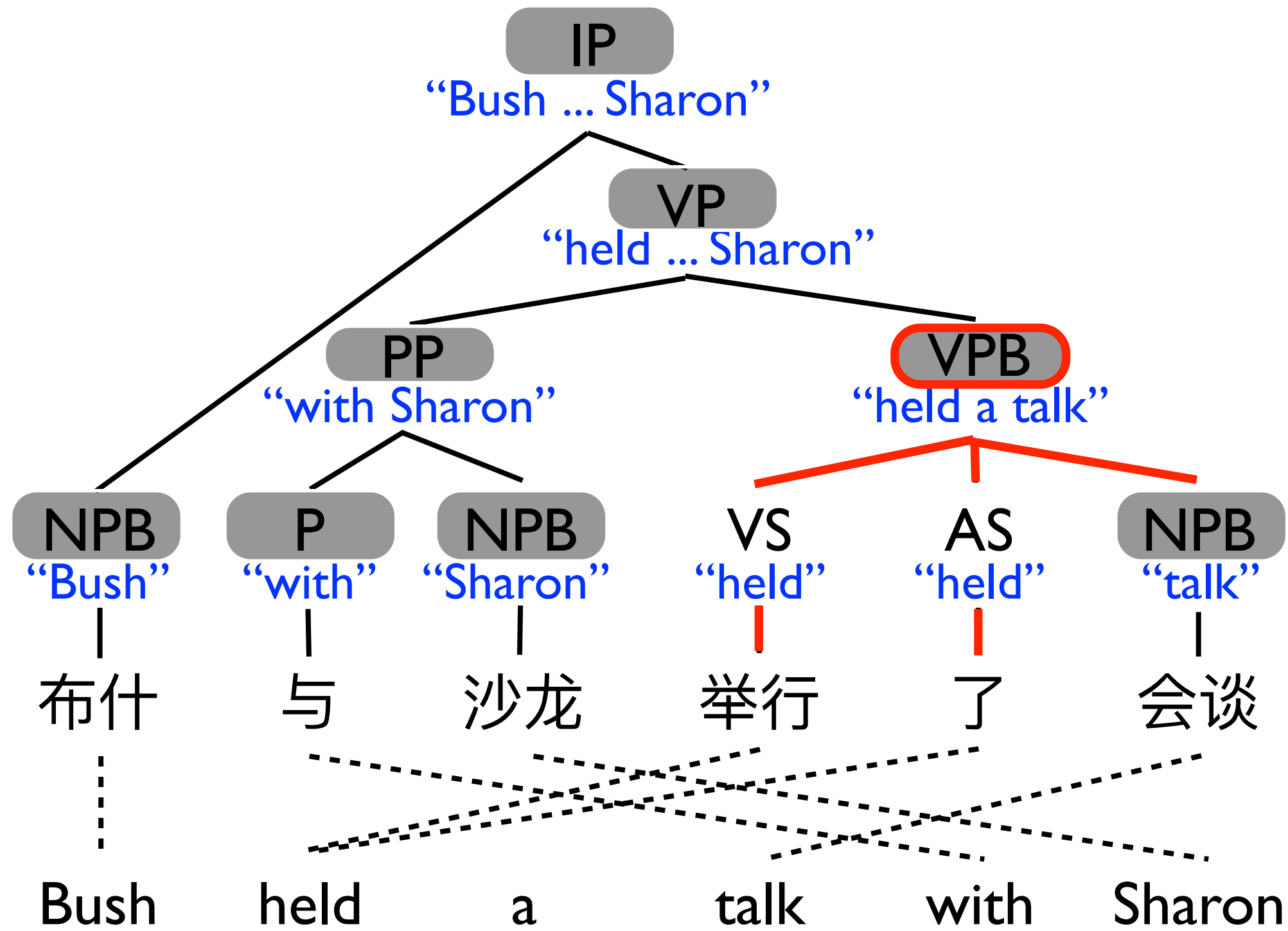
(Galley et al., 2004)

# Rule Extraction



(Galley et al., 2004)

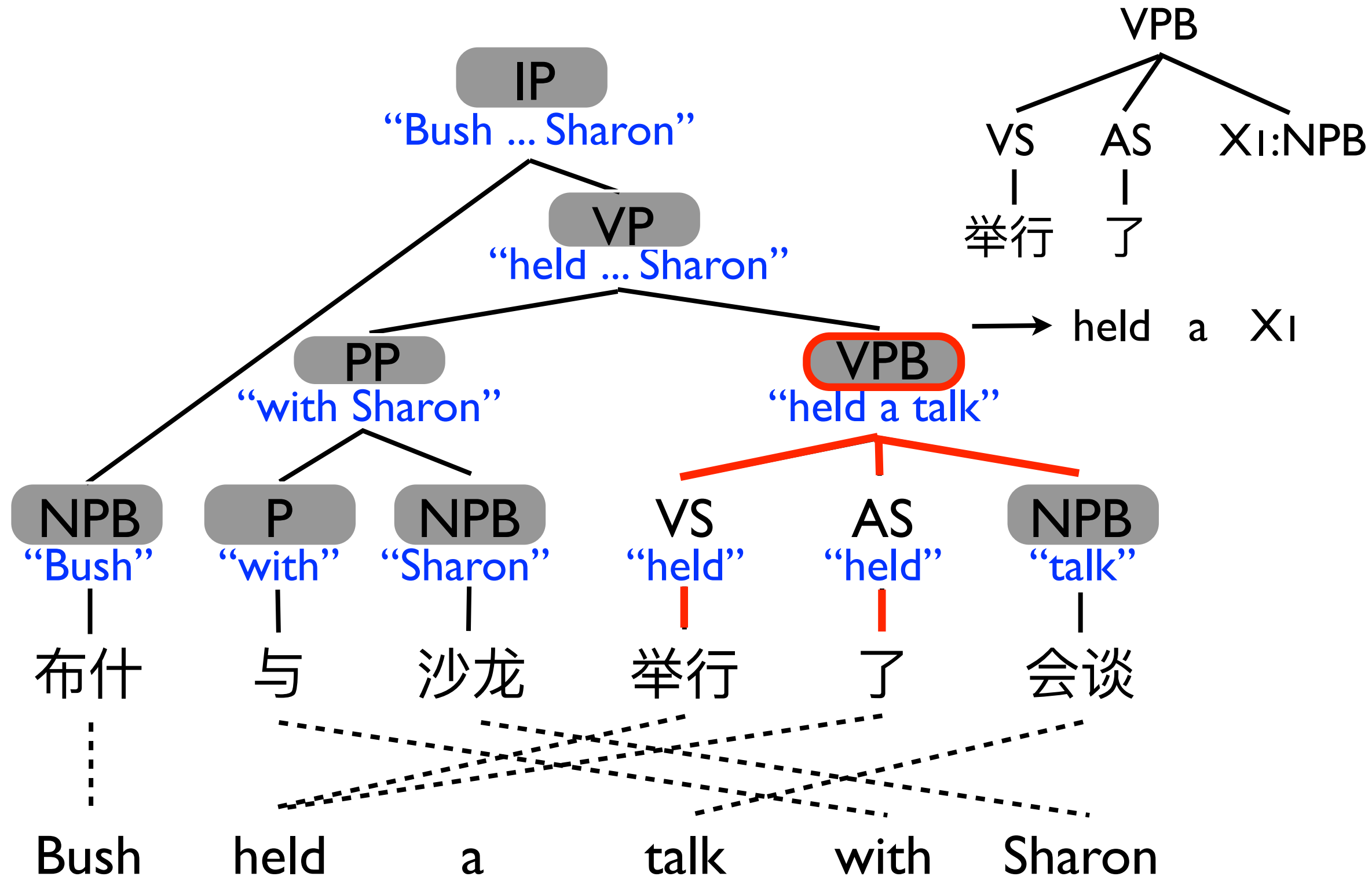
# Rule Extraction



(Galley et al., 2004)

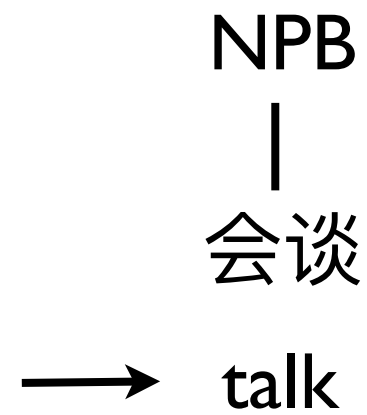
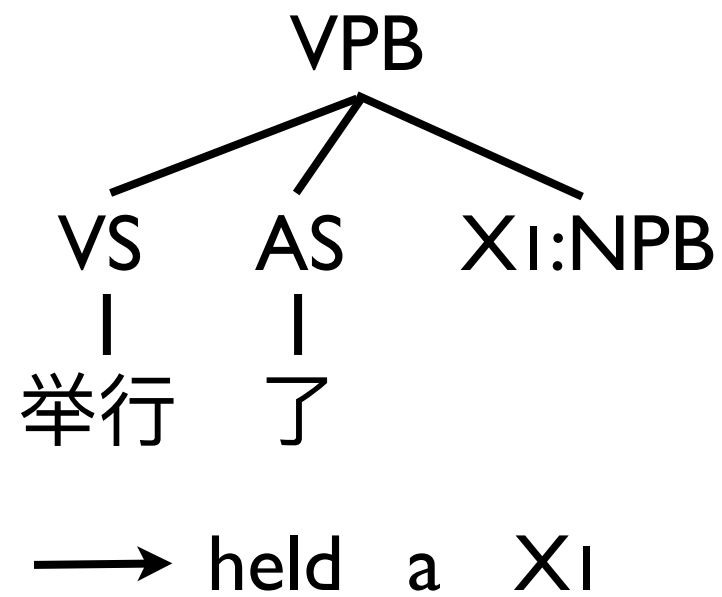


# Rule Extraction



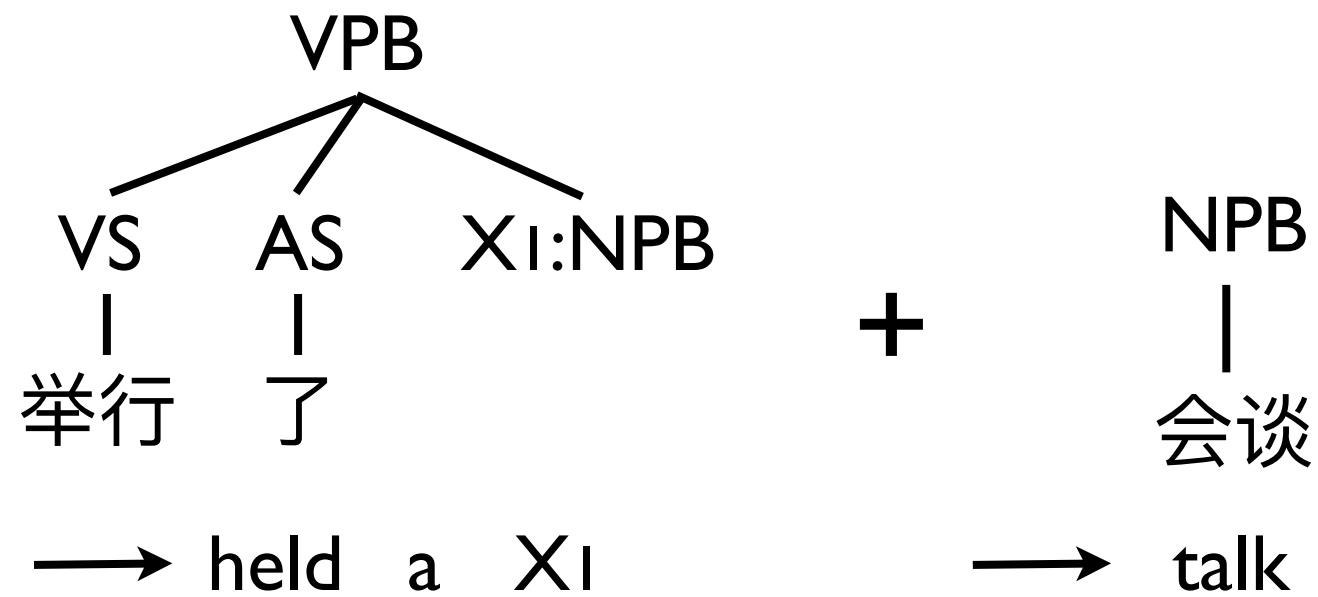
(Galley et al., 2004)

# Rule Composition



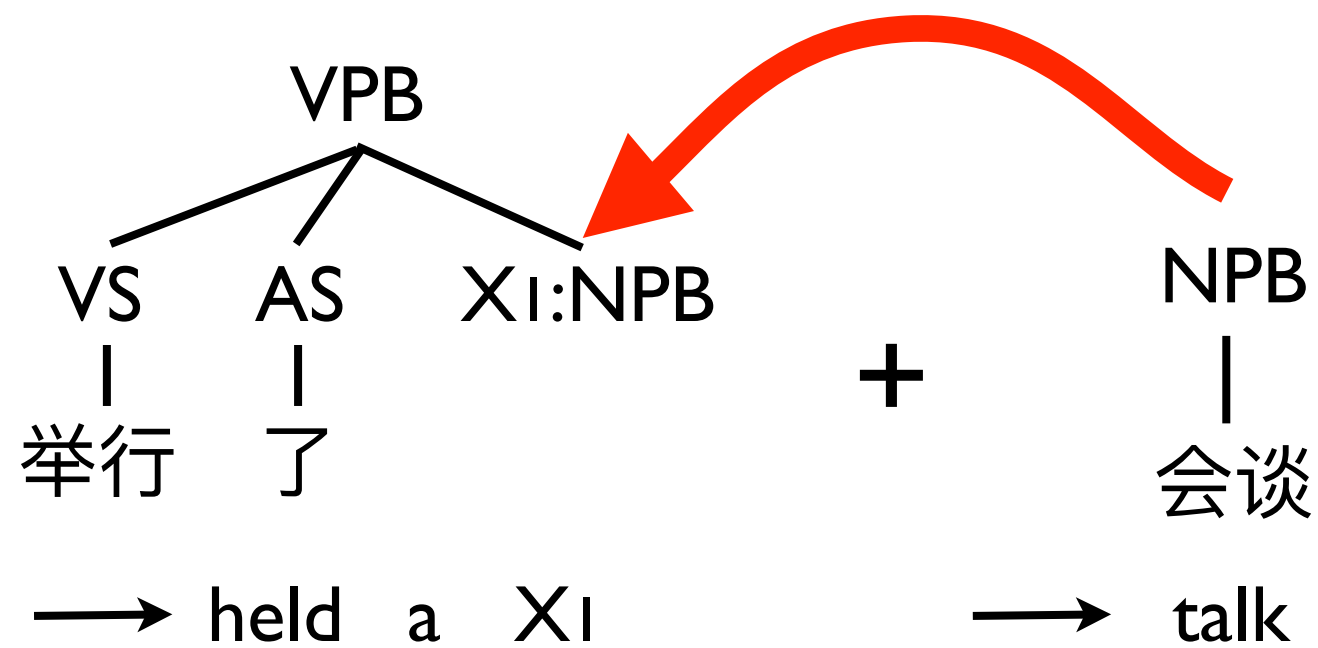
(Galley et al., 2006)

# Rule Composition



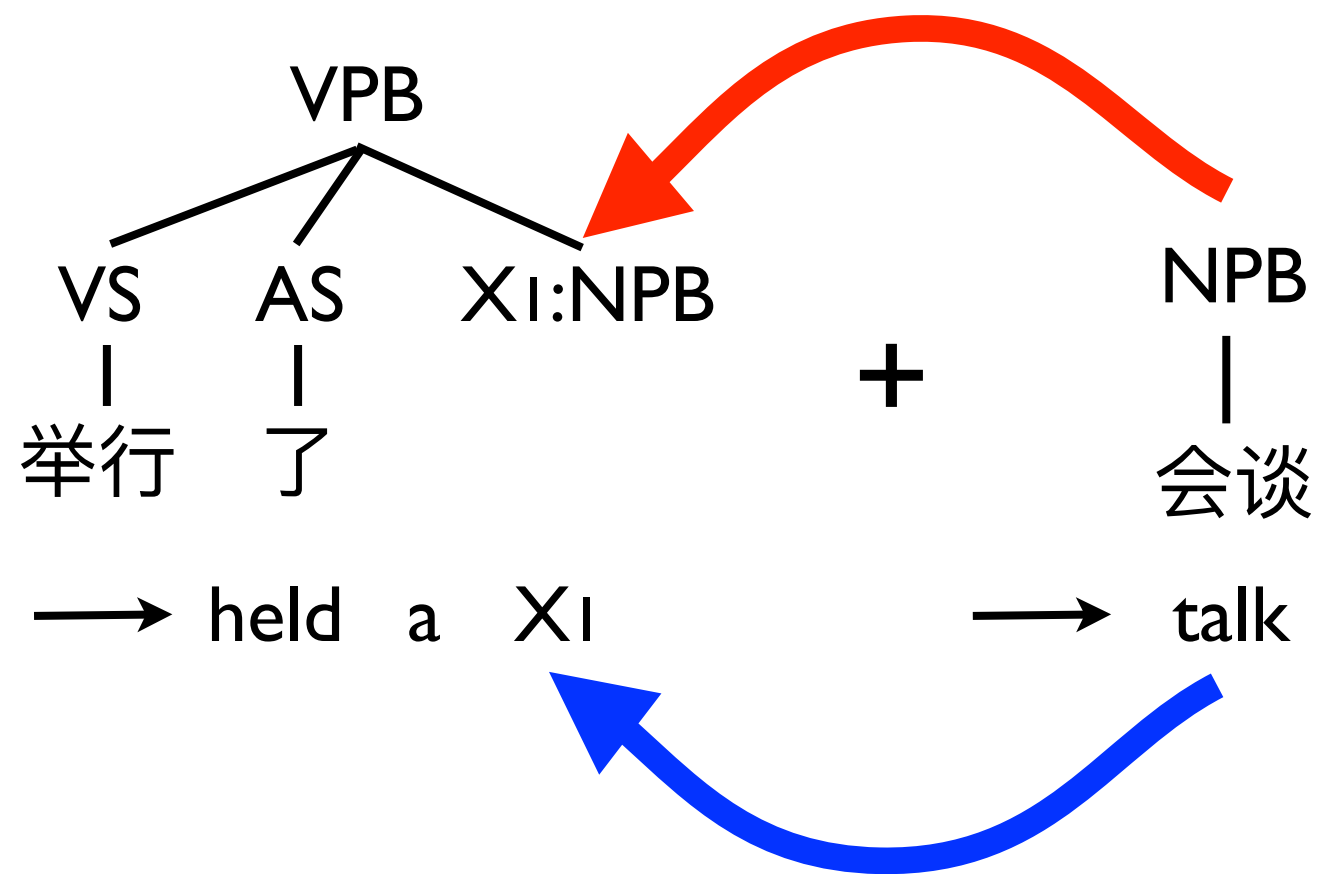
(Galley et al., 2006)

# Rule Composition



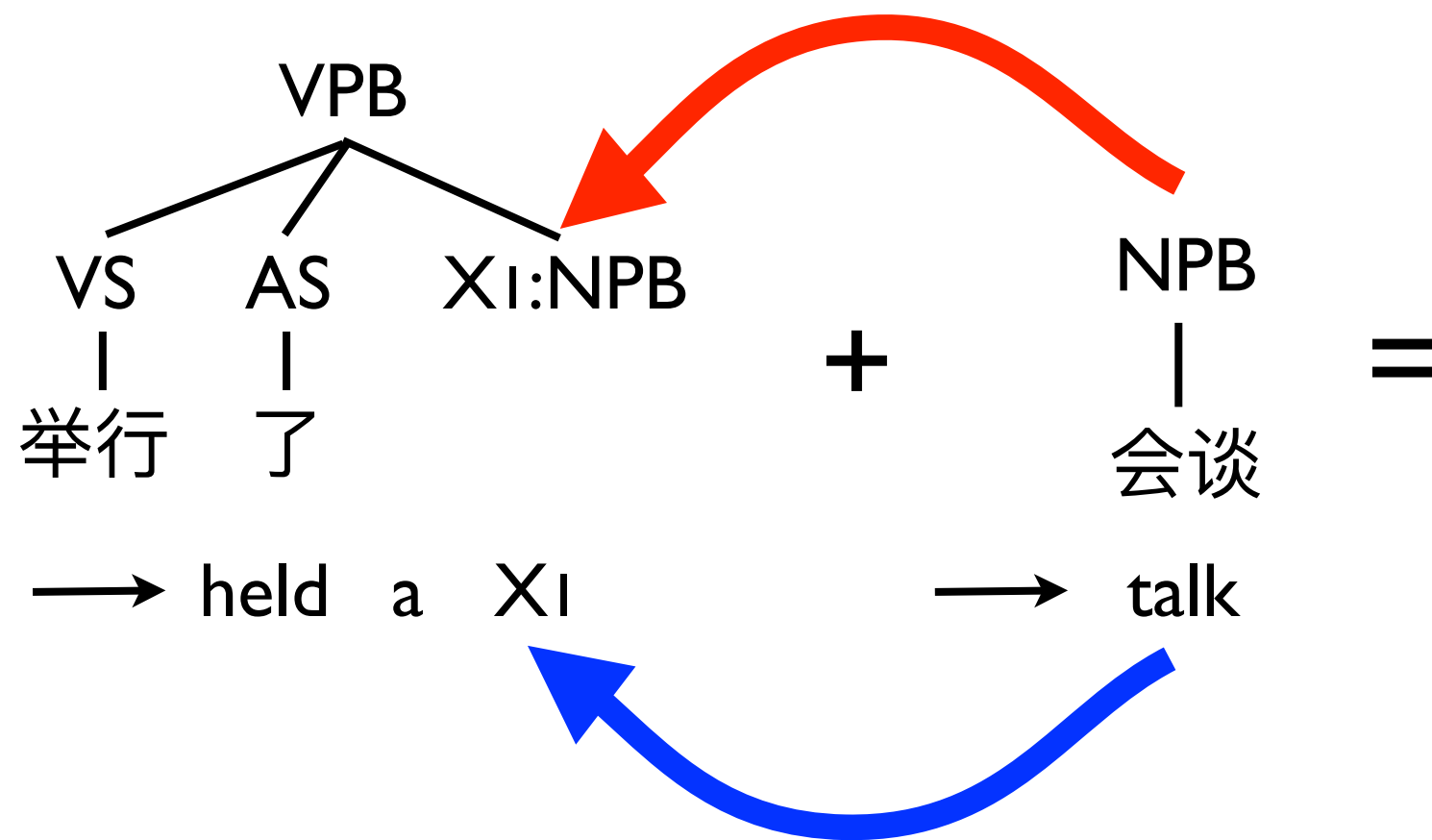
(Galley et al., 2006)

# Rule Composition



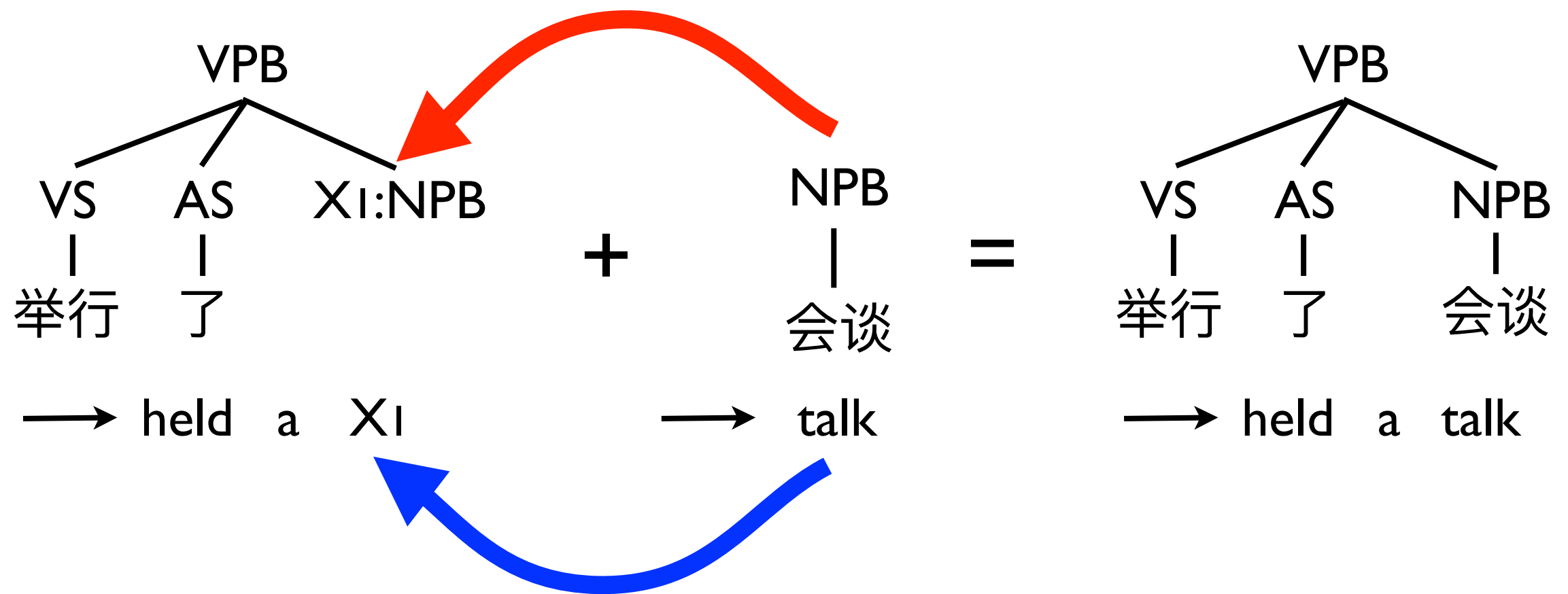
(Galley et al., 2006)

# Rule Composition



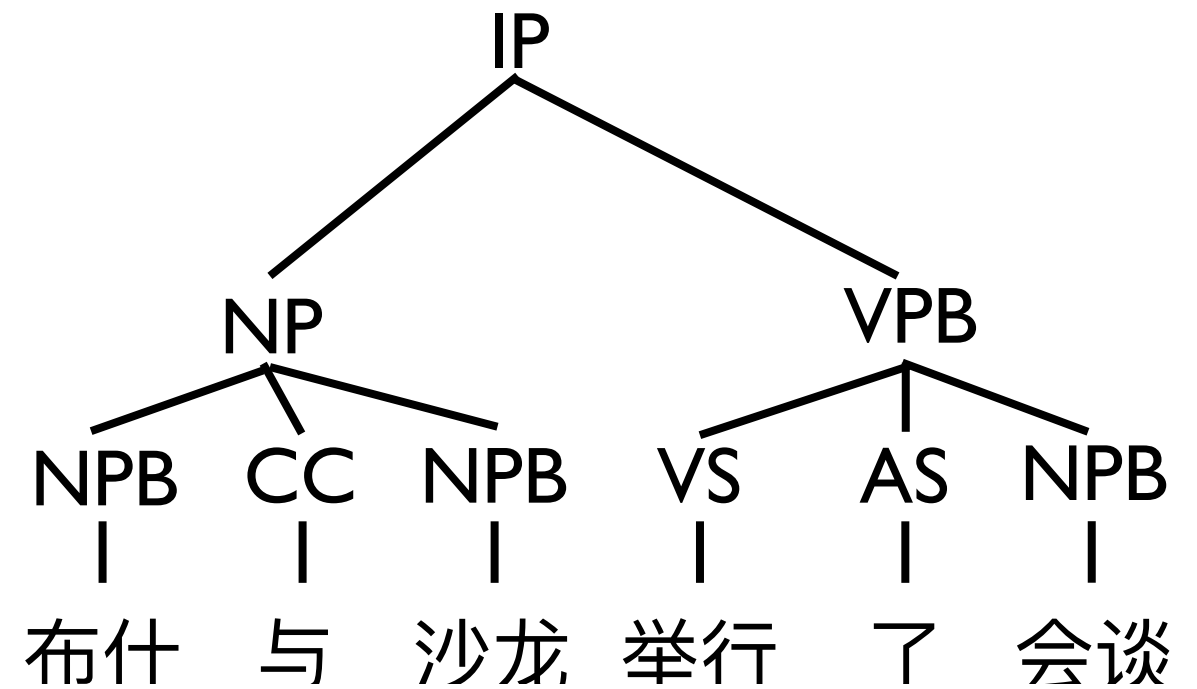
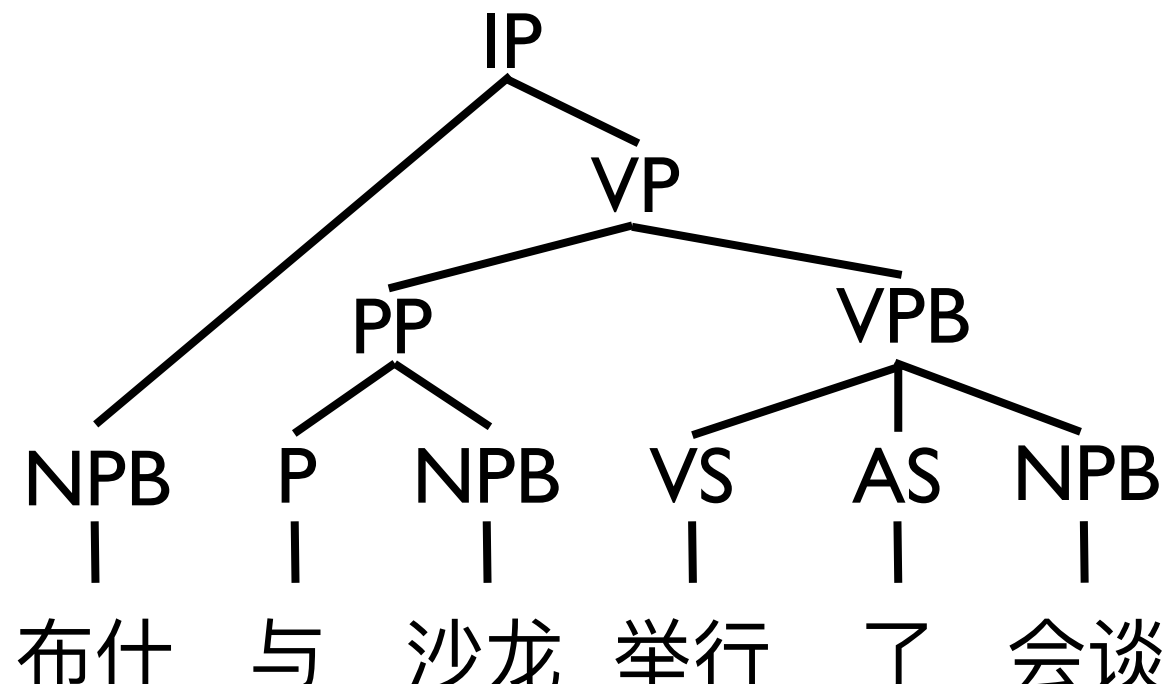
(Galley et al., 2006)

# Rule Composition



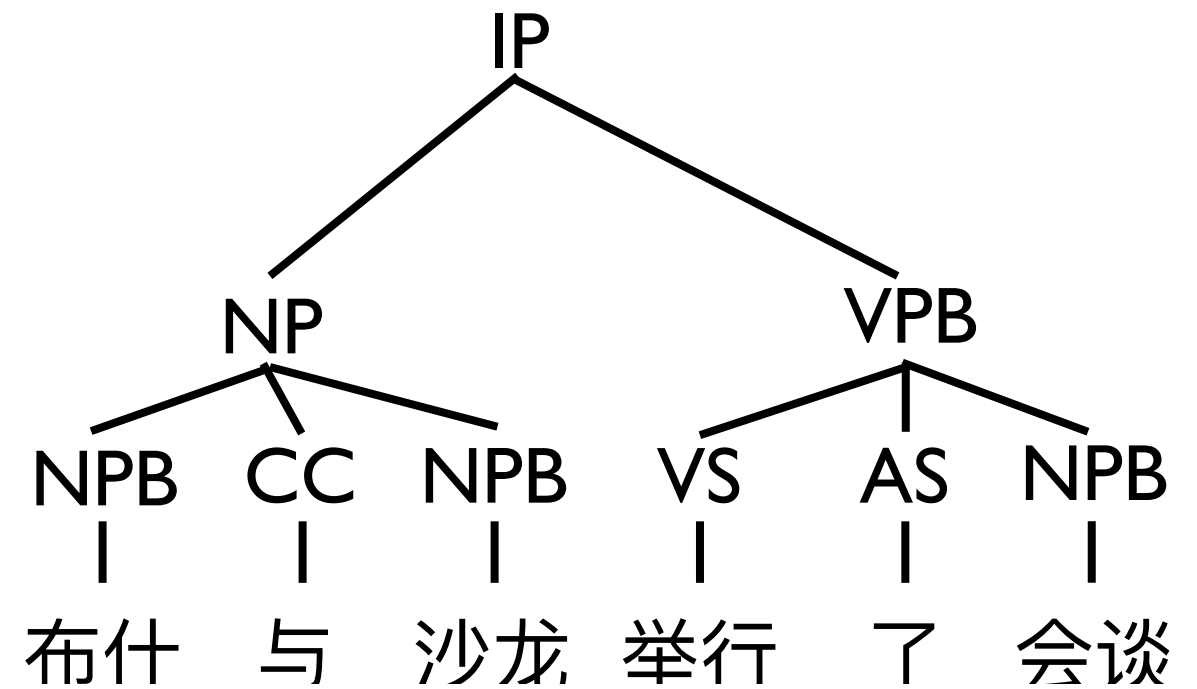
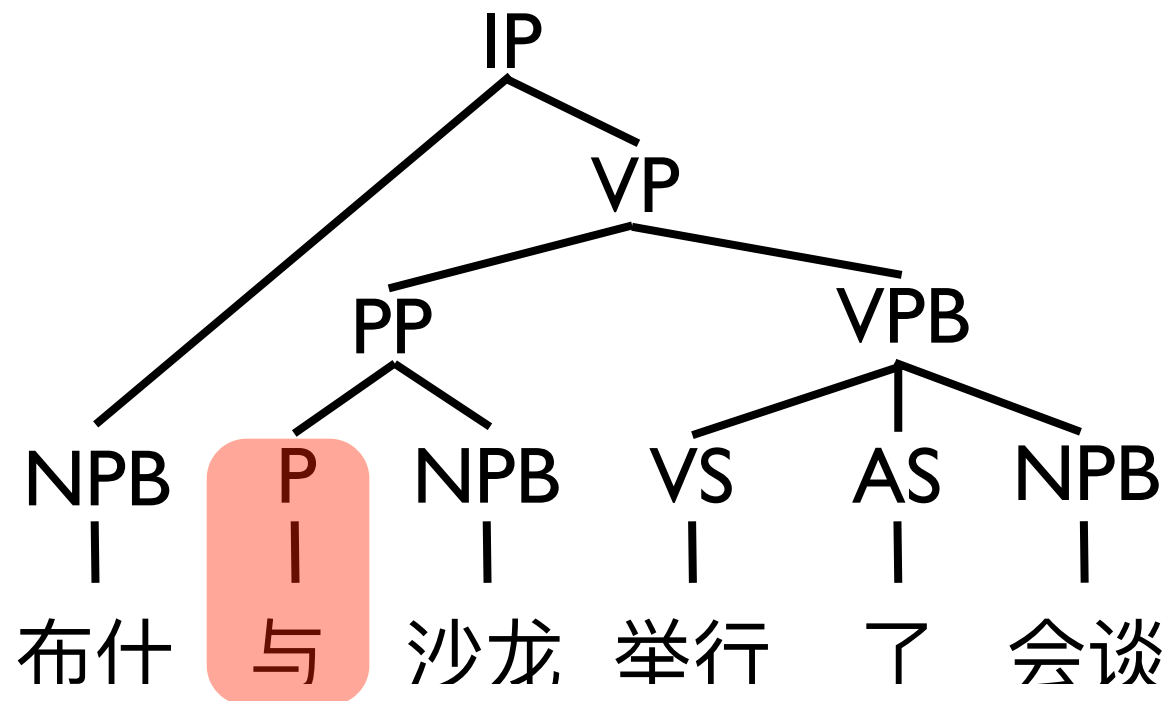
(Galley et al., 2006)

# Parsing Ambiguity

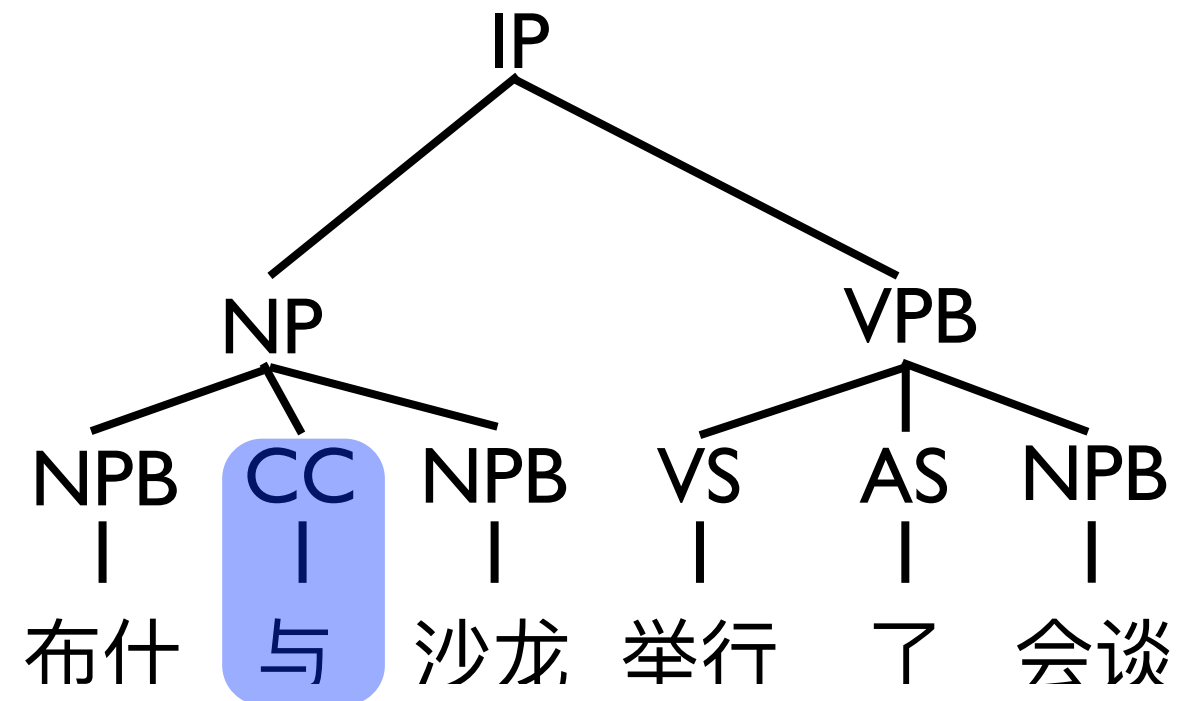
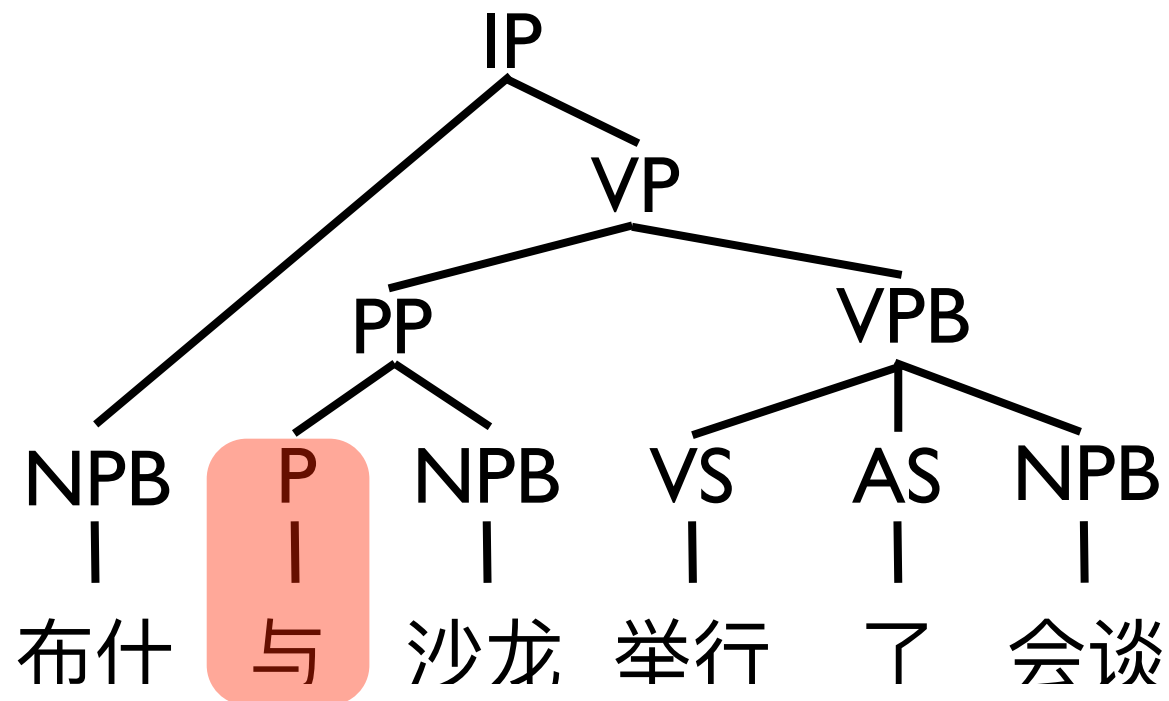




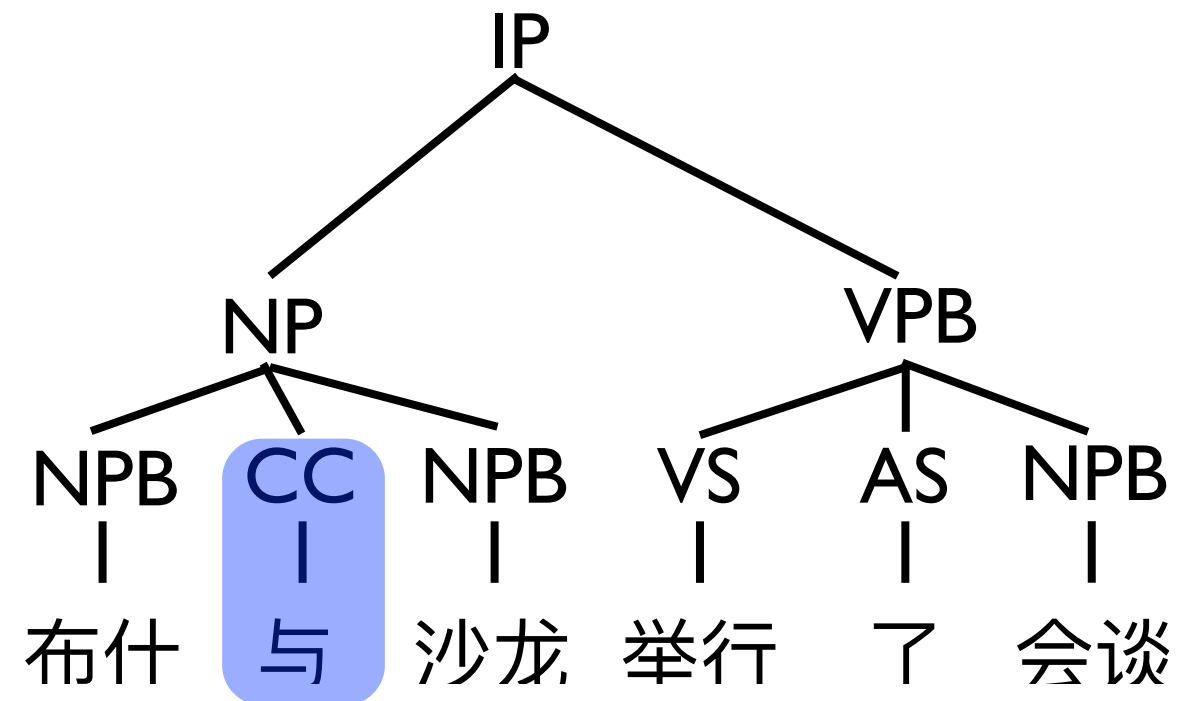
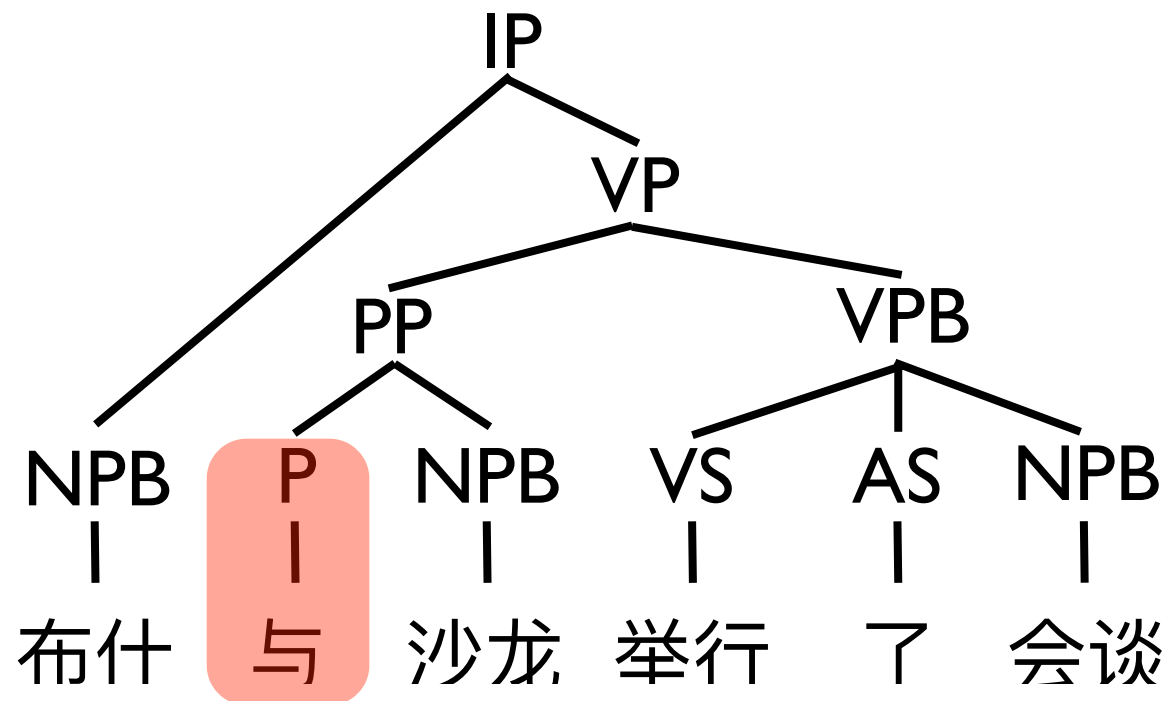
# Parsing Ambiguity



