Relations and Realization in Syntax and Parsing

Columbia University
June 15, 2010

Reut Tsarfaty
The Department of Linguistics and Philology
Uppsala University

Statistical Parsing from a Typological Perspective

Statistical Parsing

How can we learn statistical parsing models from data?

- What are the units of generalization?
- How can we exploit evidence in the data?

A Typological Perspective

Statistical modeling in the face of cross-linguistic variation

- Which probabilistic models for which languages?
- Statistical modeling for parsing rich morphosyntax

Statistical Parsing from a Typological Perspective

Statistical Parsing

How can we learn statistical parsing models from data?

- What are the units of generalization?
- How can we exploit evidence in the data?

A Typological Perspective

Statistical modeling in the face of cross-linguistic variation

- Which probabilistic models for which languages?
- Statistical modeling for parsing rich morphosyntax

The Hypothesis

Different languages $\rightsquigarrow$ different realization $\rightsquigarrow$ different modeling

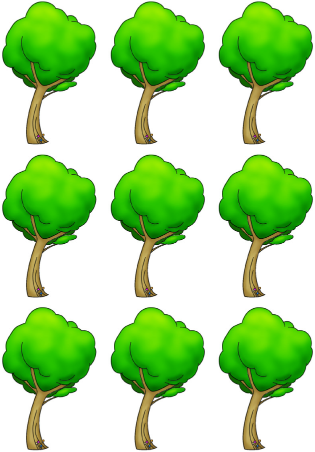
Statistical Parsing

Statistical Parsing

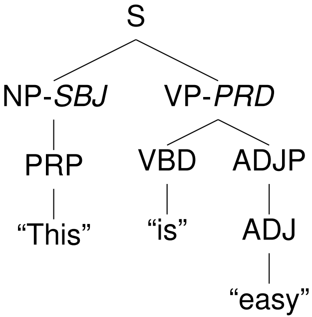
”This is easy”

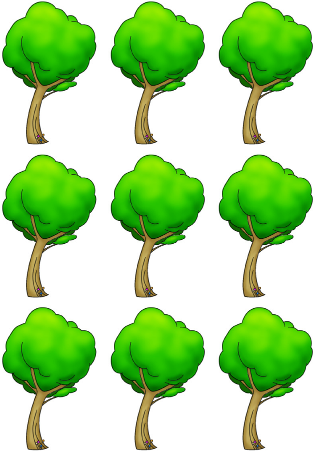
Statistical Parsing

"This is easy"

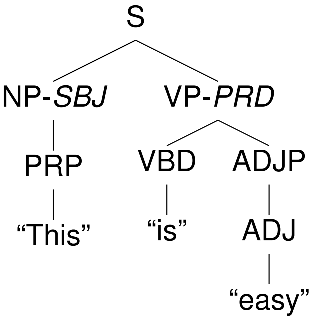


Constituency-Based Statistical Parsing





Constituency-Based Supervised Statistical Parsing



Model	Study	F-Score
Treebank Grammar	Charniak 1996	75
Head-Driven	Collins 1997	88.6
Discriminative Reranking	Collins 2000	89.7
Discriminative-Reranking	Johnson & Charniak 2005	91.0
Self-Training	McClosky 2006	92.1
State-Splits	Petrov et al 2007	90.1
Forest Reranking	Liang Huang 2008	91.7

Constituency-Based Supervised Statistical Parsing

And what about this?

將水煮開後才
使用。

And this?

إغلي الماء قبل استعماله

And this?

יש להרתיח את המים
לפני השימוש.

Language	Parser	F-Score
German	Rafferty & Manning 2008	79.2
Czech	Collins et al. 1999	79.3
Chinese	Levy & Manning 2003	78.8
Arabic	Maamouri, Bies & Kulick 2008	78.1
Hebrew	Tsarfaty & Sima'an 2007	74.4

And? ...

So What Is Going On?