# Parsing Ambiguity

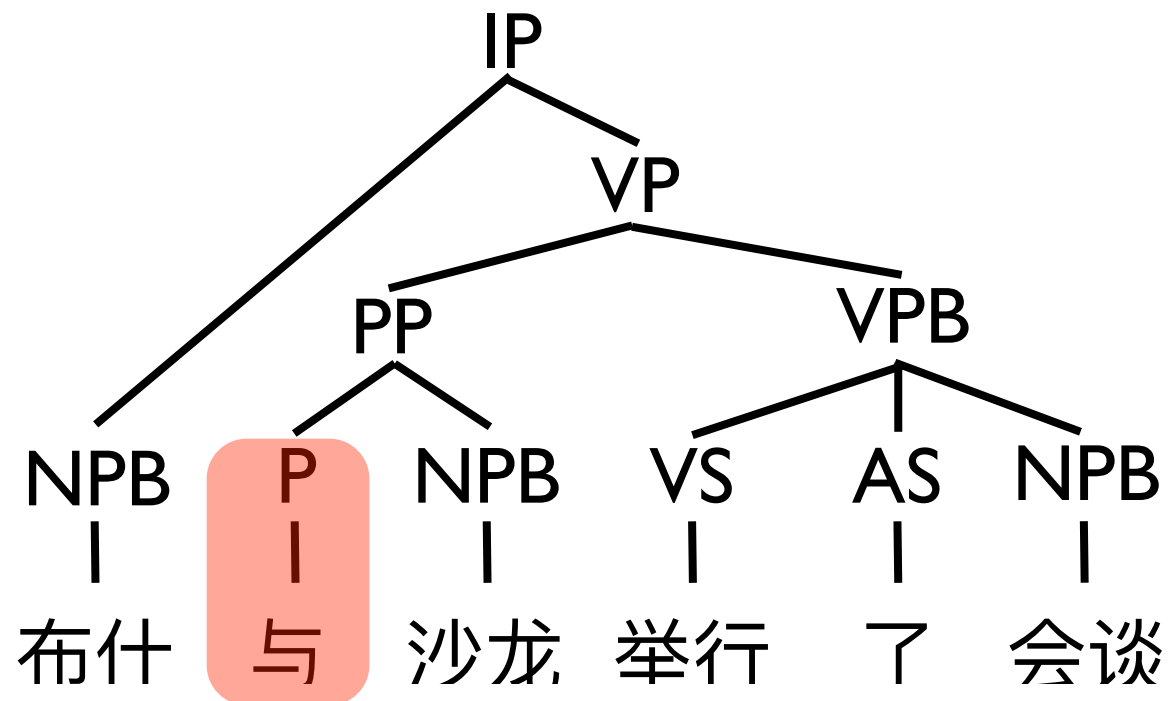


# Parsing Ambiguity

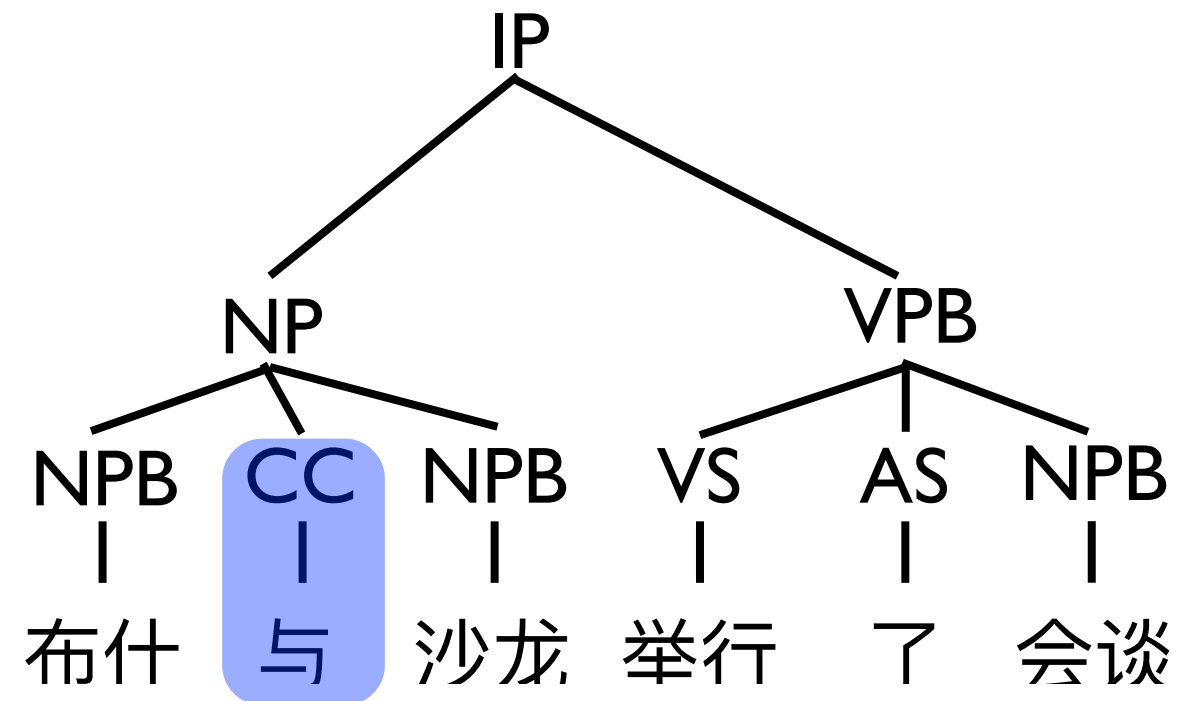


Bush held a talk **with** Sharon

# Parsing Ambiguity

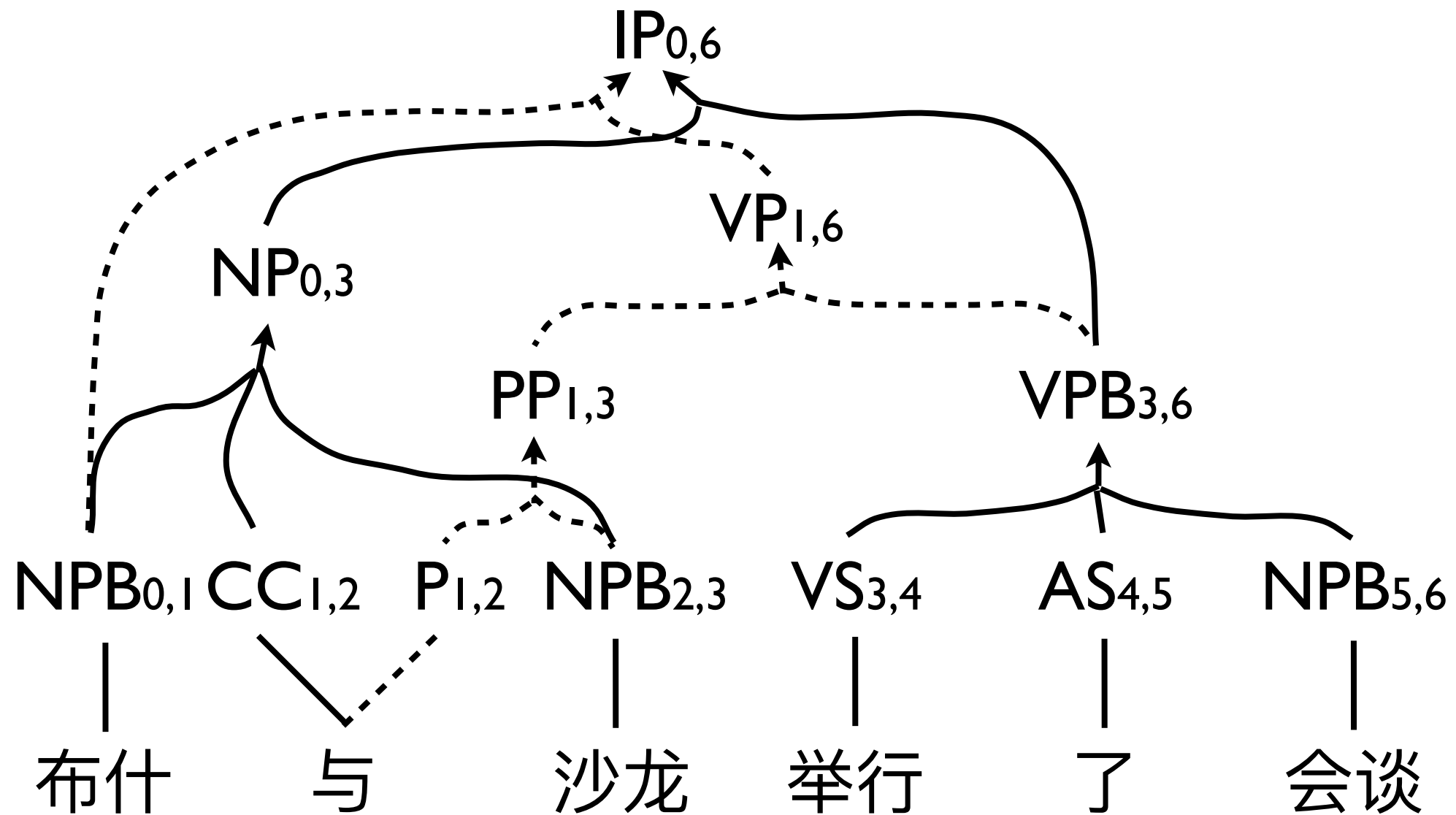


Bush held a talk **with** Sharon



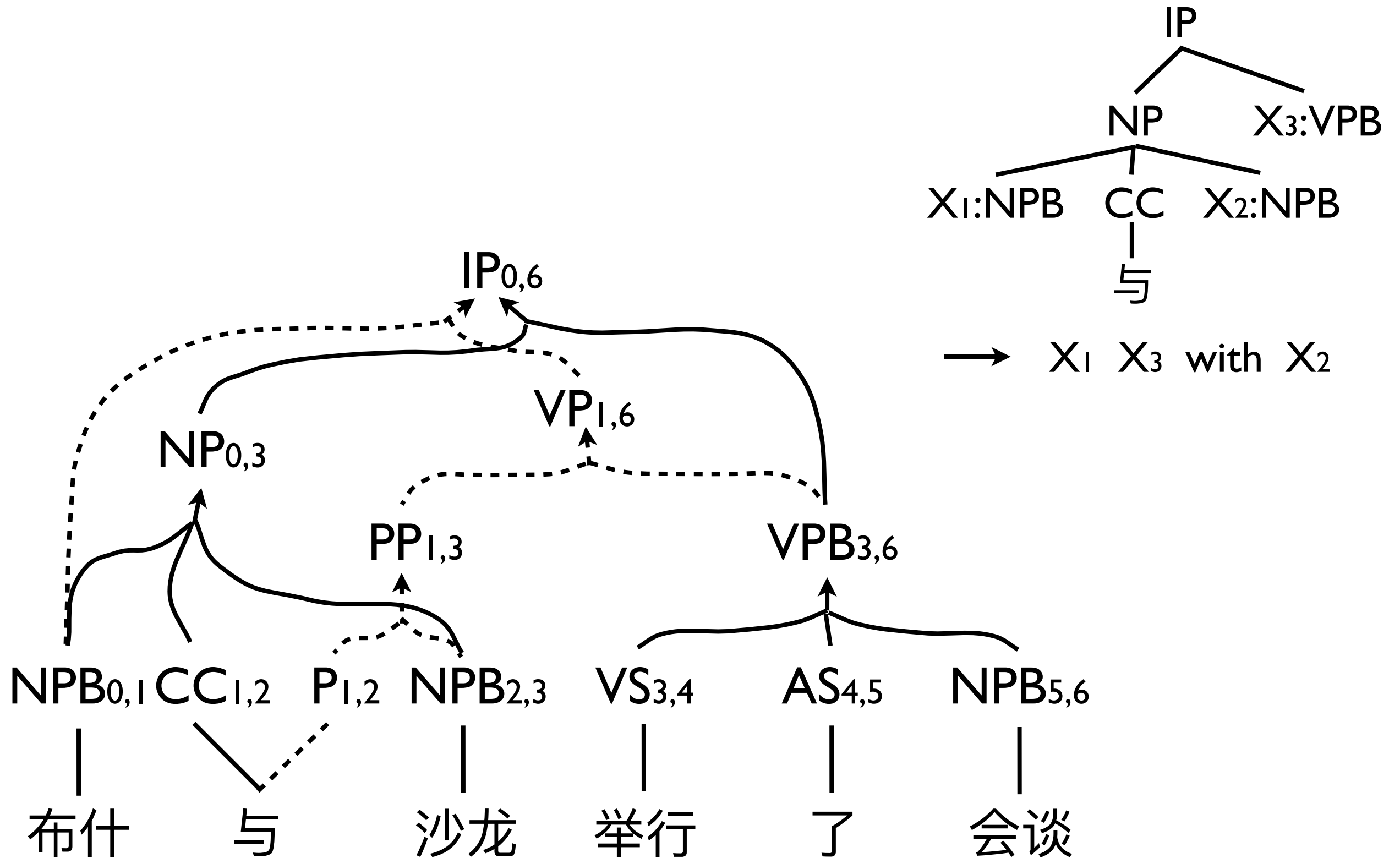
Bush **and** Sharon held a talk

# Packed Forest

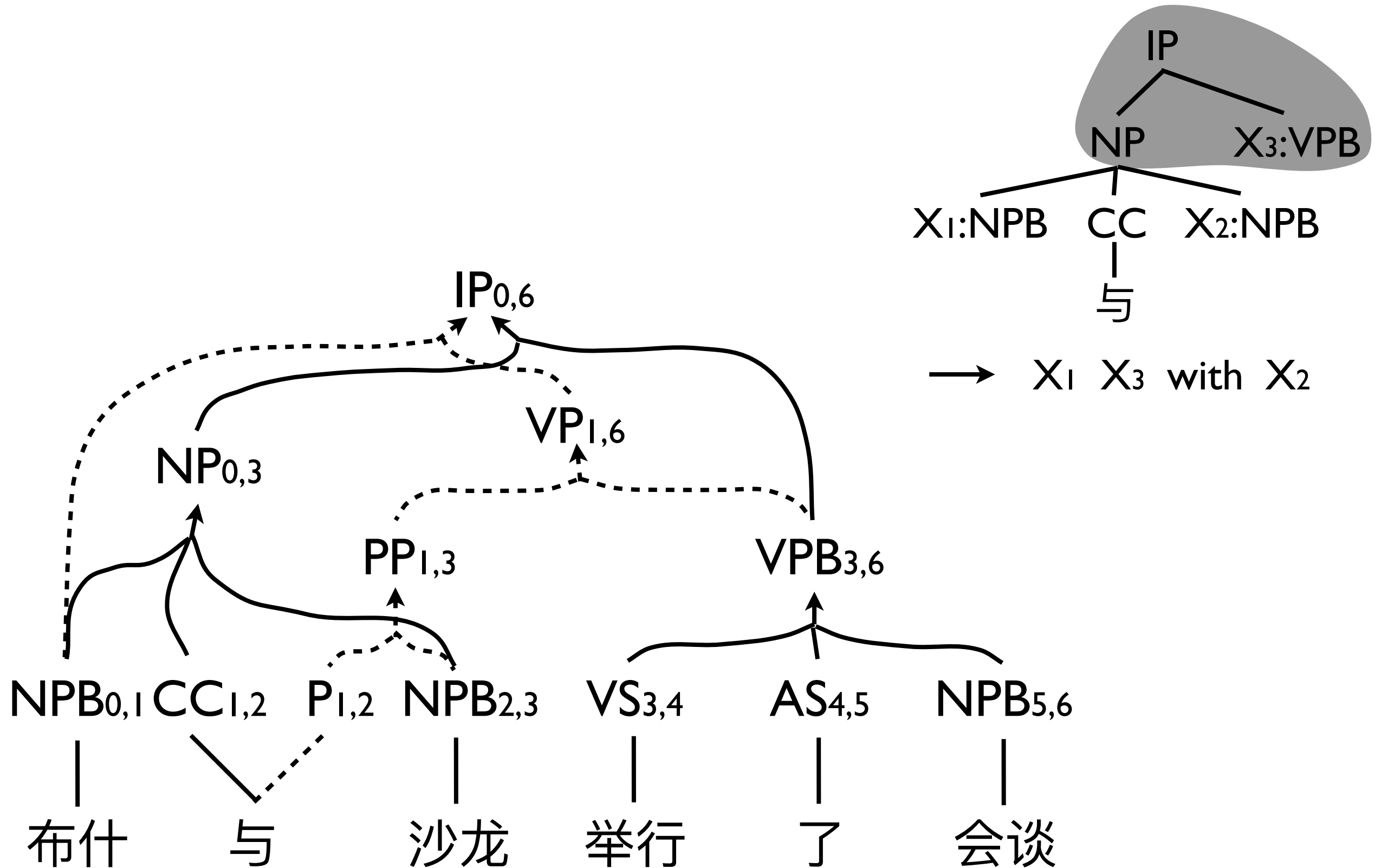


(Billot and Lang, 1989)

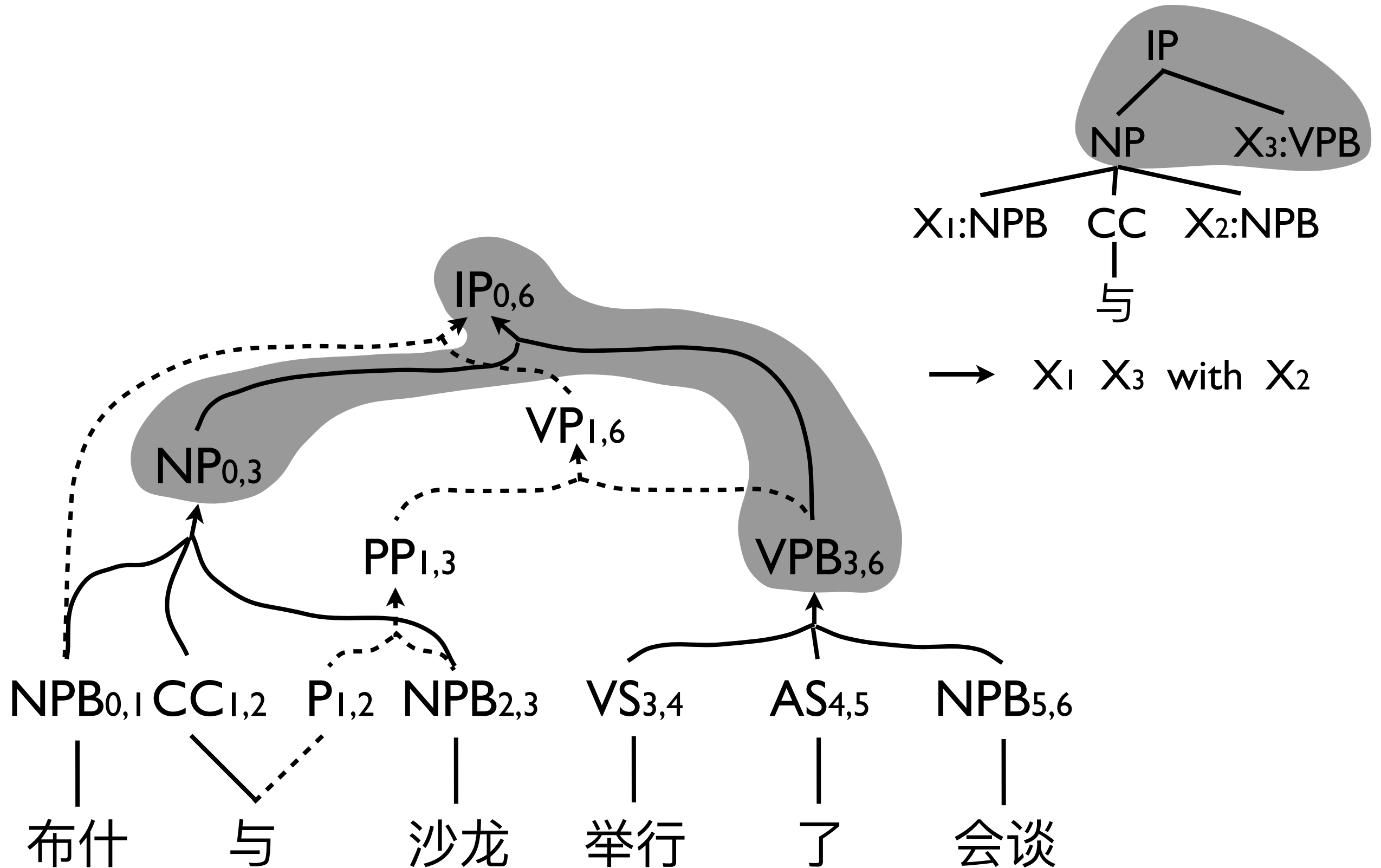
# Forest-based Decoding



# Forest-based Decoding

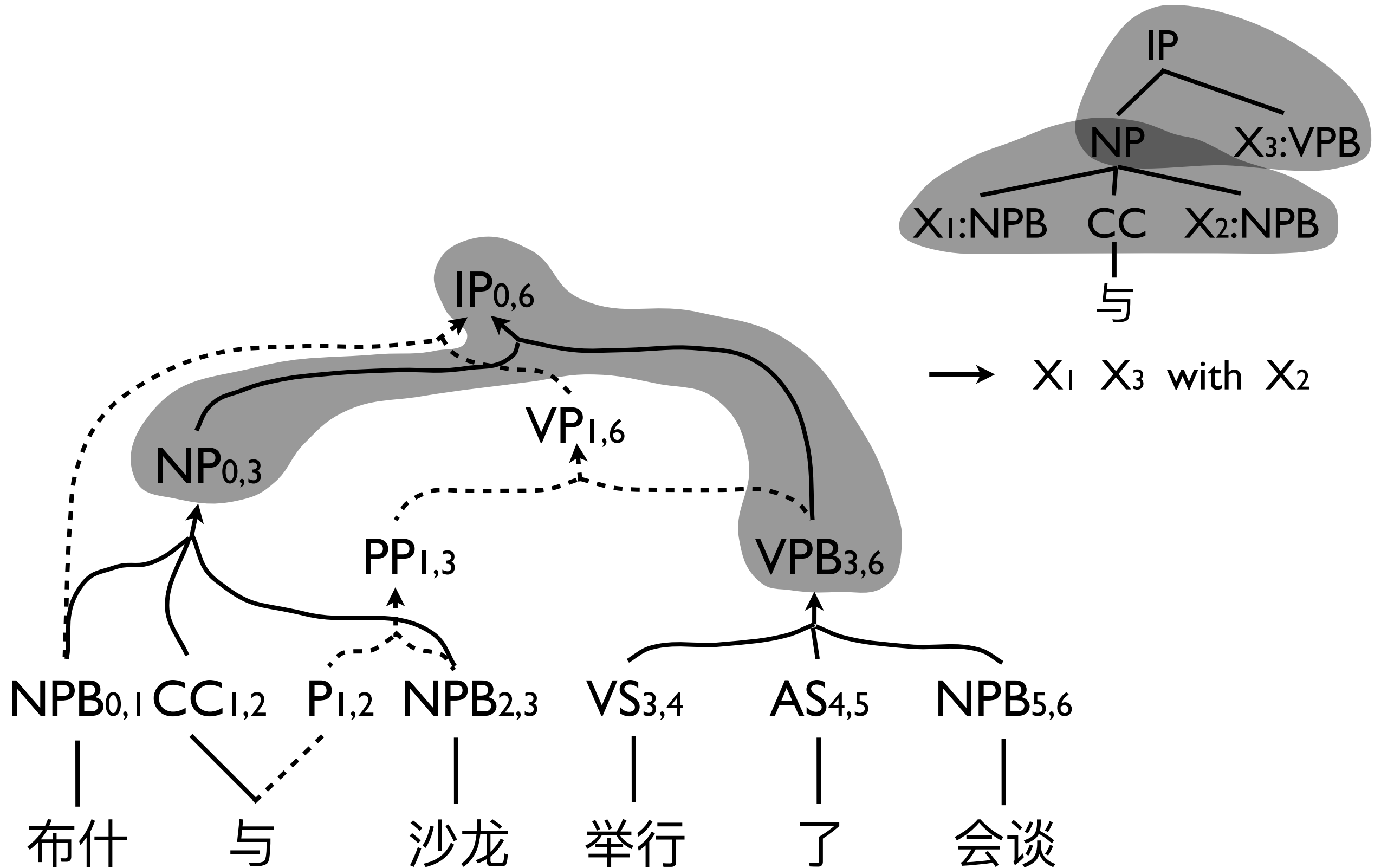


# Forest-based Decoding

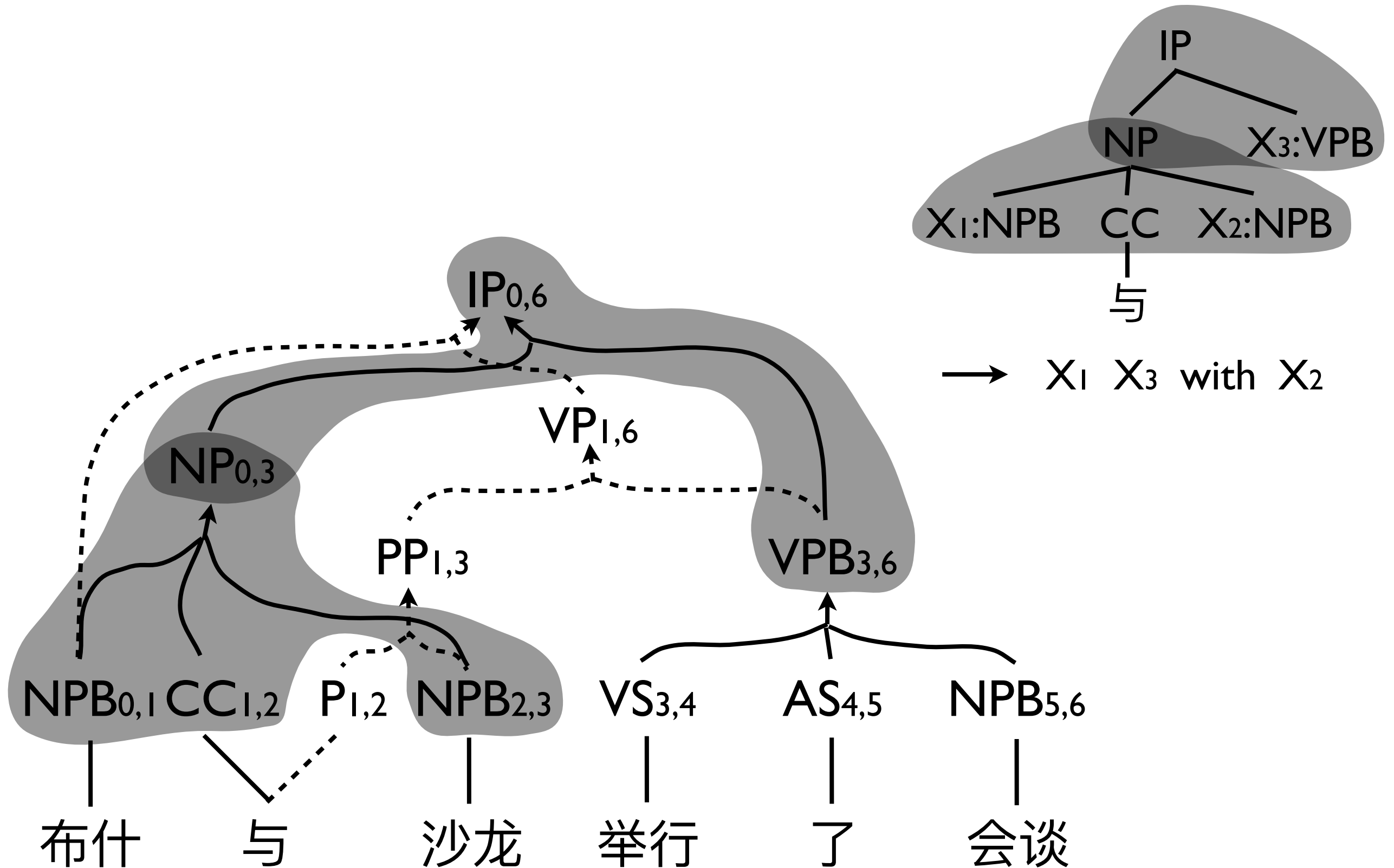




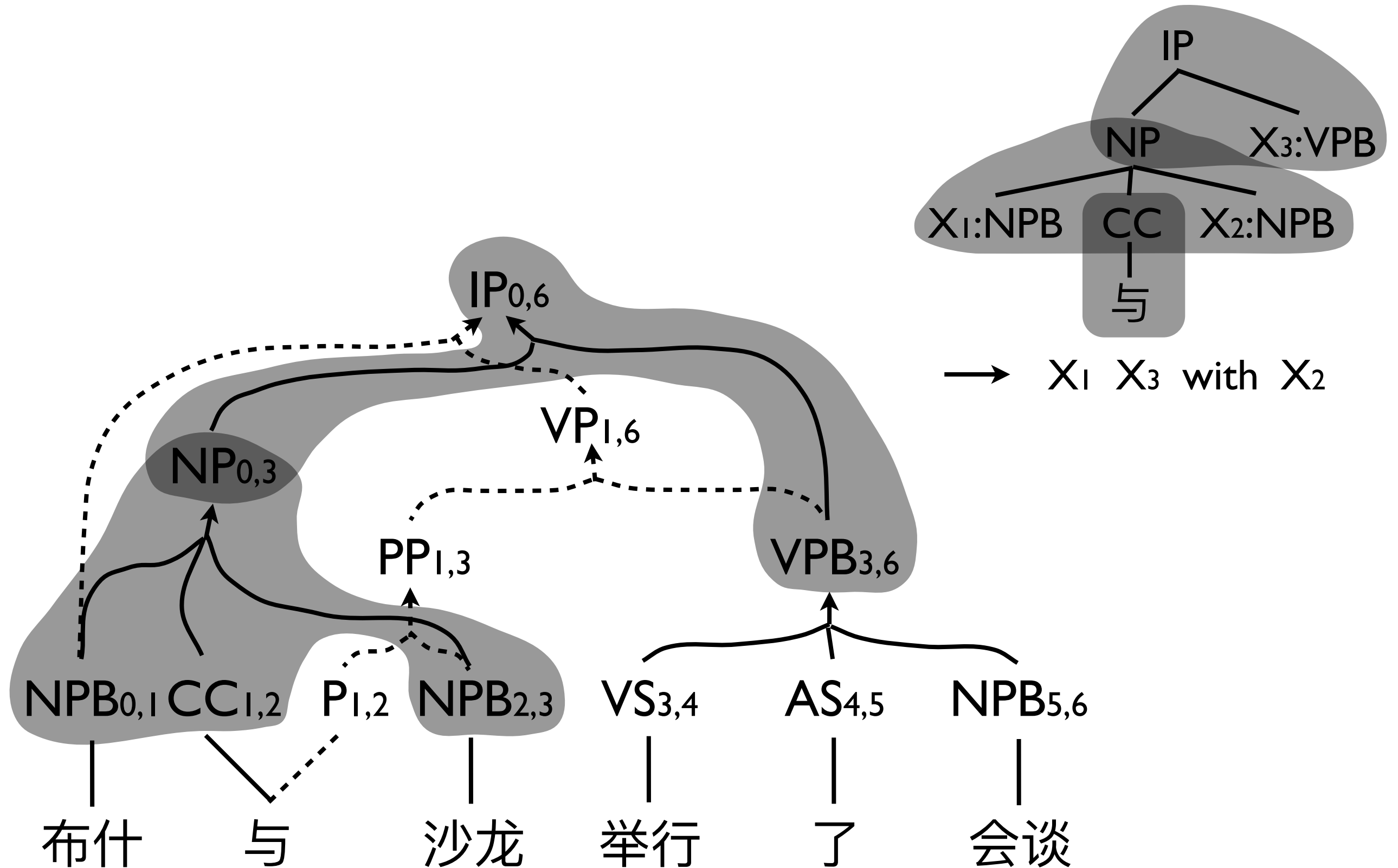
# Forest-based Decoding



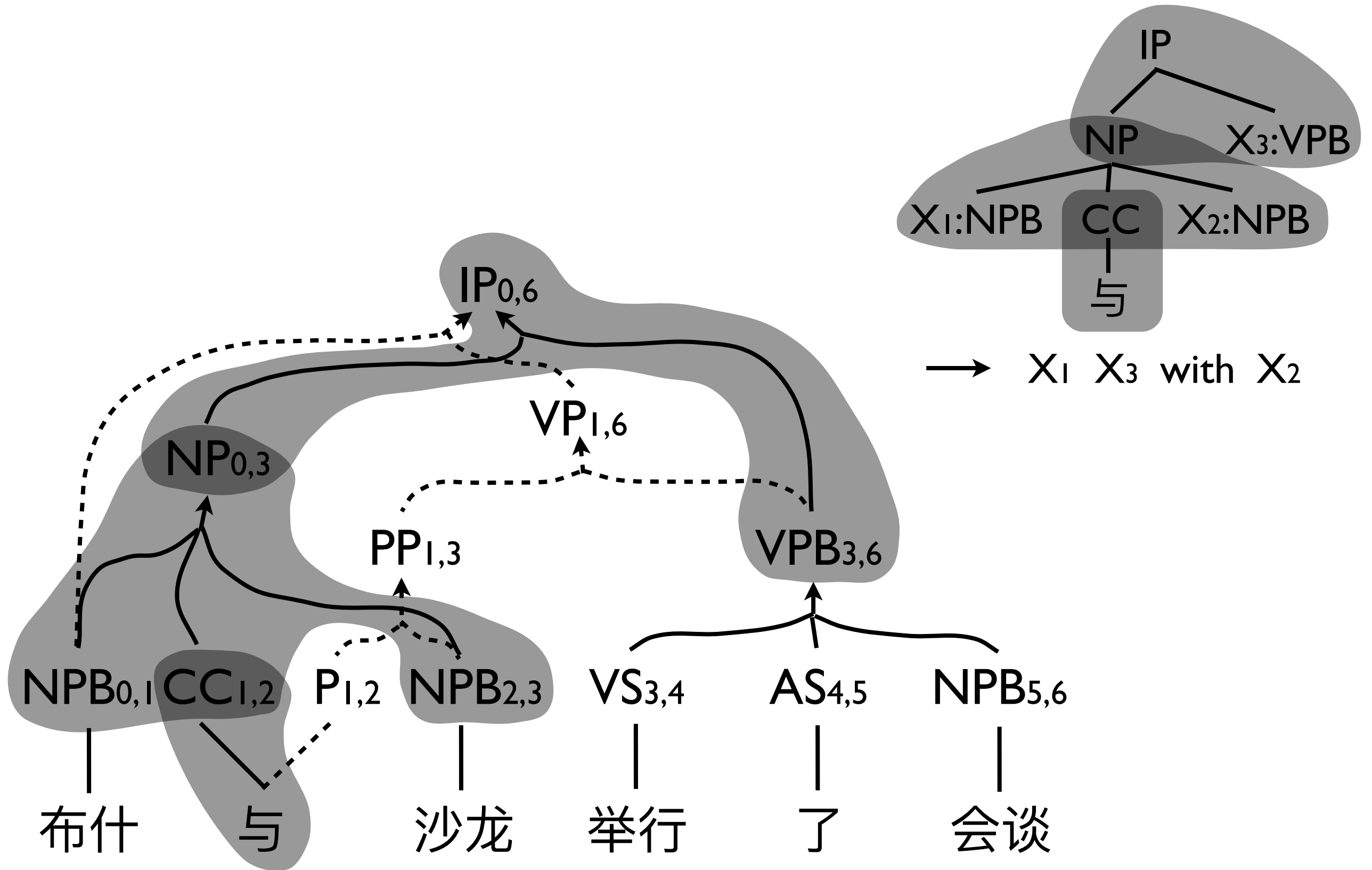
# Forest-based Decoding



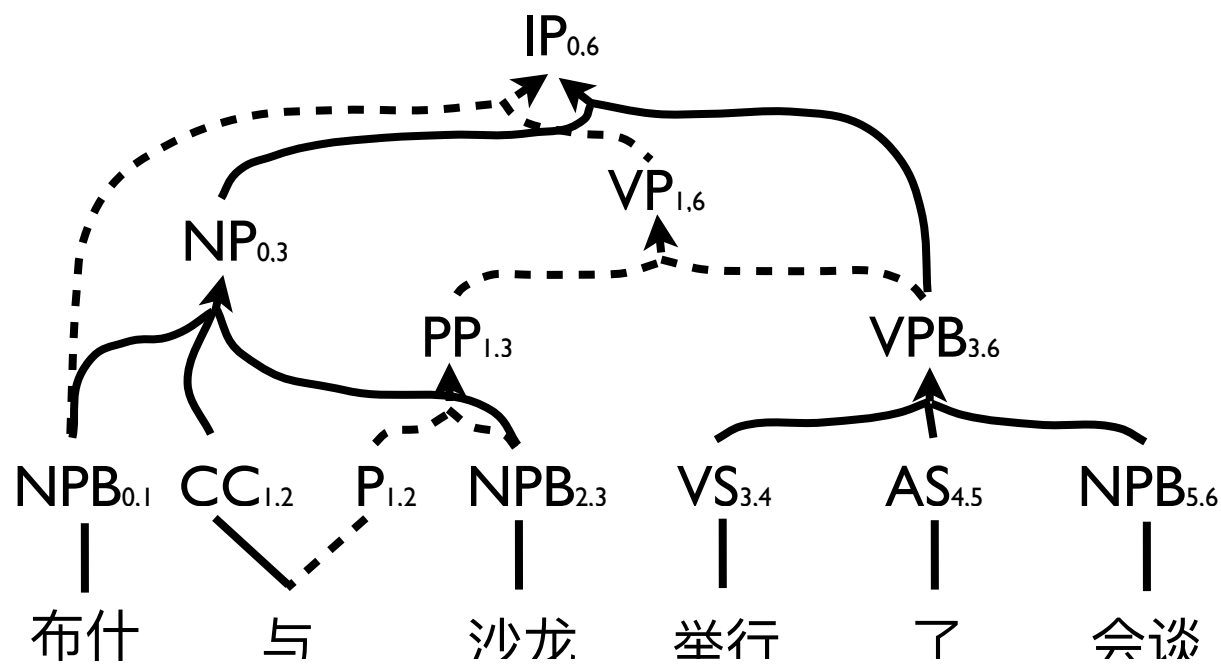
# Forest-based Decoding



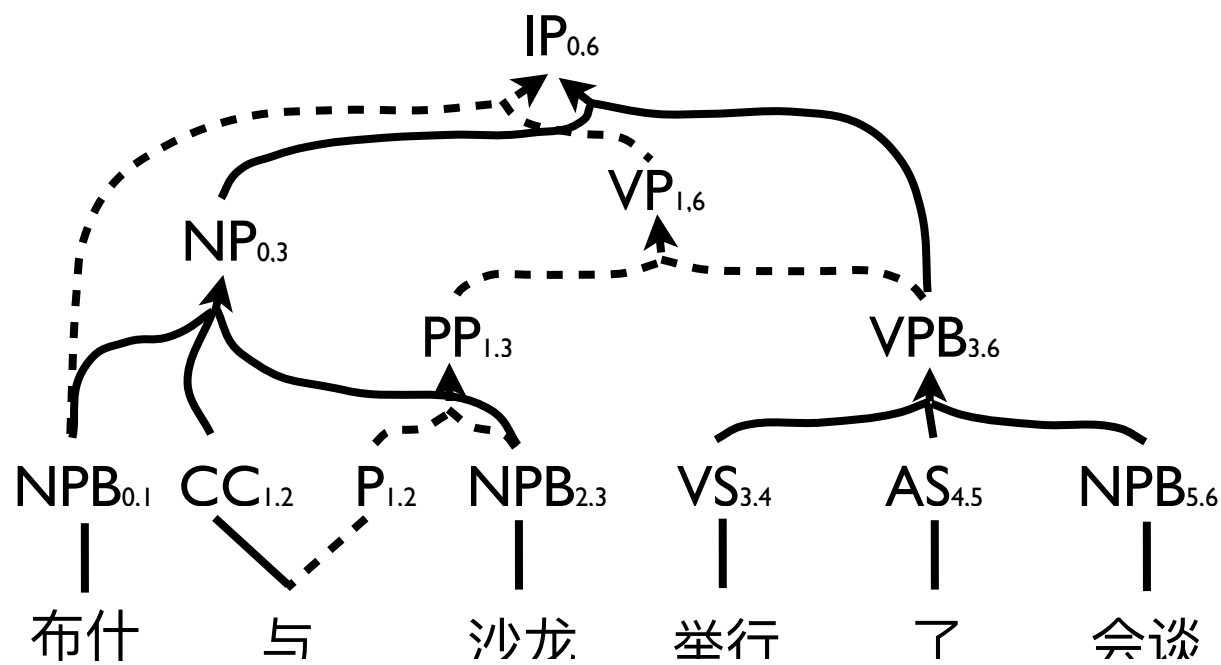
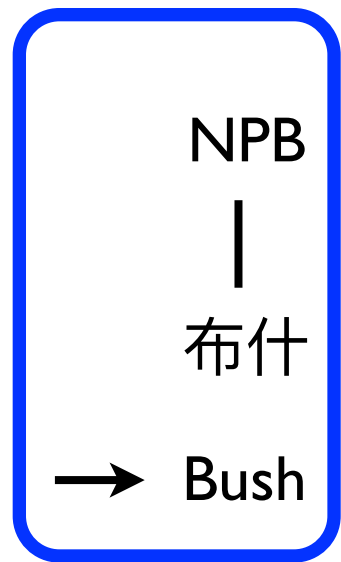
# Forest-based Decoding



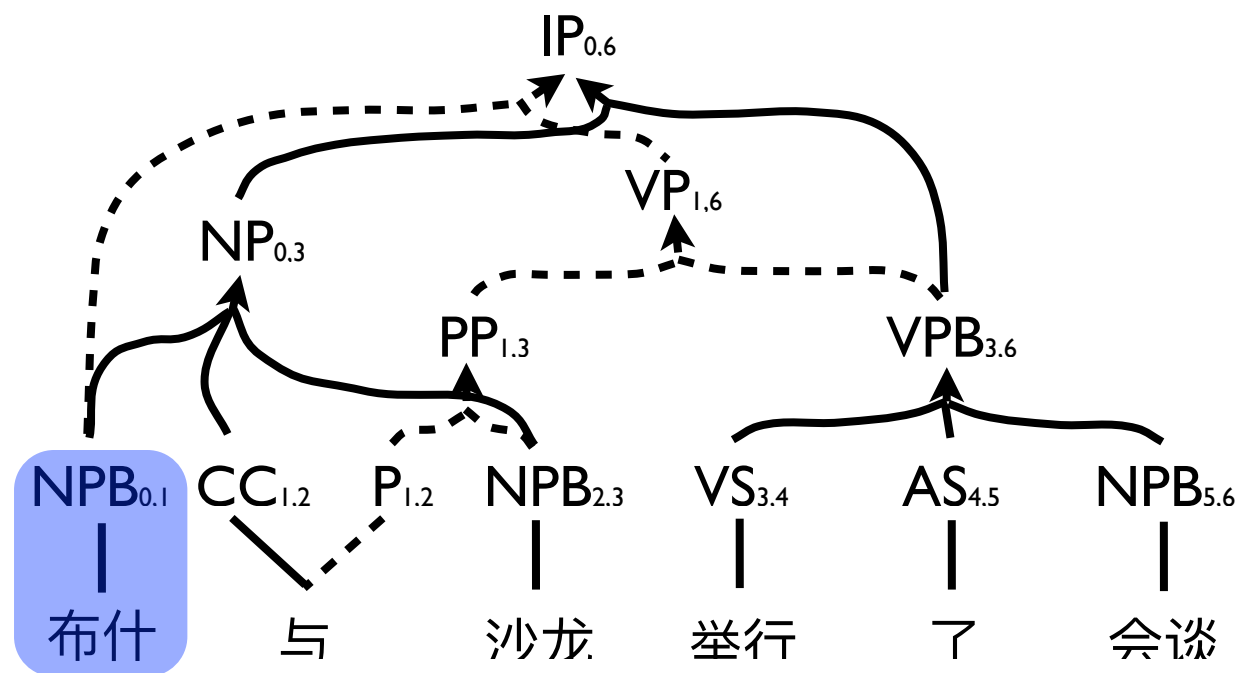
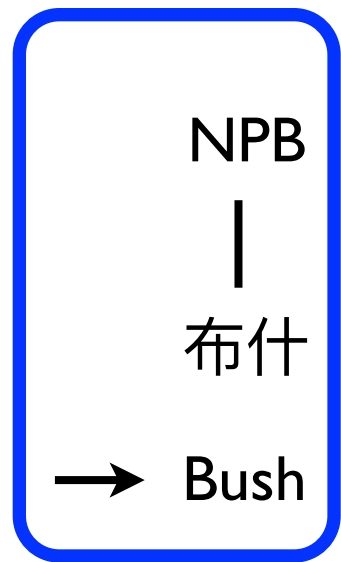
# Forest-based Decoding



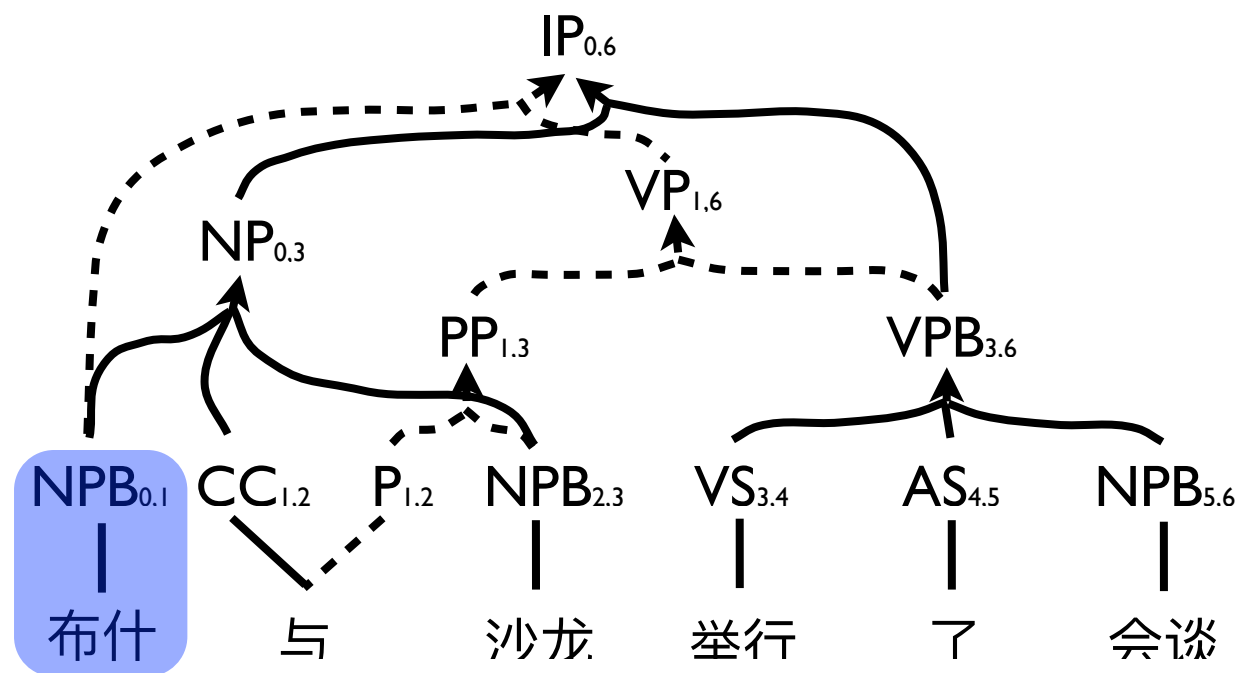
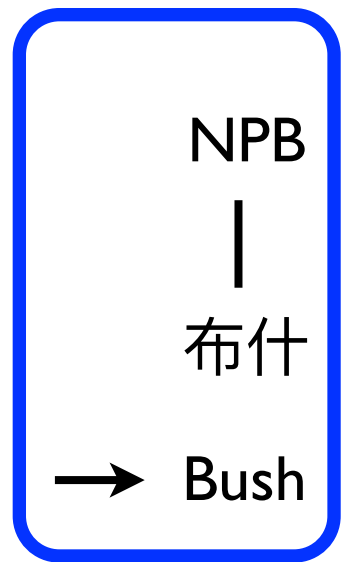
# Forest-based Decoding



# Forest-based Decoding



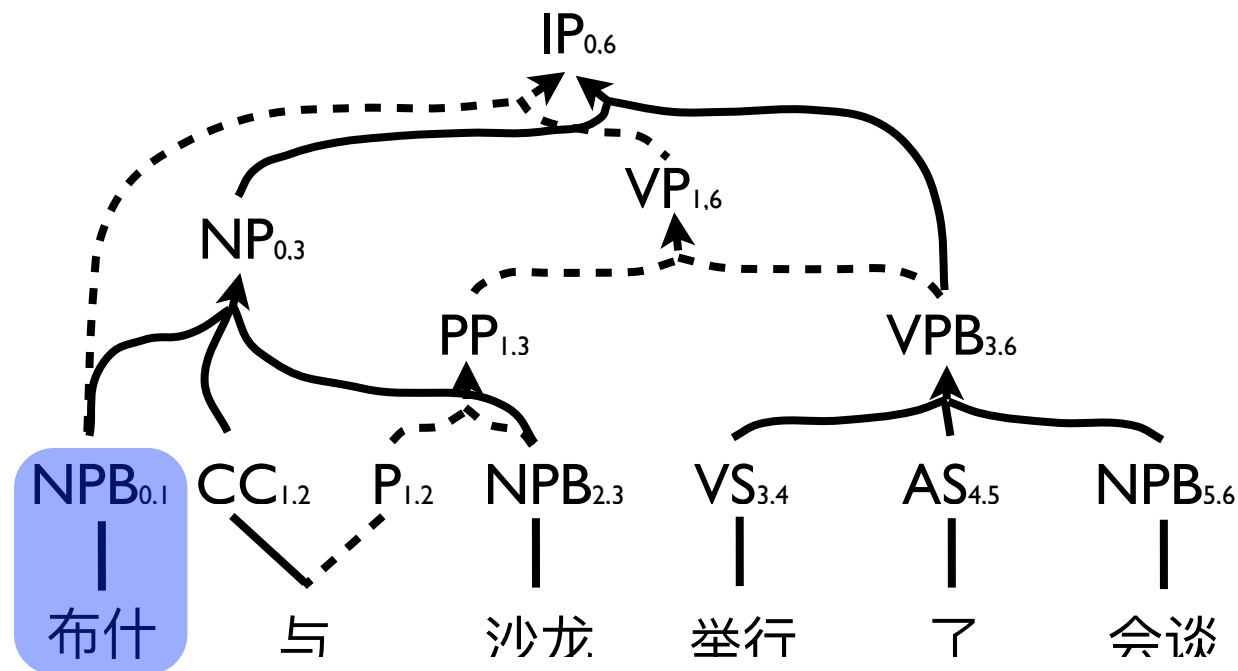
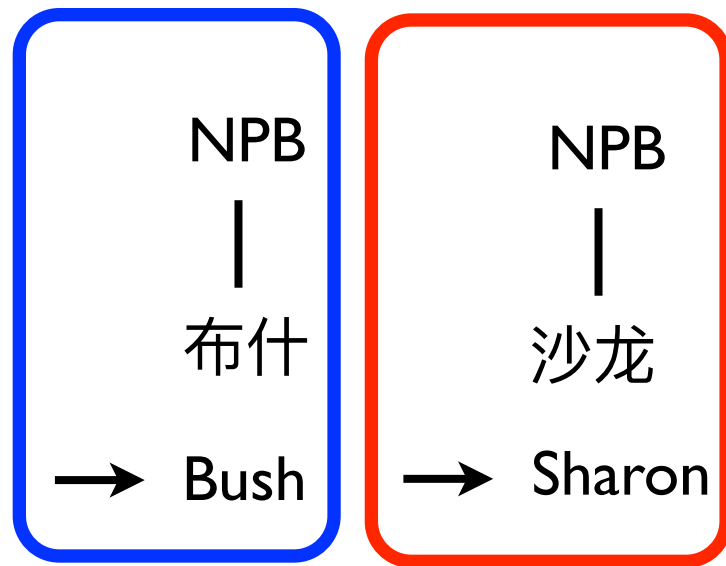
# Forest-based Decoding



NPB<sub>0,1</sub>

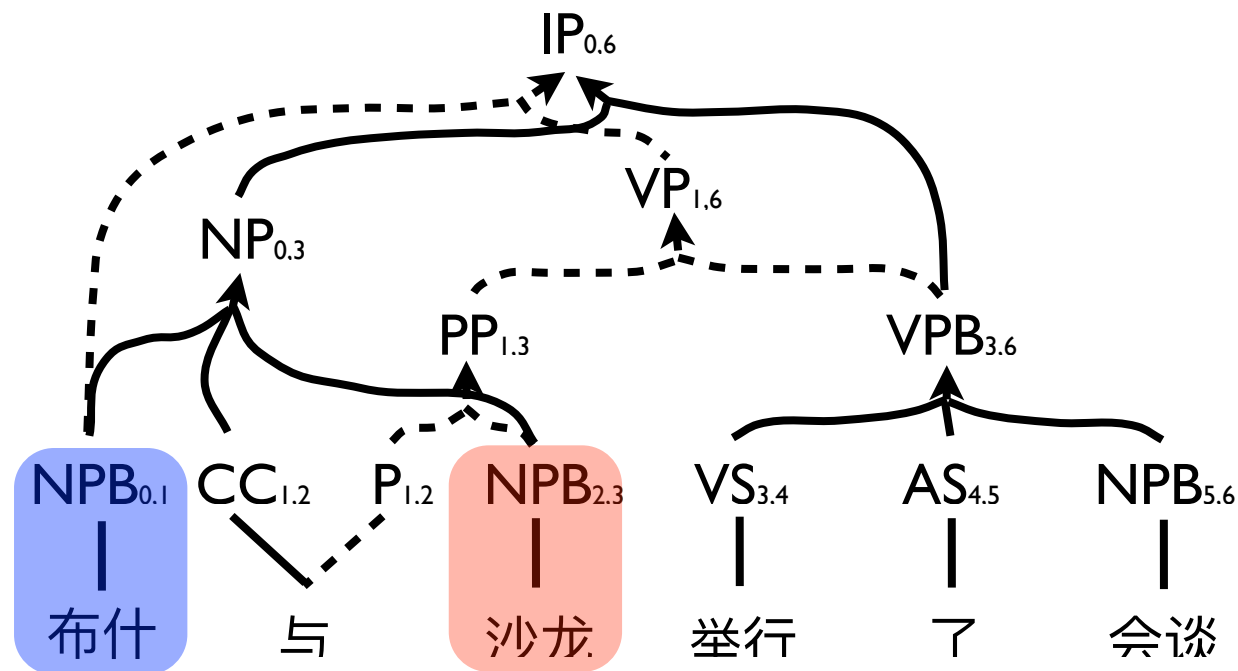
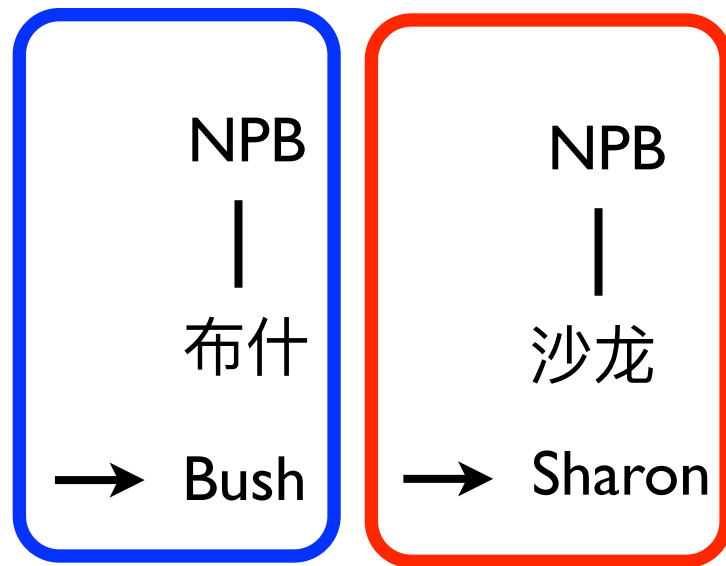


# Forest-based Decoding



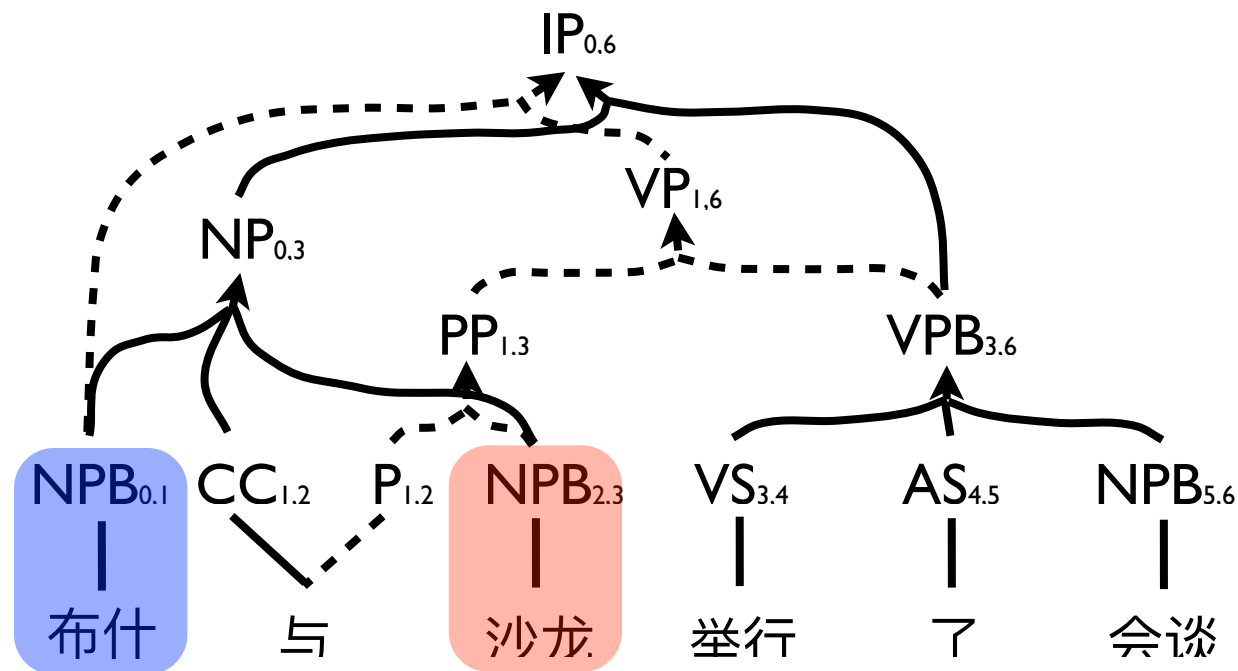
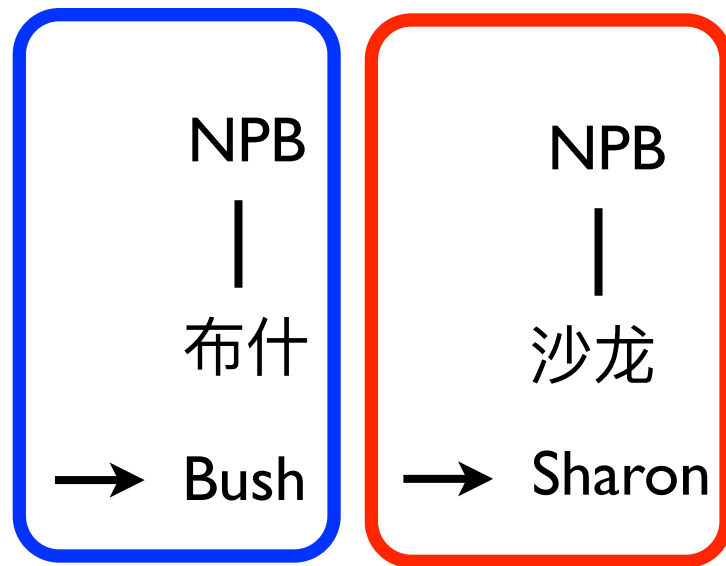
$NPB_{0,1}$

# Forest-based Decoding



NPB<sub>0,1</sub>

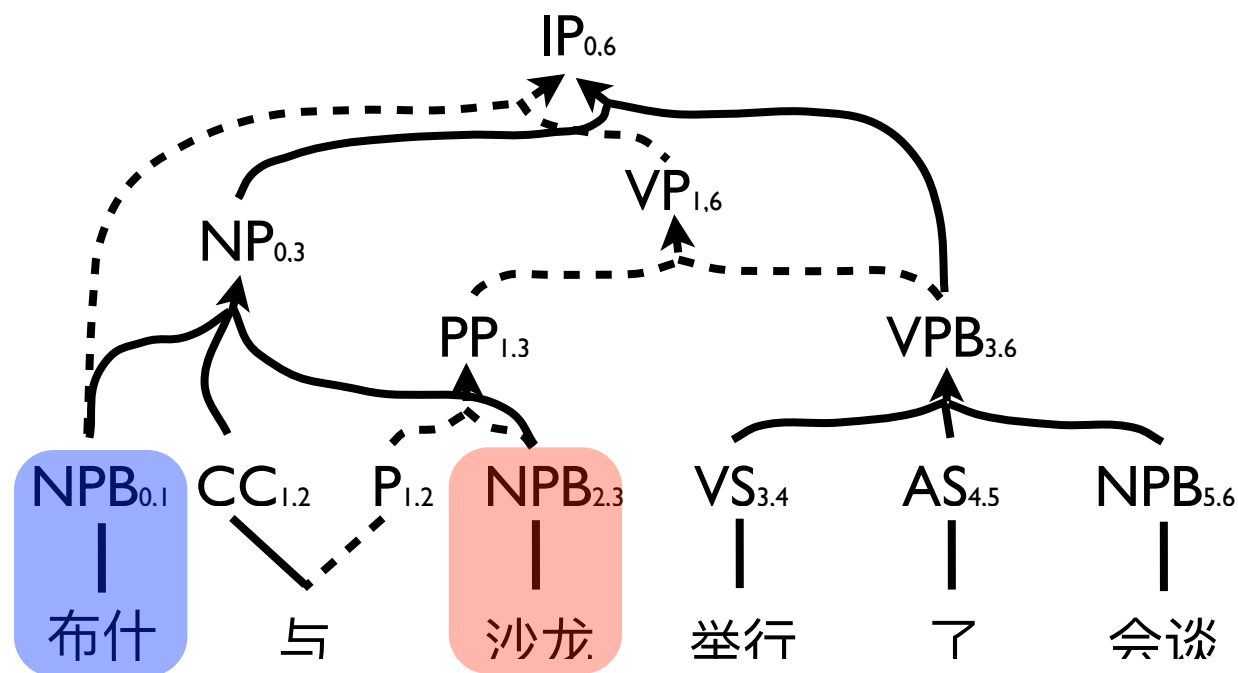
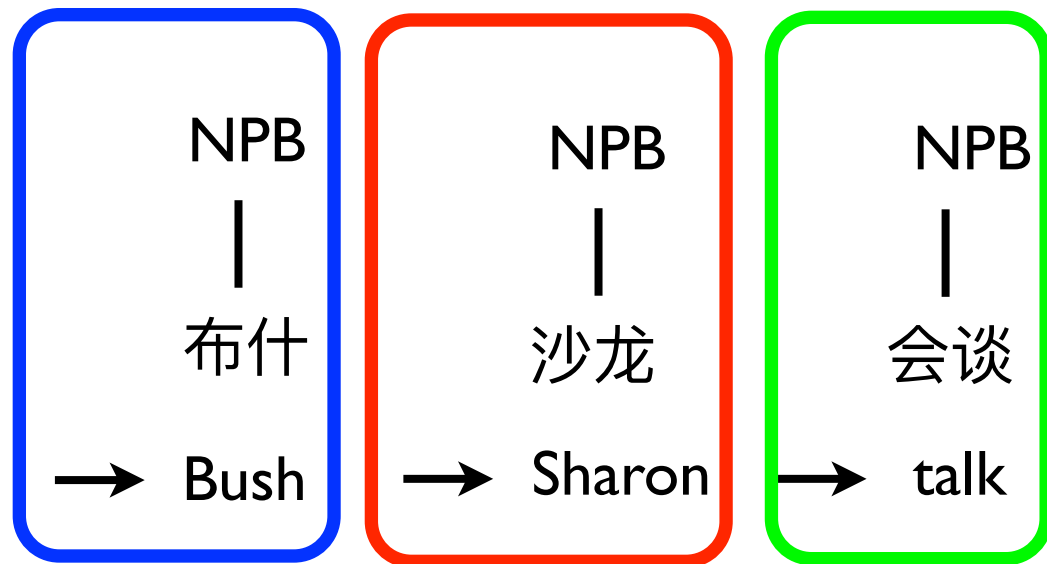
# Forest-based Decoding



NPB<sub>0,1</sub>

NPB<sub>2,3</sub>

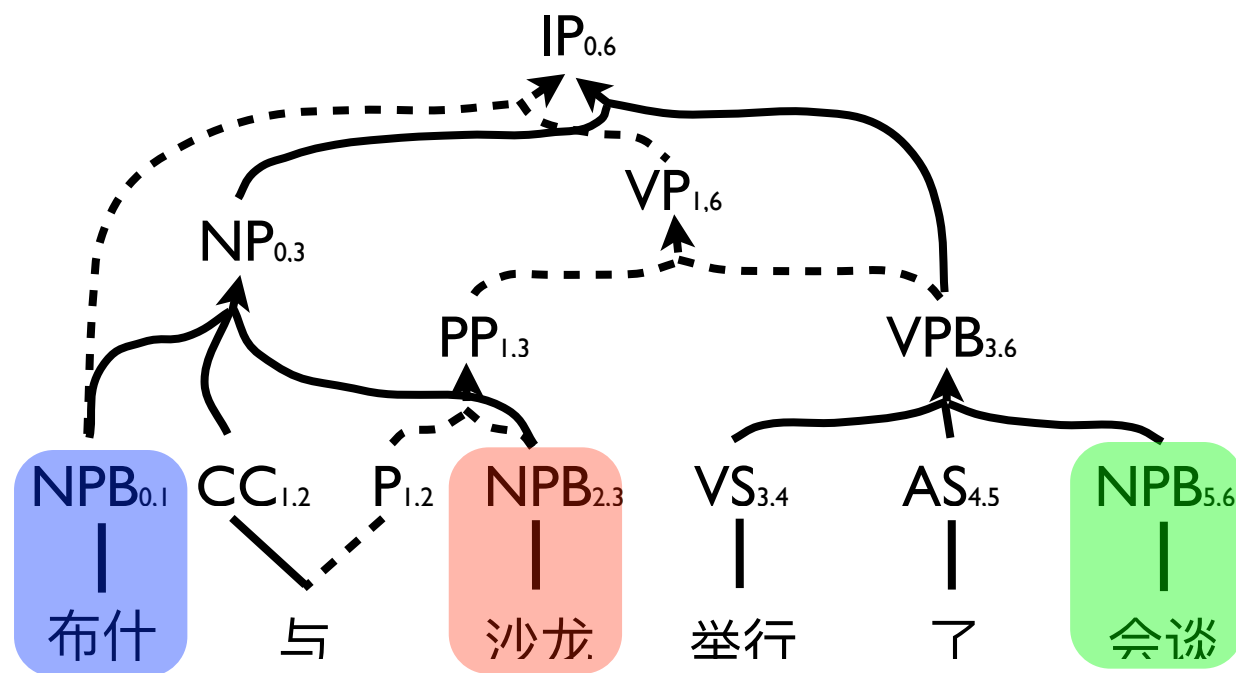
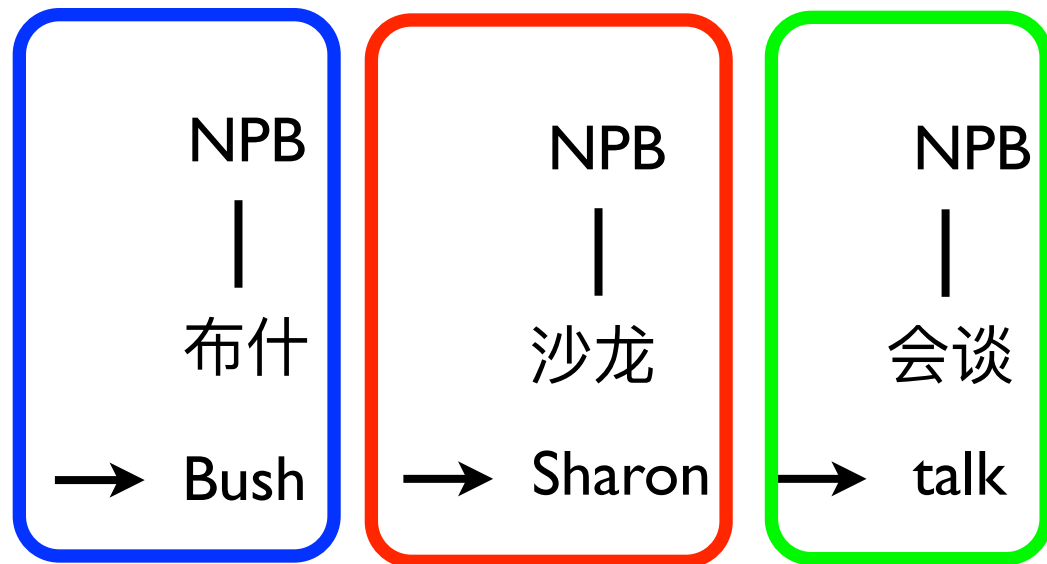
# Forest-based Decoding



NPB<sub>0,1</sub>

NPB<sub>2,3</sub>

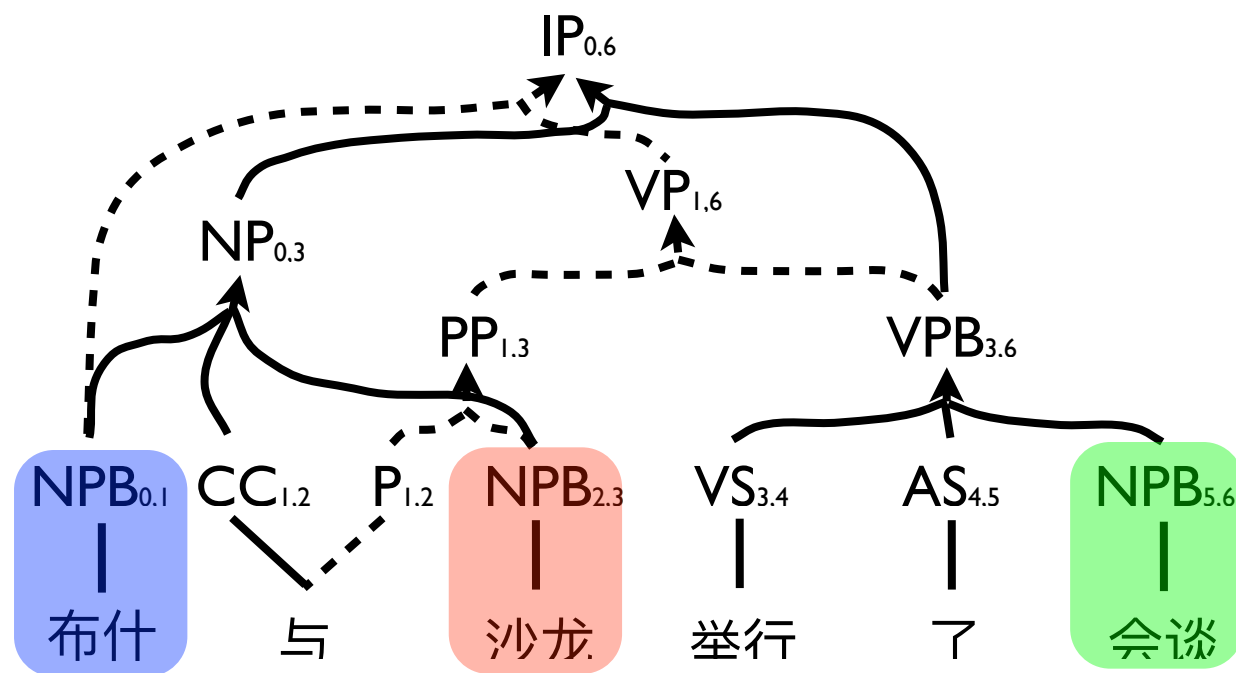
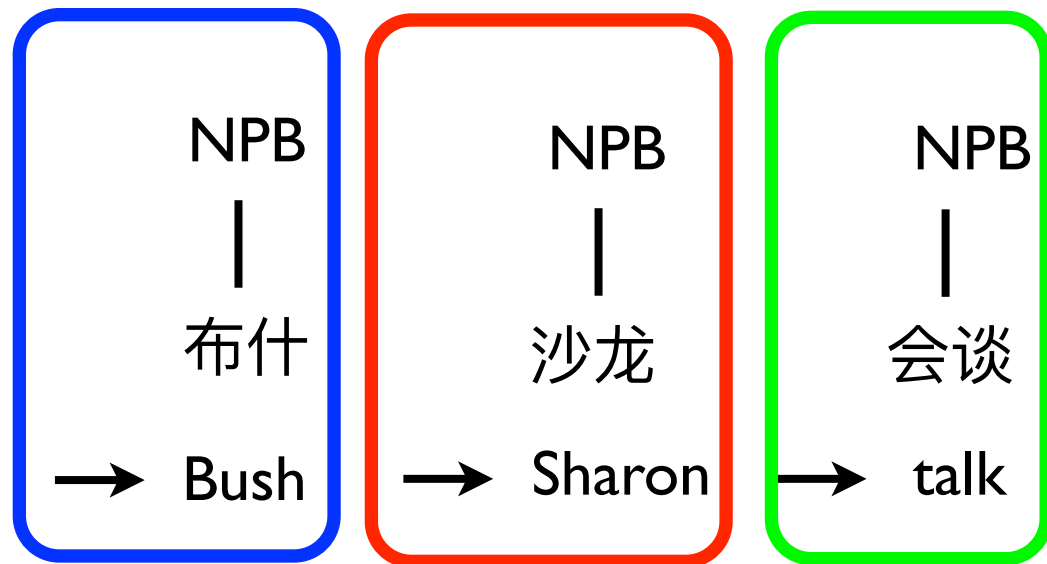
# Forest-based Decoding



NPB<sub>0,1</sub>

NPB<sub>2,3</sub>

# Forest-based Decoding

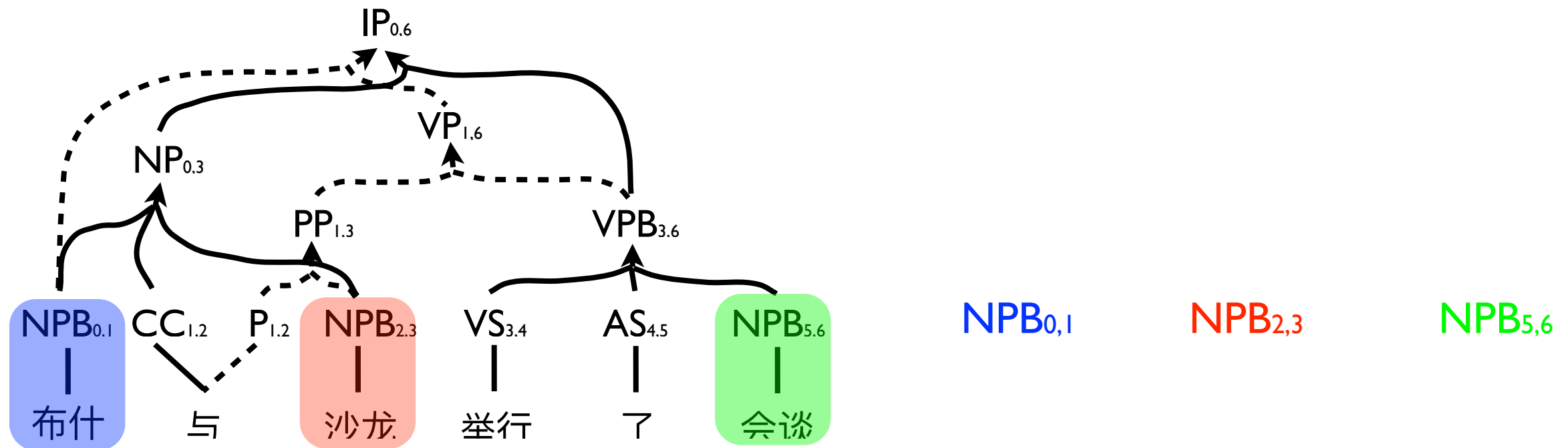
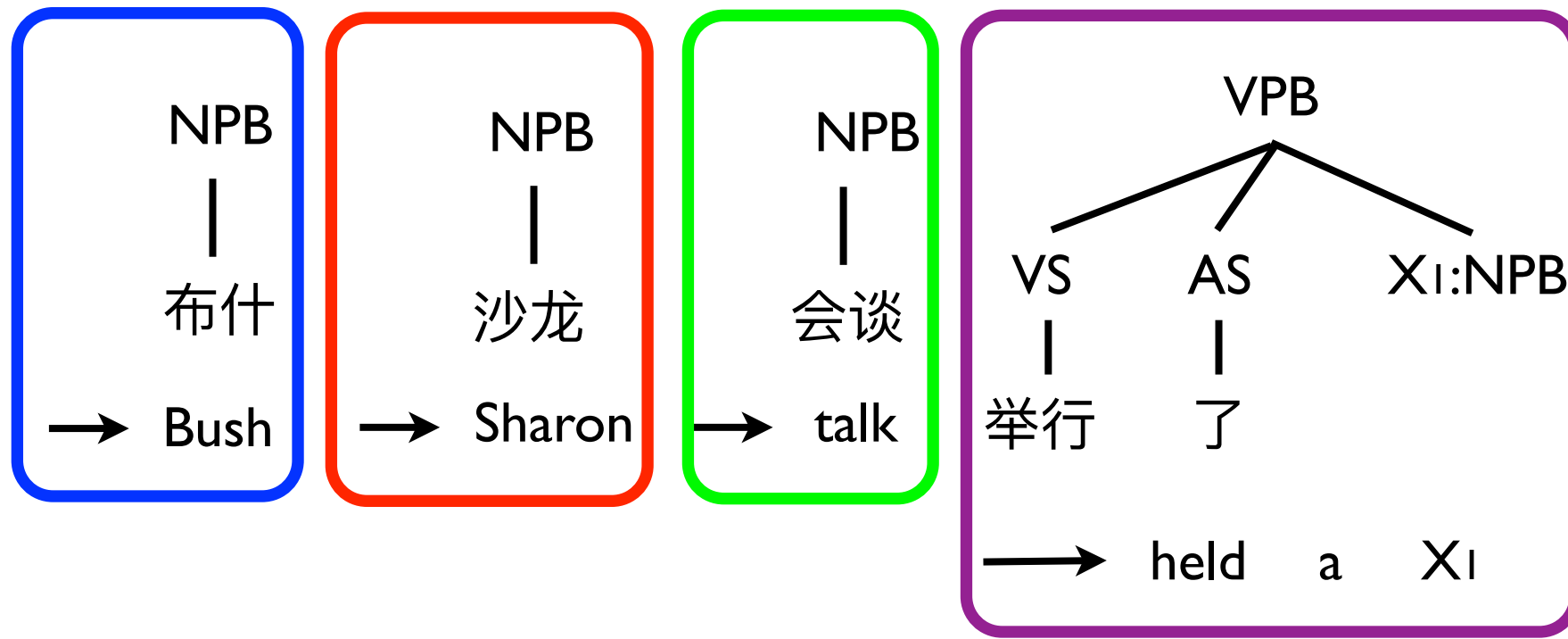


NPB<sub>0,1</sub>

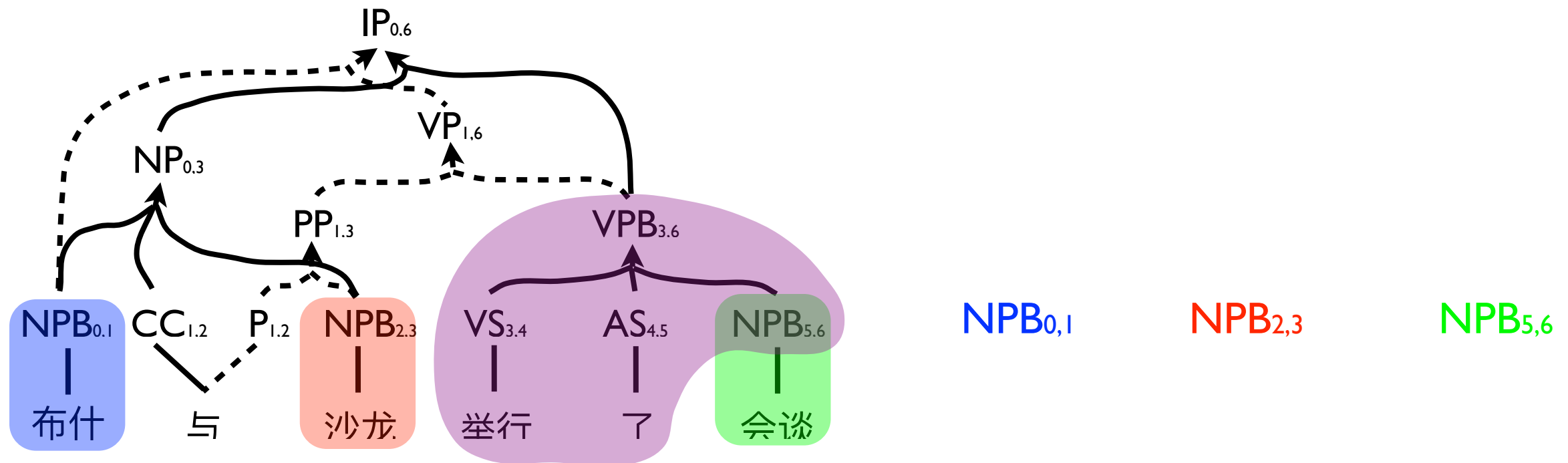
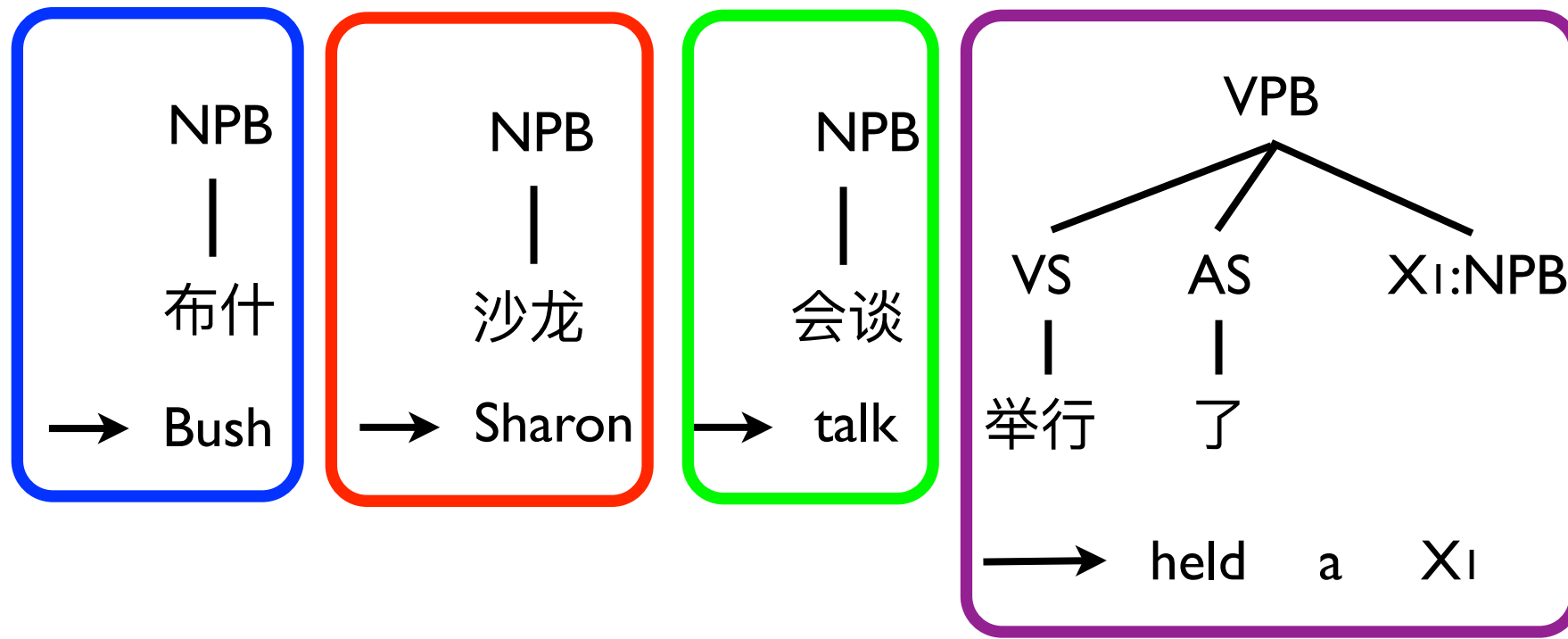
NPB<sub>2,3</sub>

NPB<sub>5,6</sub>

# Forest-based Decoding

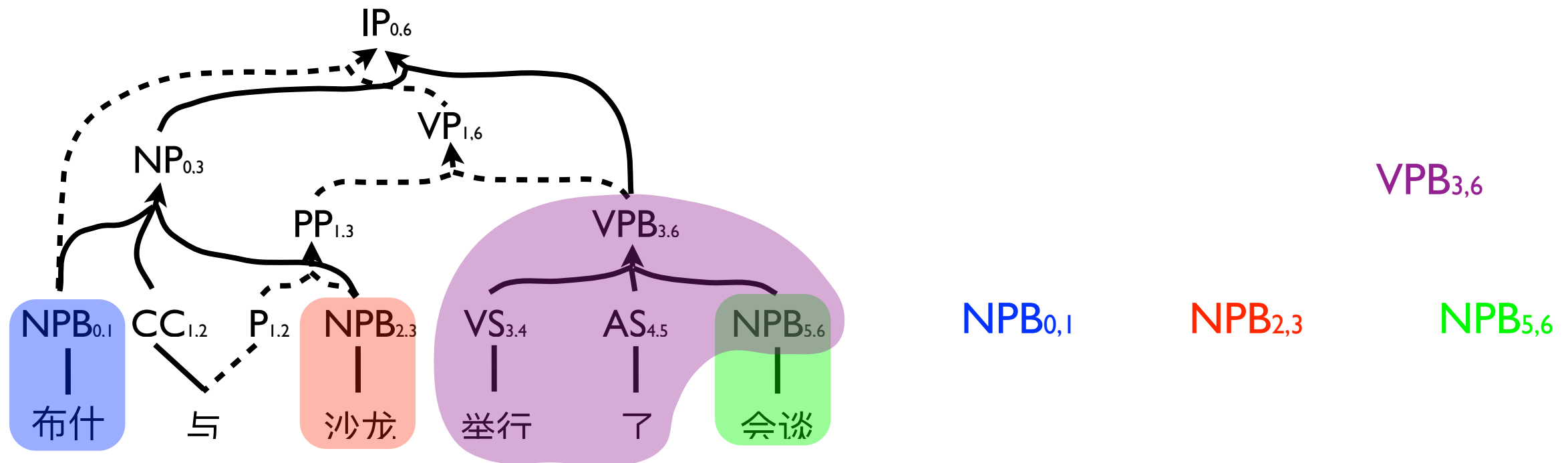
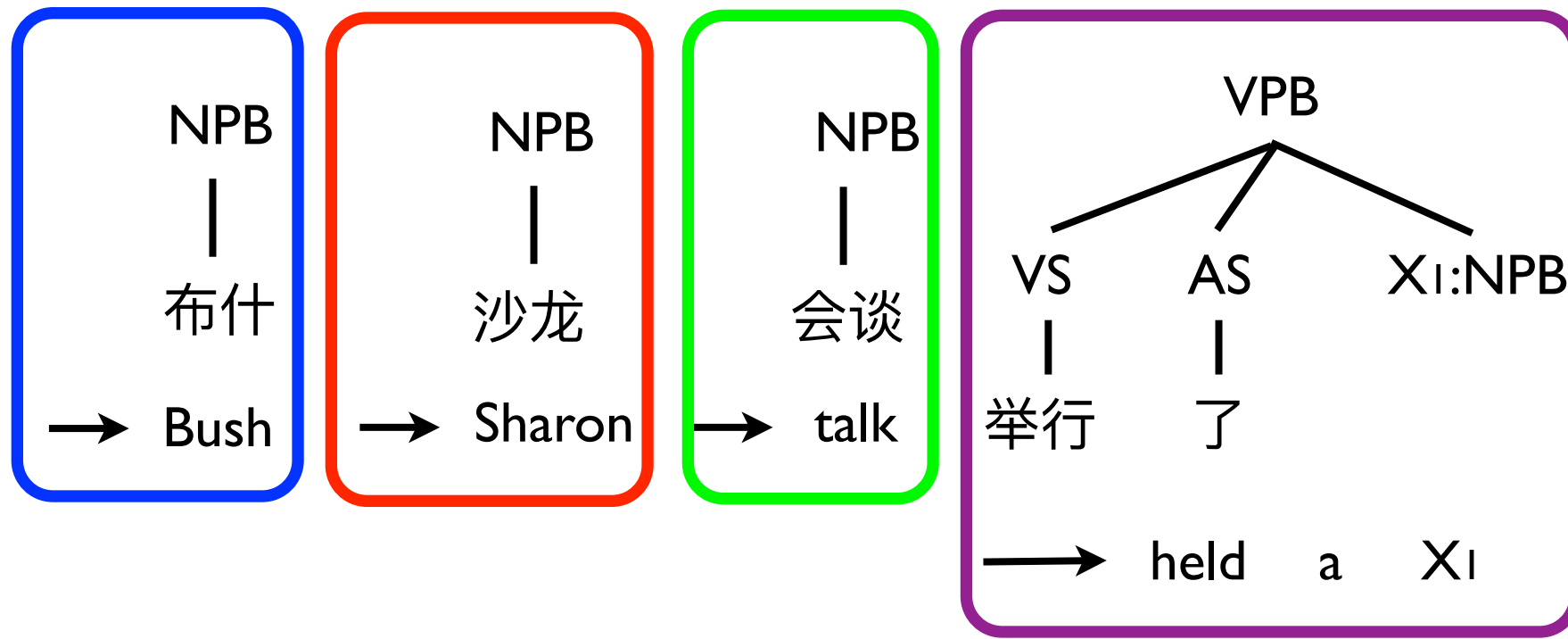


# Forest-based Decoding

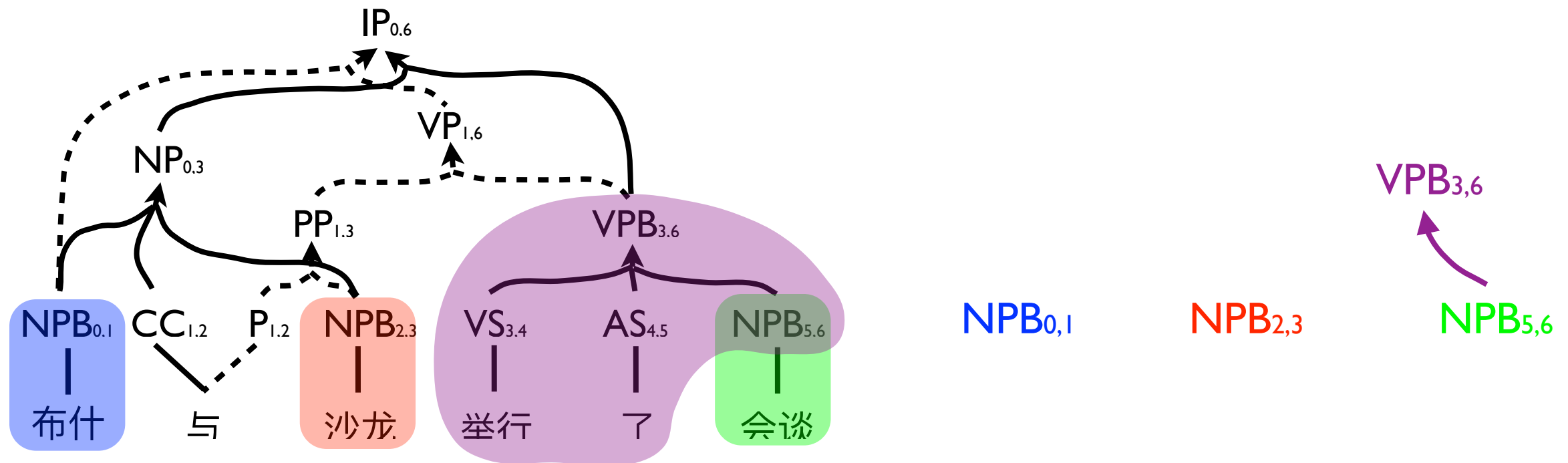
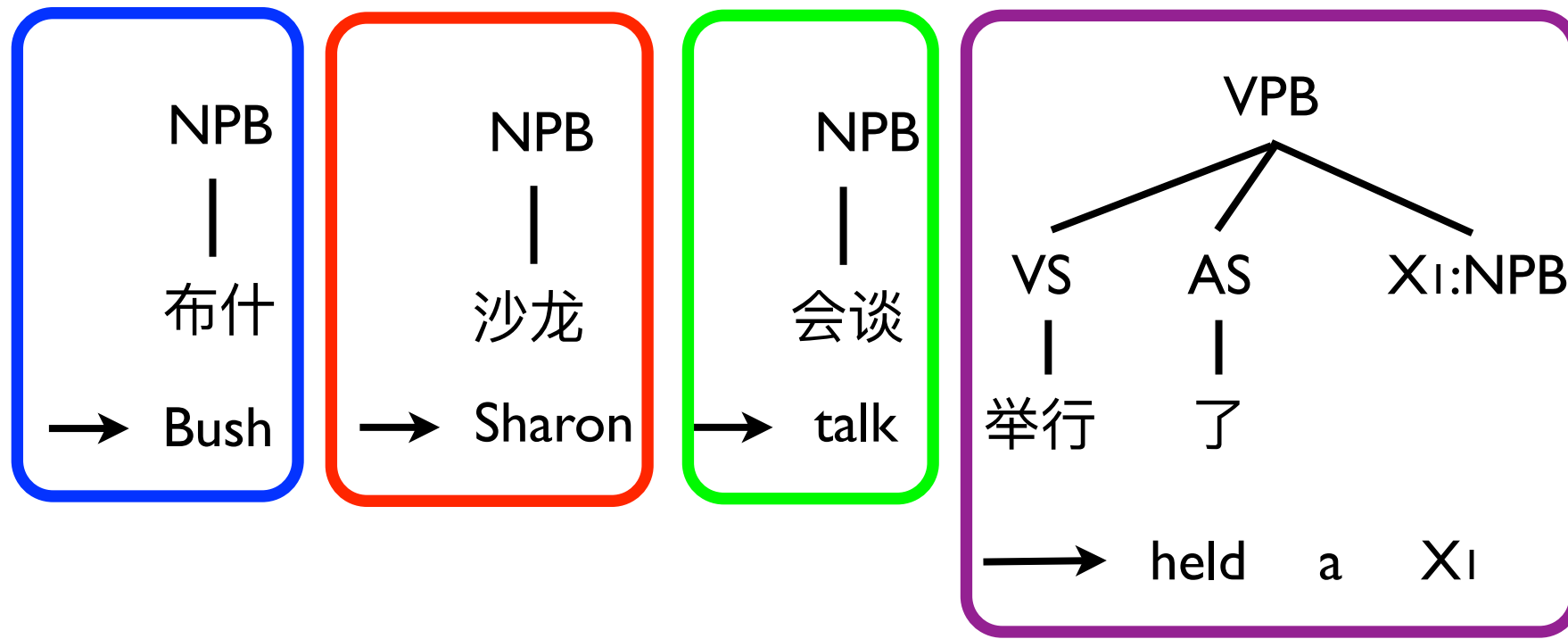




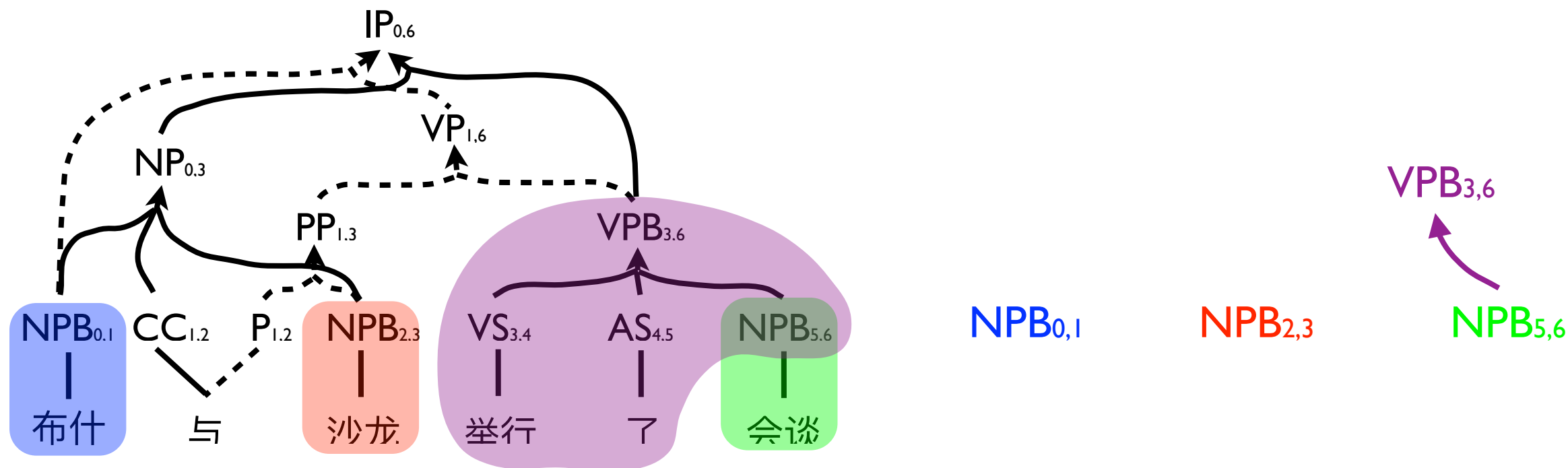
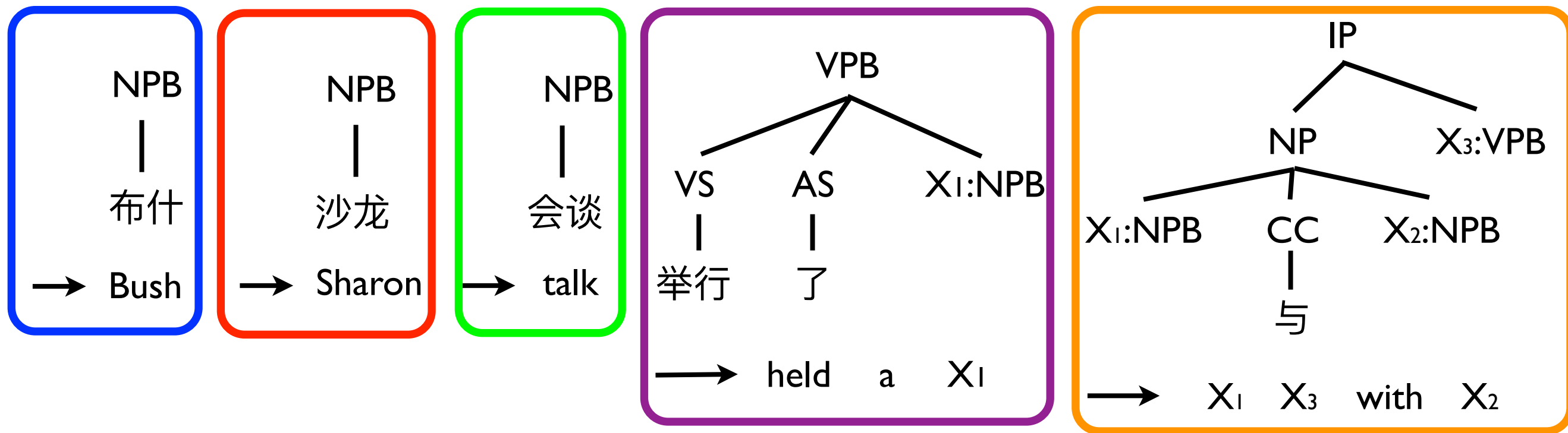
# Forest-based Decoding



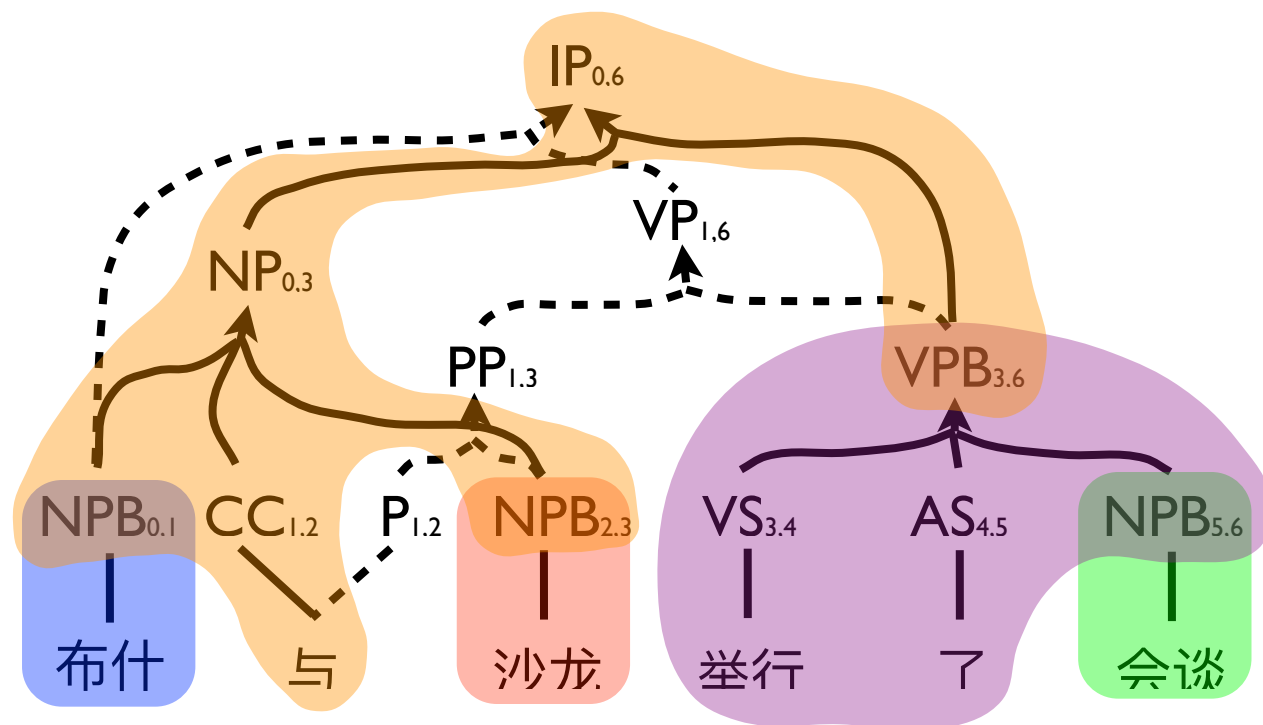
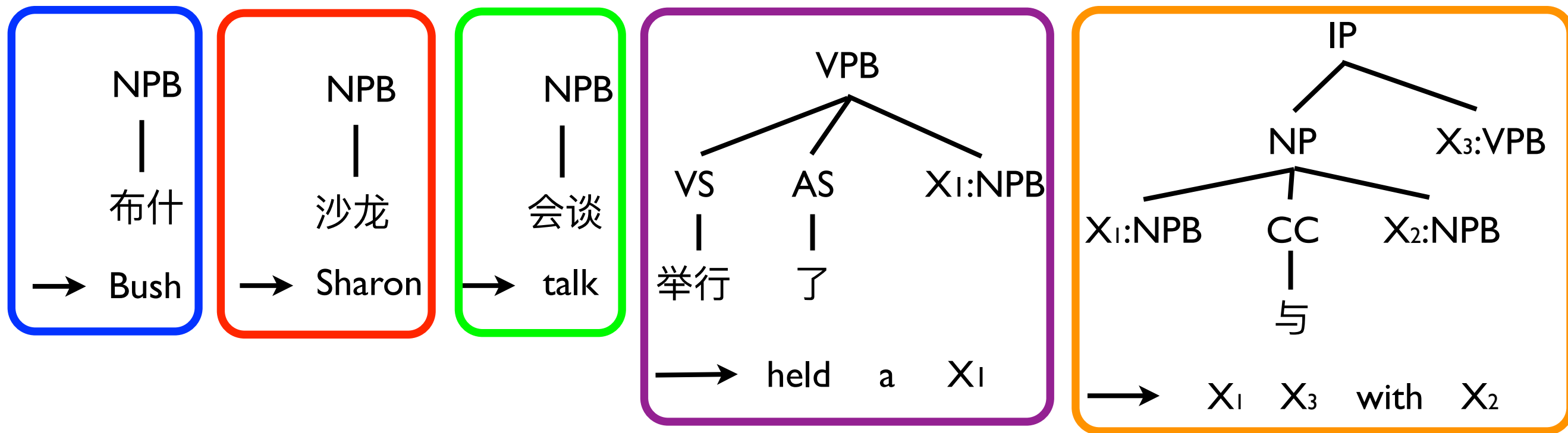
# Forest-based Decoding



# Forest-based Decoding



# Forest-based Decoding



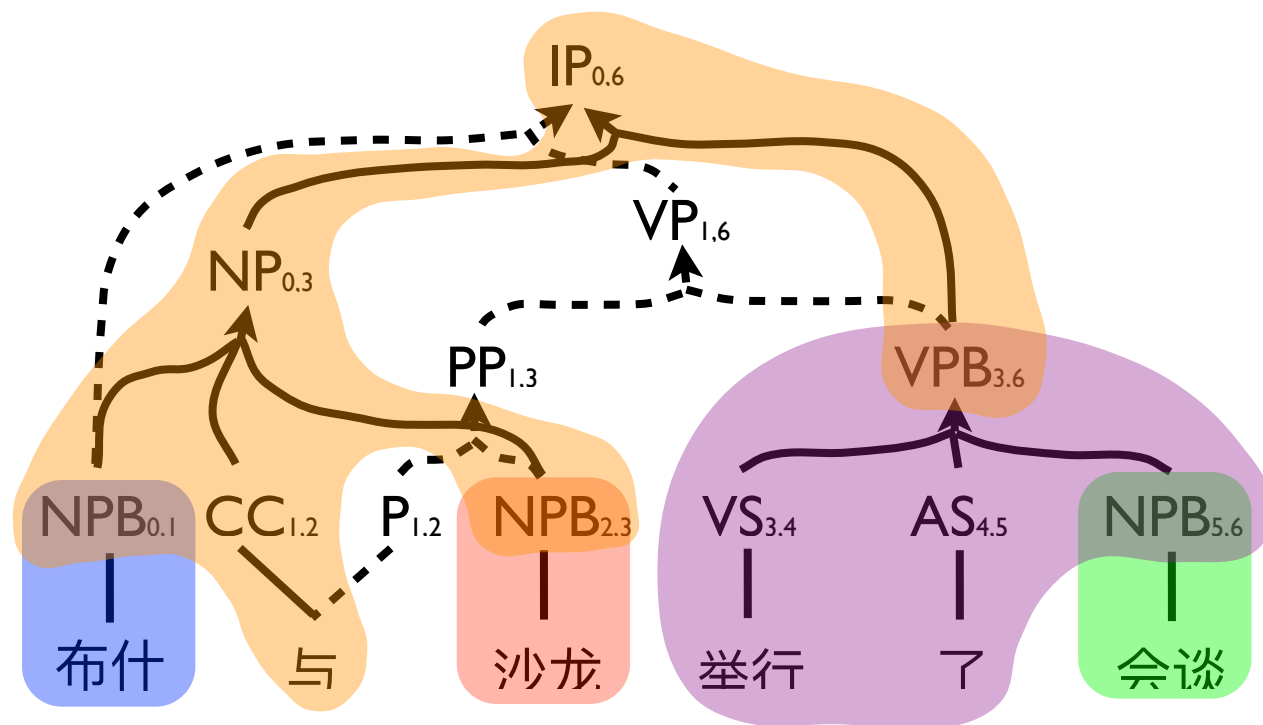
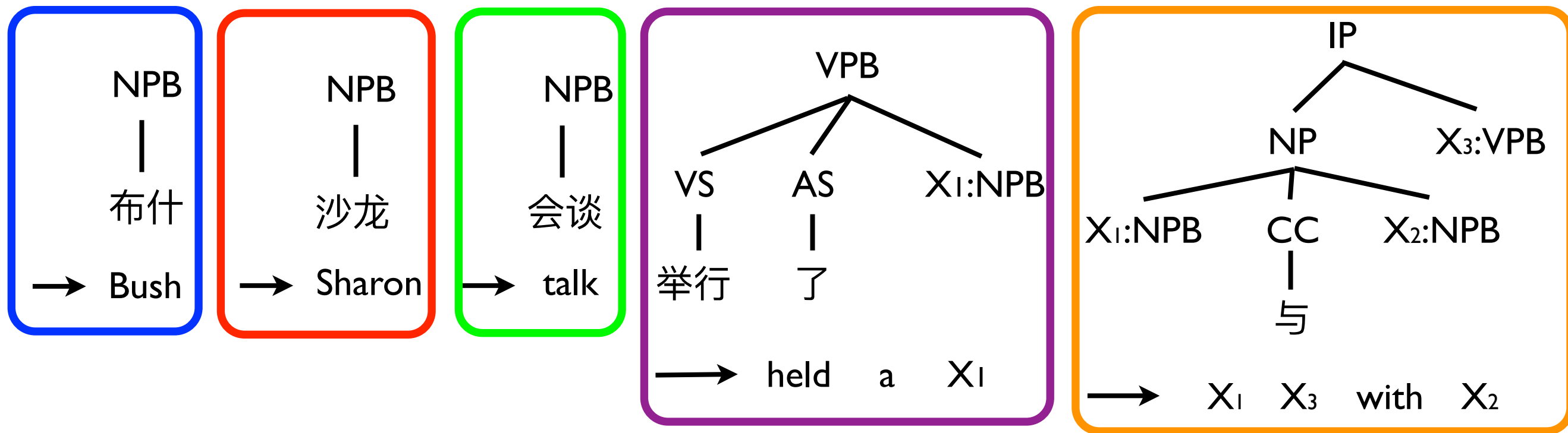
NPB<sub>0,1</sub>