Often Considered

- Corpora Size
  E.g., For *Chinese* (Bikel & Chiang 2000)
- Annotation Idiosyncrasies
  E.g., For *Arabic* (Maamouri, Bies & Kulick 2008, 2009)
- Evaluation Matters
  E.g., For *German* (Rehiben & van Genabith 2007, Kübler 2008)

So What Is Going On?

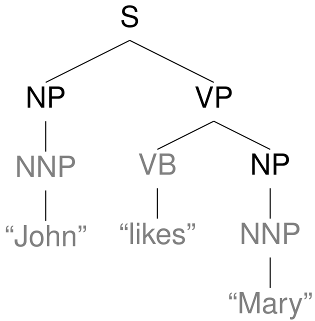
Often Considered

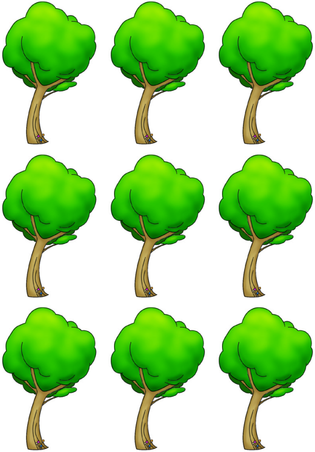
- Corpora Size
  E.g., For *Chinese* (Bikel & Chiang 2000)
- Annotation Idiosyncrasies
  E.g., For *Arabic* (Maamouri, Bies & Kulick 2008, 2009)
- Evaluation Matters
  E.g., For *German* (Rehiben & van Genabith 2007, Kübler 2008)

In This Talk

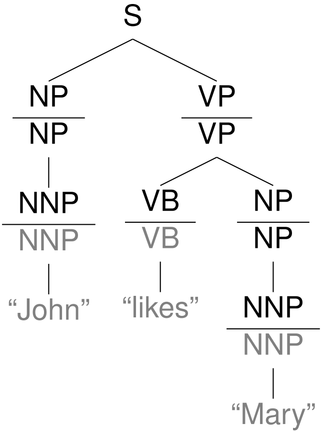
- Modeling Strategy
- Language Type

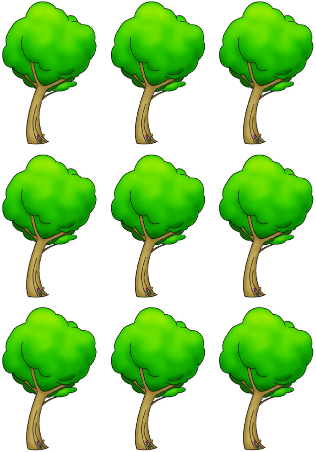
Modeling Strategies



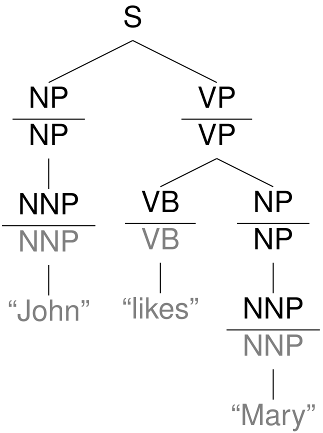


Modeling Strategies





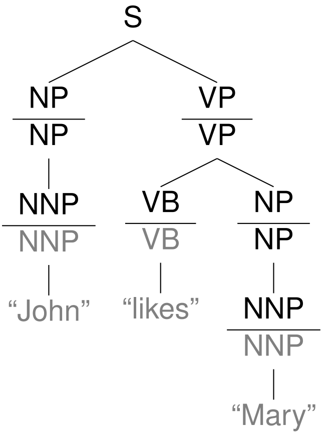
Modeling Strategies



$\Rightarrow$

P(NP VP|S)	1
P(NNP|NP)	1
P(VB NP|VP)	1
P("John"|NNP)	0.5
P("likes"|VB)	1
P("Mary"|NNP)	0.5