NPB<sub>2,3</sub>

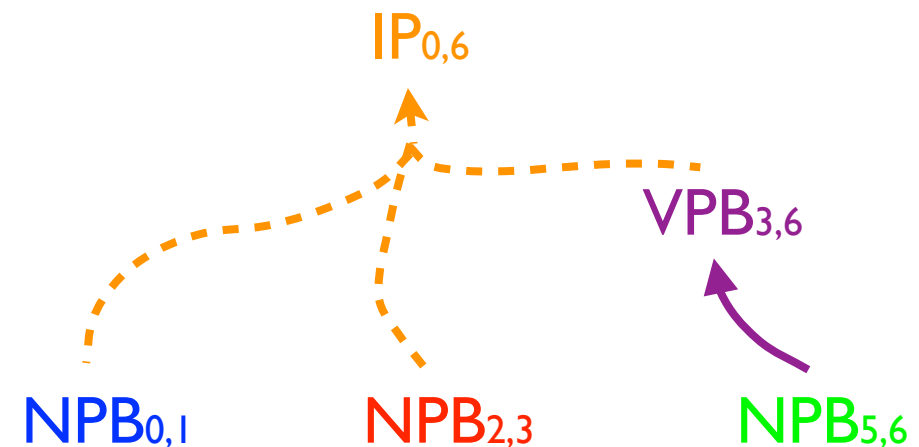
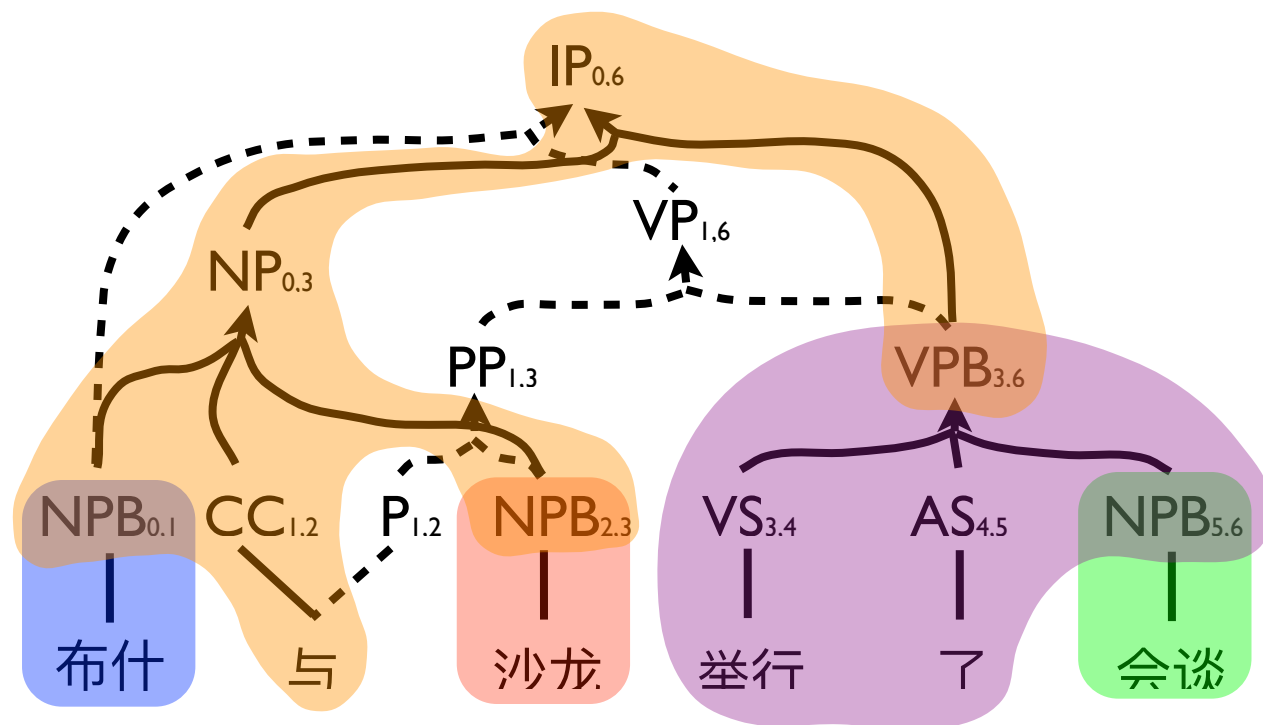
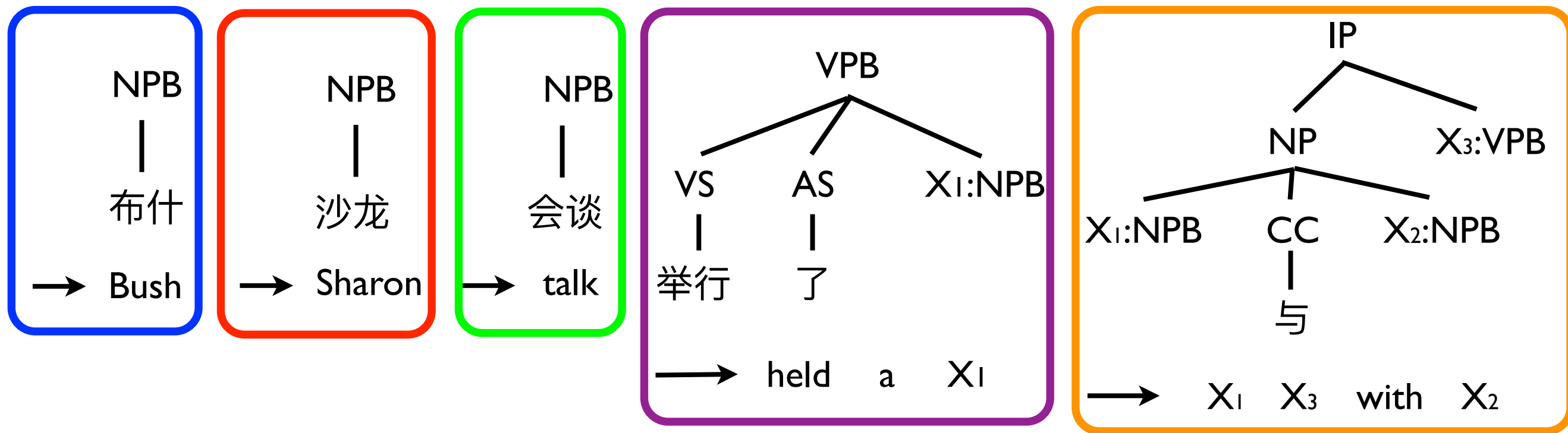
NPB<sub>5,6</sub>

VPB<sub>3,6</sub>

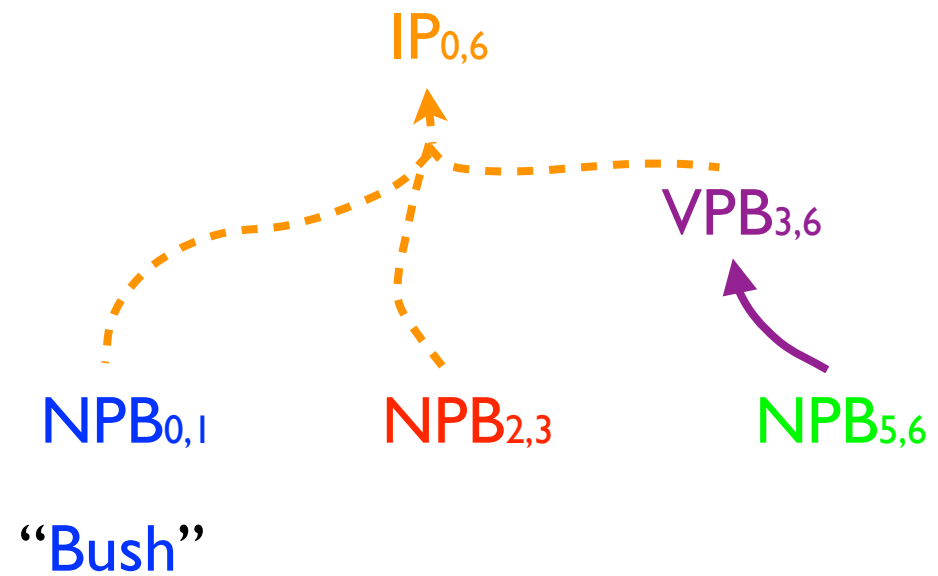
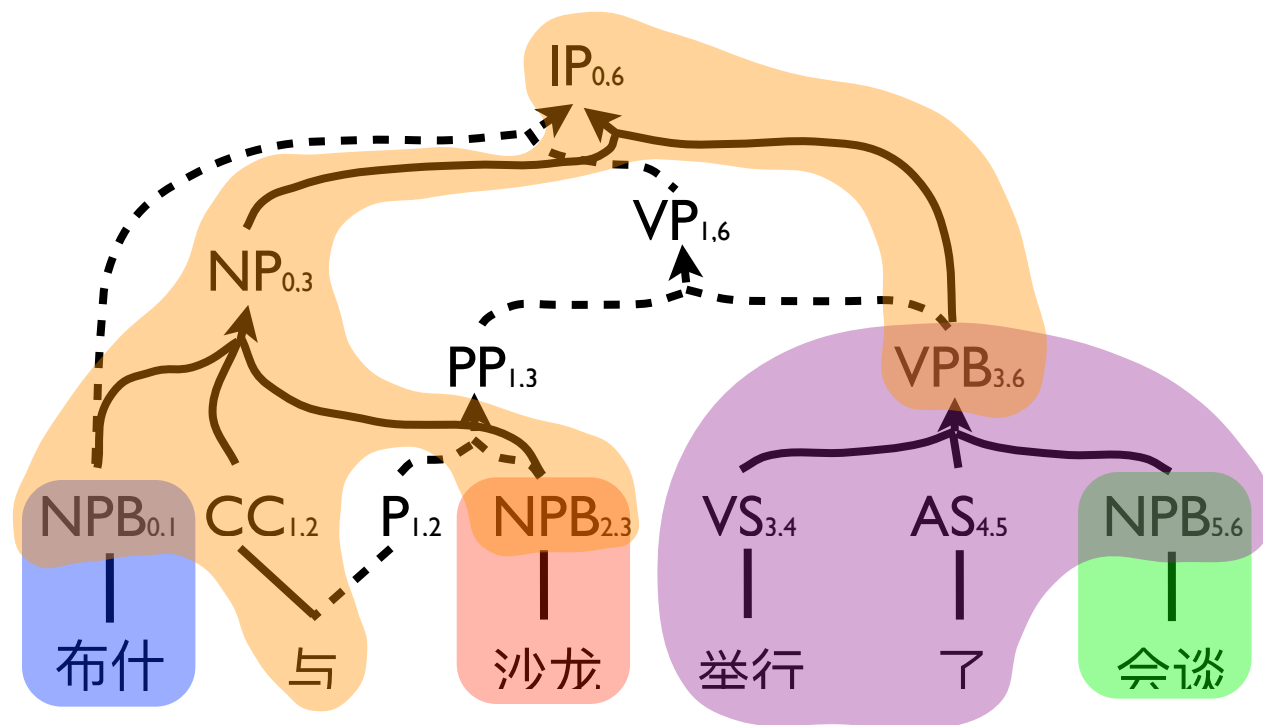
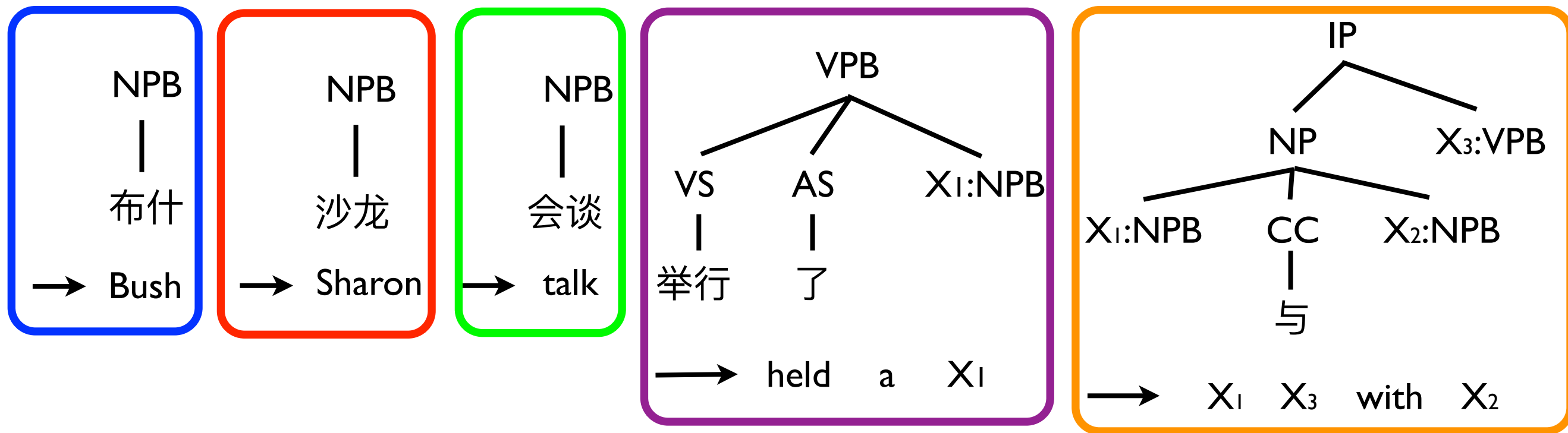
# Forest-based Decoding



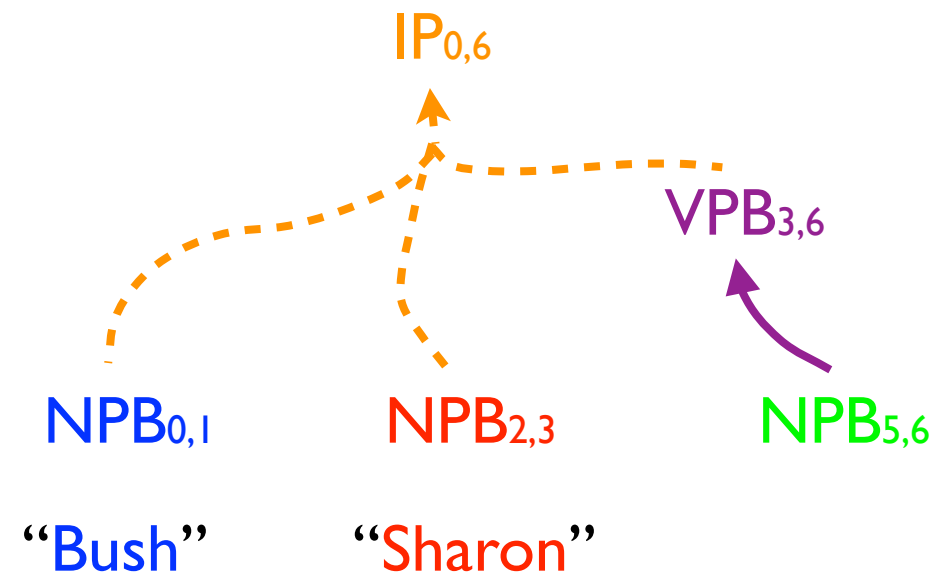
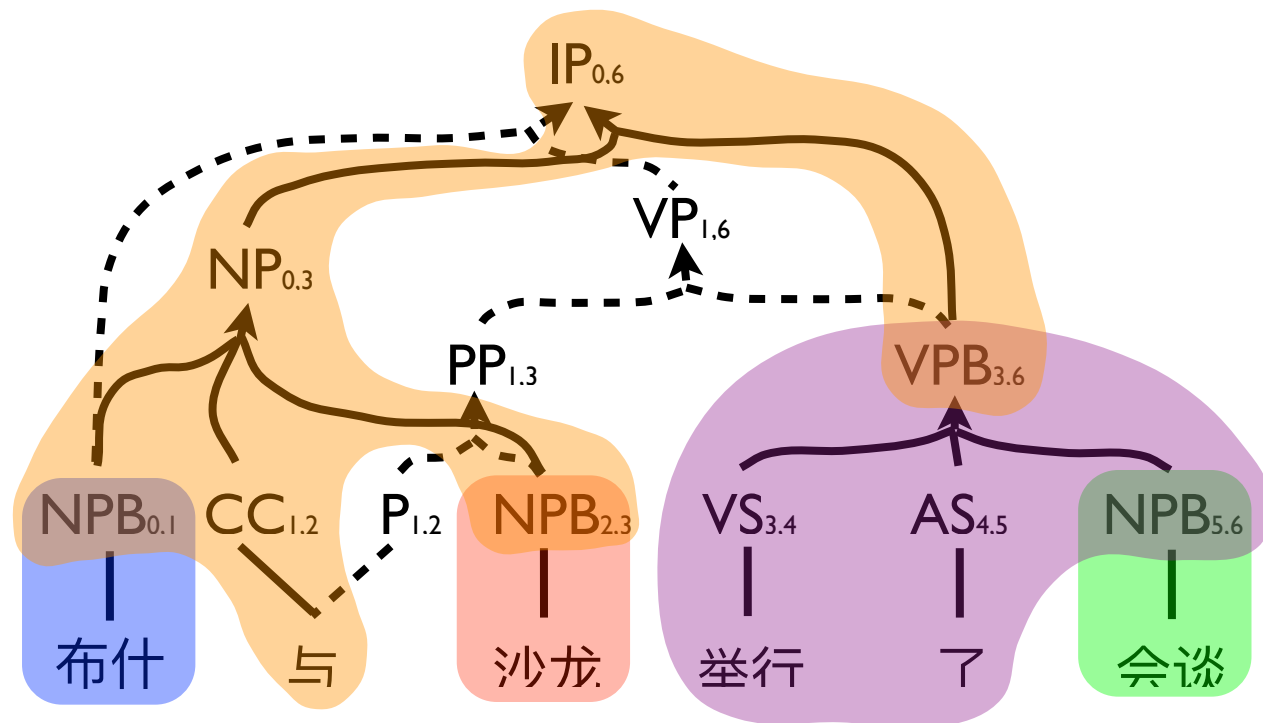
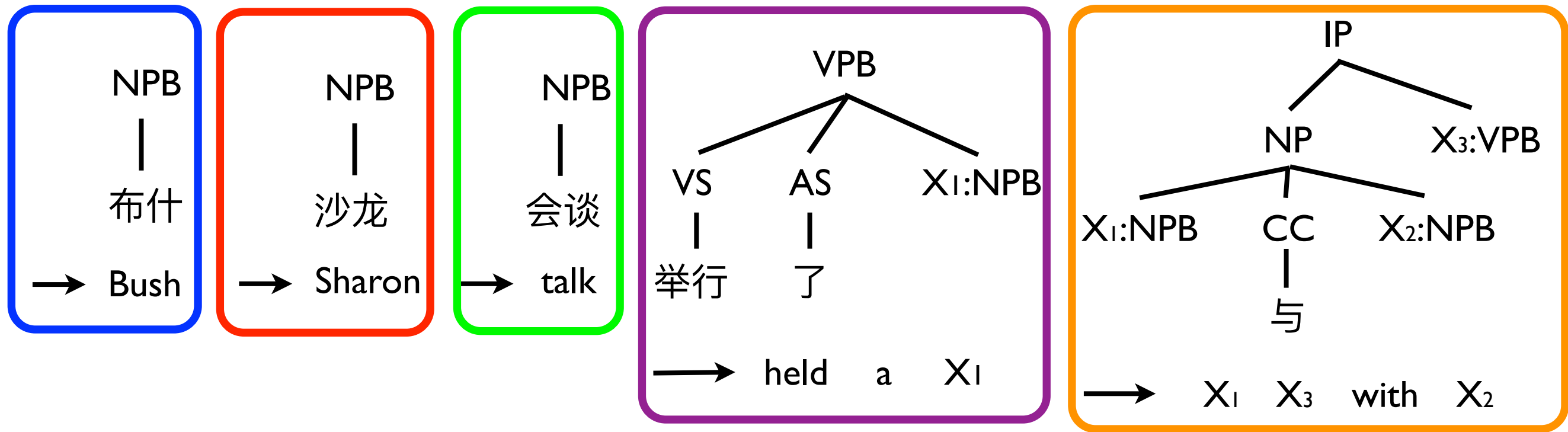
# Forest-based Decoding



# Forest-based Decoding

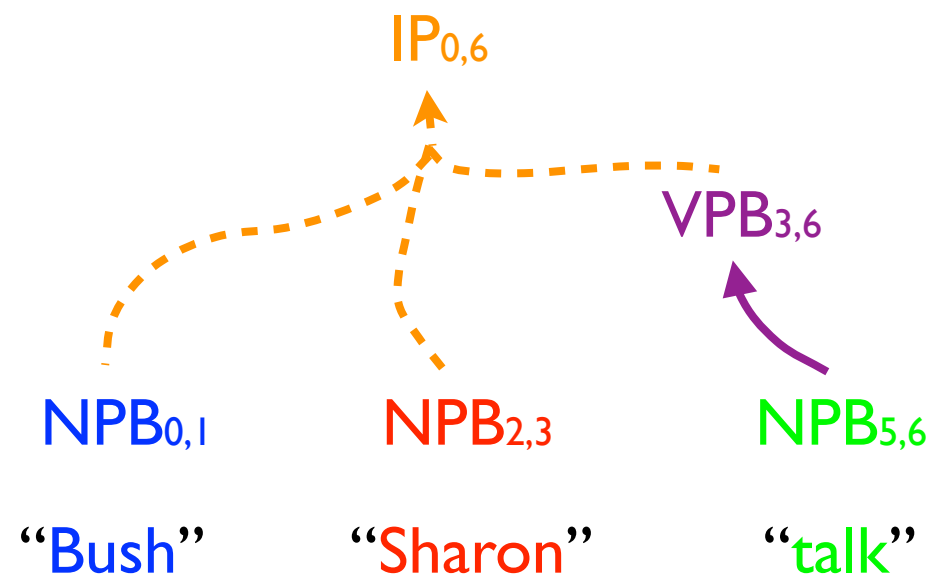
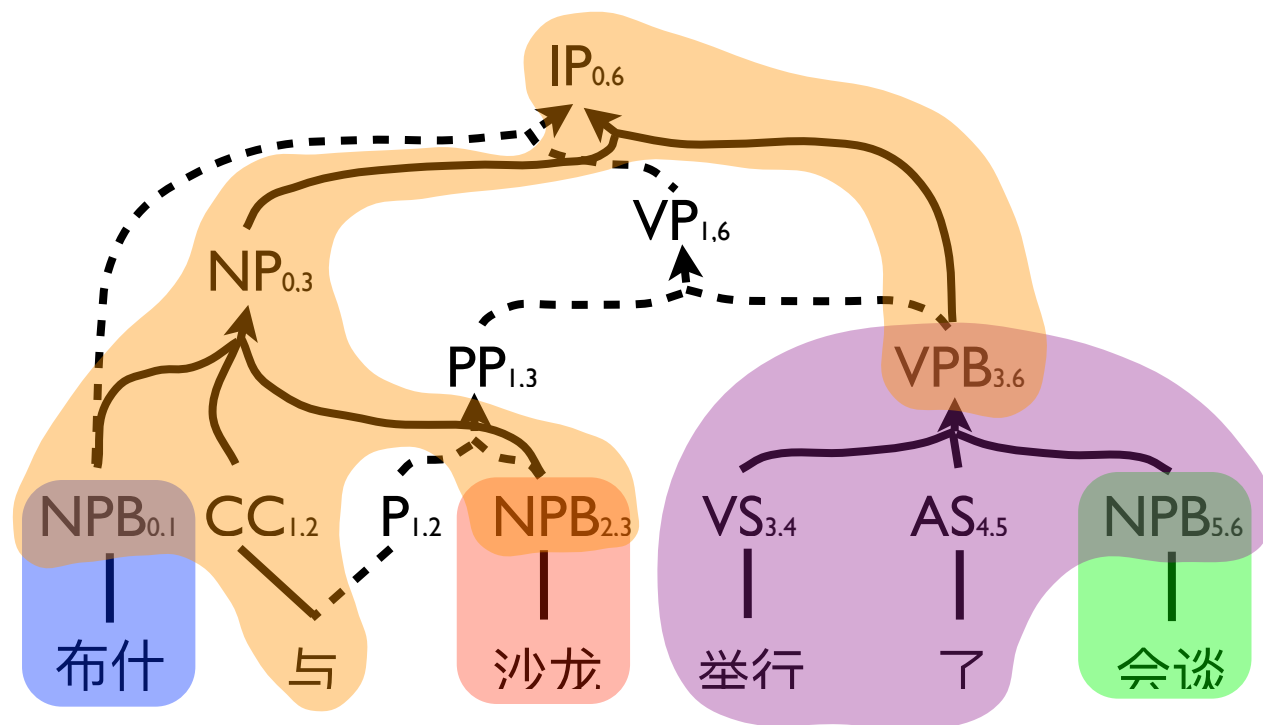
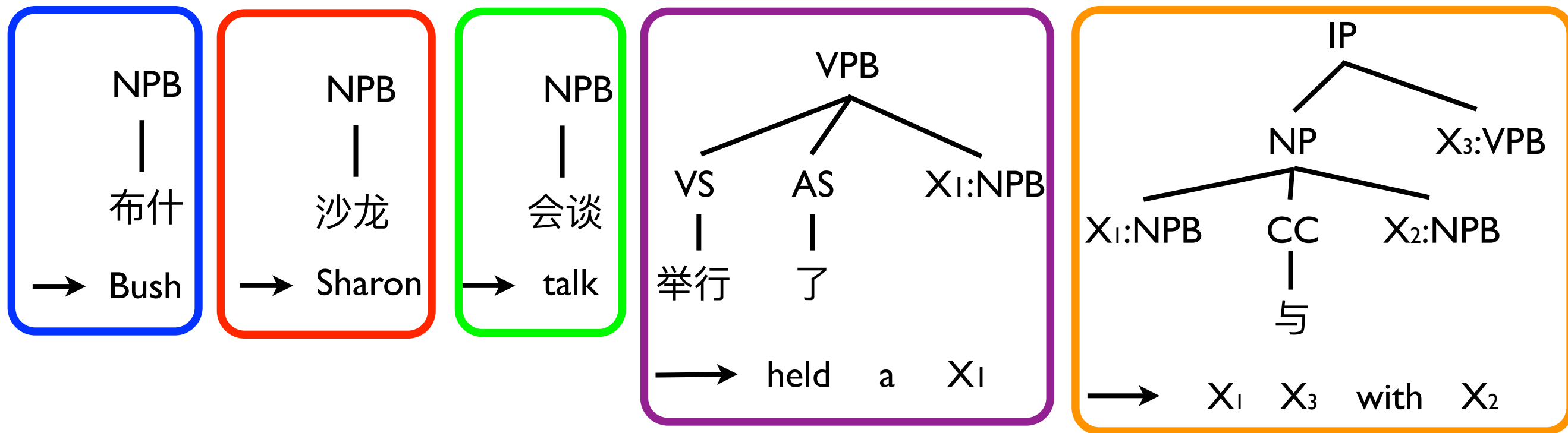


# Forest-based Decoding

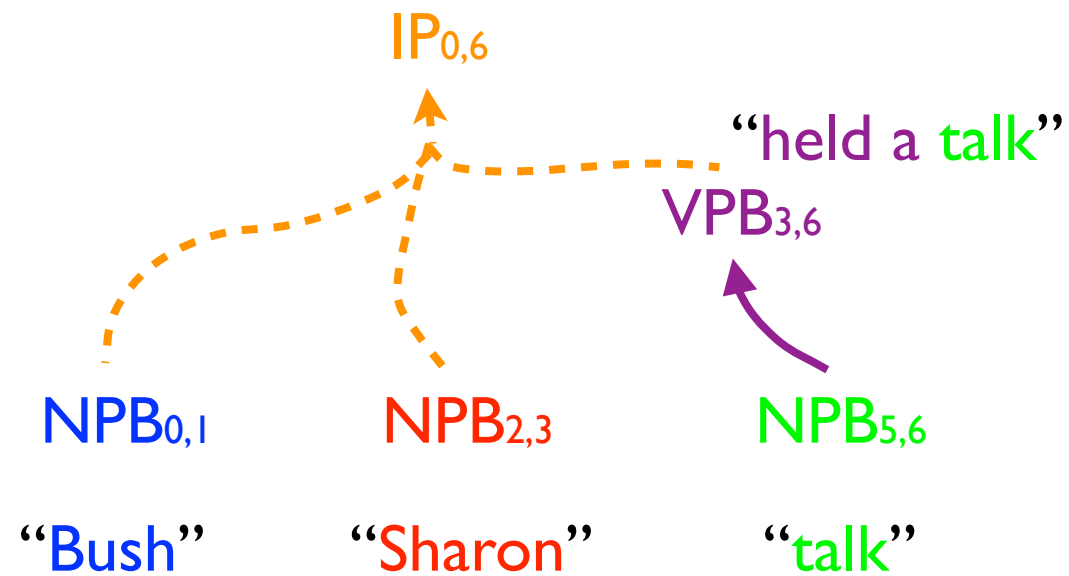
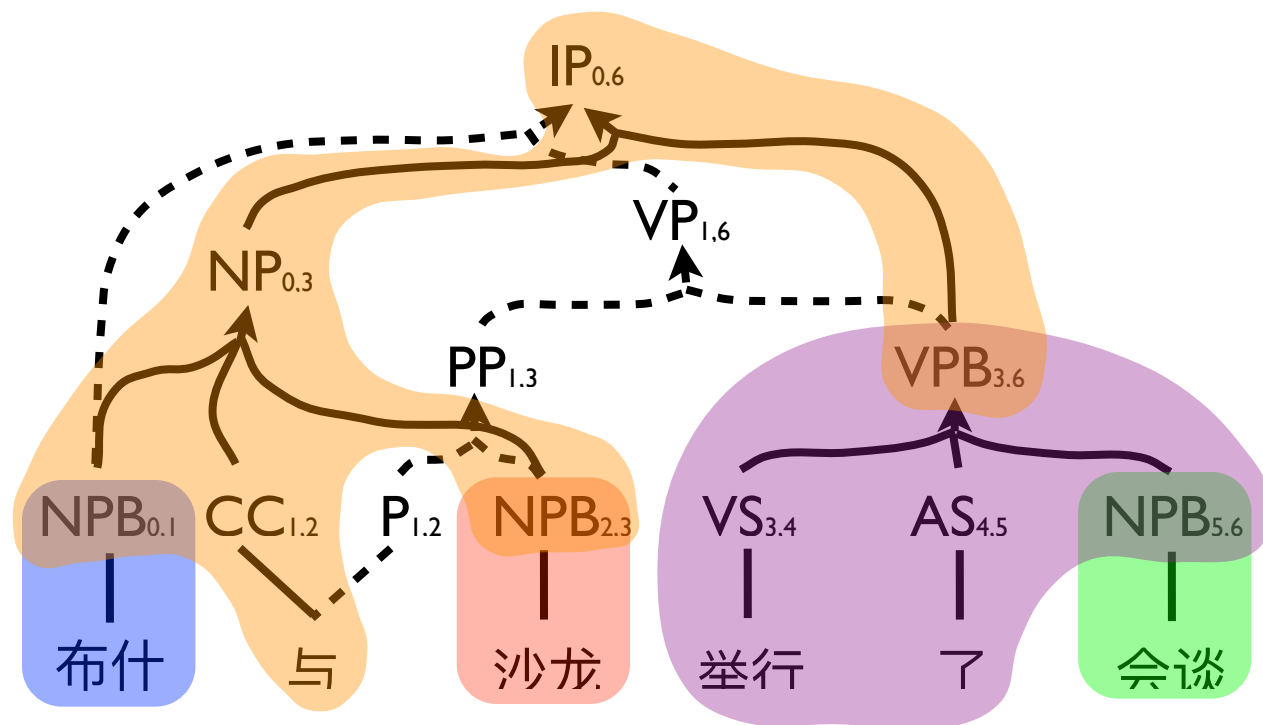
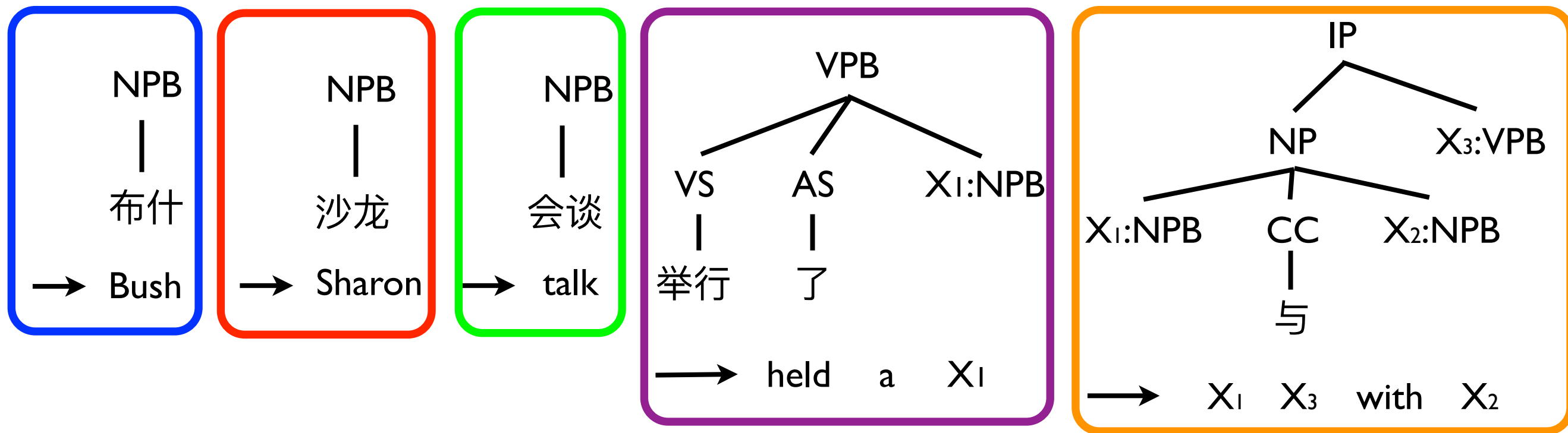




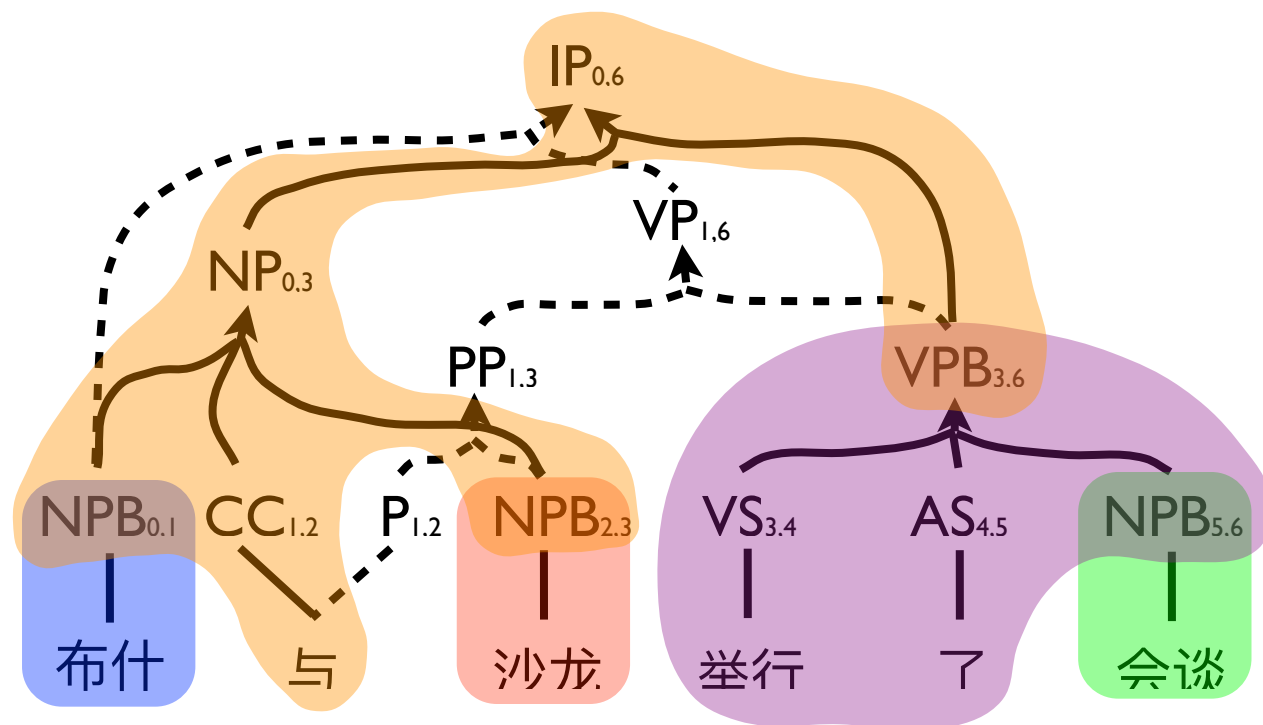
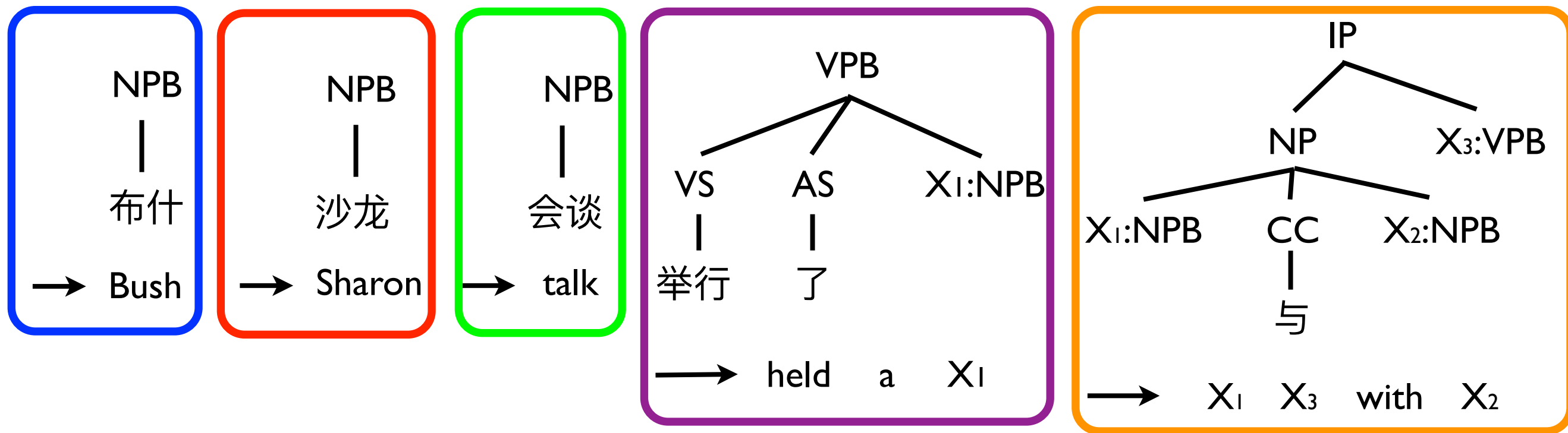
# Forest-based Decoding



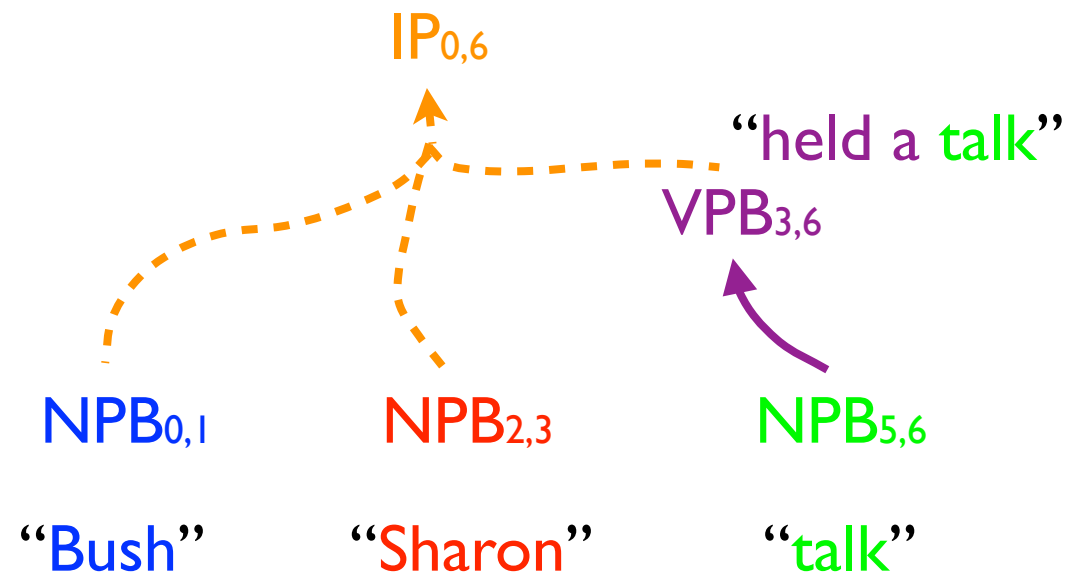
# Forest-based Decoding



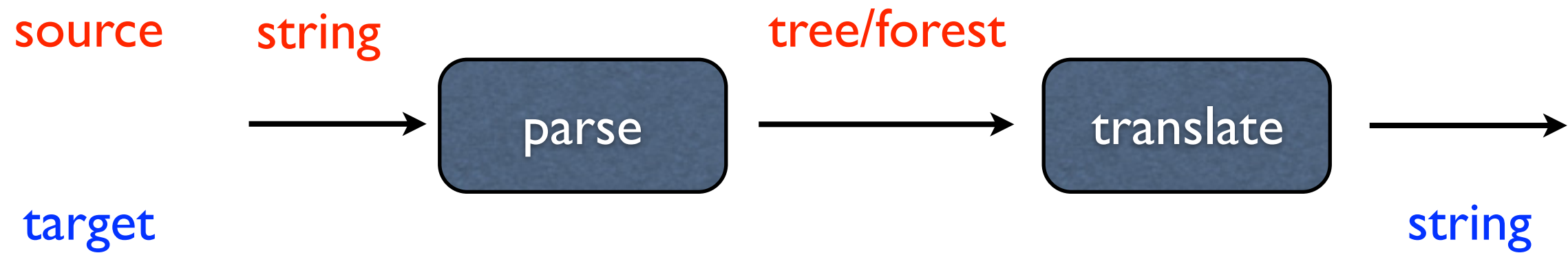
# Forest-based Decoding



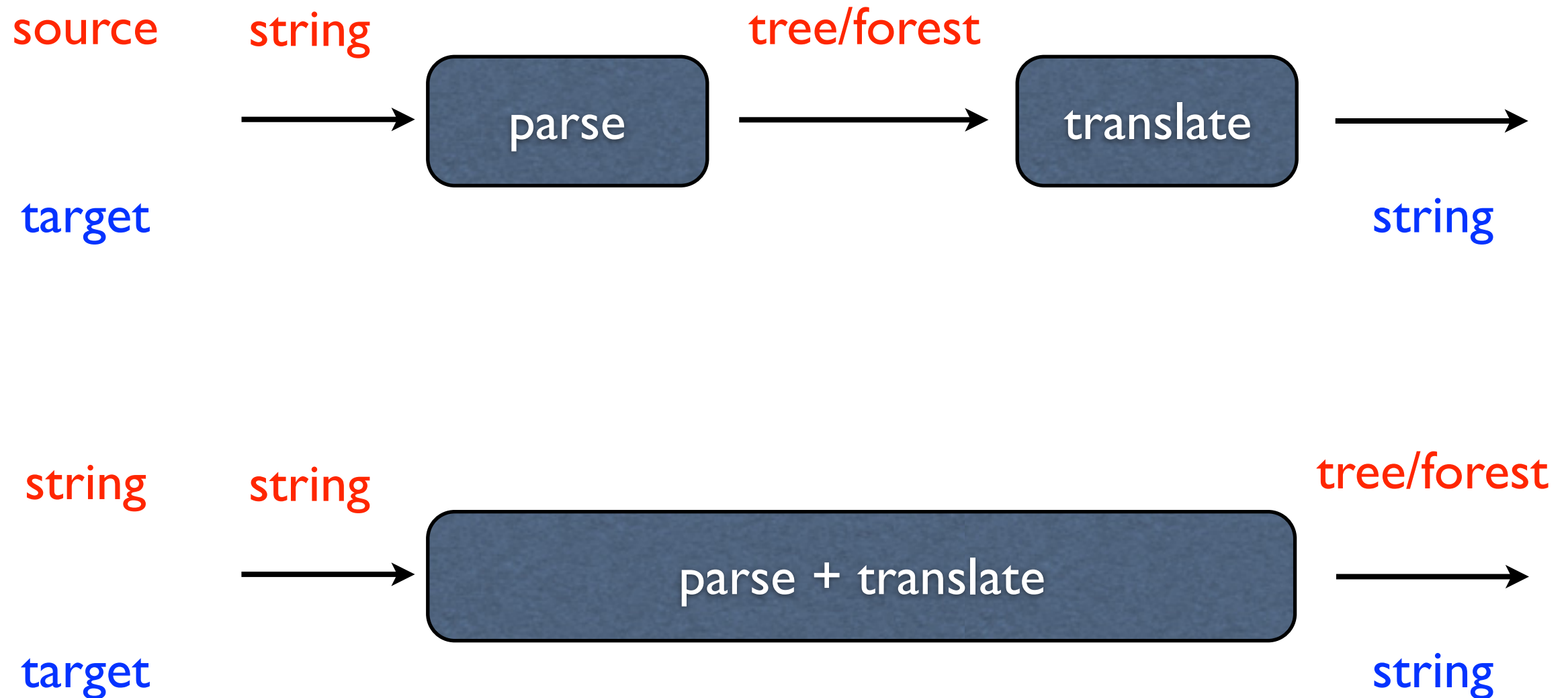
“Bush held a talk with Sharon”



# Forest-based Decoding



# Forest-based Decoding



# Joint Parsing and Translation

布什 与 沙龙 举行 了 会谈

(Liu and Liu, 2010)

# Joint Parsing and Translation

NPB  
|  
布什  
→ Bush

布什 与 沙龙 举行 了 会谈

(Liu and Liu, 2010)

# Joint Parsing and Translation

NPB  
|  
布什  
→ Bush

NPB  
|  
布什      与      沙龙      举行      了      会谈

(Liu and Liu, 2010)



# Joint Parsing and Translation

NPB  
|  
布什  
→ Bush

NPB  
|  
布什    与    沙龙    举行    了    会谈

Bush

(Liu and Liu, 2010)

# Joint Parsing and Translation

NPB

|

布什

与

沙龙

举行

了

会谈

Bush

(Liu and Liu, 2010)

# Joint Parsing and Translation

NPB  
|  
沙龙  
→ Sharon

NPB  
|  
布什      与      沙龙      举行      了      会谈

Bush

(Liu and Liu, 2010)

# Joint Parsing and Translation

NPB  
|  
沙龙  
→ Sharon

NPB                      NPB  
|                            |  
布什                      沙龙                      与                      举行                      了                      会谈

Bush

(Liu and Liu, 2010)

# Joint Parsing and Translation

NPB  
|  
沙龙  
→ Sharon

NPB                  NPB  
|                      |  
布什                  沙龙                  与                  举行                  了                  会谈

Bush

Sharon

(Liu and Liu, 2010)

# Joint Parsing and Translation

NPB                      NPB  
|                            |  
布什                      沙龙                      与                      举行                      了                      会谈

Bush

Sharon

(Liu and Liu, 2010)

# Joint Parsing and Translation

NPB  
|  
会谈  
→ talk

NPB  
|  
布什

与

NPB  
|  
沙龙

举行

了

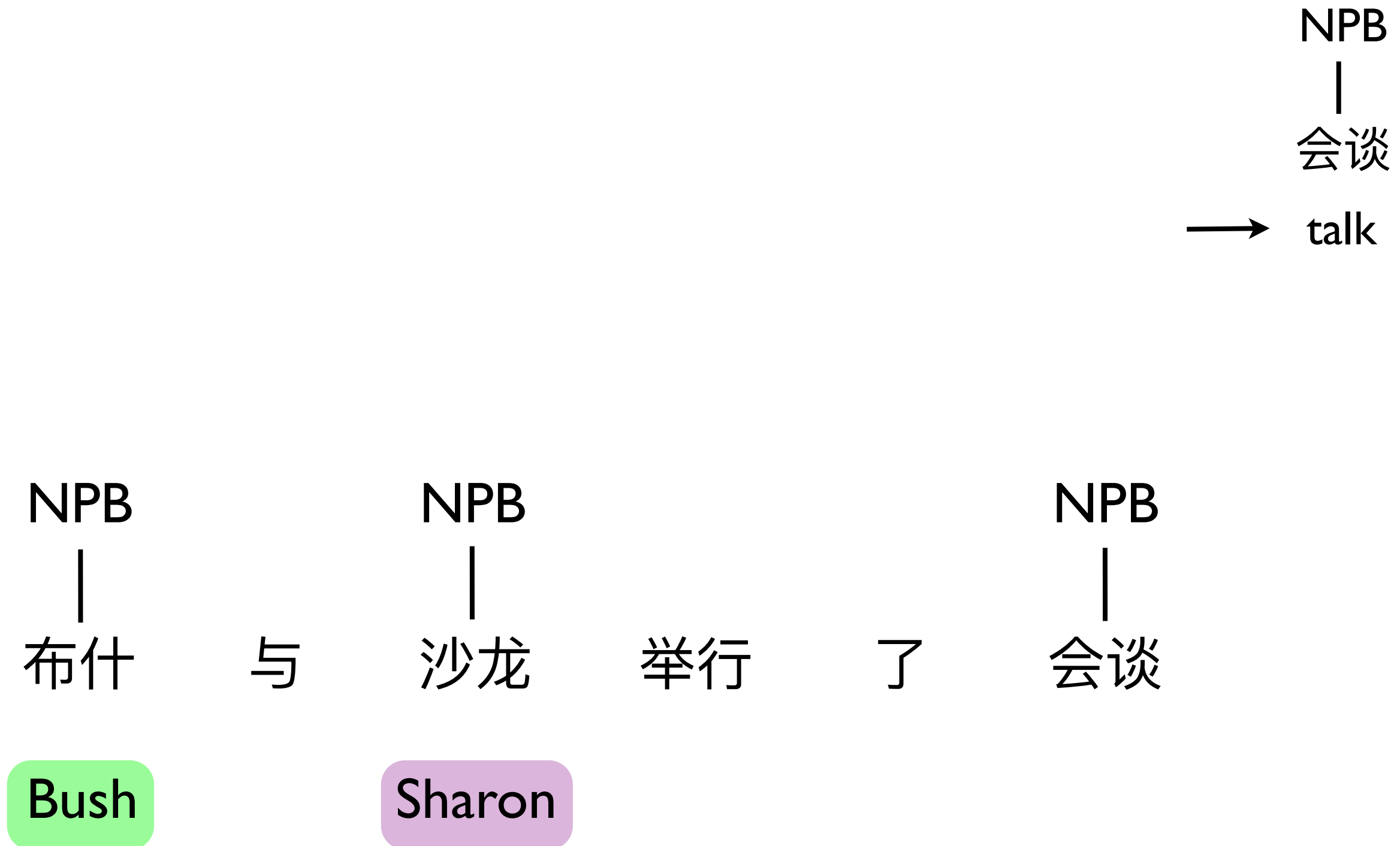
会谈

Bush

Sharon

(Liu and Liu, 2010)

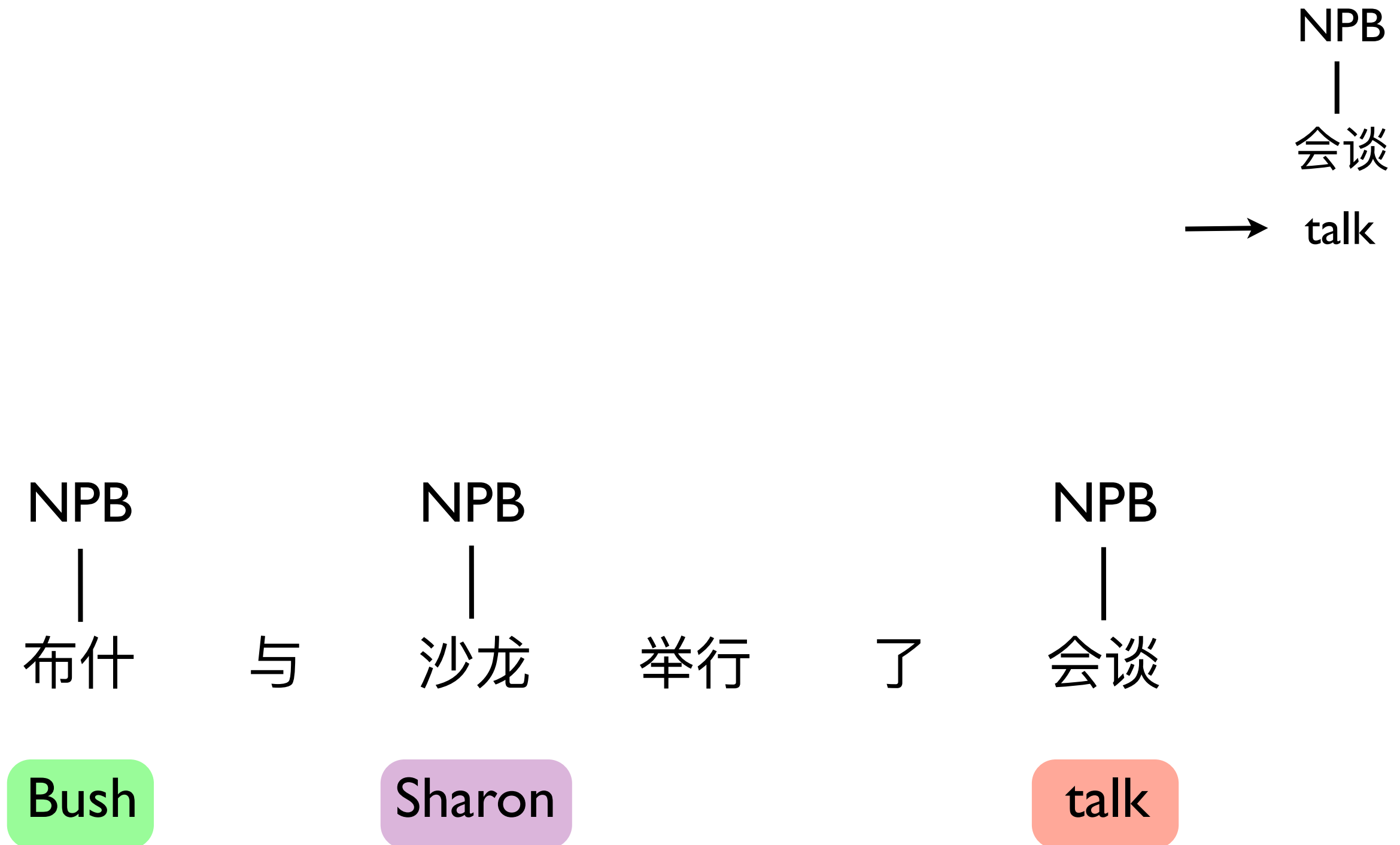
# Joint Parsing and Translation



(Liu and Liu, 2010)

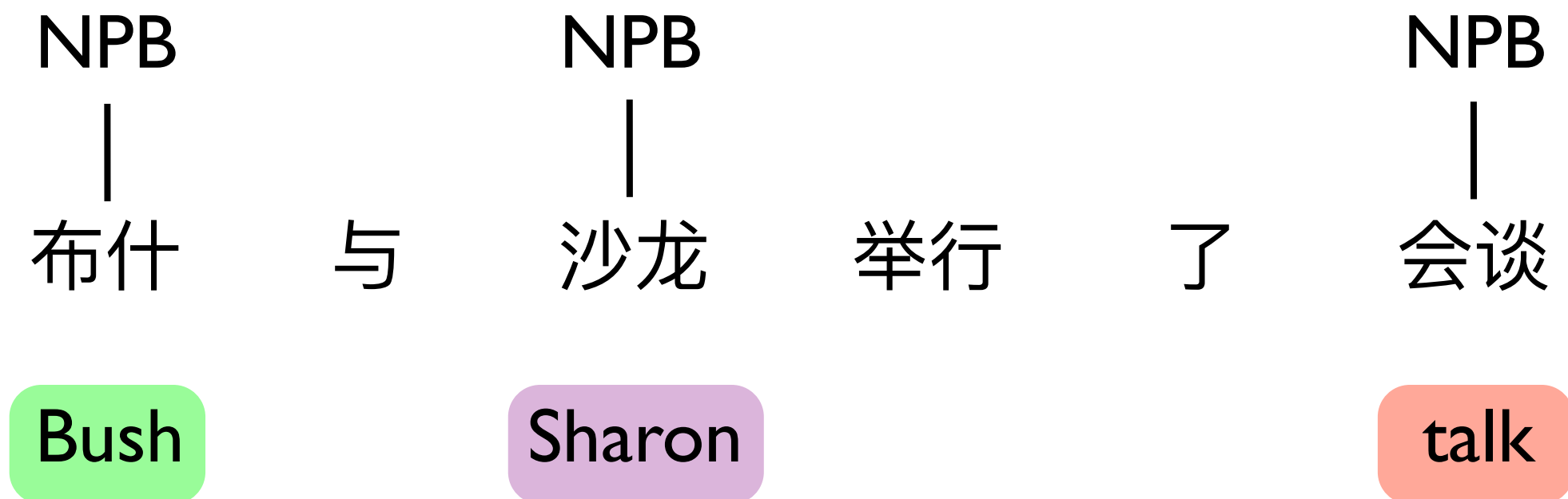


# Joint Parsing and Translation



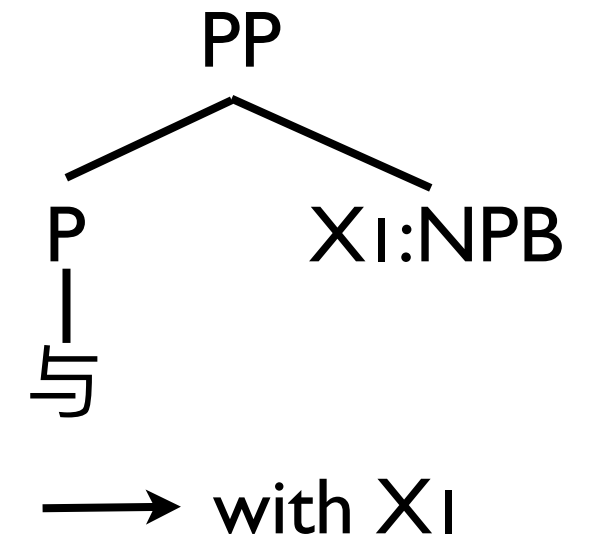
(Liu and Liu, 2010)

# Joint Parsing and Translation



(Liu and Liu, 2010)

# Joint Parsing and Translation



NPB  
|  
布什

Bush

与

NPB  
|  
沙龙

Sharon

举行

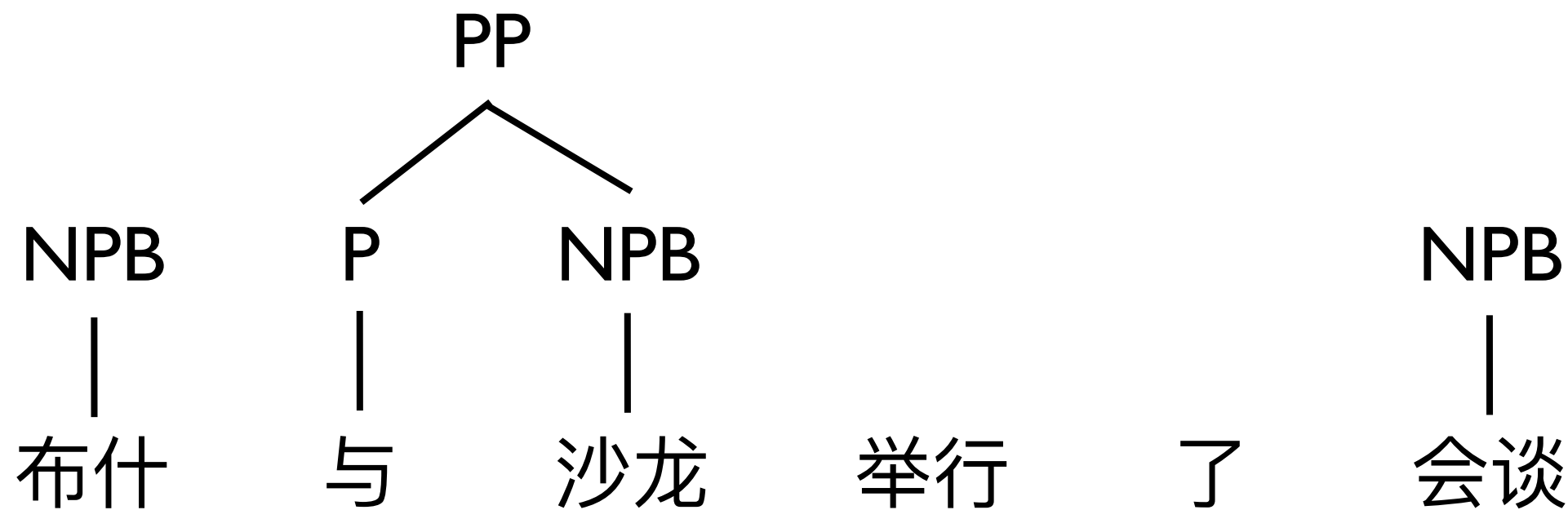
了

NPB  
|  
会谈

talk

(Liu and Liu, 2010)

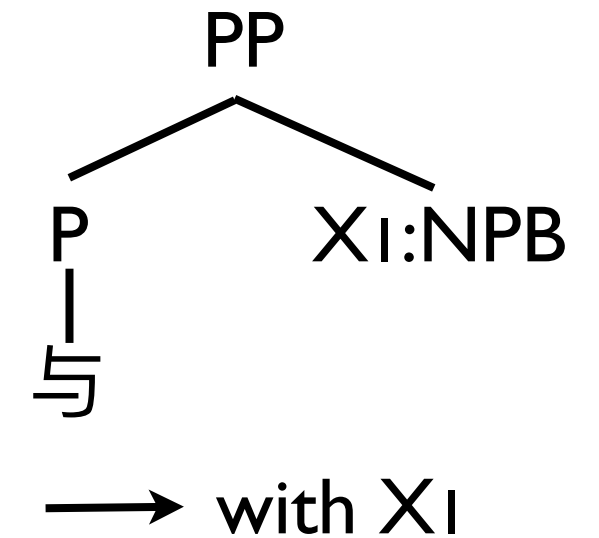
# Joint Parsing and Translation



Bush

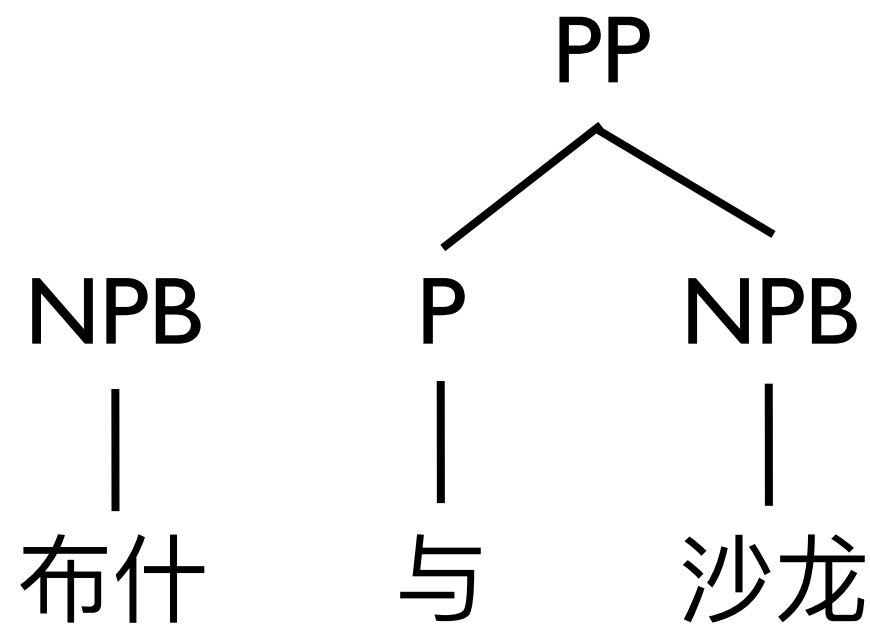
Sharon

talk



(Liu and Liu, 2010)

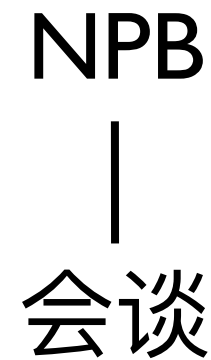
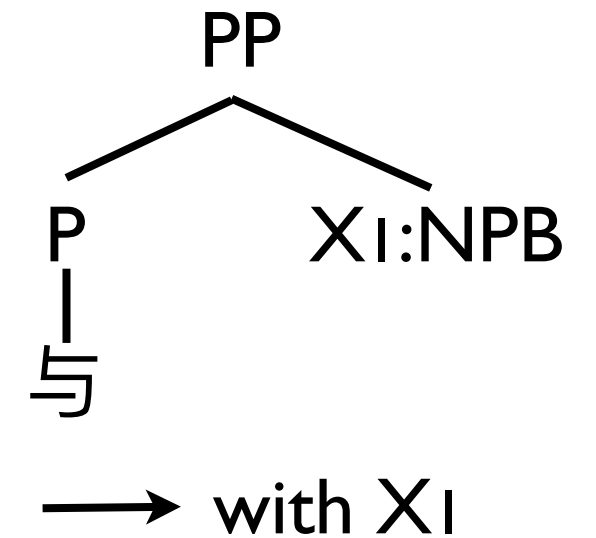
# Joint Parsing and Translation



Bush

with

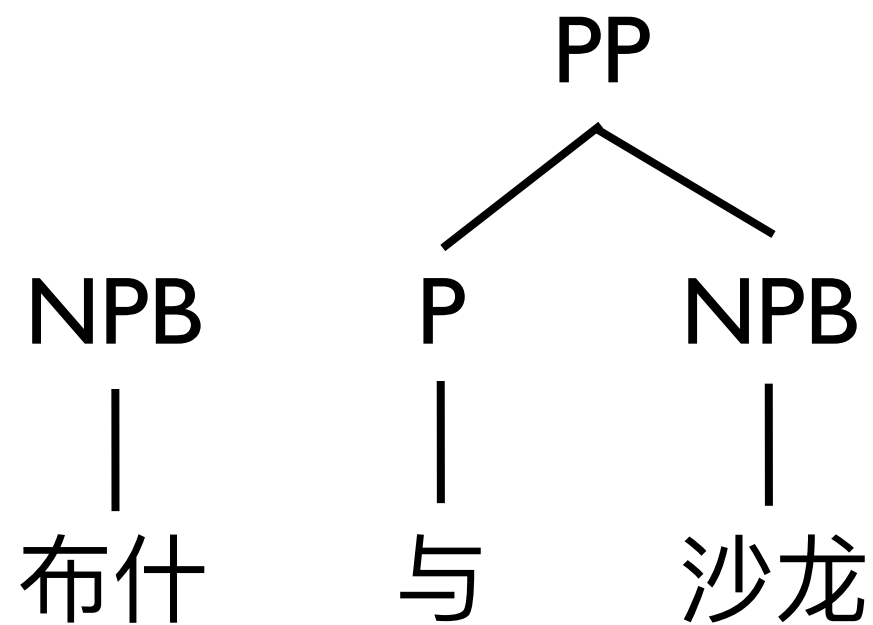
Sharon



talk

(Liu and Liu, 2010)

# Joint Parsing and Translation

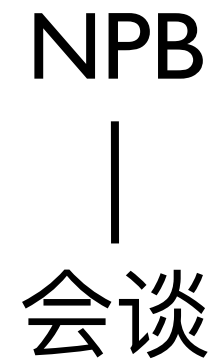


Bush

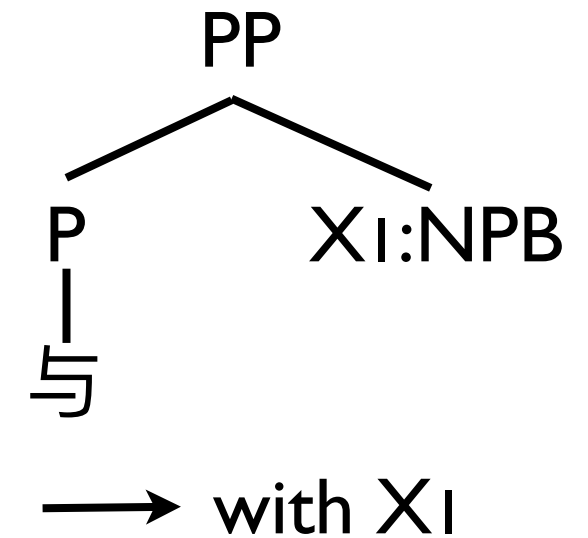
with Sharon

举行

了

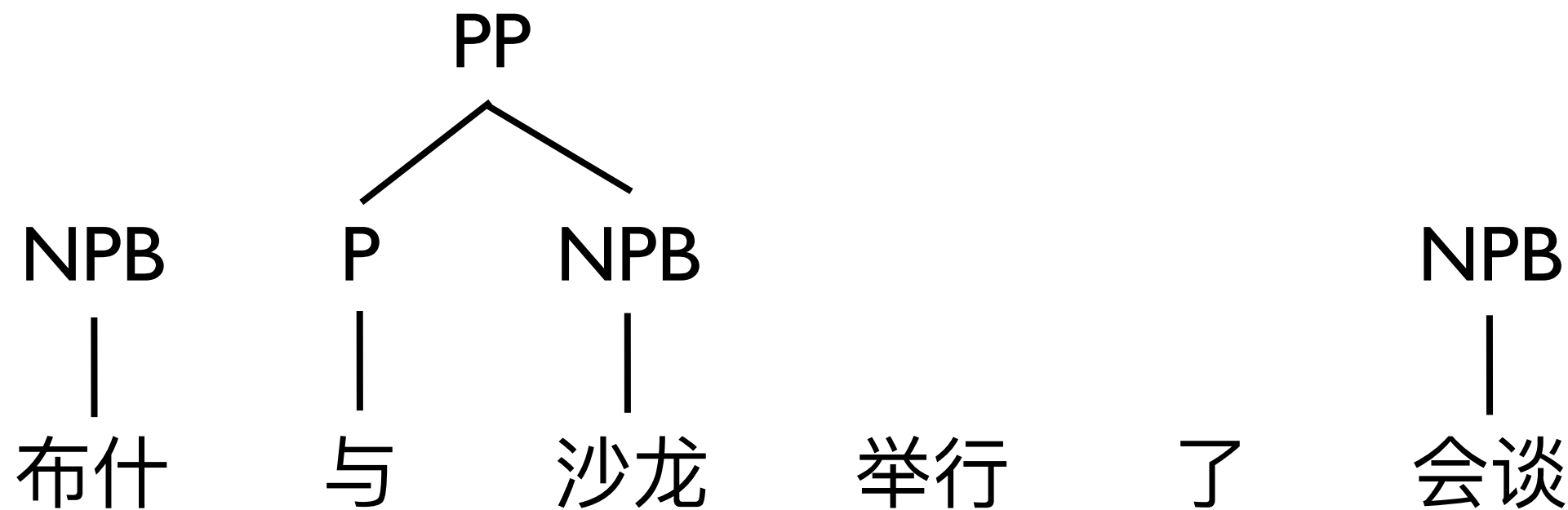


talk



(Liu and Liu, 2010)

# Joint Parsing and Translation



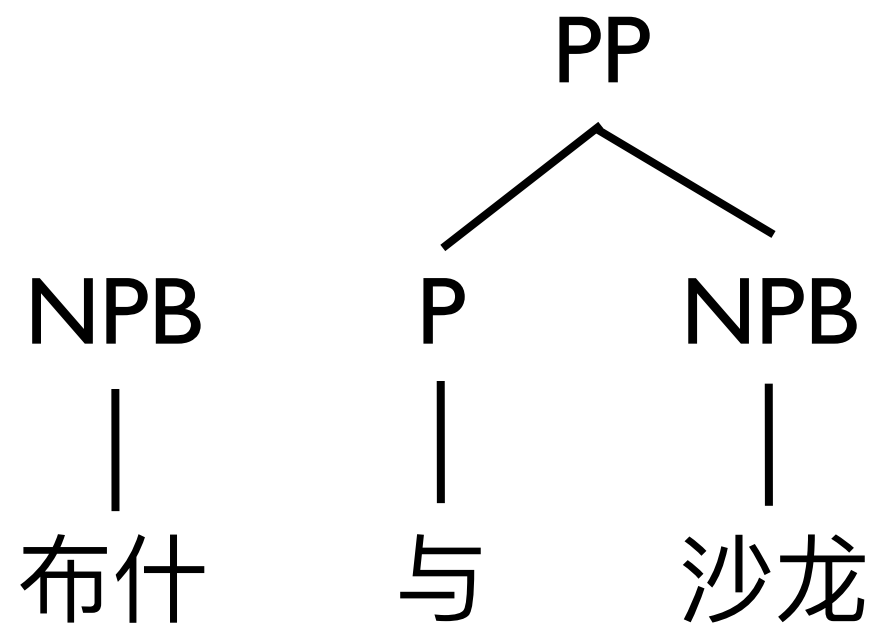
Bush

with Sharon

talk

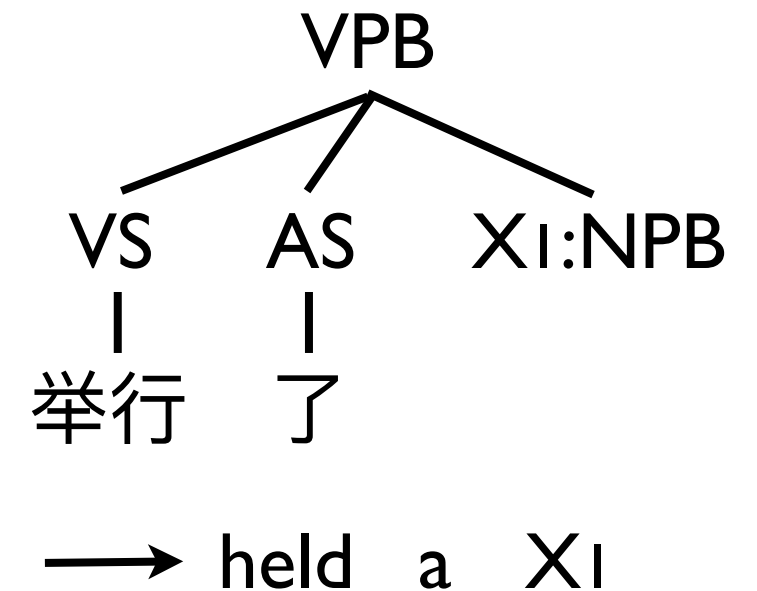
(Liu and Liu, 2010)

# Joint Parsing and Translation



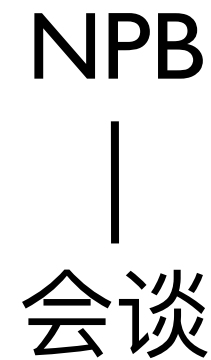
Bush

with Sharon



举行

了

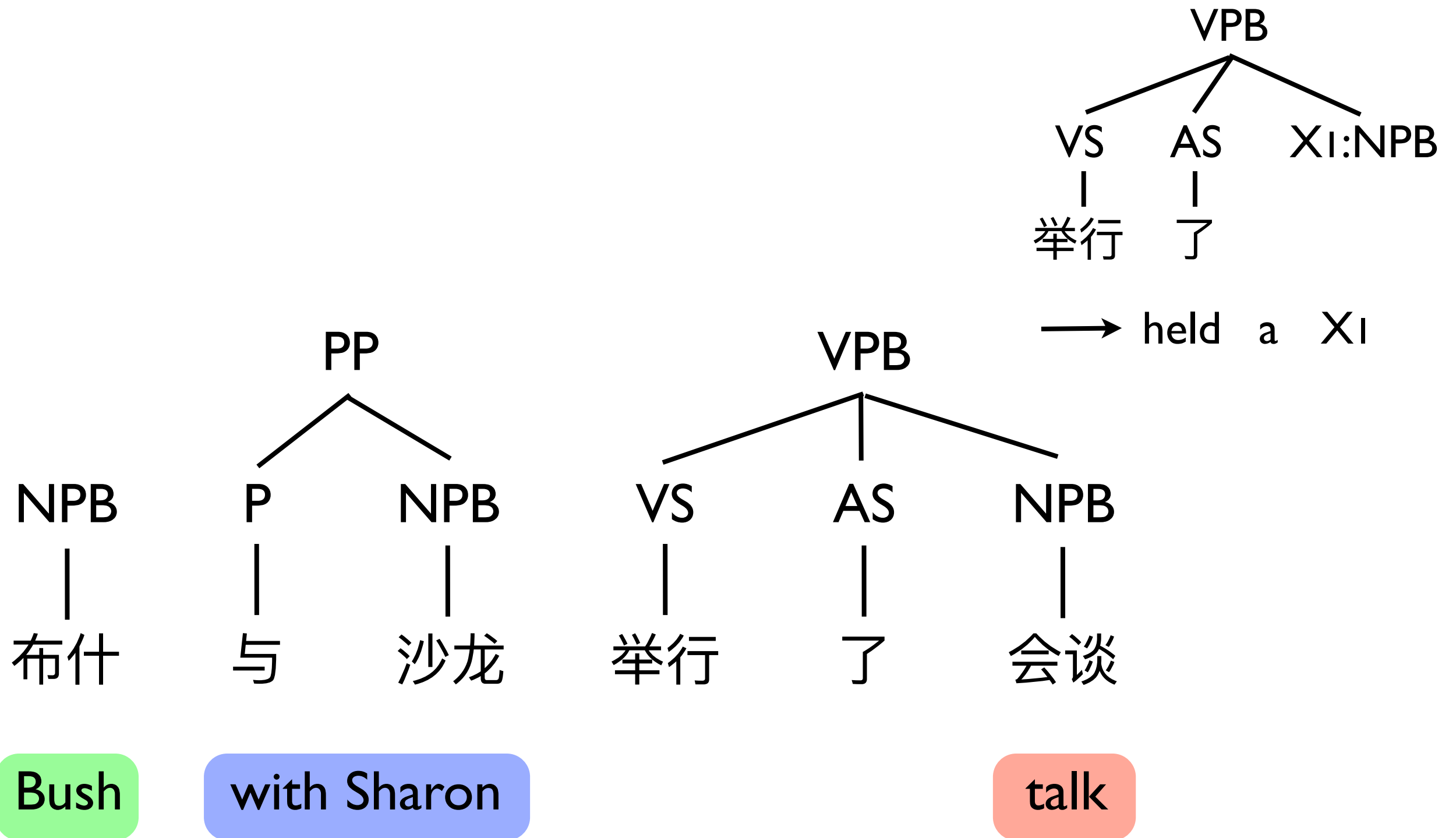


talk

(Liu and Liu, 2010)

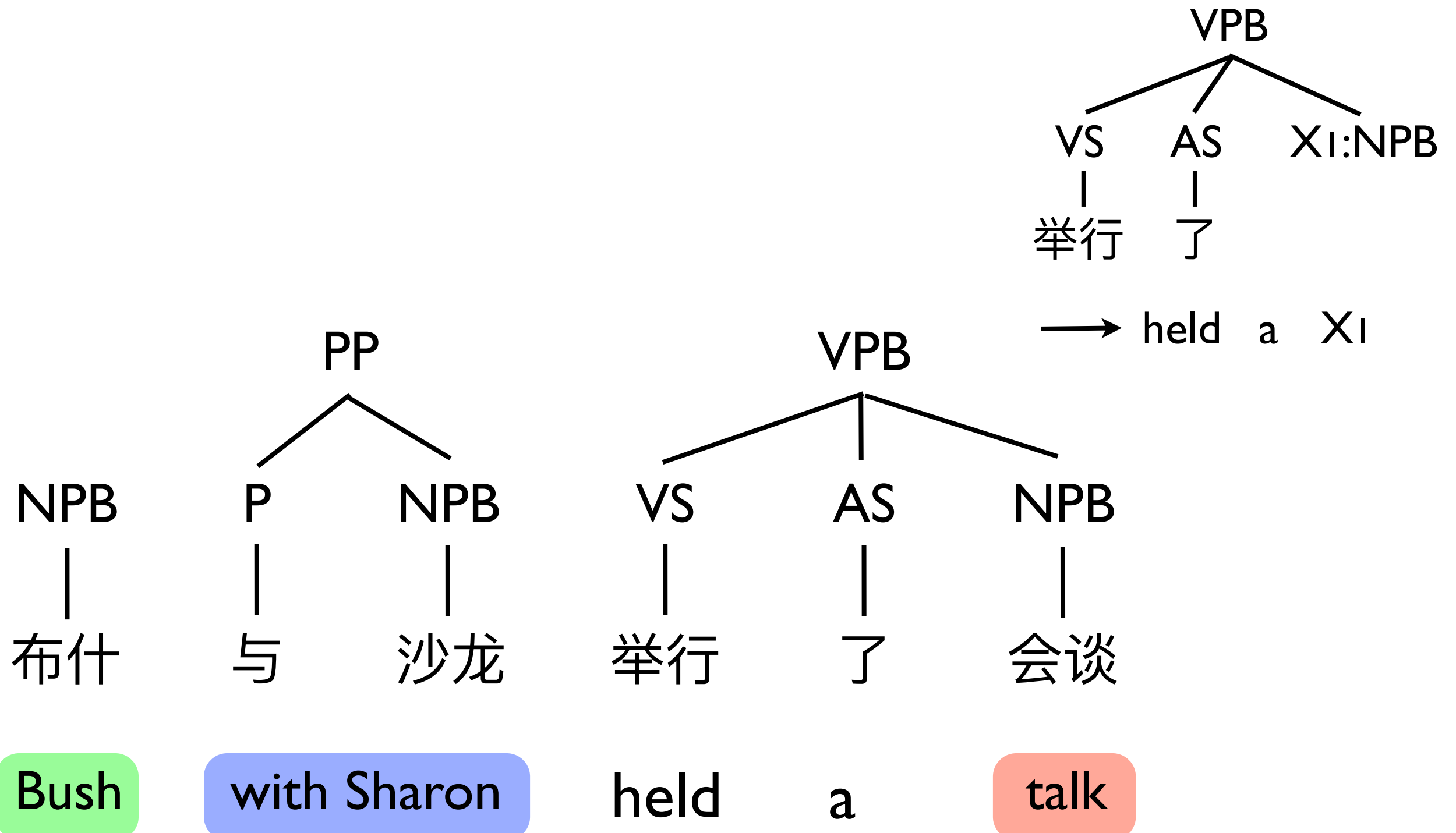


# Joint Parsing and Translation



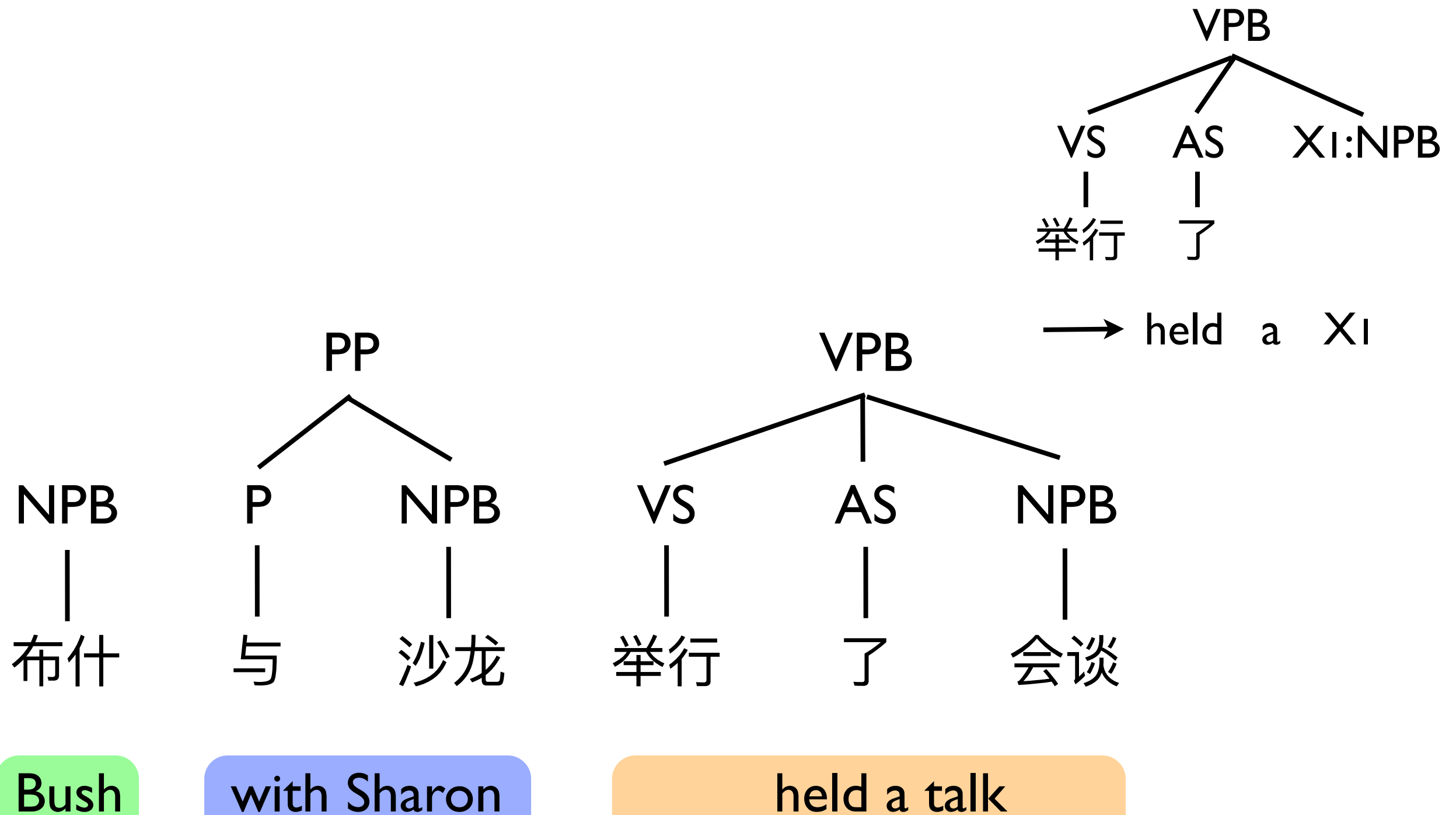
(Liu and Liu, 2010)

# Joint Parsing and Translation



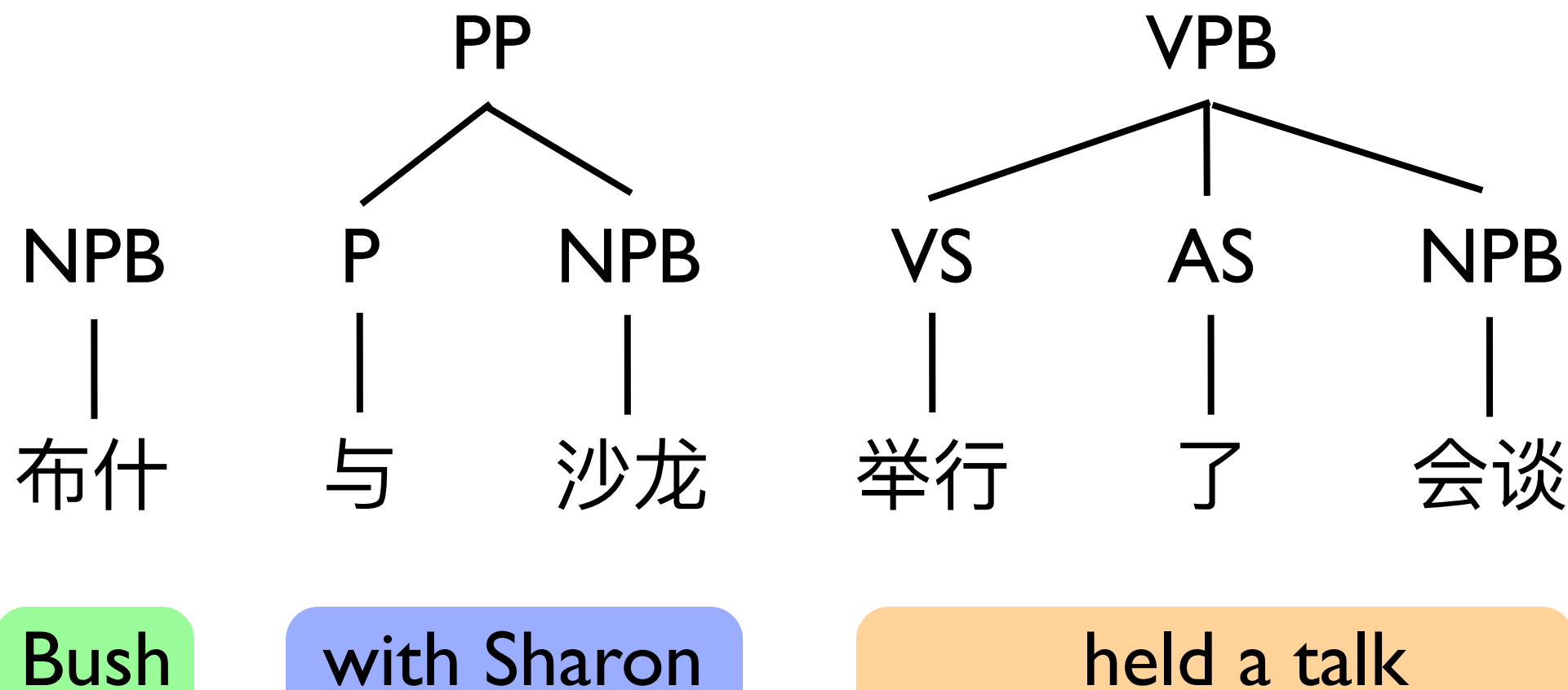
(Liu and Liu, 2010)

# Joint Parsing and Translation



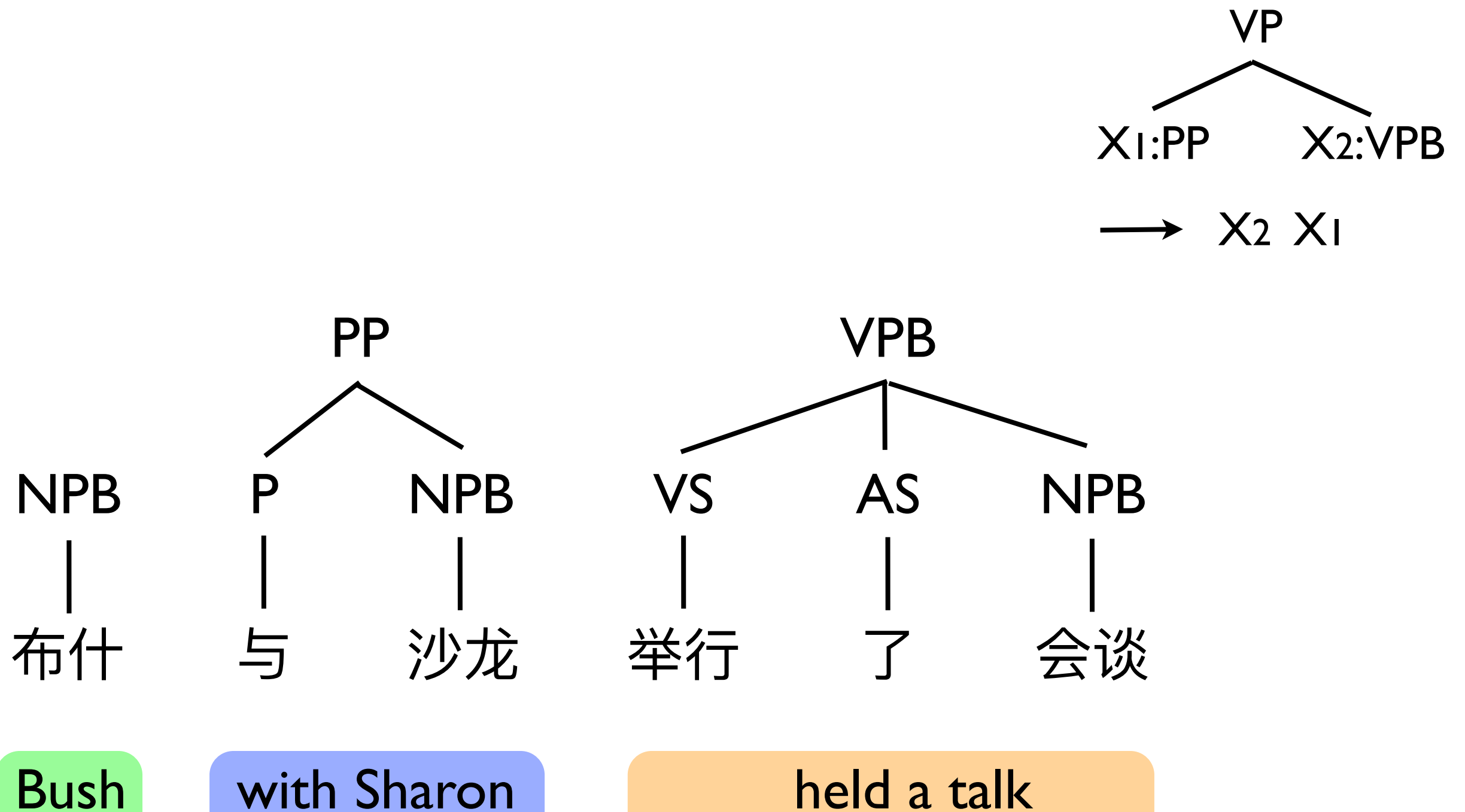
(Liu and Liu, 2010)

# Joint Parsing and Translation



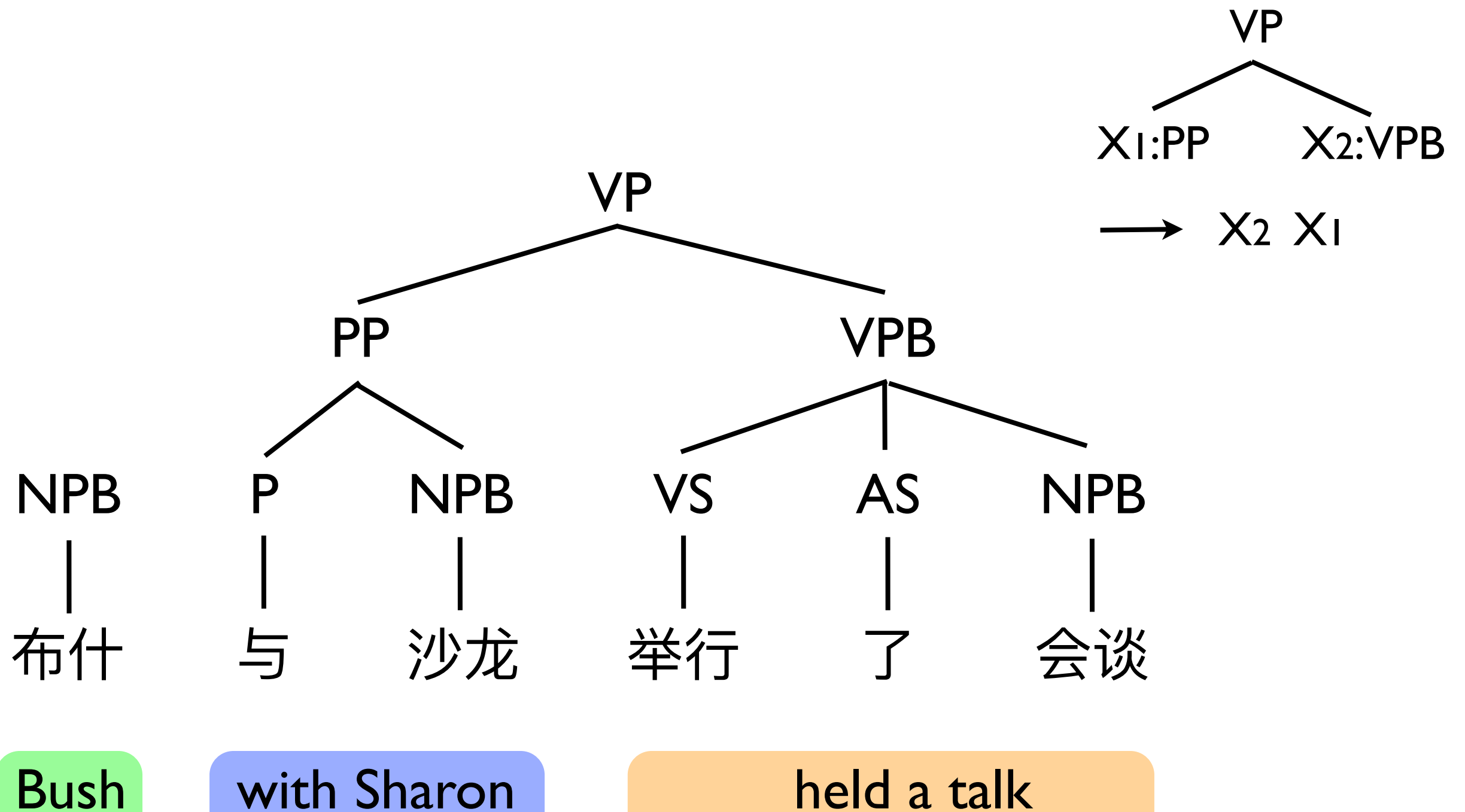
(Liu and Liu, 2010)

# Joint Parsing and Translation



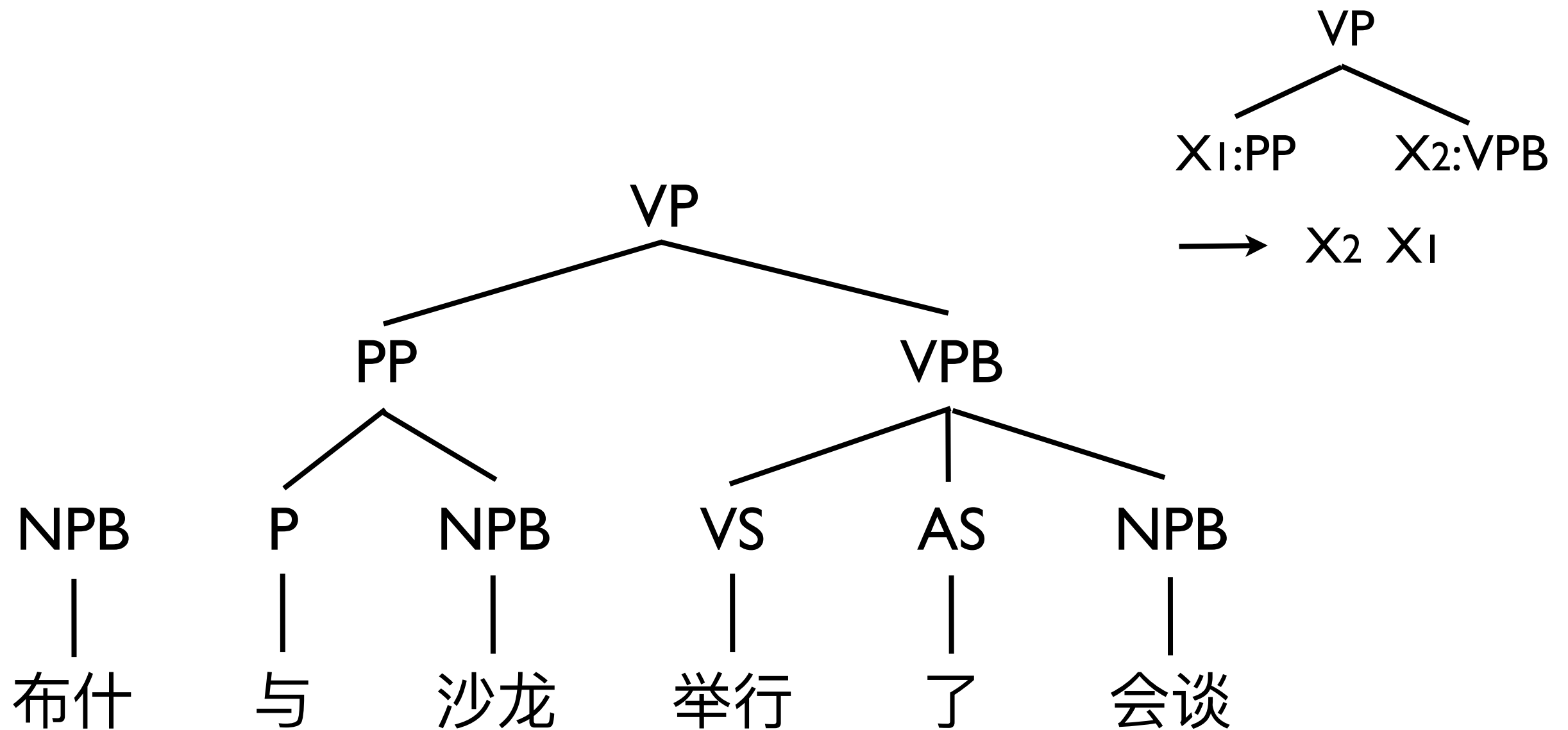
(Liu and Liu, 2010)

# Joint Parsing and Translation



(Liu and Liu, 2010)

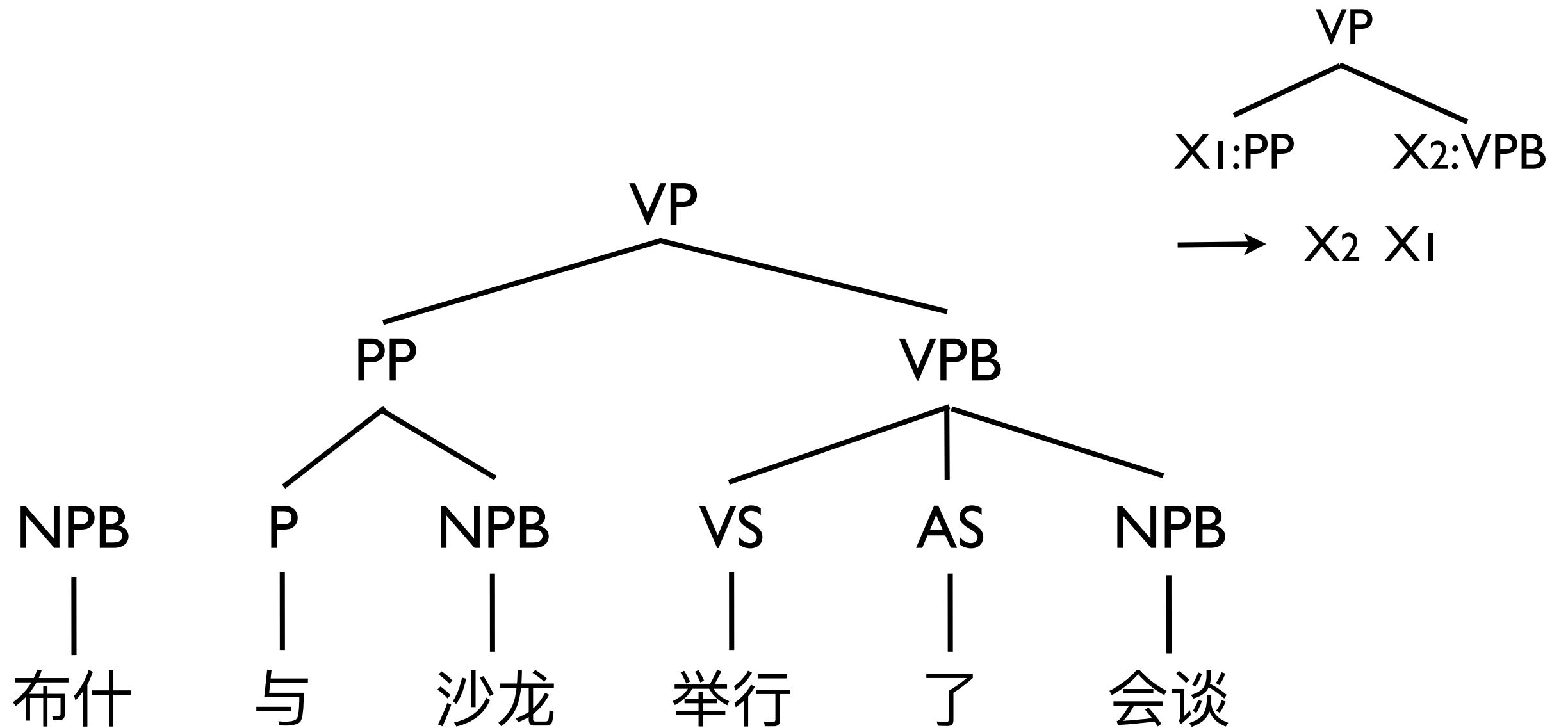
# Joint Parsing and Translation



Bush

(Liu and Liu, 2010)

# Joint Parsing and Translation



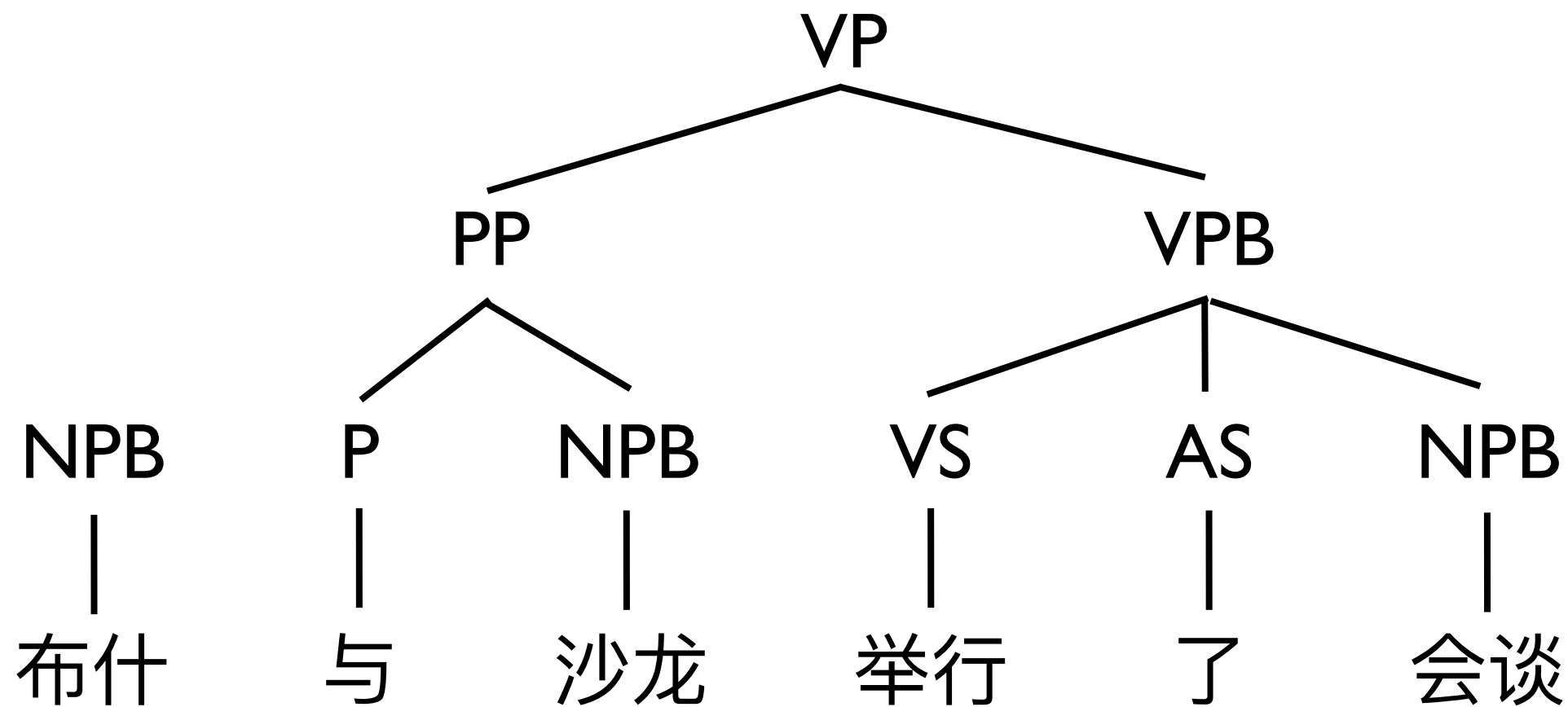
Bush

held a talk with Sharon

(Liu and Liu, 2010)



# Joint Parsing and Translation

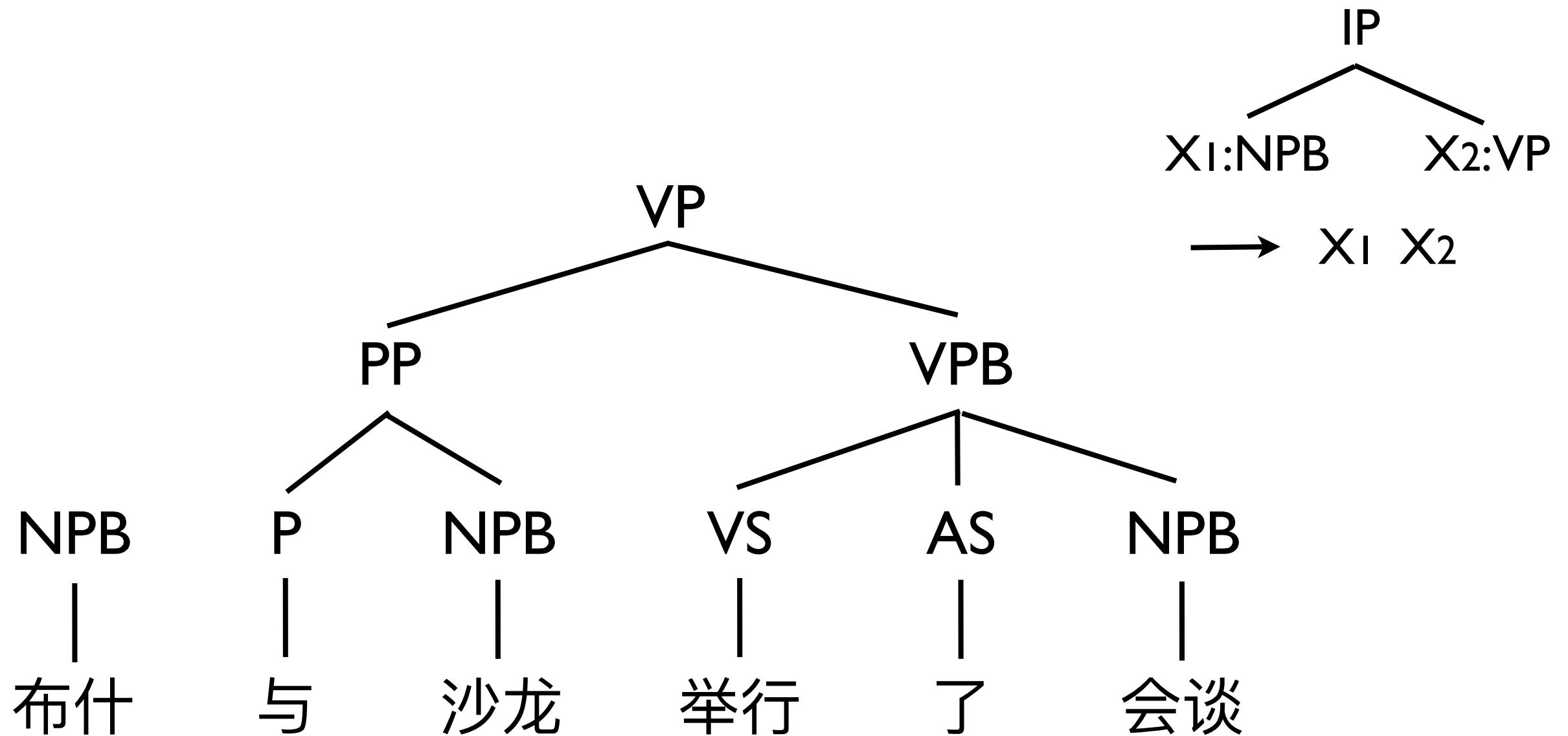


Bush

held a talk with Sharon

(Liu and Liu, 2010)

# Joint Parsing and Translation

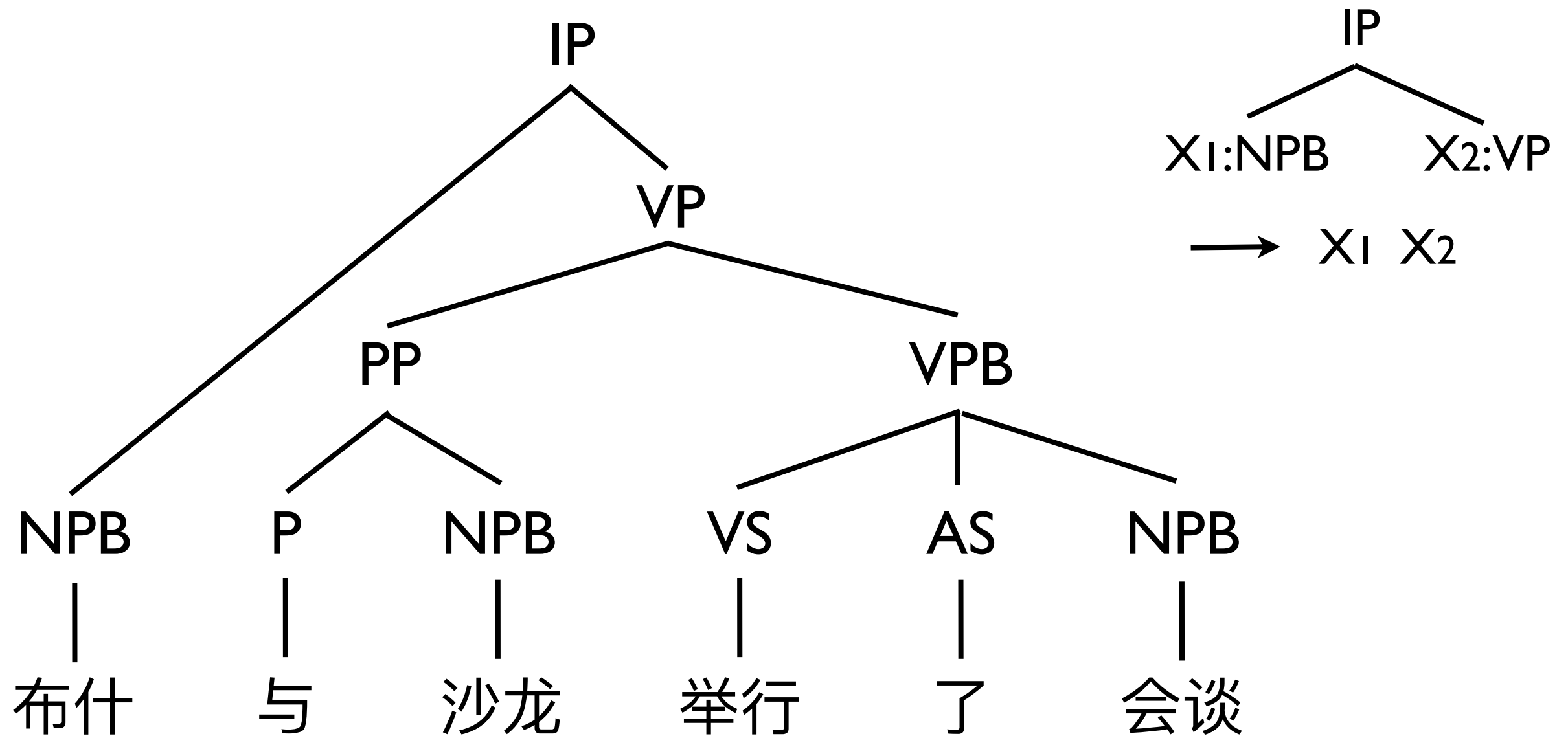


Bush

held a talk with Sharon

(Liu and Liu, 2010)

# Joint Parsing and Translation

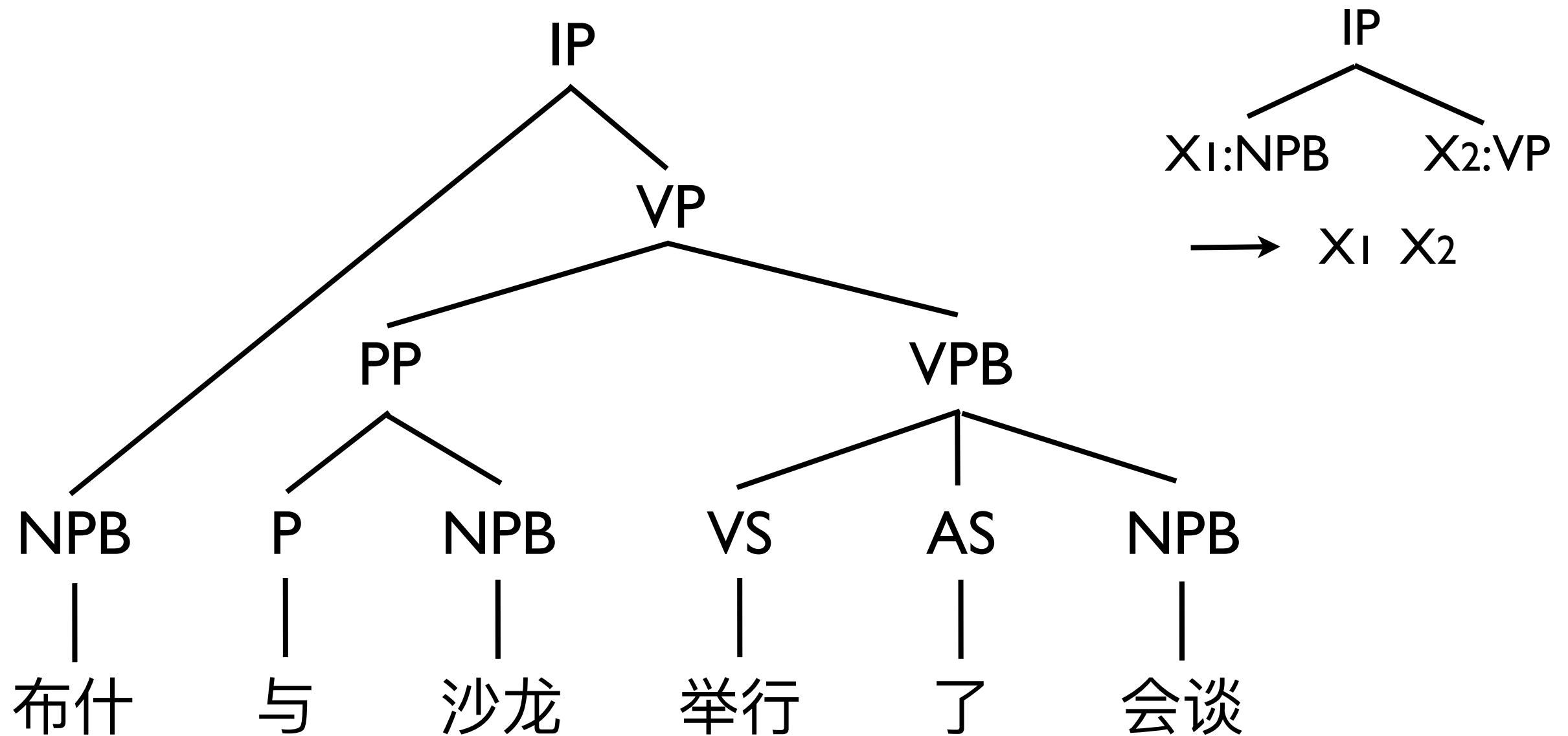


Bush

held a talk with Sharon

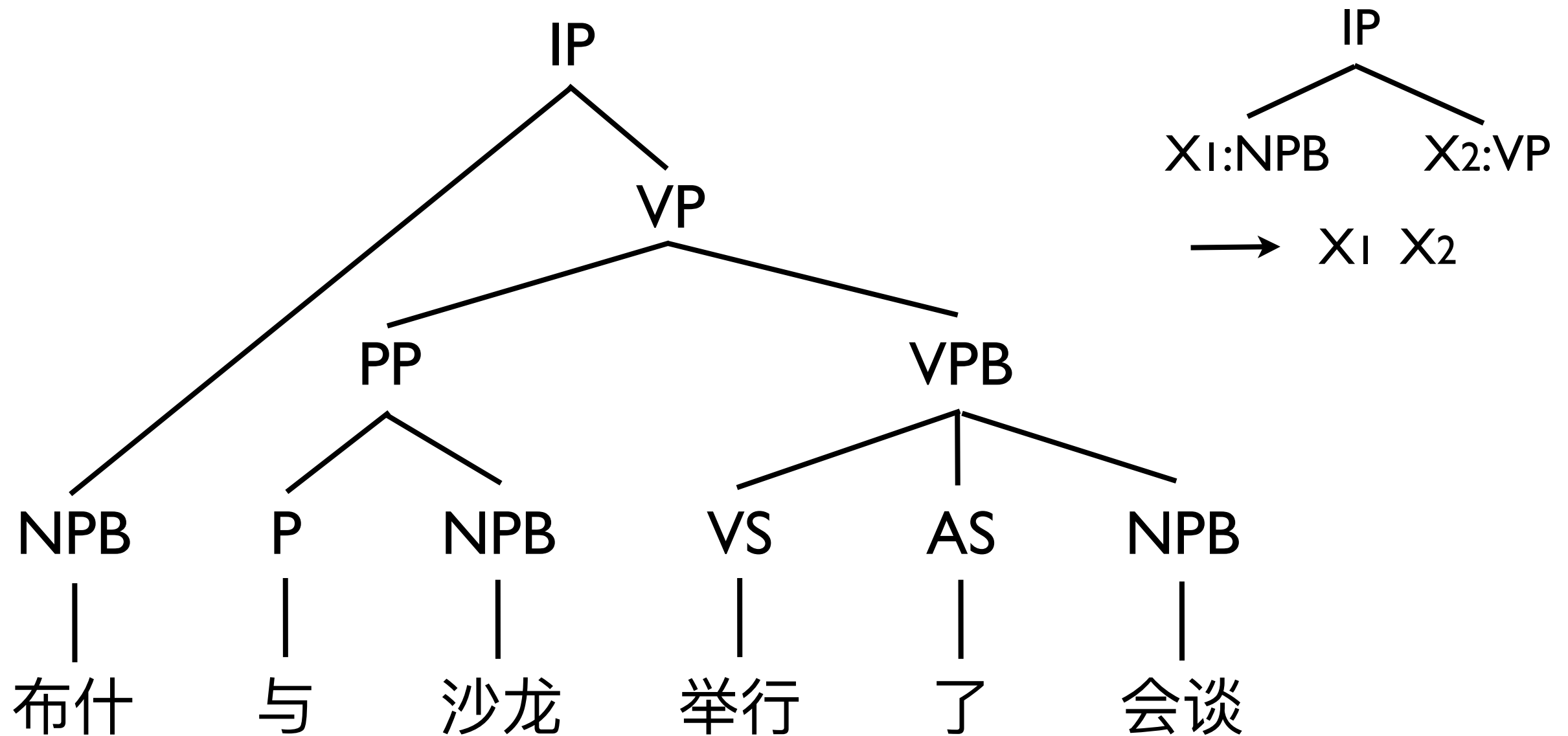
(Liu and Liu, 2010)

# Joint Parsing and Translation



(Liu and Liu, 2010)

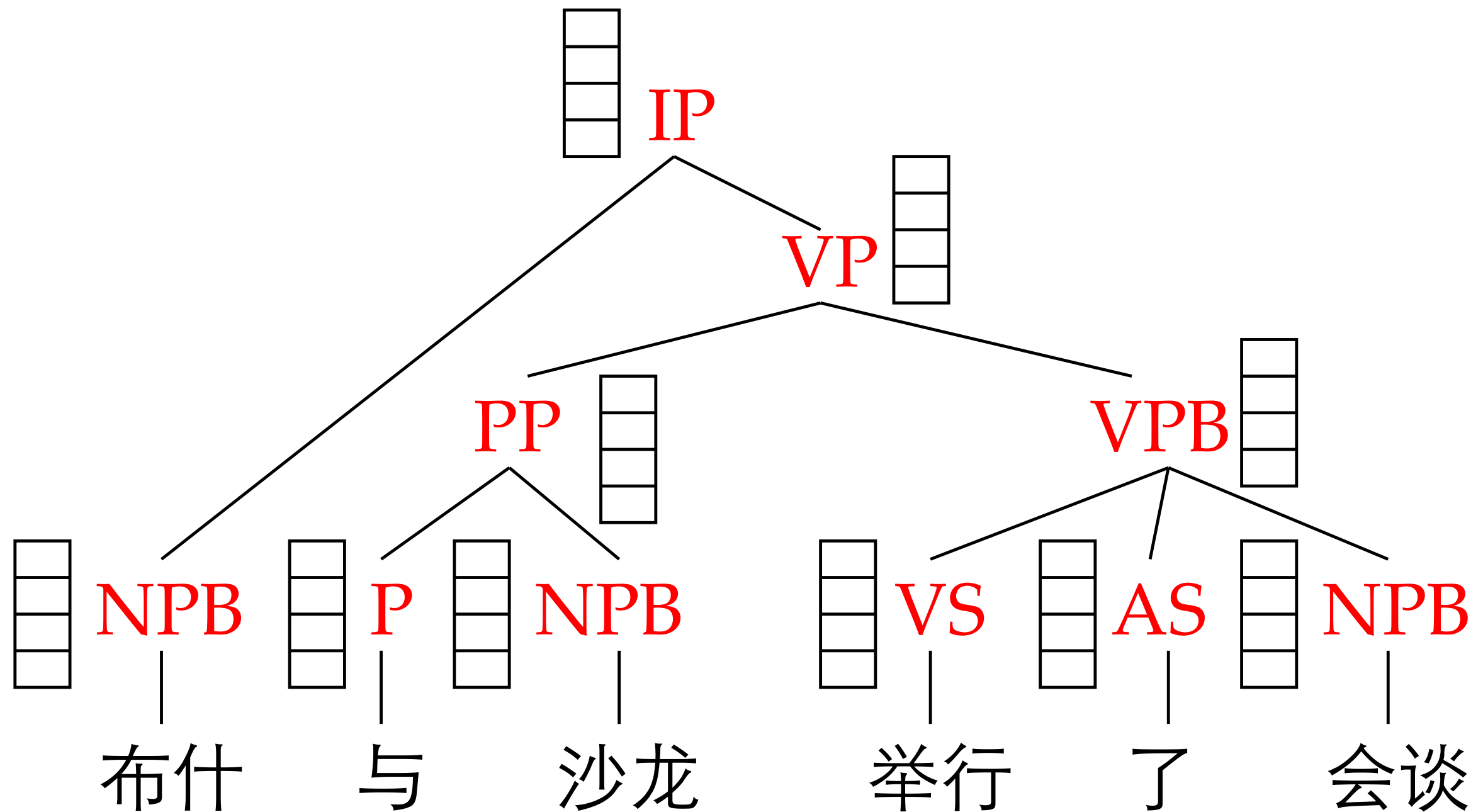
# Joint Parsing and Translation



Bush held a talk with Sharon

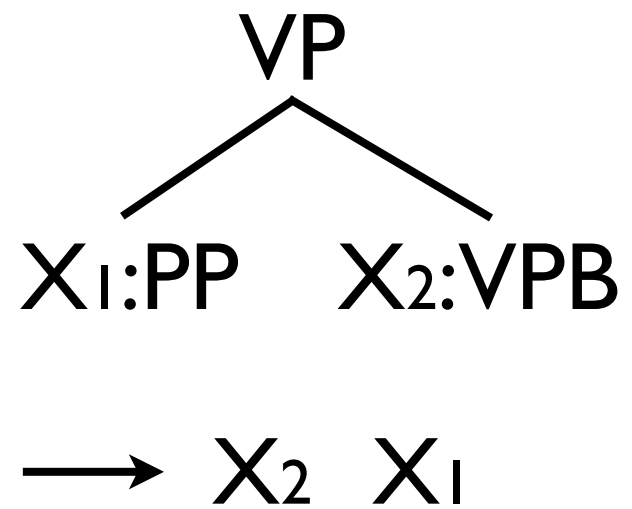
(Liu and Liu, 2010)

# Stacks in Tree-to-String Translation



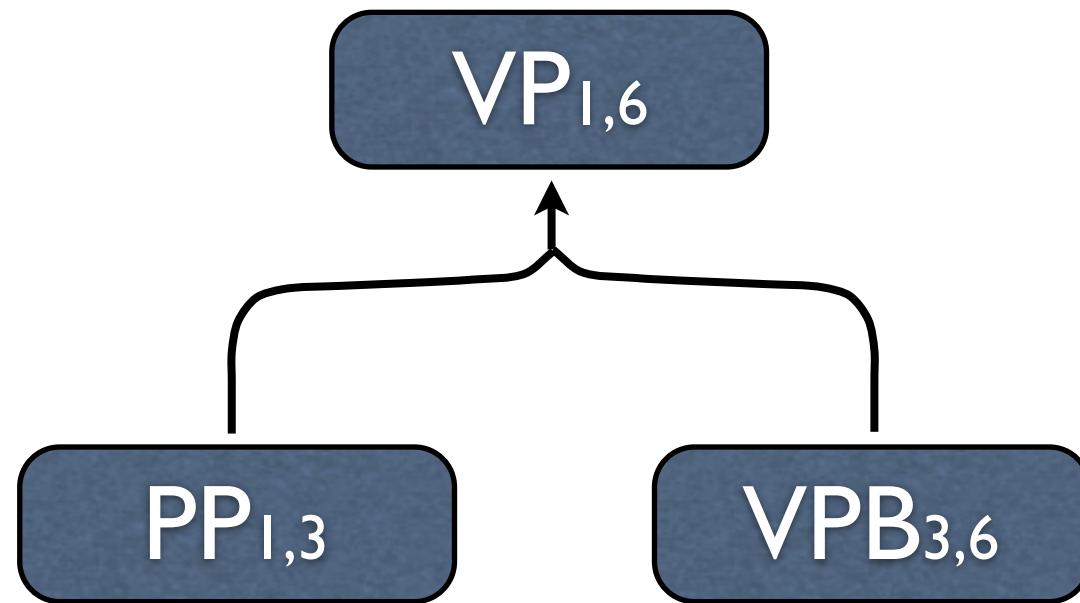
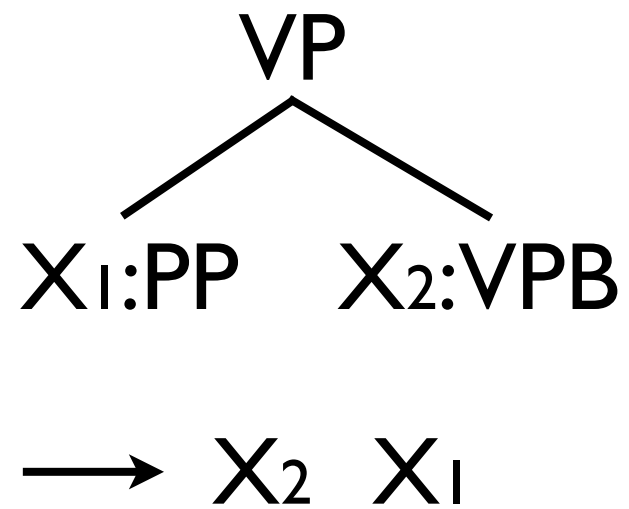
# Exhaustive Search

# Exhaustive Search

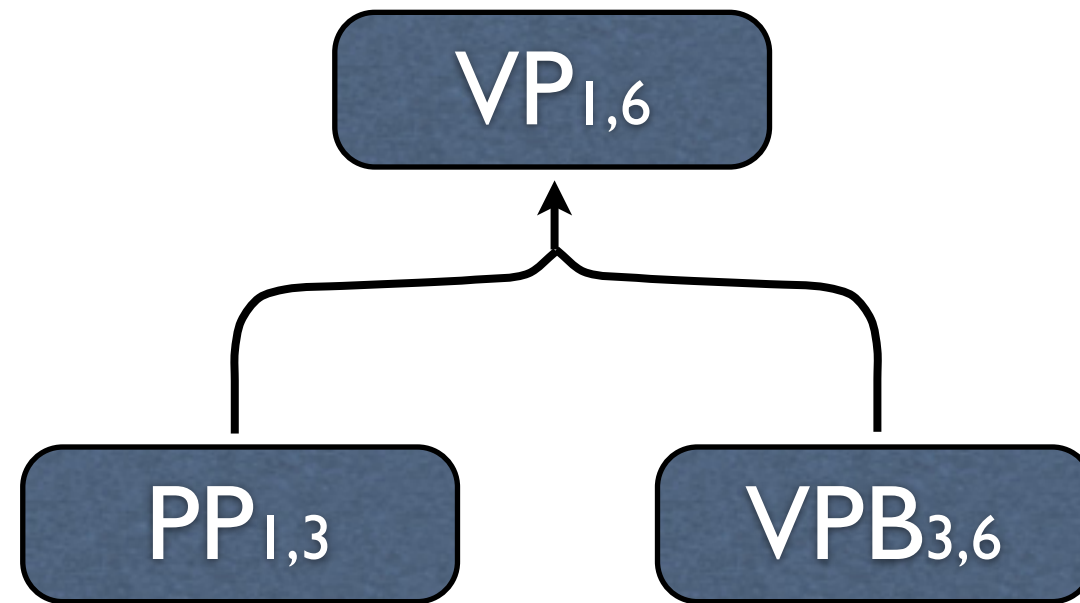
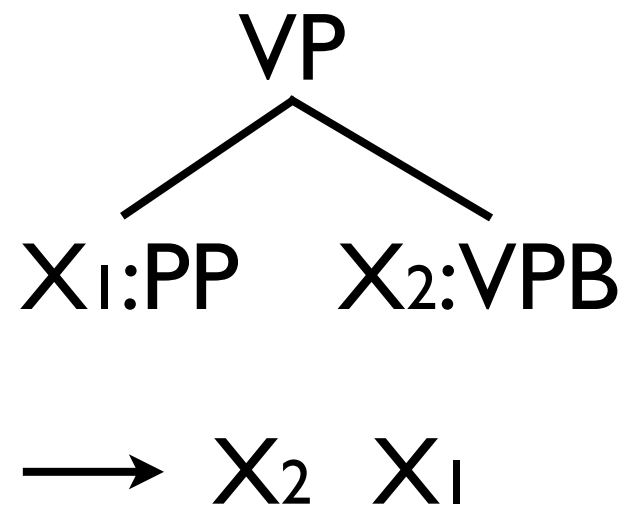




# Exhaustive Search

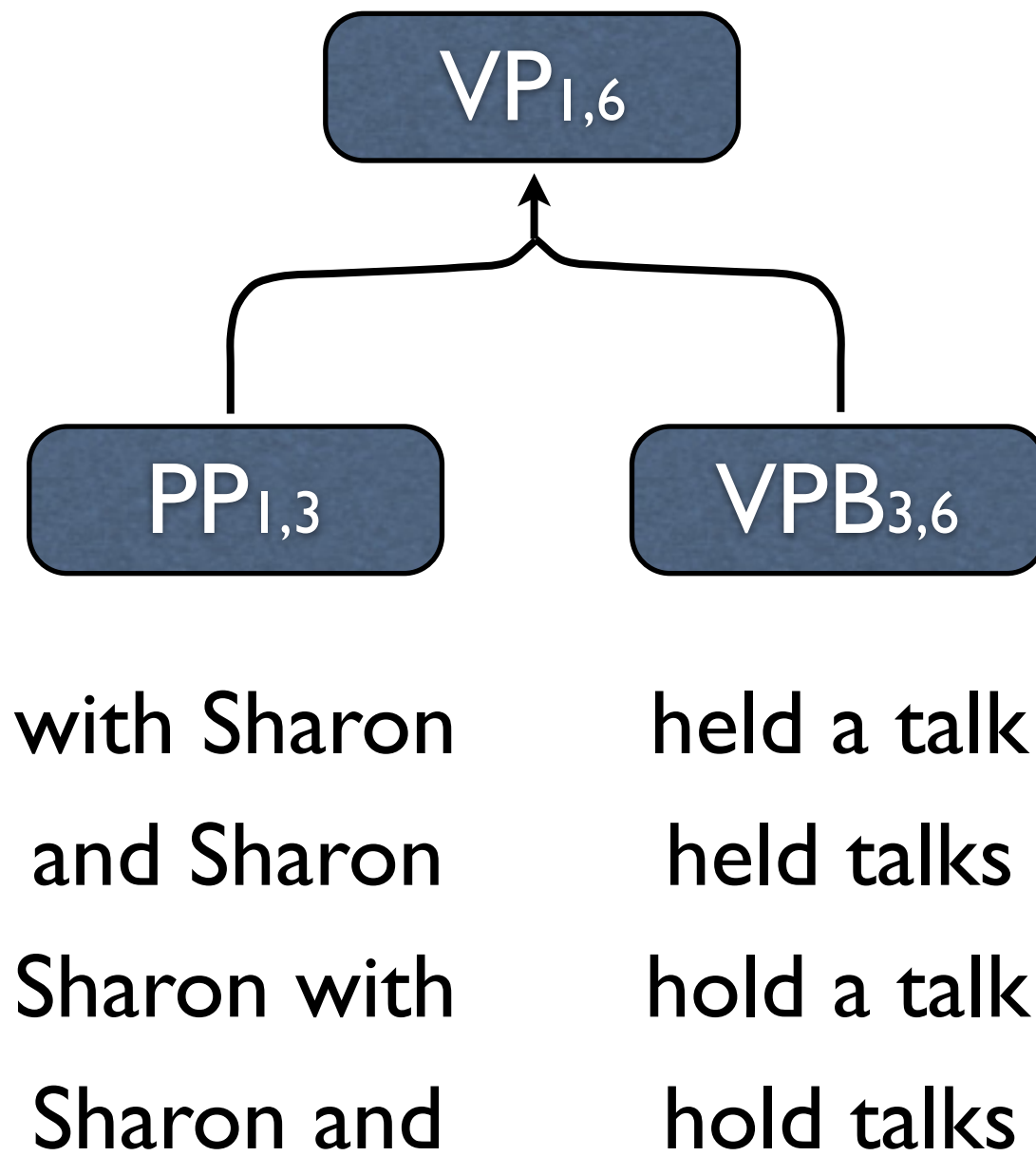
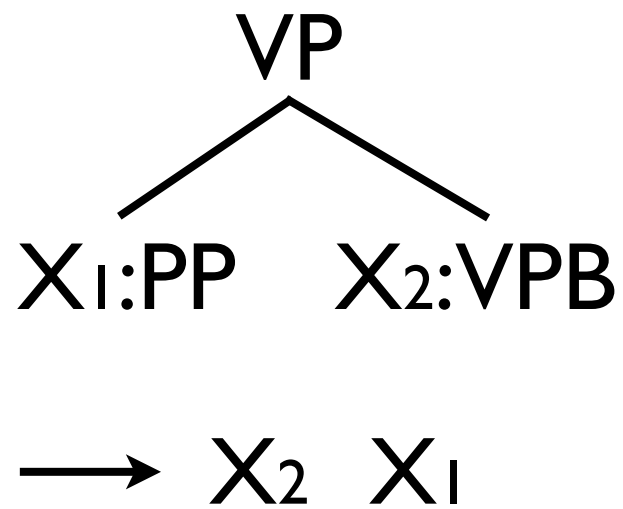


# Exhaustive Search

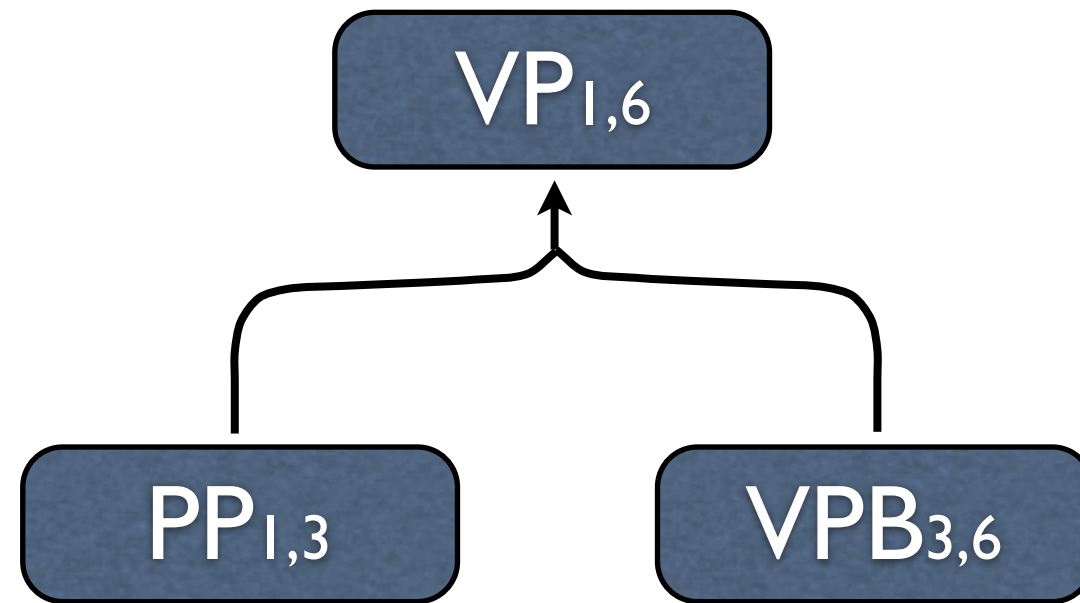
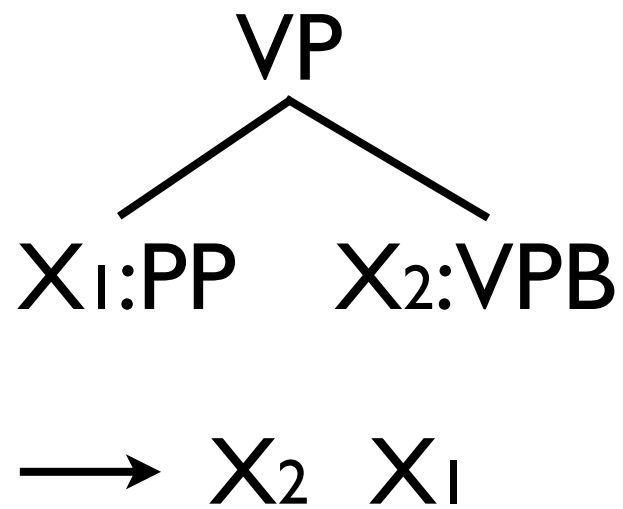


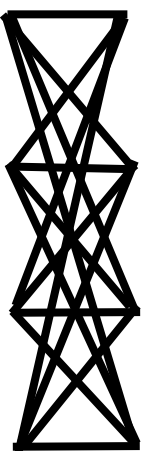
with Sharon  
and Sharon  
Sharon with  
Sharon and

# Exhaustive Search

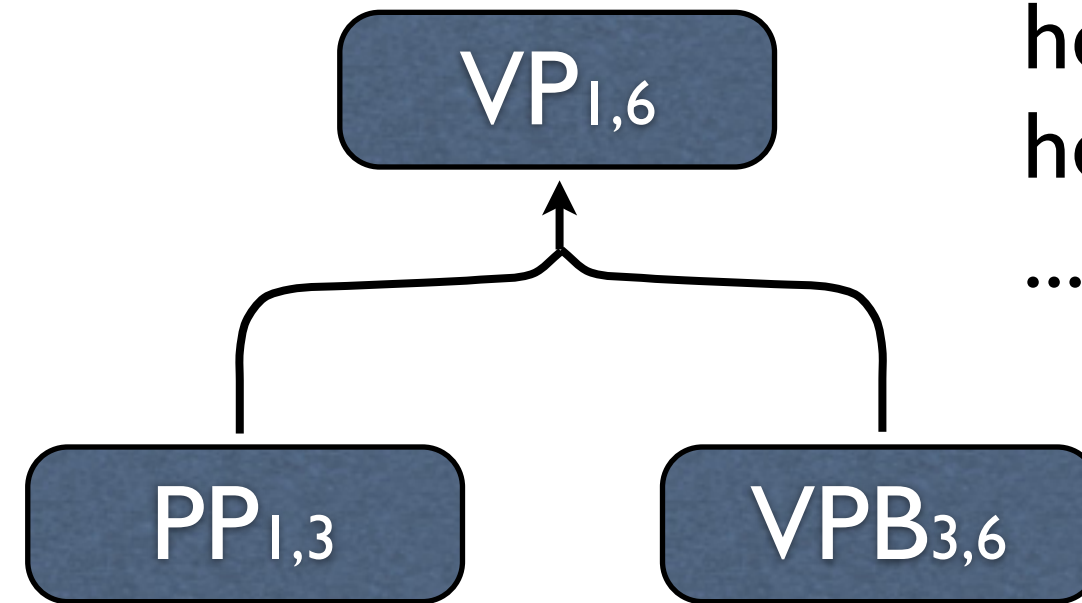
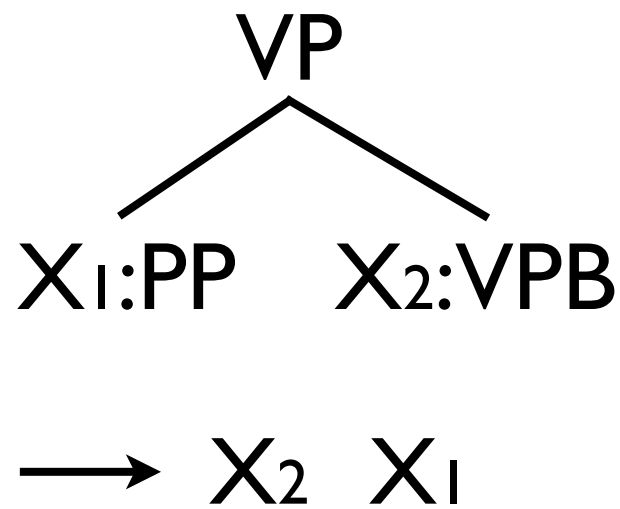


# Exhaustive Search



with Sharon		held a talk
and Sharon		held talks
Sharon with		hold a talk
Sharon and		hold talks

# Exhaustive Search



held a talk with Sharon  
 held a talk and Sharon  
 held talks with Sharon  
 held talks and Sharon  
 ...

with Sharon  
 and Sharon  
 Sharon with  
 Sharon and

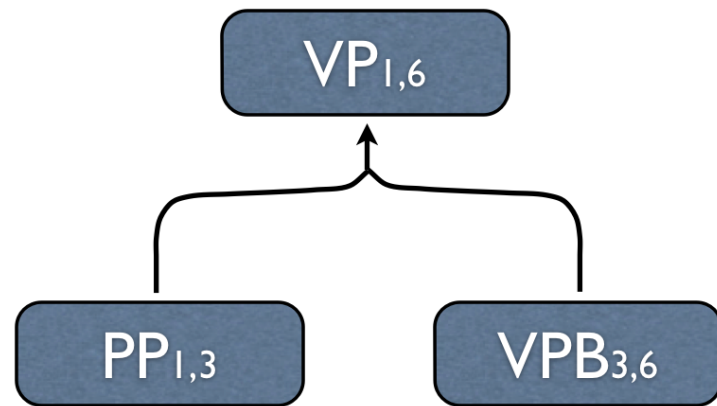
~~held a talk  
 held talks  
 hold a talk  
 hold talks~~

# Calculating N-best List

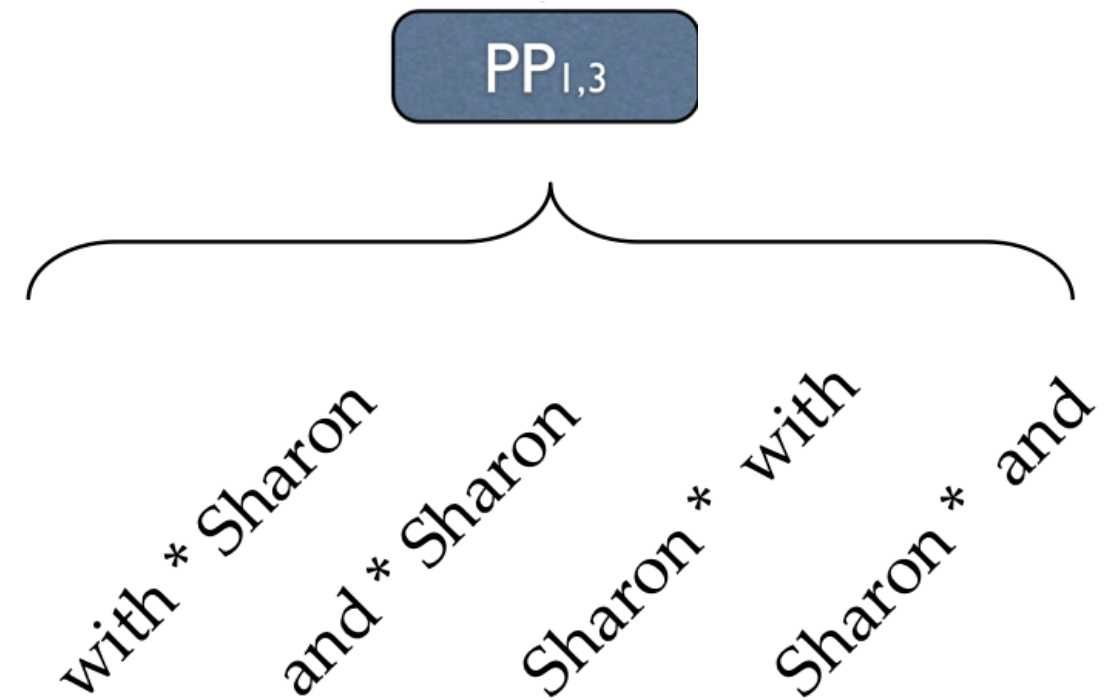
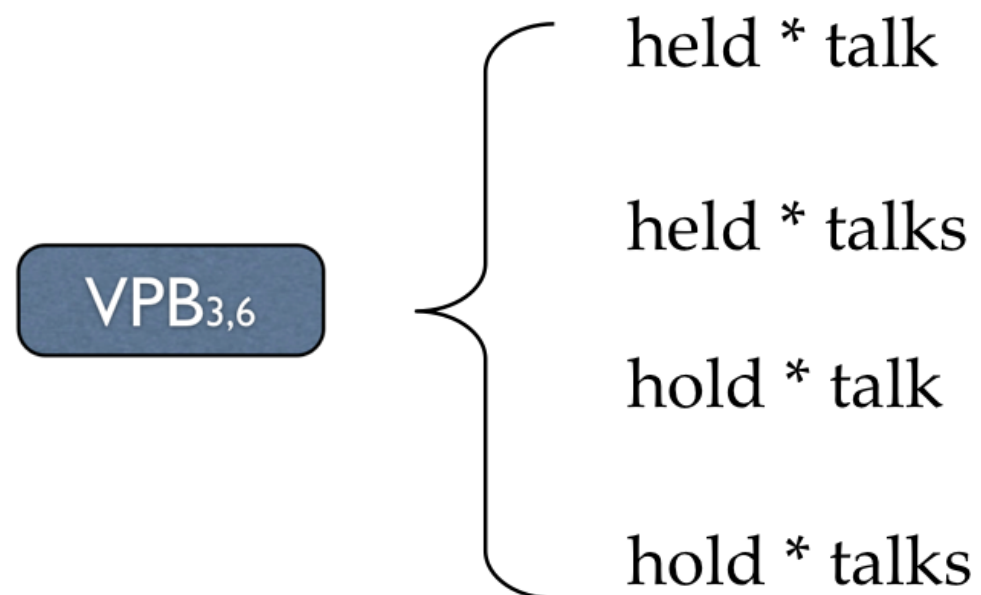
1.0	1.3
2.9	1.5
3.4	4.2
5.6	5.1
6.0	6.3
<b>a</b>	<b>b</b>

What's the N-best list of  $\mathbf{a}_i + \mathbf{b}_j$ 's ?

# Monotonicity

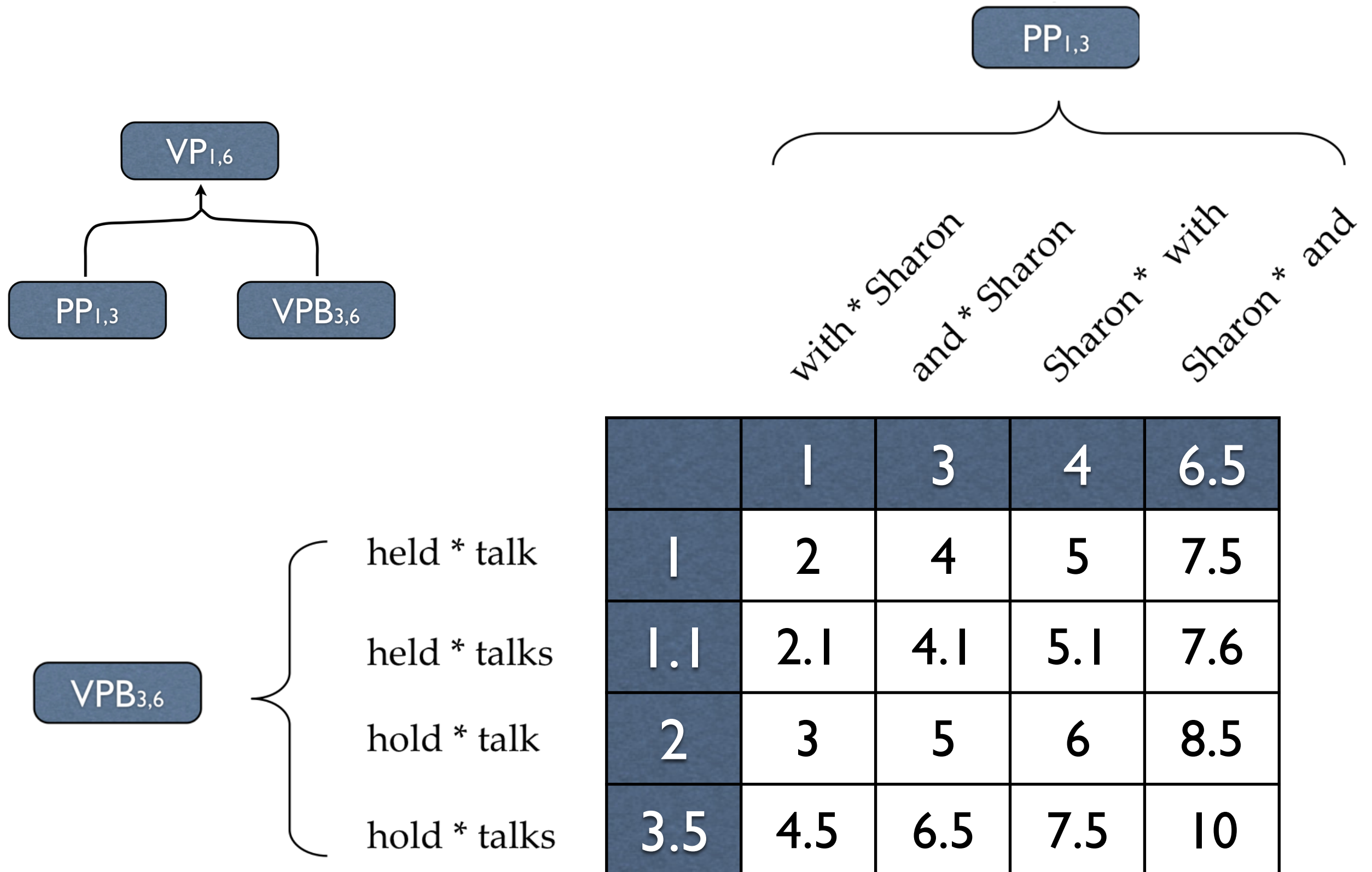


monotonic



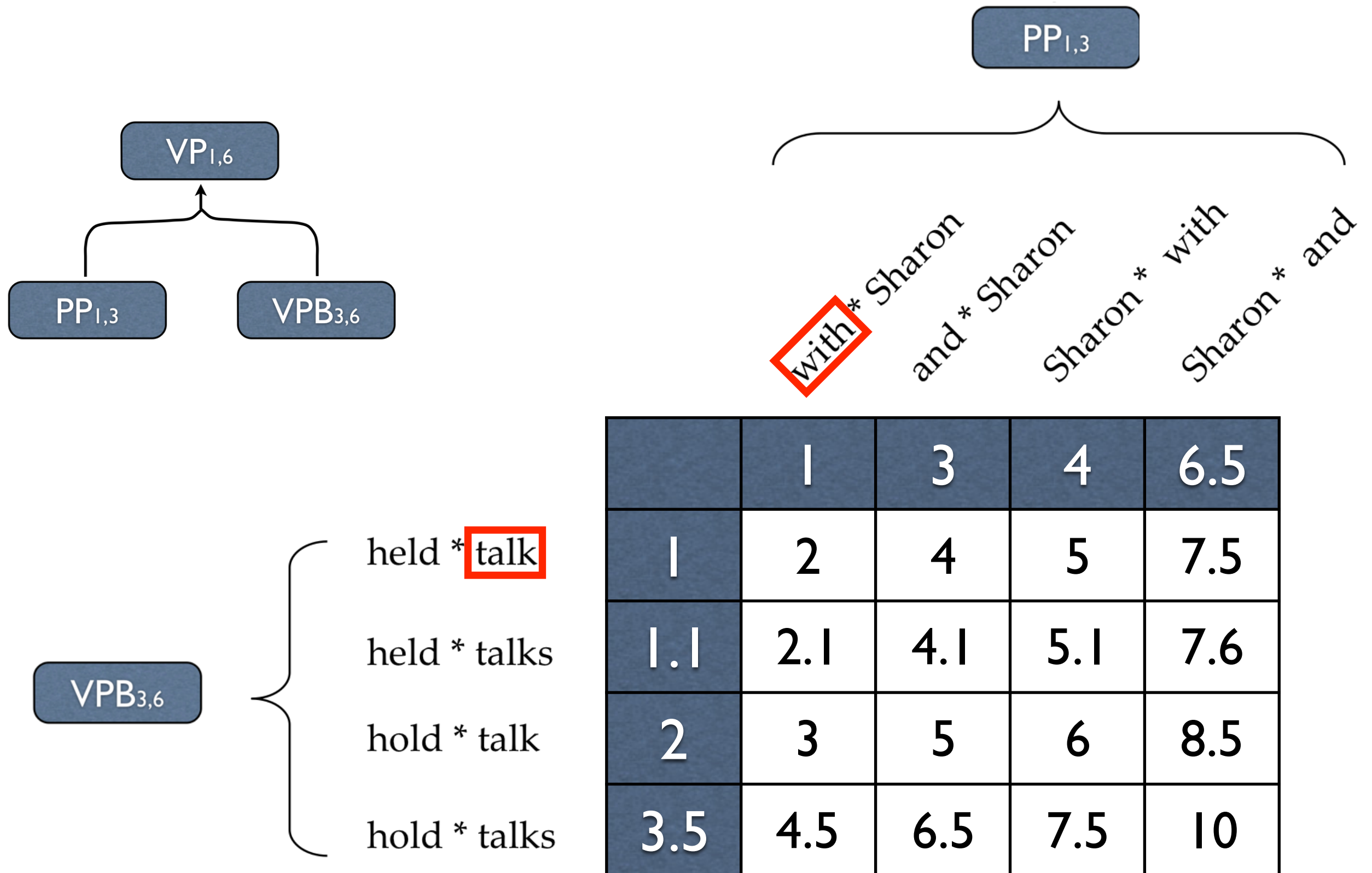
	1	3	4	6.5
1	2	4	5	7.5
1.1	2.1	4.1	5.1	7.6
2	3	5	6	8.5
3.5	4.5	6.5	7.5	10

# Non-Monotonicity



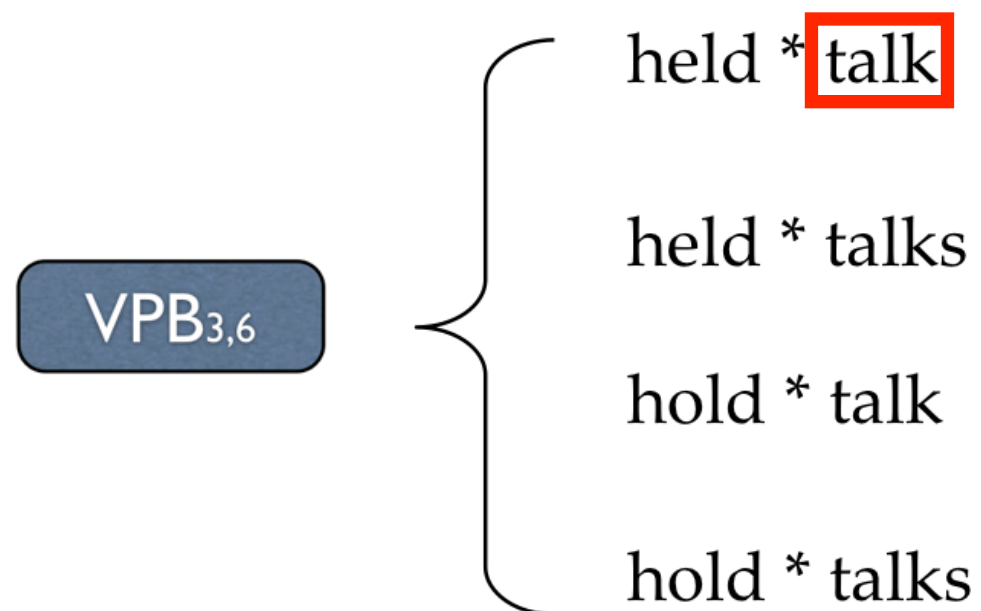
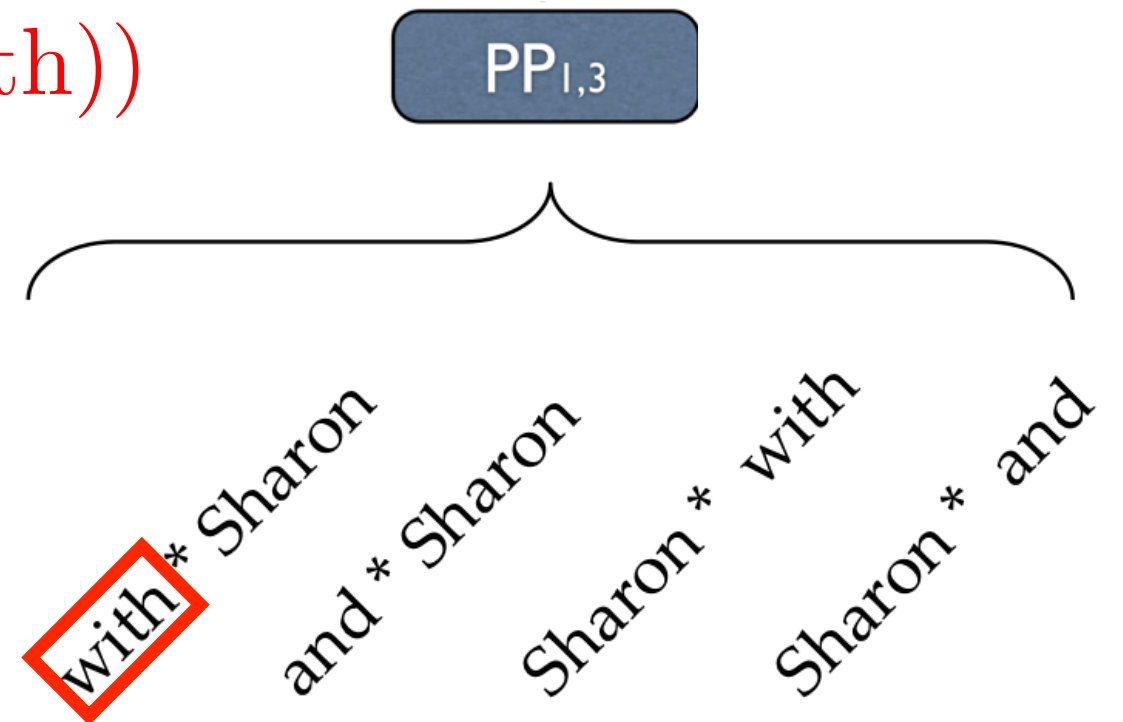
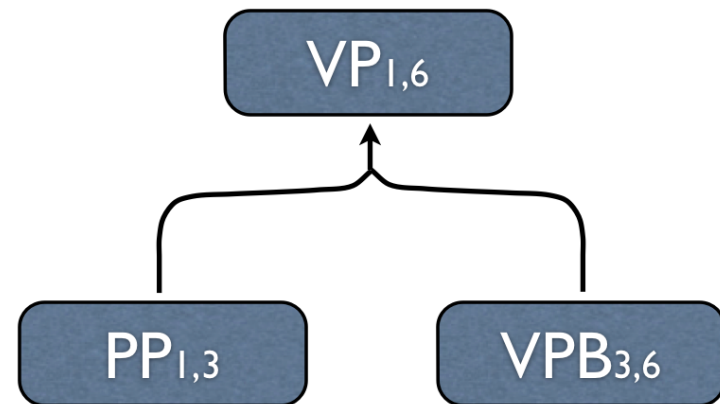


# Non-Monotonicity



# Non-Monotonicity

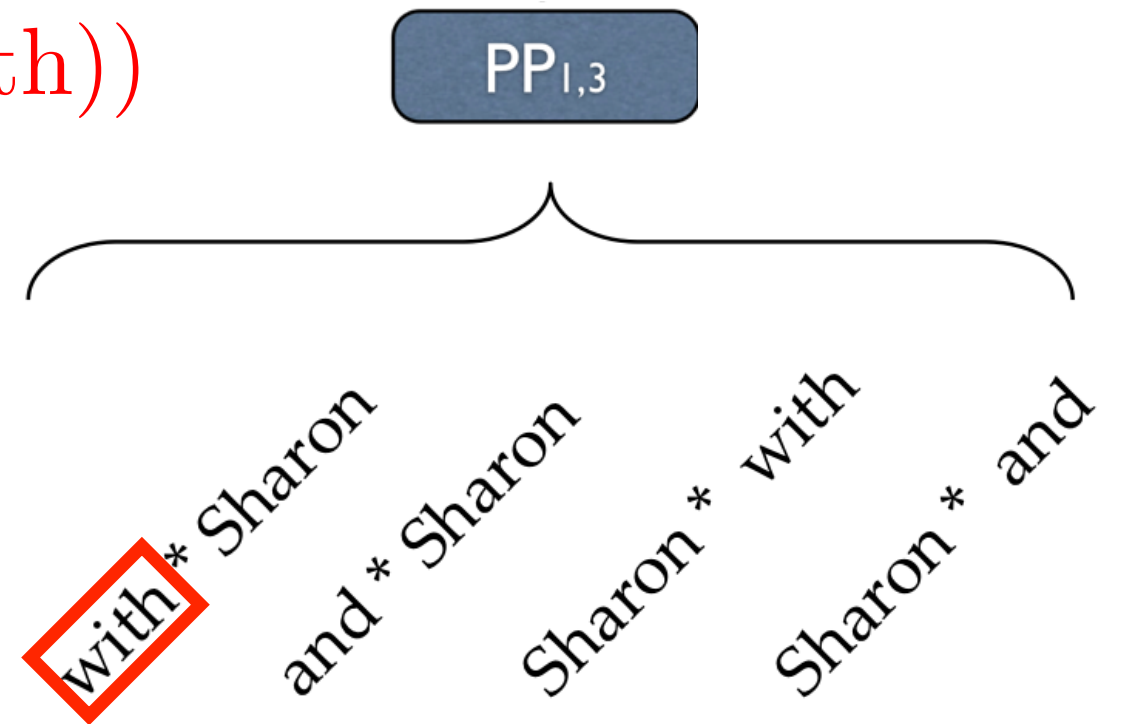
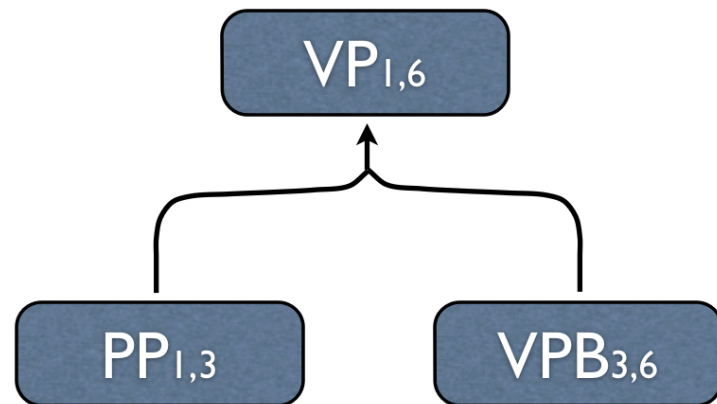
$$\log(p(\text{with}|\text{talk})) - \log(p(\text{with}))$$



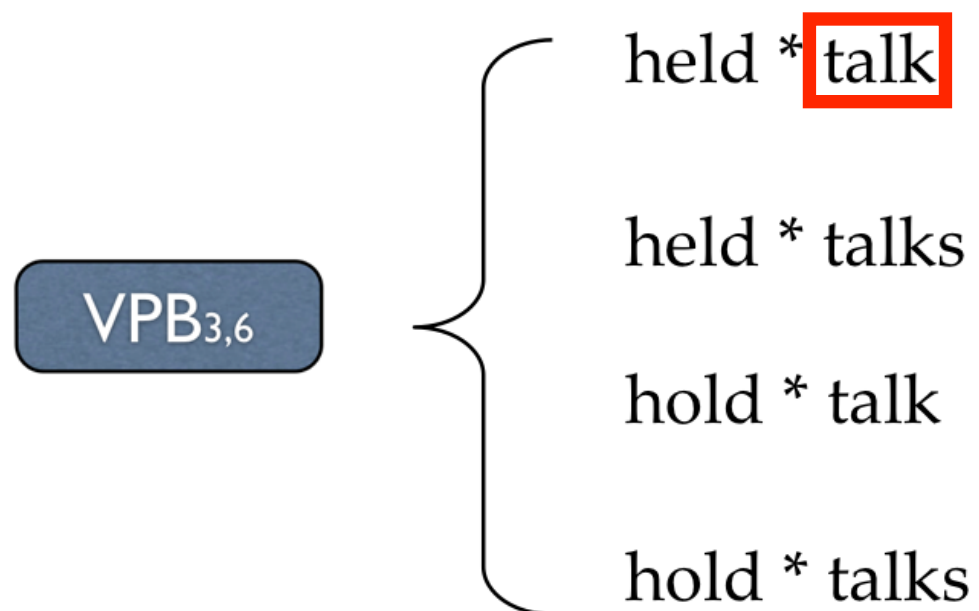
	1	3	4	6.5
1	2	4	5	7.5
1.1	2.1	4.1	5.1	7.6
2	3	5	6	8.5
3.5	4.5	6.5	7.5	10

# Non-Monotonicity

$$\log(p(\text{with}|\text{talk})) - \log(p(\text{with}))$$

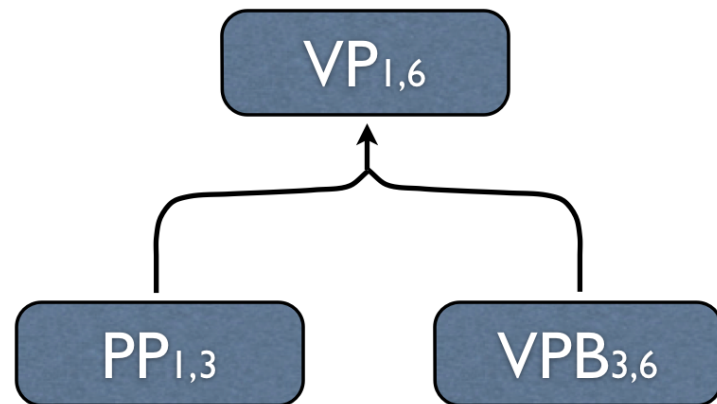


	1	3	4	6.5
1	2.0+0.5	4.0+2.0	5.0+4.0	7.5+4.0
1.1	2.1+0.3	4.1+1.5	5.1+3.5	7.6+3.0
2	3.0+0.5	5.0+2.0	6.0+4.0	8.5+4.0
3.5	4.5+0.3	6.5+1.5	7.5+3.5	10+3.5



# Non-Monotonicity

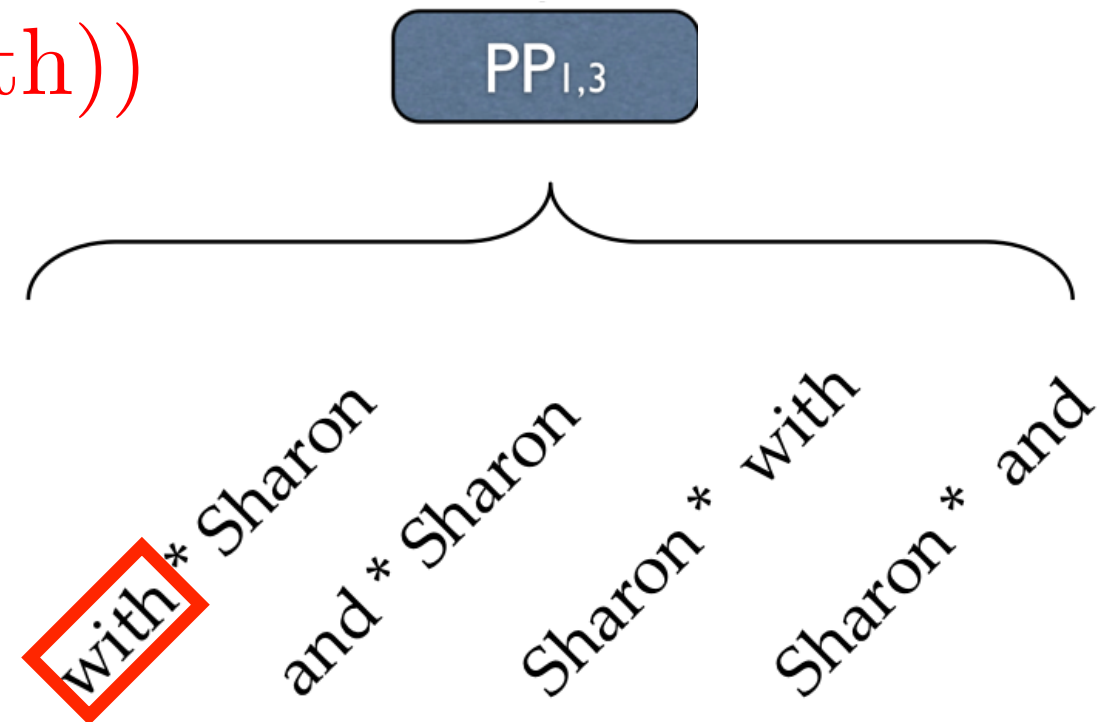
$$\log(p(\text{with}|\text{talk})) - \log(p(\text{with}))$$



LM introduces  
non-monotonicity

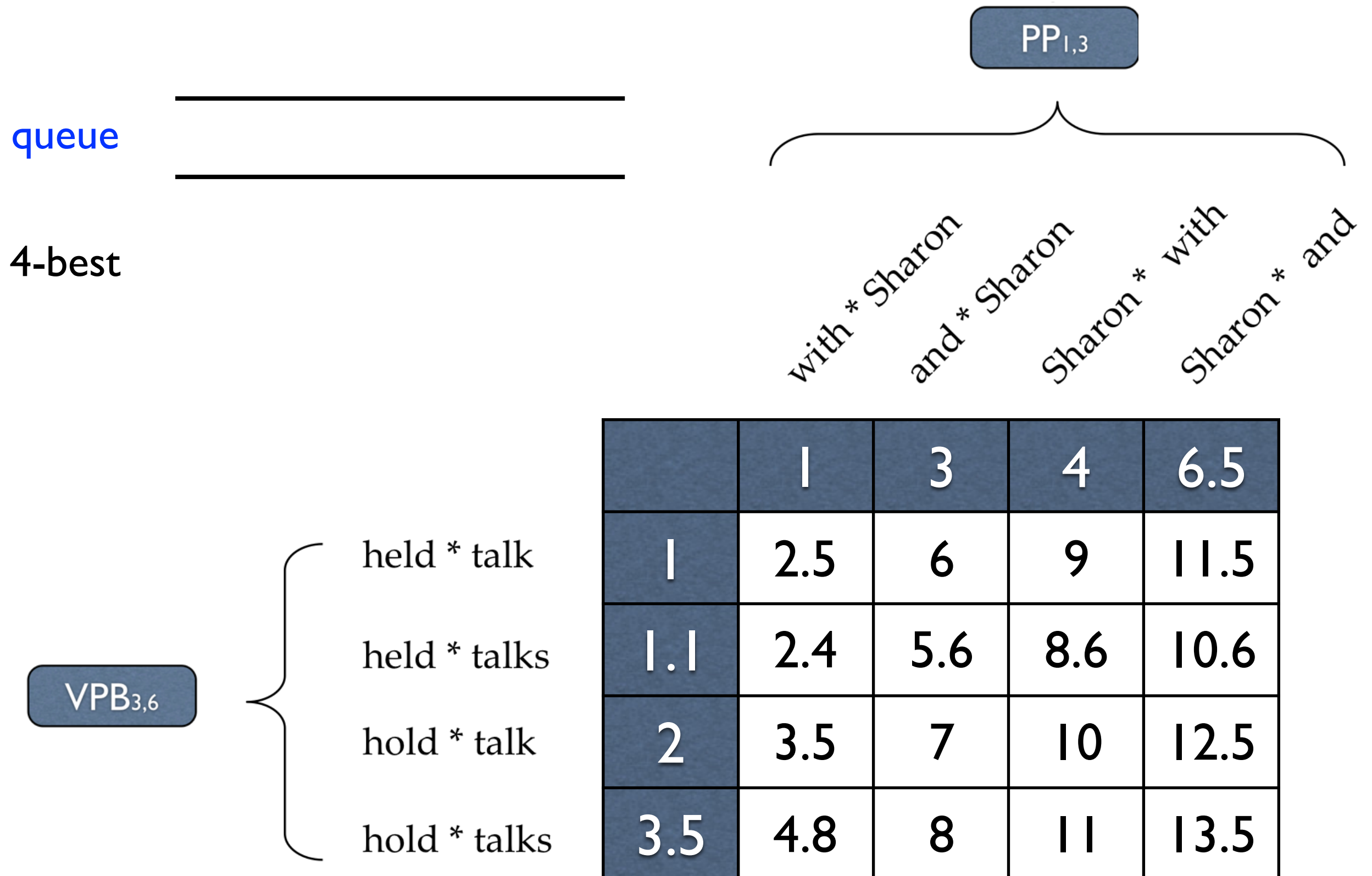
VPB<sub>3,6</sub>

held \* **talk**  
held \* talks  
hold \* talk  
hold \* talks



	1	3	4	6.5
1	2.0+0.5	4.0+2.0	5.0+4.0	7.5+4.0
1.1	2.1+0.3	4.1+1.5	5.1+3.5	7.6+3.0
2	3.0+0.5	5.0+2.0	6.0+4.0	8.5+4.0
3.5	4.5+0.3	6.5+1.5	7.5+3.5	10+3.5

# Cube Pruning



# Cube Pruning

queue

4-best

PP<sub>1,3</sub>

with \* Sharon  
and \* Sharon  
Sharon \* with  
Sharon \* and

VPB<sub>3,6</sub>

held \* talk

held \* talks

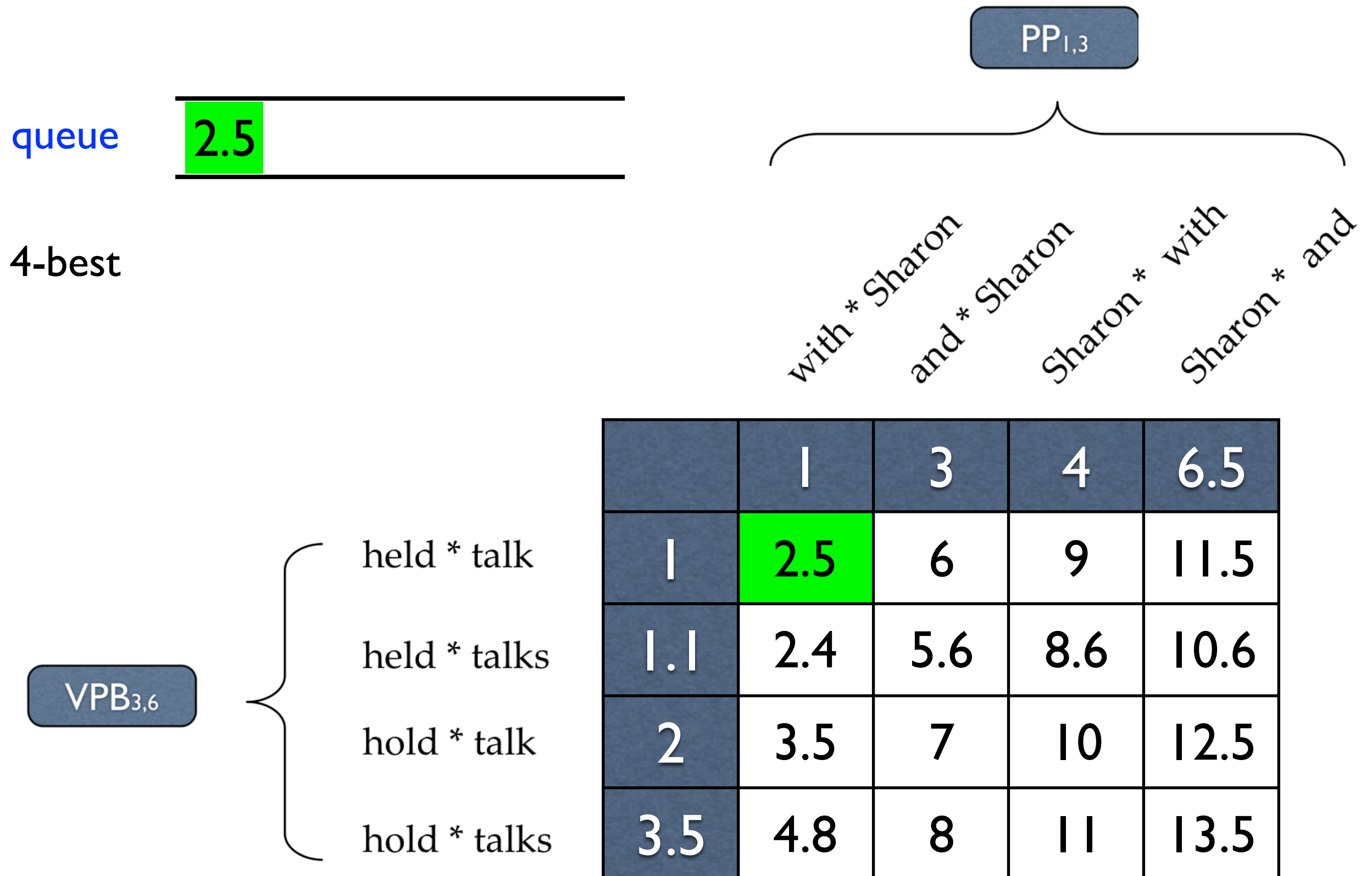
hold \* talk

hold \* talks

	1	3	4	6.5
1	2.5	6	9	11.5
1.1	2.4	5.6	8.6	10.6
2	3.5	7	10	12.5
3.5	4.8	8	11	13.5



# Cube Pruning



# Cube Pruning

queue

4-best

2.5

PP<sub>1,3</sub>

with \* Sharon  
and \* Sharon  
Sharon \* with  
Sharon \* and

VPB<sub>3,6</sub>

held \* talk

held \* talks

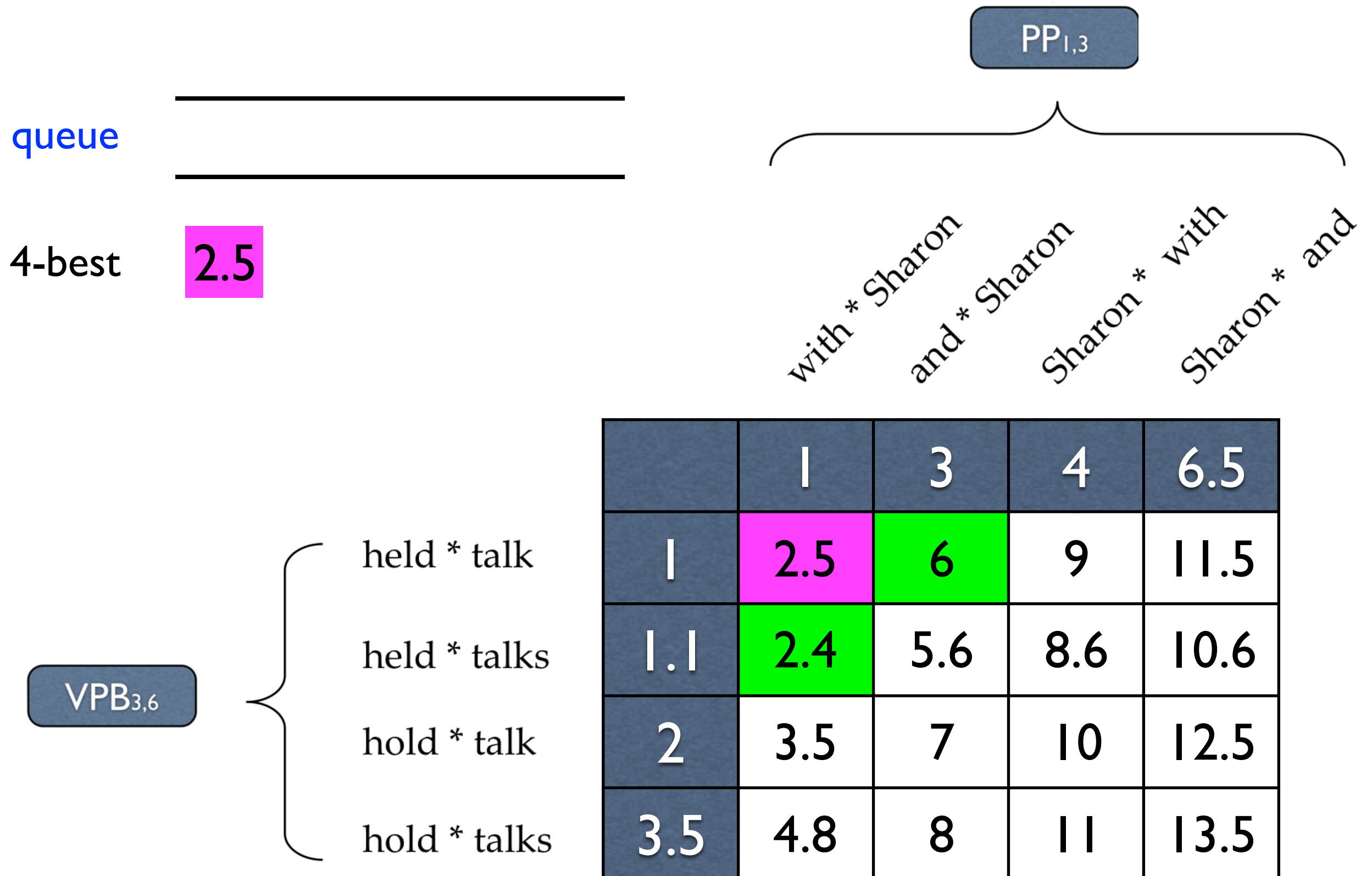
hold \* talk

hold \* talks

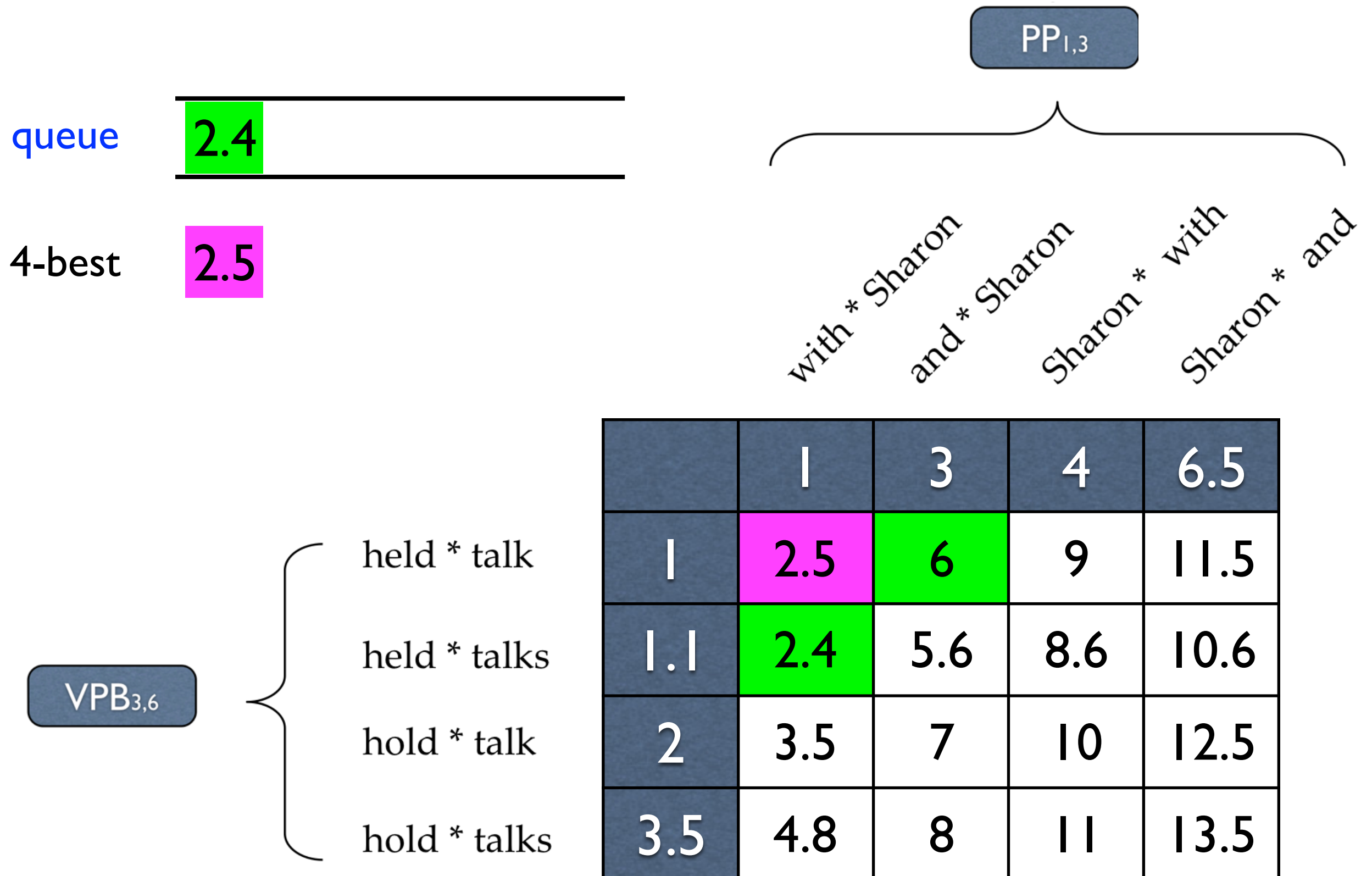
	1	3	4	6.5
1	2.5	6	9	11.5
1.1	2.4	5.6	8.6	10.6
2	3.5	7	10	12.5
3.5	4.8	8	11	13.5



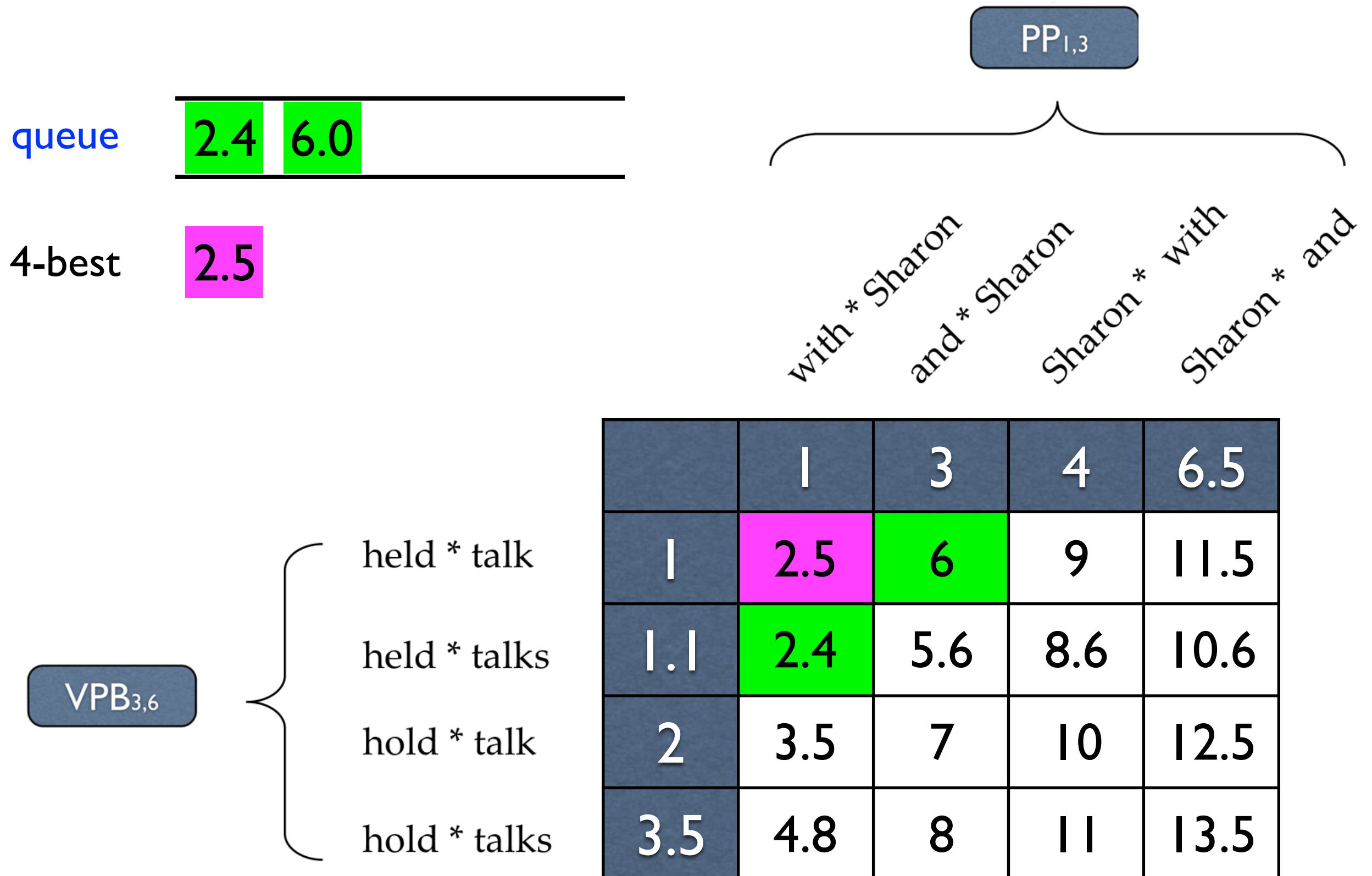
# Cube Pruning



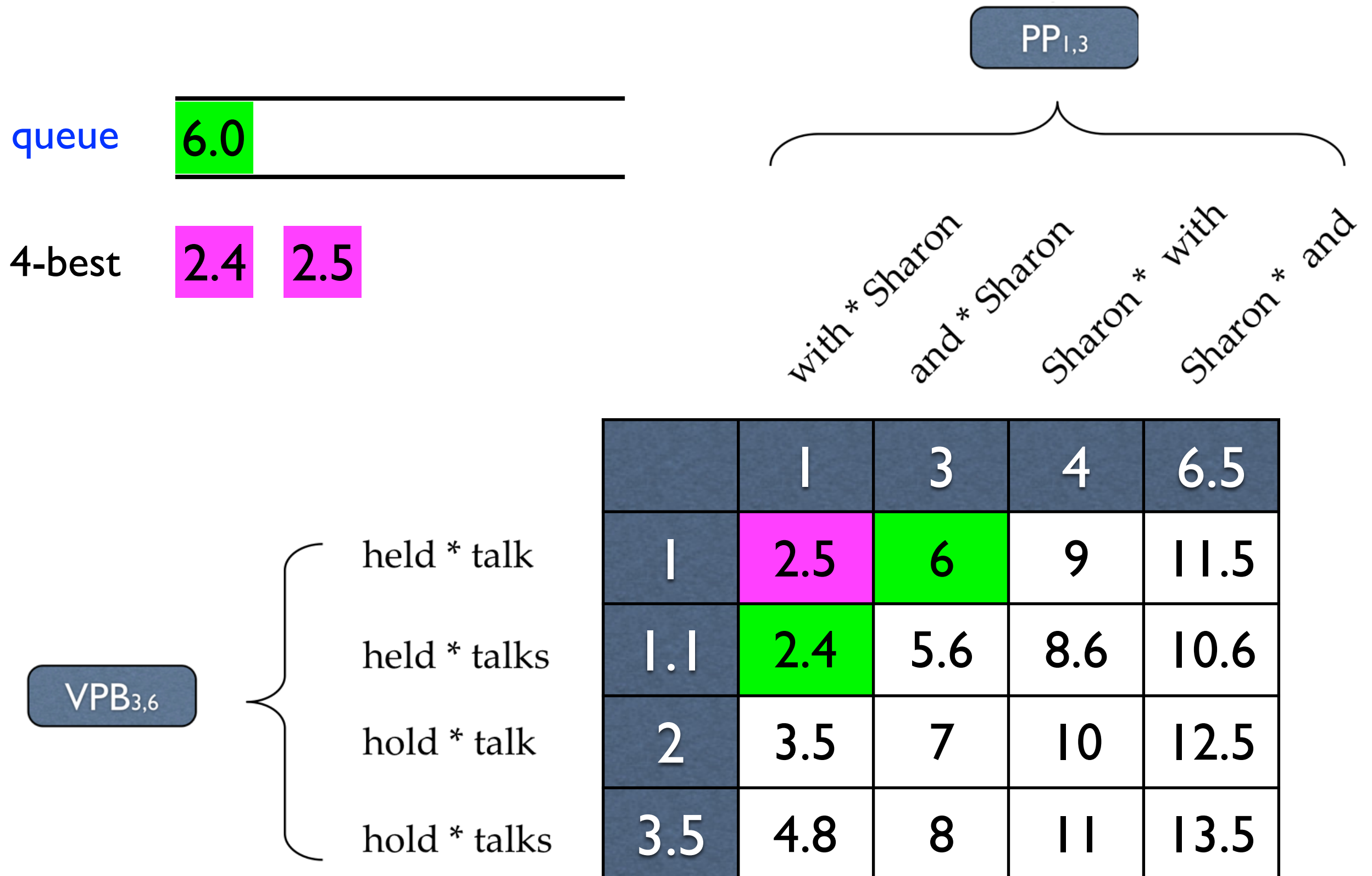
# Cube Pruning



# Cube Pruning



# Cube Pruning



# Cube Pruning

queue

6.0



4-best

2.4

2.5

PP<sub>1,3</sub>

with \* Sharon  
and \* Sharon  
Sharon \* with  
Sharon \* and

VPB<sub>3,6</sub>

held \* talk

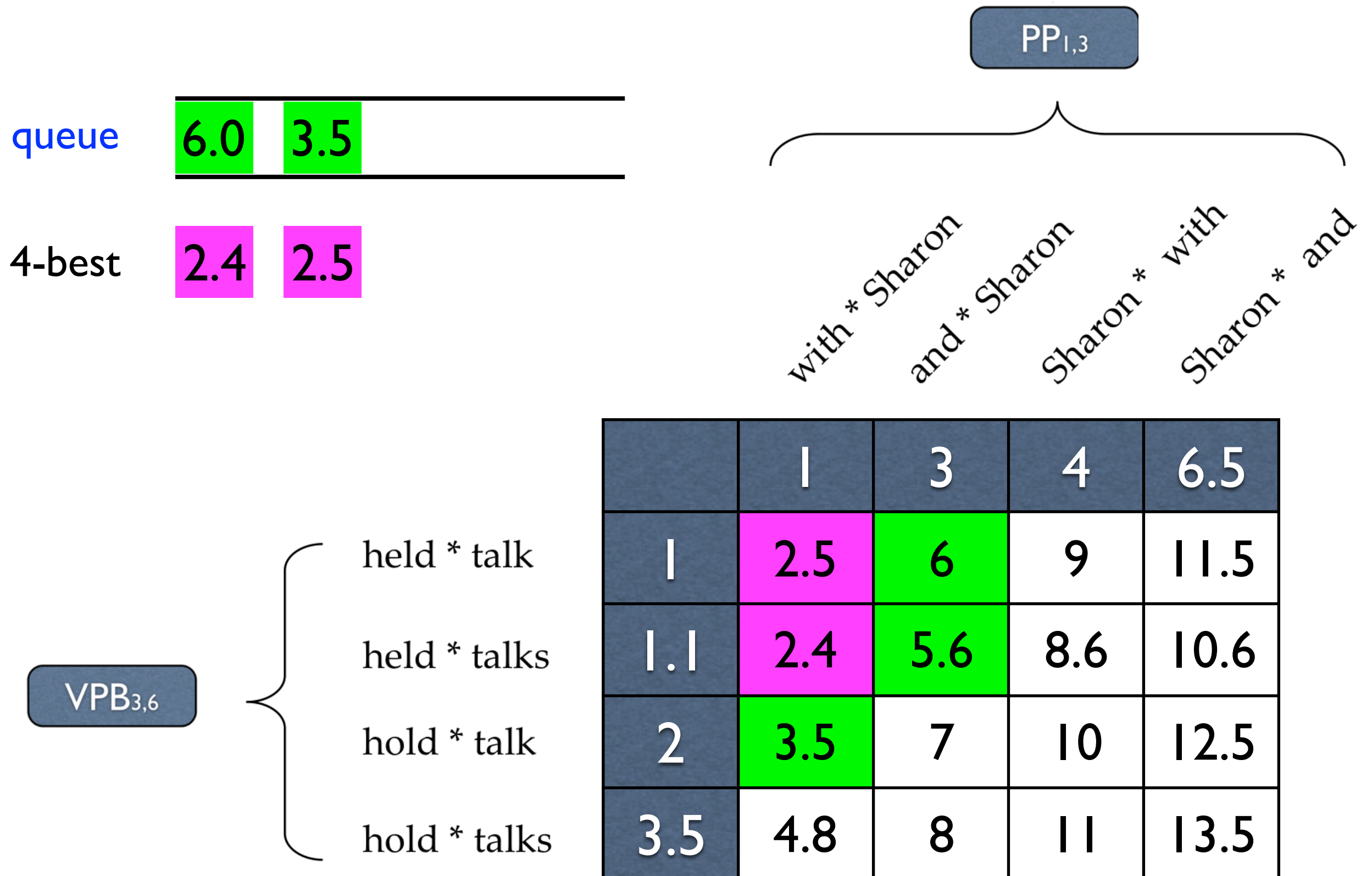
held \* talks

hold \* talk

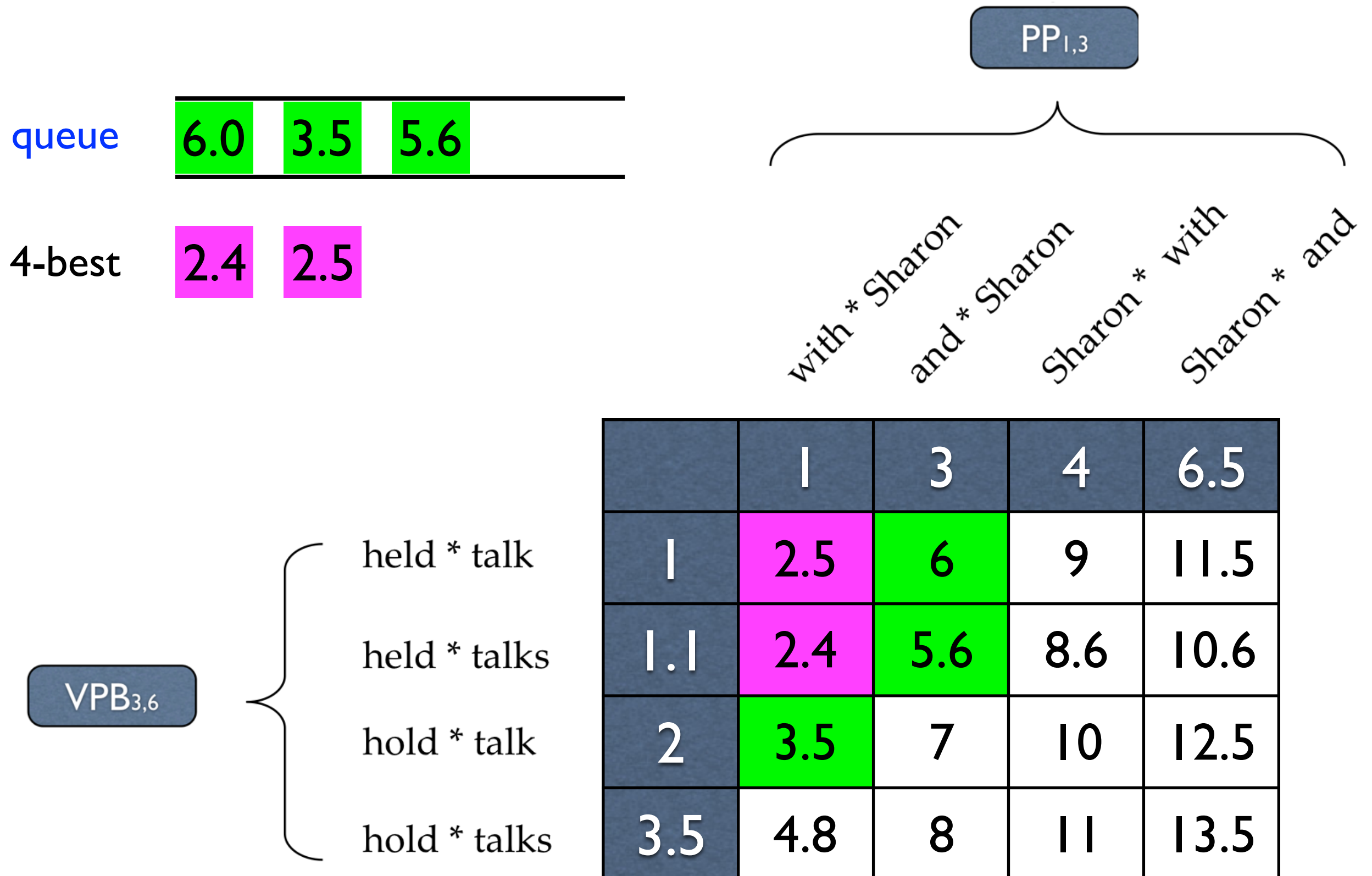
hold \* talks

	1	3	4	6.5
1	2.5	6	9	11.5
1.1	2.4	5.6	8.6	10.6
2	3.5	7	10	12.5
3.5	4.8	8	11	13.5

# Cube Pruning

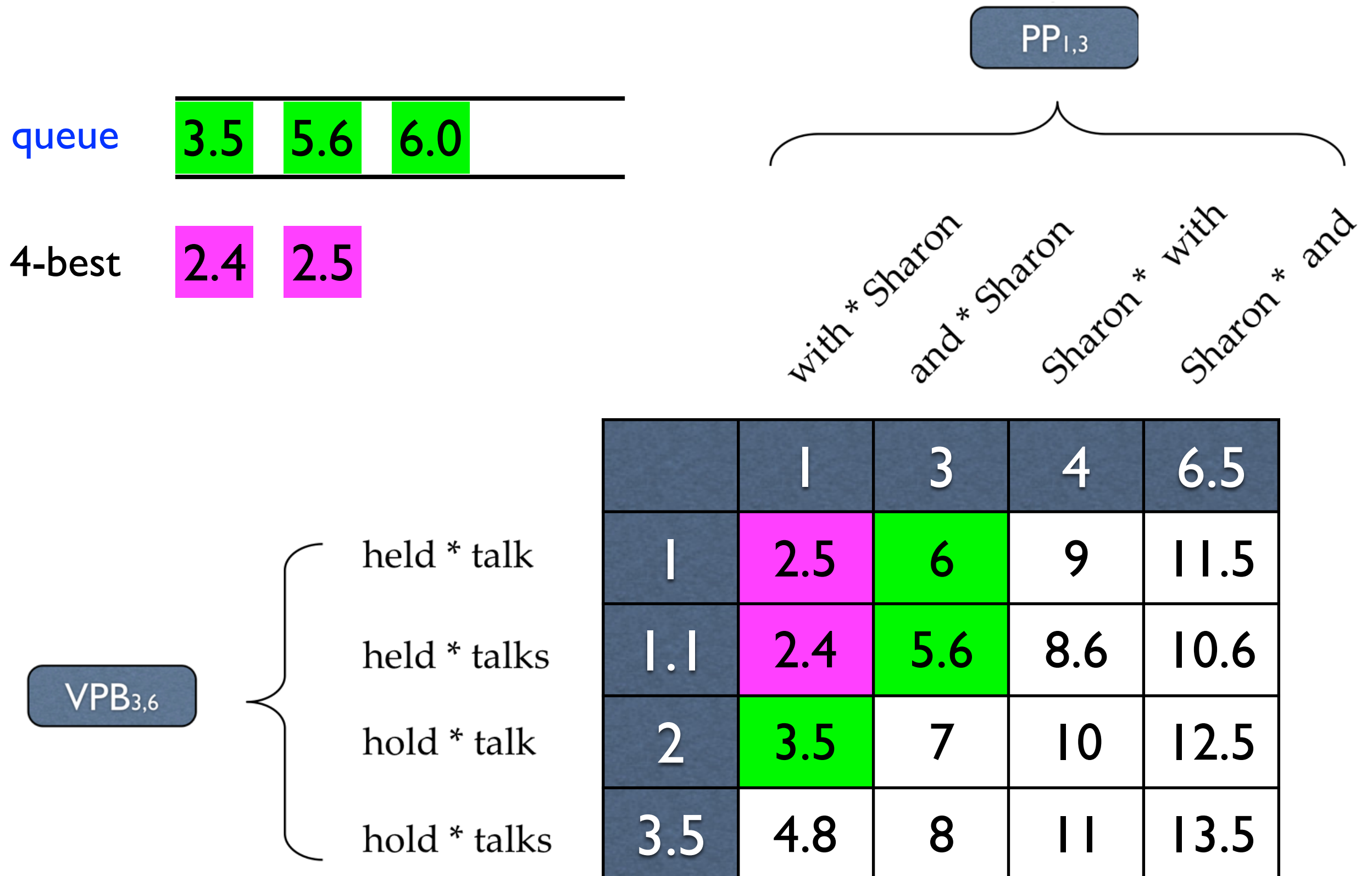


# Cube Pruning



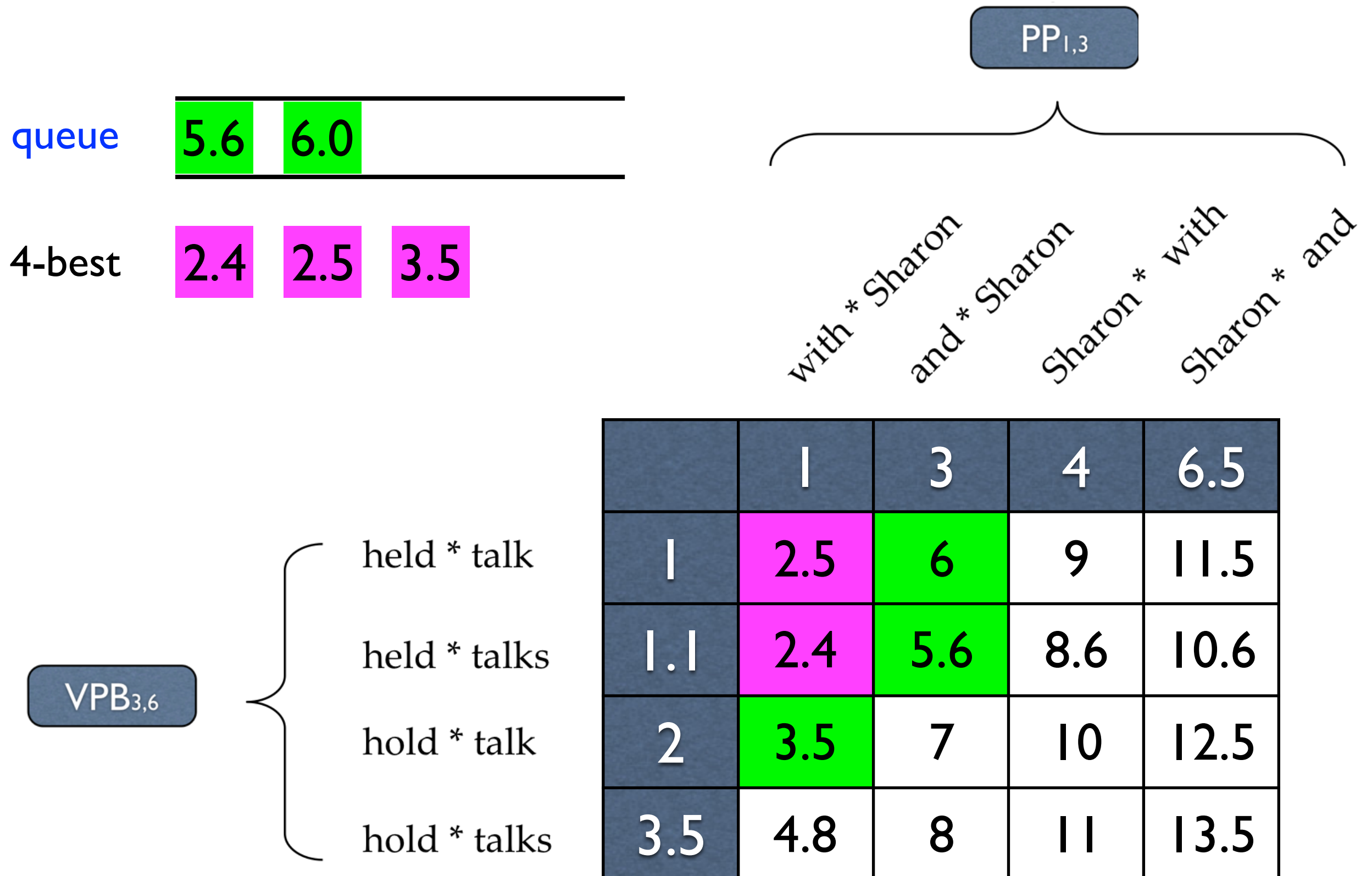


# Cube Pruning

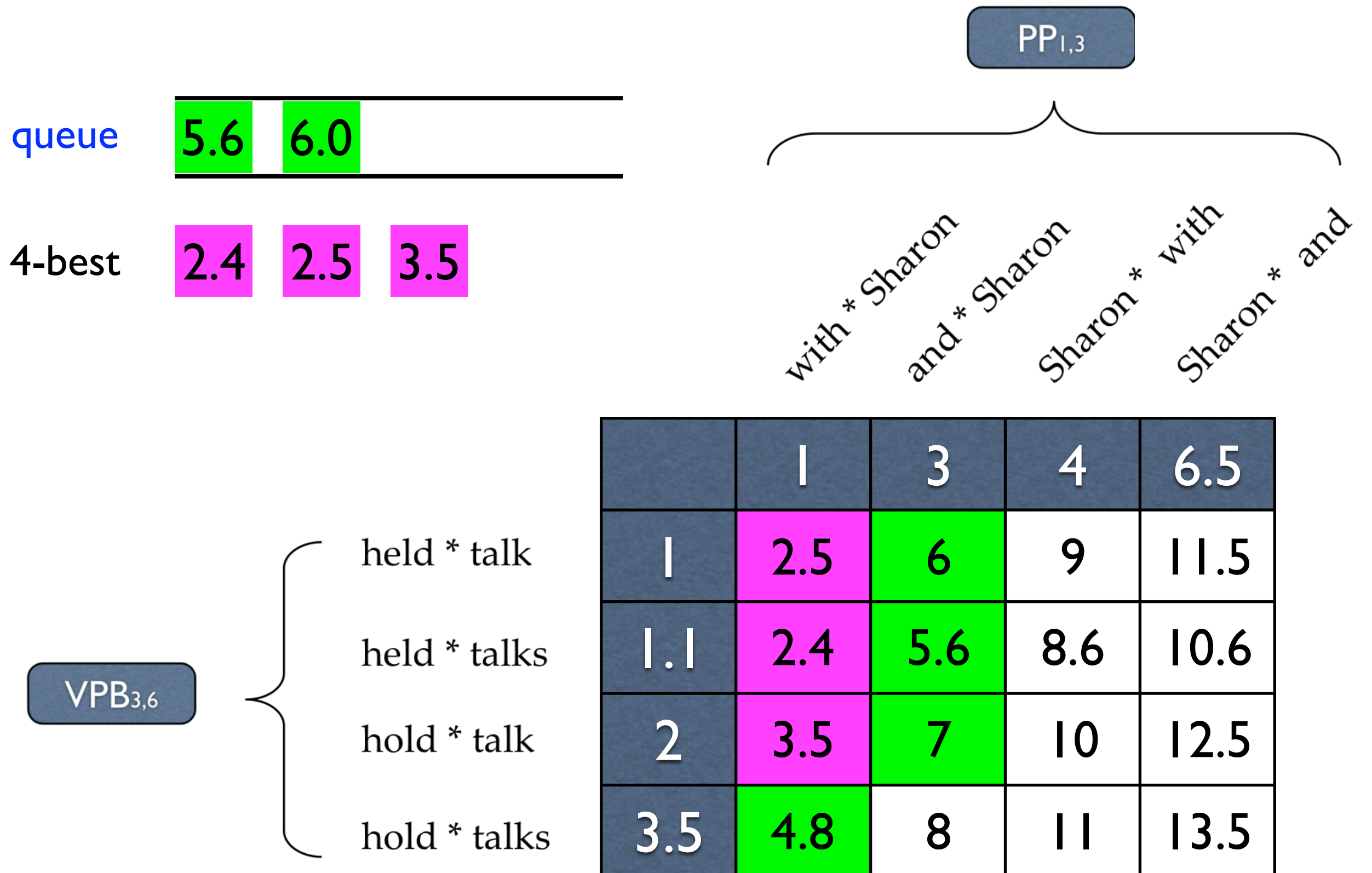




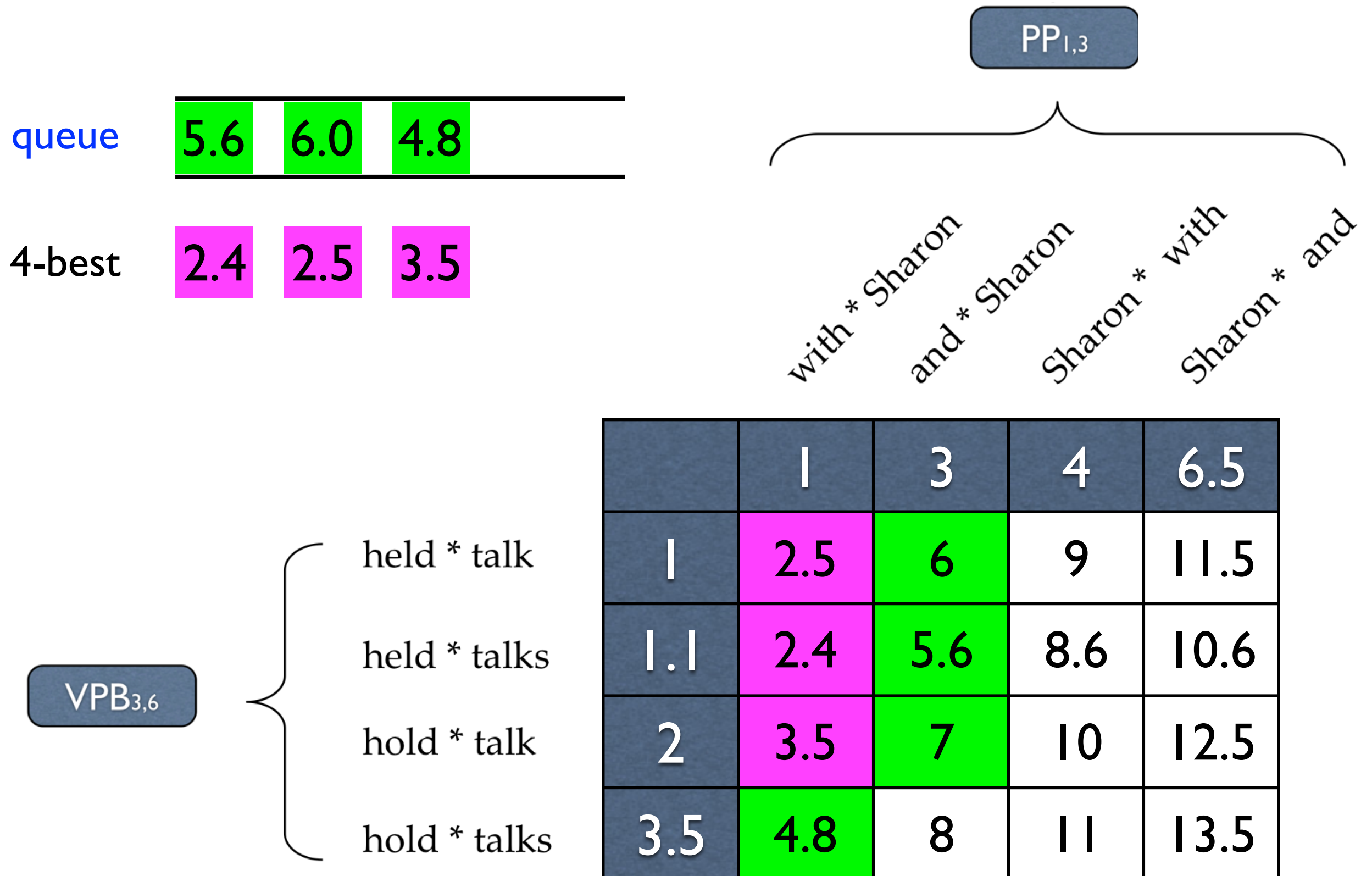
# Cube Pruning



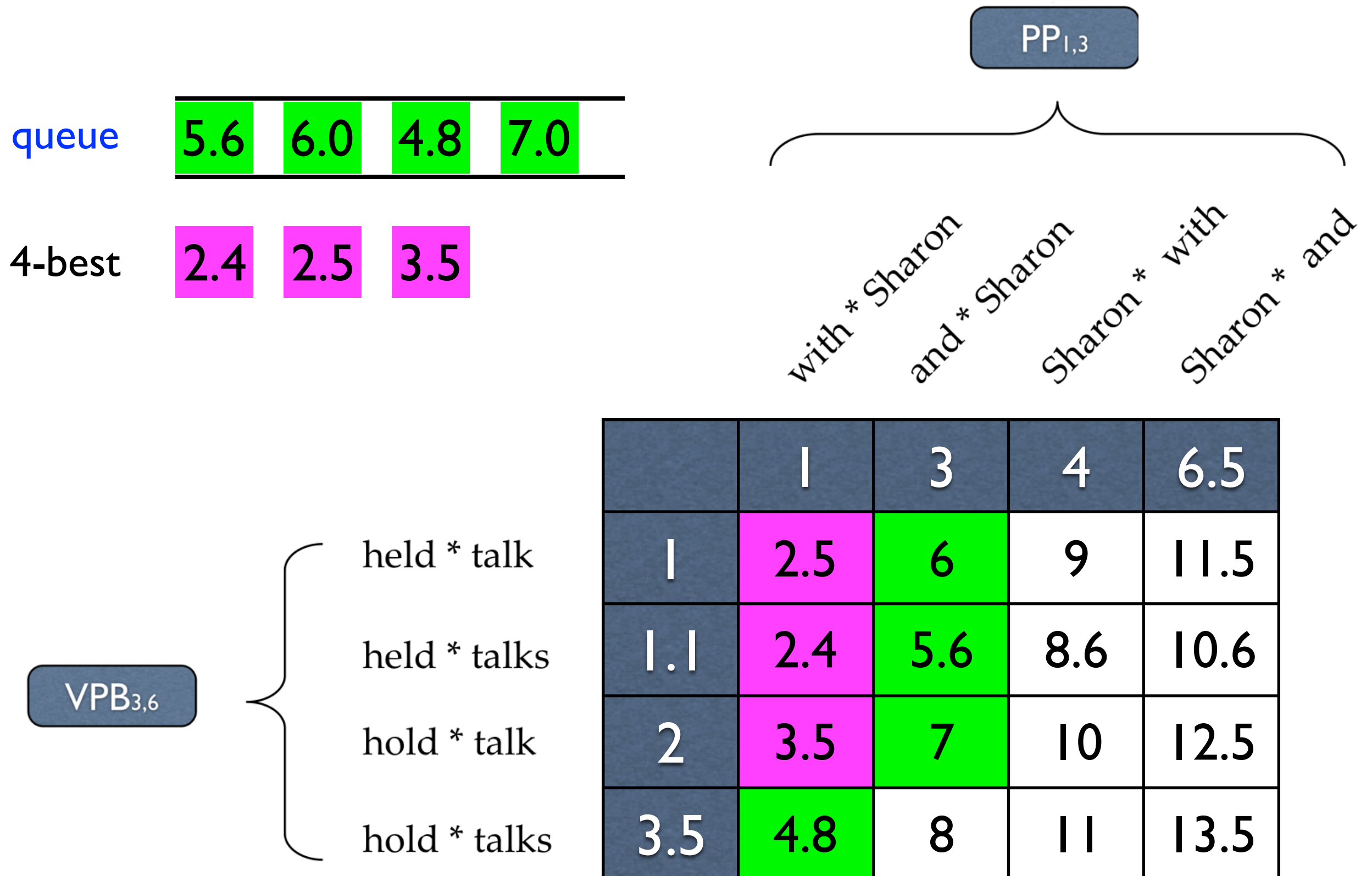
# Cube Pruning



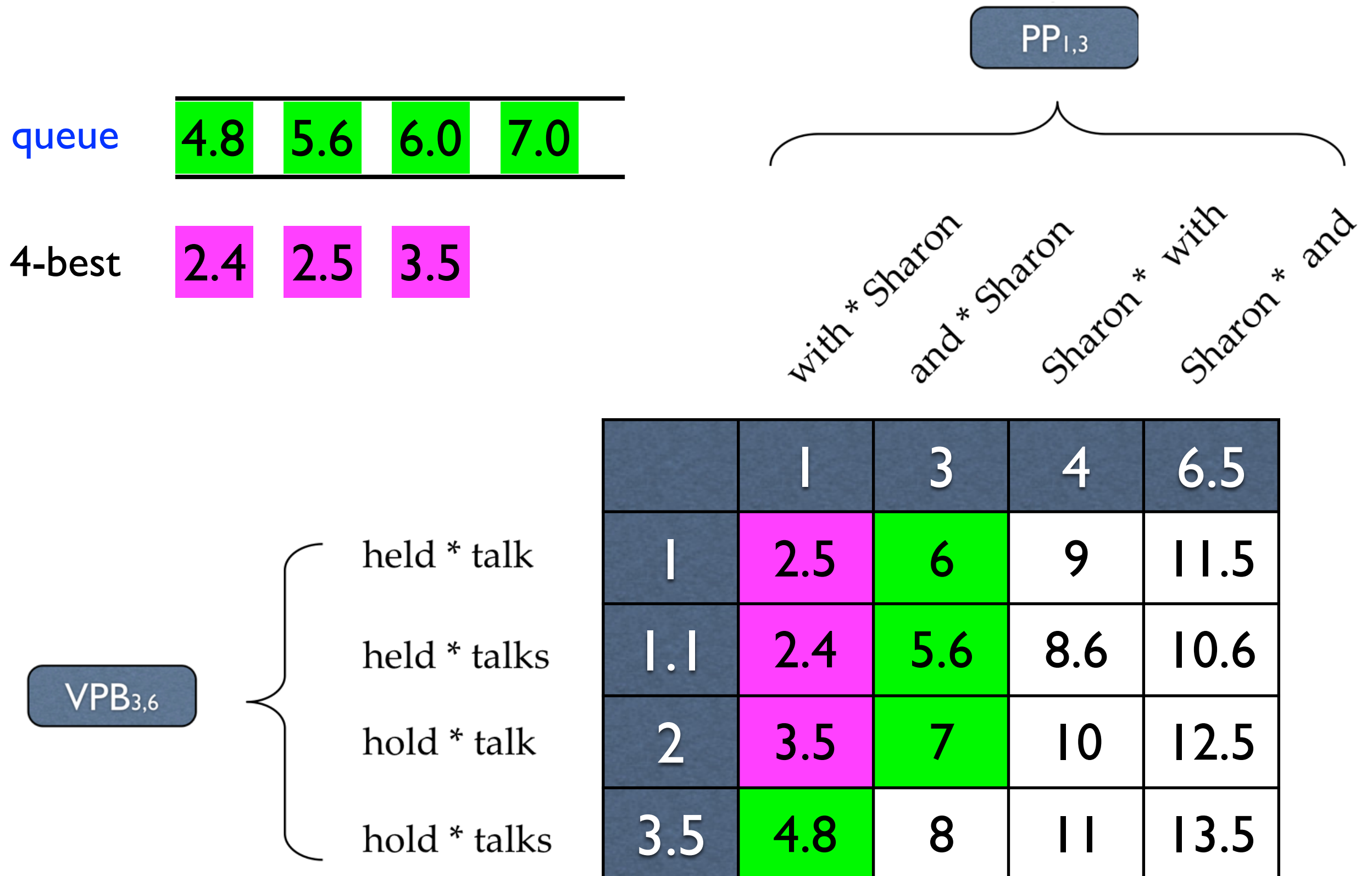
# Cube Pruning



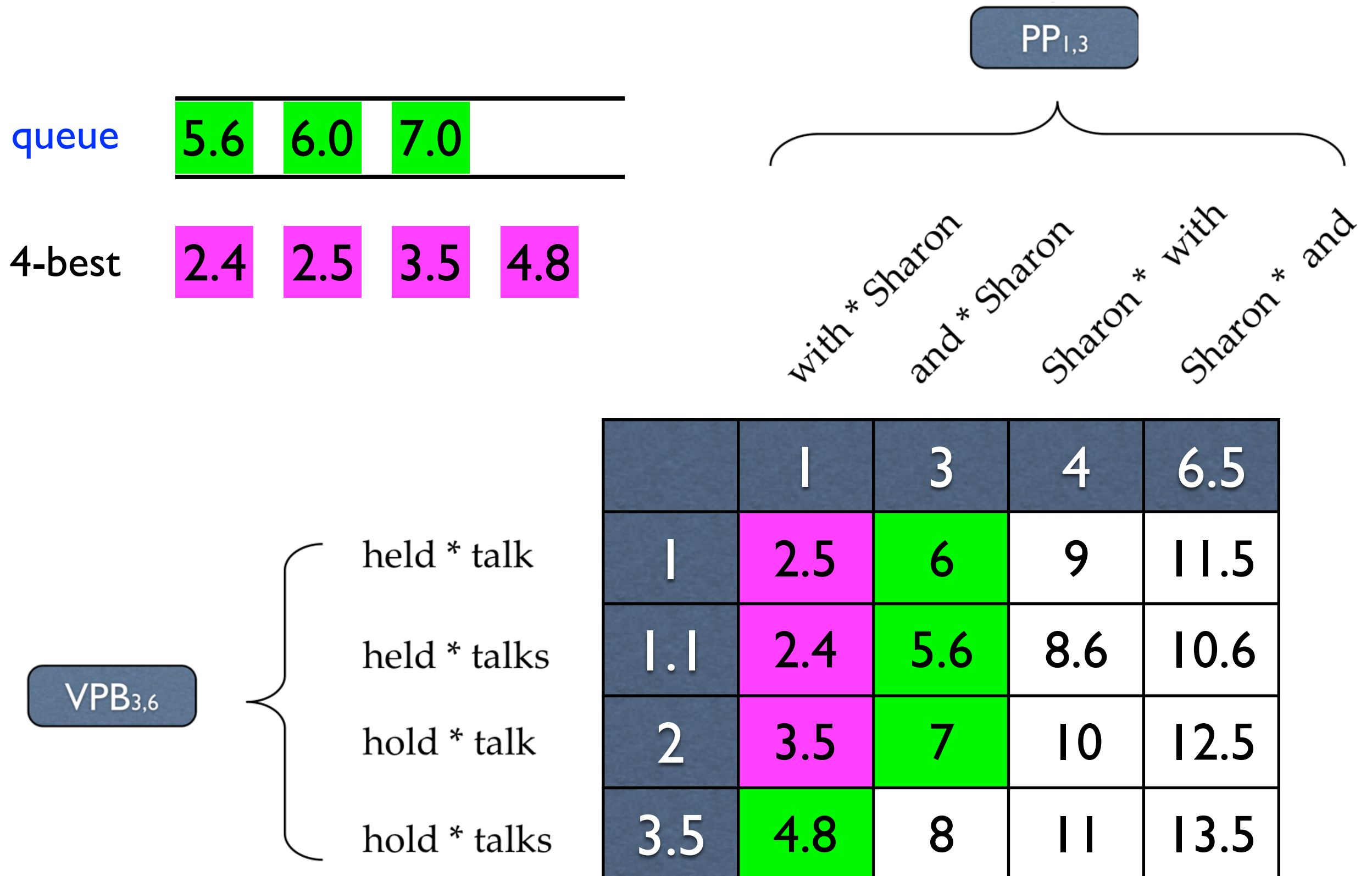
# Cube Pruning



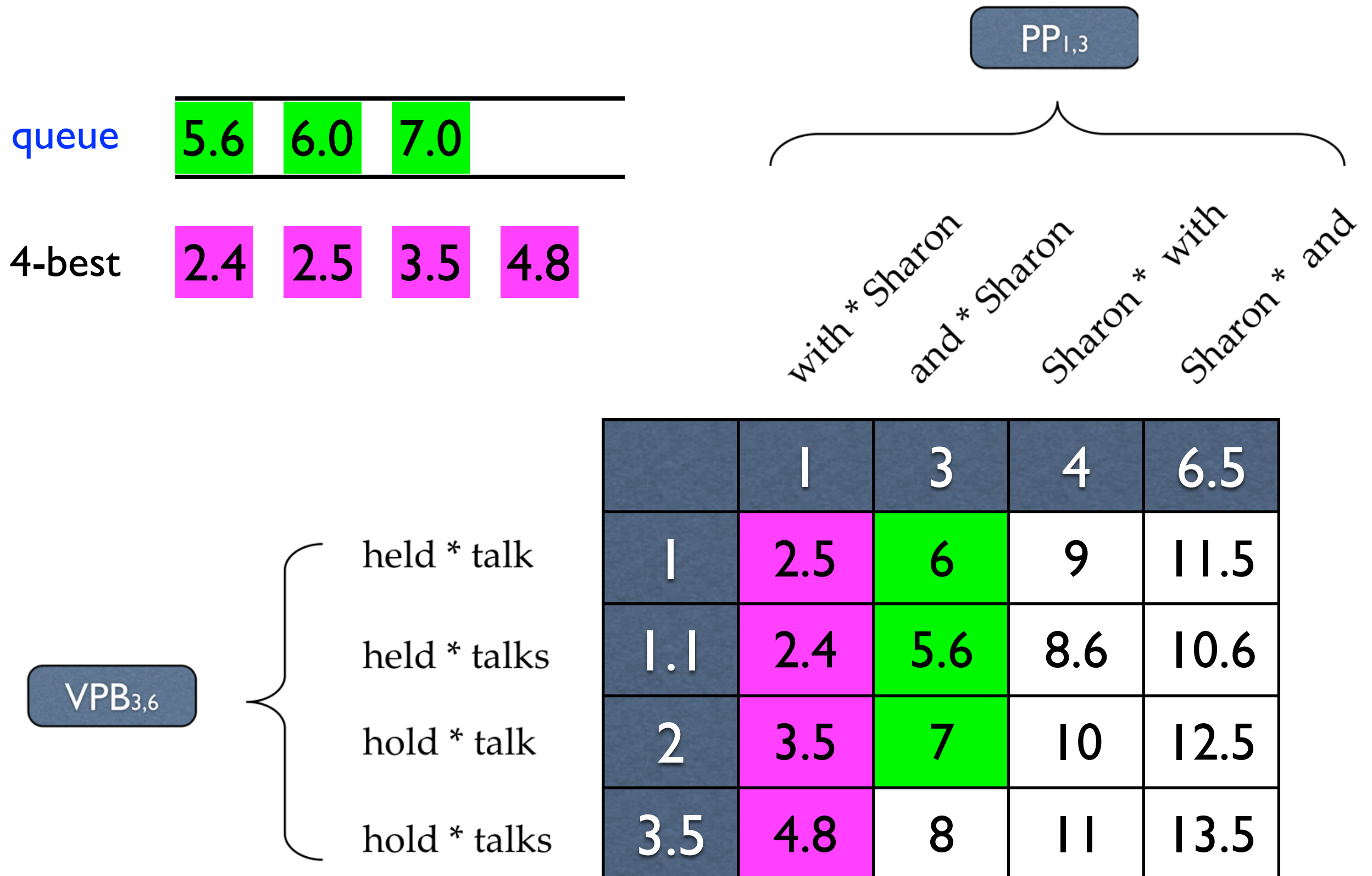
# Cube Pruning



# Cube Pruning

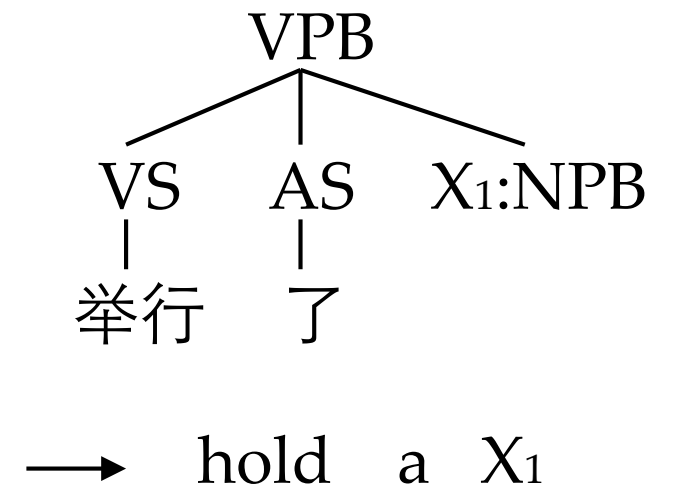
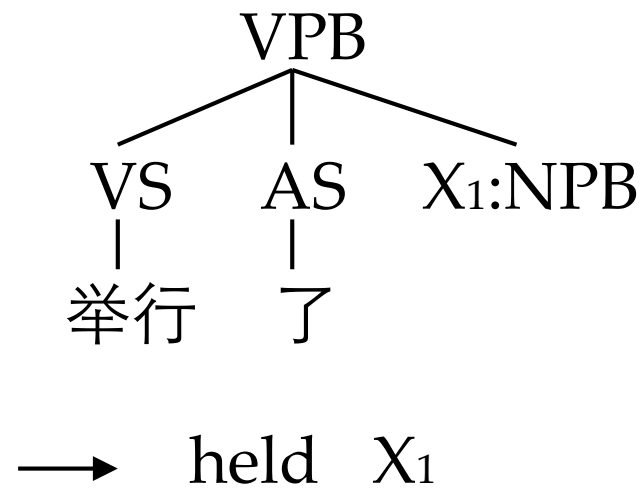
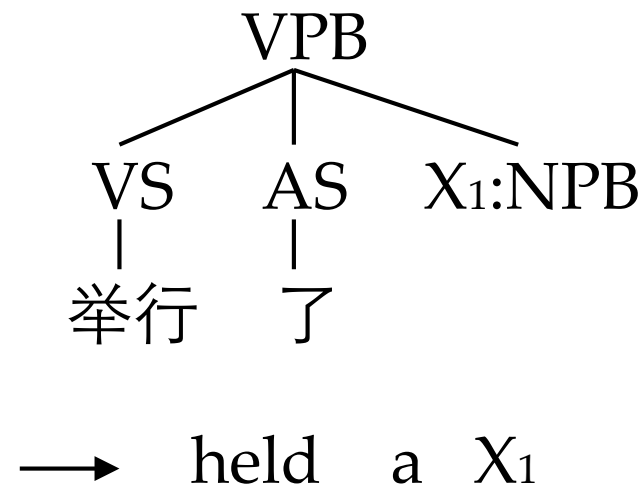


# Cube Pruning





# Cube Pruning with Rule Group



NPB<sub>5,6</sub>

talk talks meeting

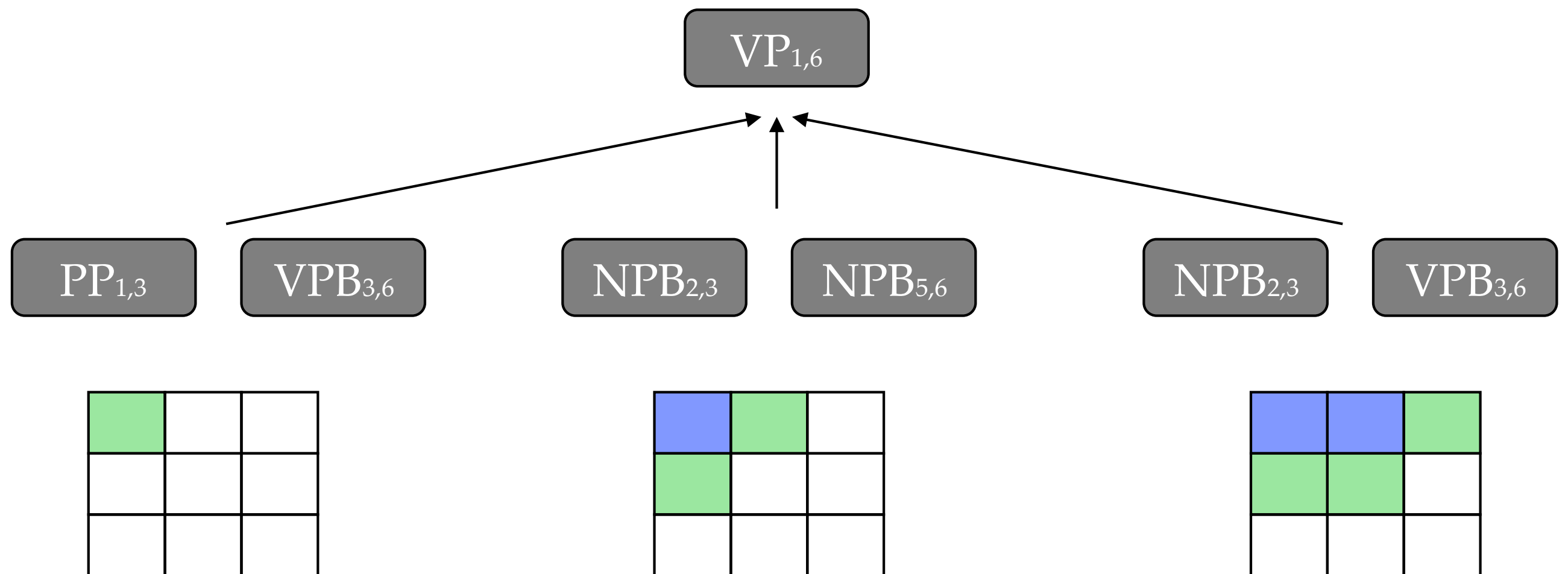
Group rules  
that have the  
same LHS

held a  $X_1$   
held  $X_1$   
hold a  $X_1$

	1.0	2.0	2.5
1.0	2.1	5.0	3.7
1.4	3.2	4.0	5.0
2.0	3.1	6.0	4.7



# Cube Pruning within Node



# Syntax-based MT

## SCFGs without linguistic syntax

*inverted transduction grammar*

*hierarchical phrase-based model*

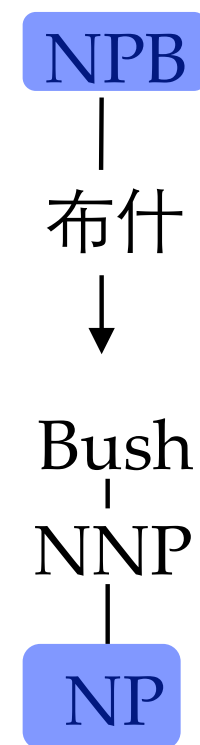
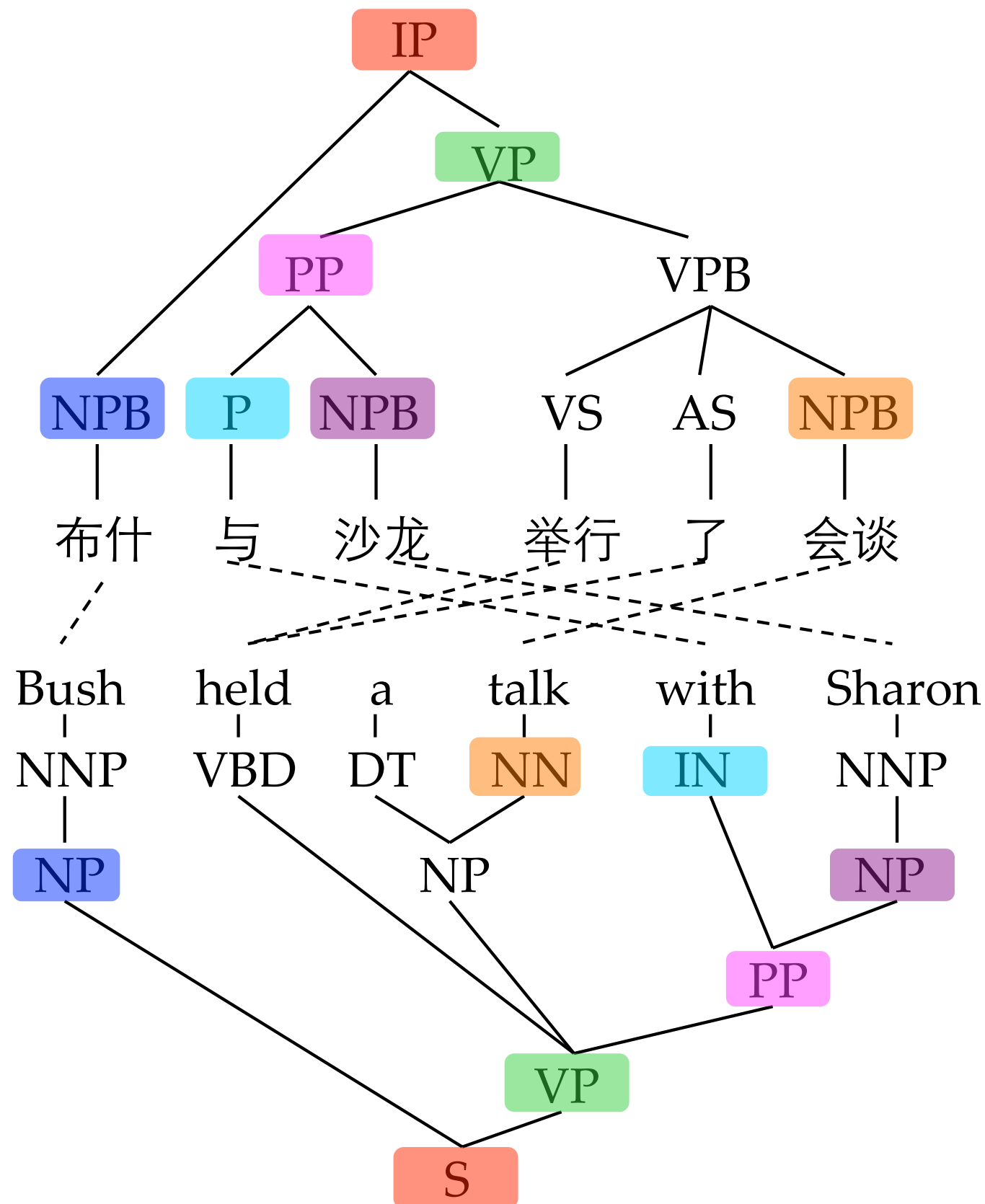
## STSGs with linguistic syntax

*string-to-tree*

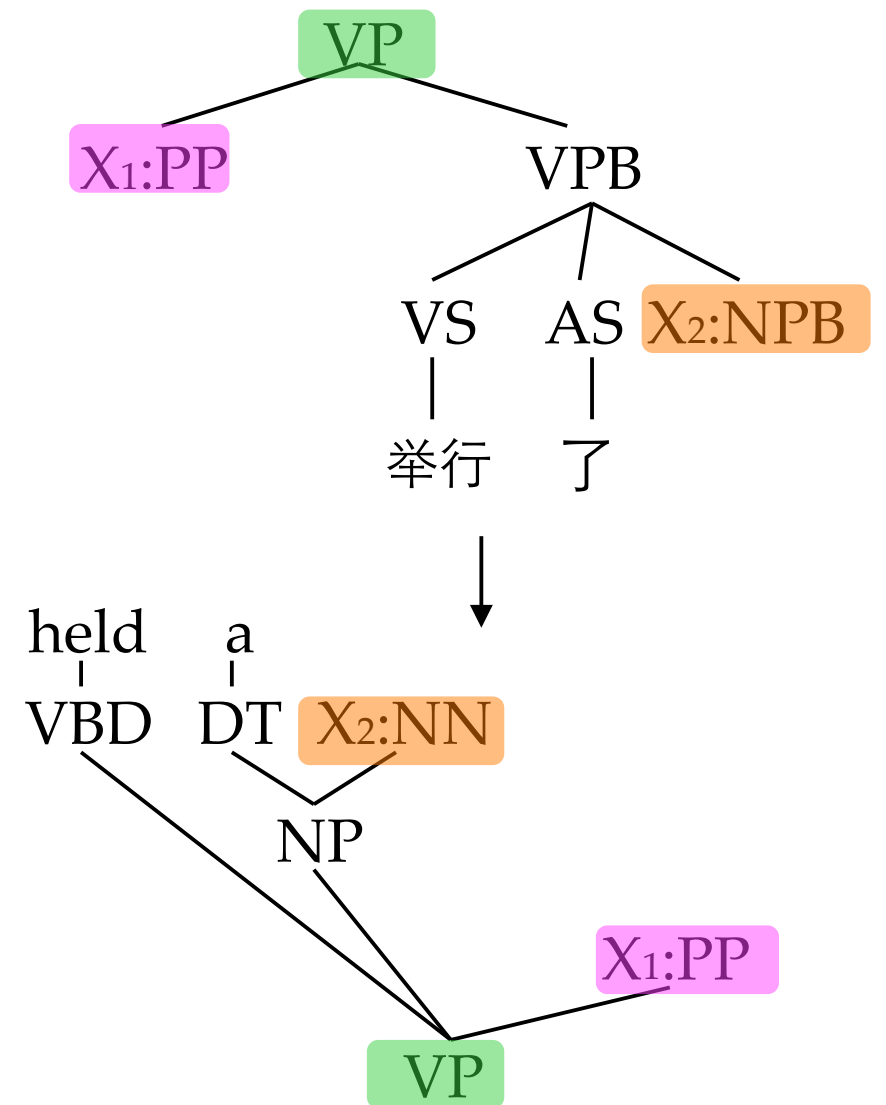
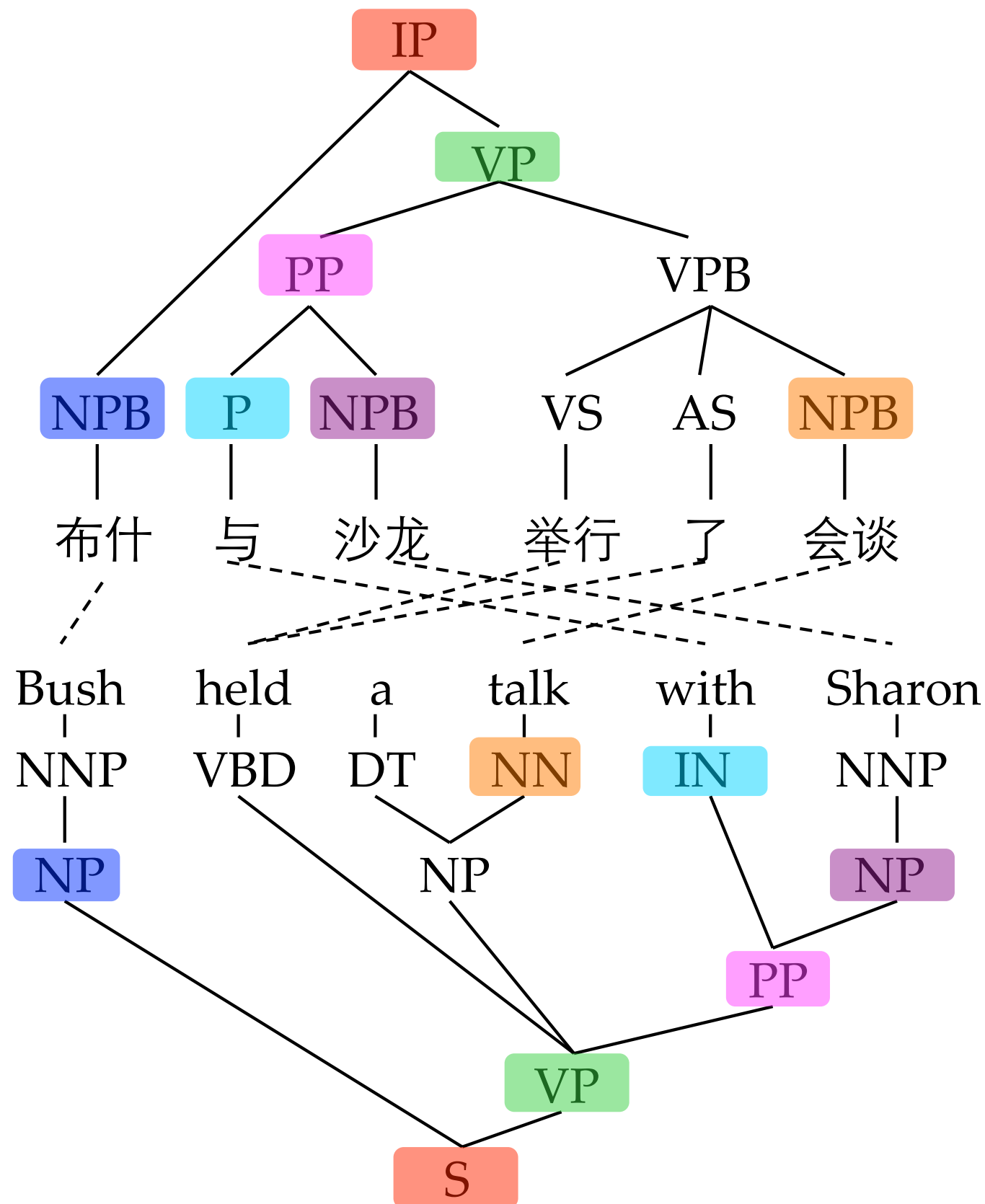
*tree-to-string*

*tree-to-tree*

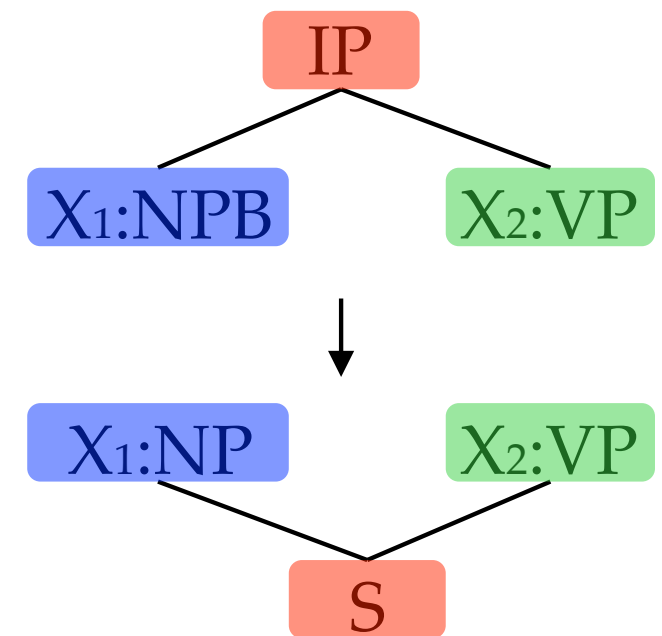
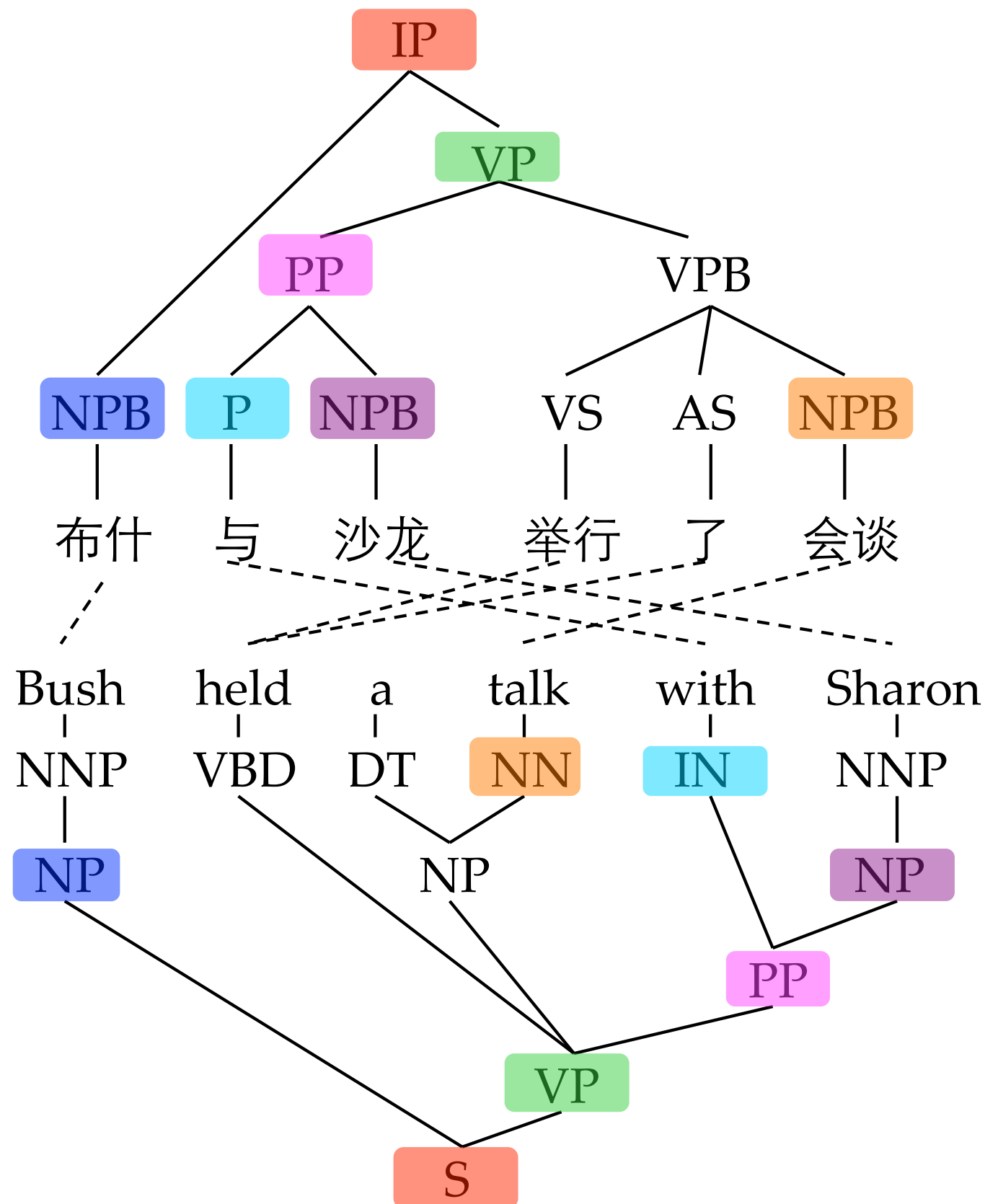
# Tree-to-Tree Translation



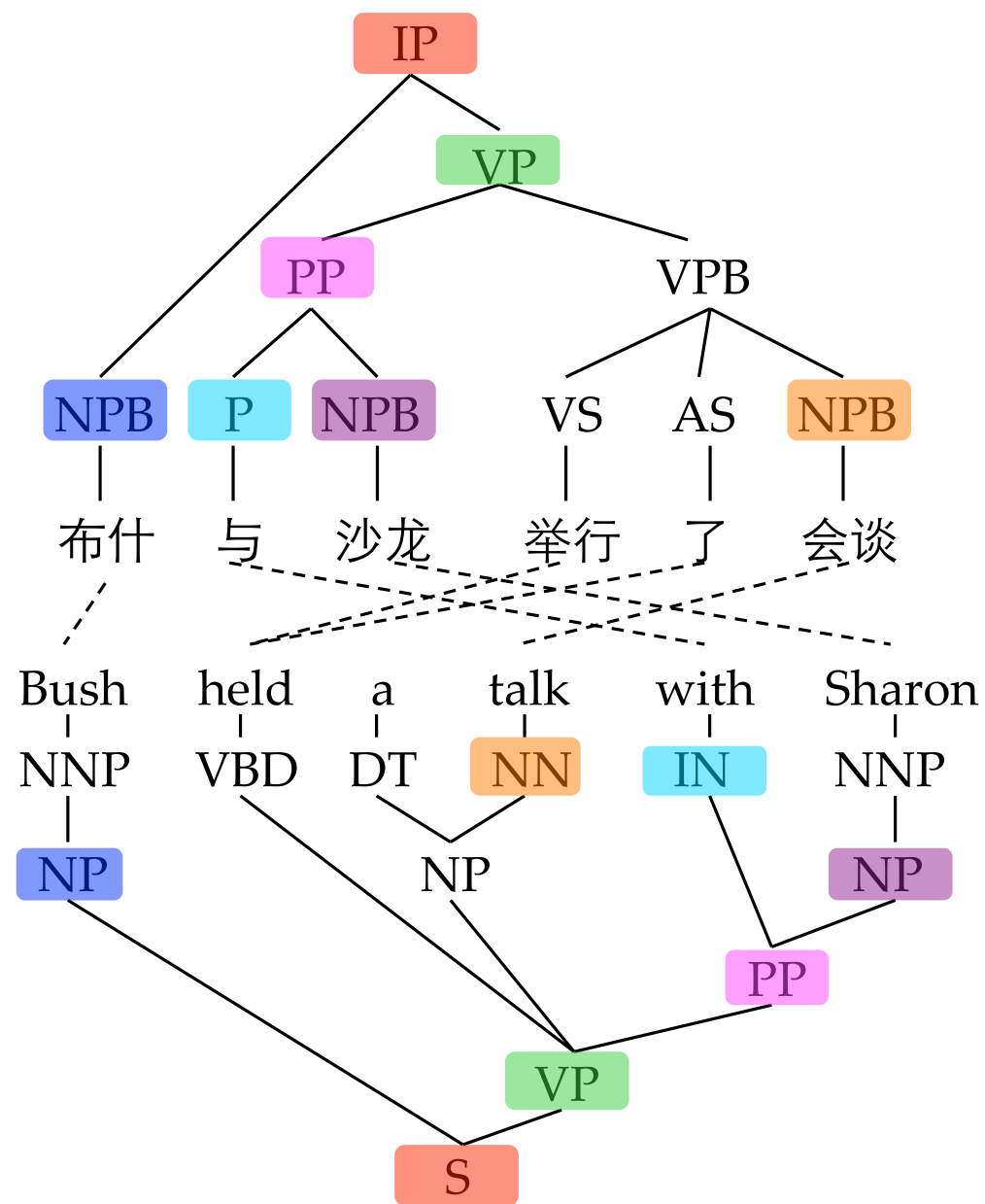
# Tree-to-Tree Translation



# Tree-to-Tree Translation



# Rule Coverage



phrase pair	s2s	t2s	s2t	t2t
(布什, Bush)	✓	✓	✓	✓
(与, with)	✓	✓	✓	✓
(沙龙, Sharon)	✓	✓	✓	✓
(会谈, talk)	✓	✓	✓	✓
(与 沙龙, with Sharon)	✓	✓	✓	✓
(举行 了, held)	✓	✗	✓	✗
(举行... 会谈, held ... talk)	✓	✓	✗	✗
(与 ... 会谈, held ... Sharon)	✓	✓	✓	✓
(布什 ... 会谈, Bush ... Sharon)	✓	✓	✓	✓
	100%	89%	89%	78%

# Rule Coverage

model	human	automatic
string-to-string	100%	100%
tree-to-string	78%	75%
string-to-tree	76%	72%
tree-to-tree	68%	60%

(Chiang, 2010)

# Summary of Syntax-based Models

source	target	model	examples
N/A	N/A	string-to-string	Chiang (2005) Wu (1997)
N/A	syntax	string-to-tree	Galley et al. (2006) Shen et al. (2008)
syntax	N/A	tree-to-string	Liu et al. (2006) Huang et al. (2006)
syntax	syntax	tree-to-tree	Eisner (2003) Zhang et al. (2008)



# Thanks

<http://nlp.csai.tsinghua.edu.cn/~ly/>