Modeling Strategies

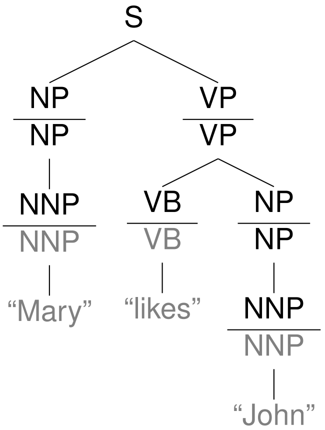


$\Rightarrow$

P(NP VP|S)	1
P(NNP|NP)	1
P(VB NP|VP)	1
P("John"|NNP)	0.5
P("likes"|VB)	1
P("Mary"|NNP)	0.5

$\Rightarrow$P("John likes Mary")$=$P(NP VP$|$S)$\times \ldots \times$P("Mary"$|$NNP)$= 0.25$

Modeling Strategies

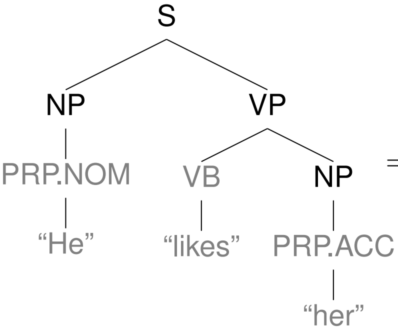


$\Leftarrow$

P(NP VP|S)	1
P(NNP|NP)	1
P(VBD NP|VP)	1
P("John"|NNP)	0.5
P("likes"|VB)	1
P("Mary"|NNP)	0.5

$\Rightarrow$P("Mary likes John")$=$P(NP VP$|$S)$\times$...$\times$P("Mary"$|$NNP)$=$ 0.25

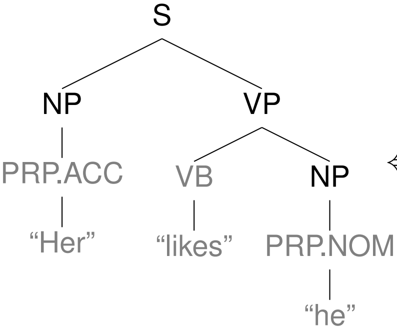
Modeling Strategies



$\Rightarrow$

P(NP VP|S)	1
P(PRP.NOM|NP)	0.5
P(PRP.ACC|NP)	0.5
P(VB NP|VP)	1
P("He"|PRP.NOM)	1
P("likes"|VB)	1
P("her"|PRP.ACC)	1

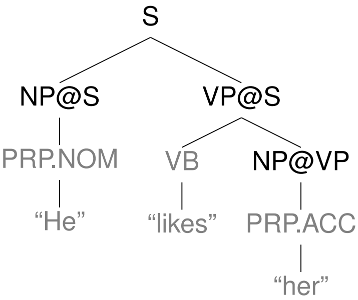
Modeling Strategies



P(NP VP|S)	1
P(PRP.NOM|NP)	0.5
P(PRP.ACC|NP)	0.5
P(VB NP|VP)	1
P("He"|PRP.NOM)	1
P("likes"|VB)	1
P("her"|PRP.ACC)	1

$$P(\text{"Her likes he"}) = P(\text{NP VP}|\text{S}) \times \ldots \times P(\text{"her"}|\text{PRP.ACC}) = 0.25$$

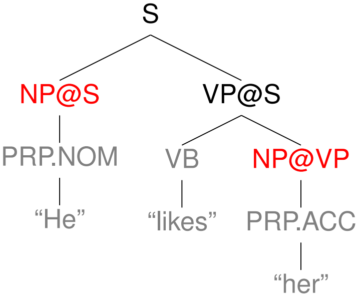
Example 1: Parent Encoding (Johnson 1998)



⇒

P(NP@S VP@S|S)	1
P(PRP.NOM |NP@S)	1
P(PRP.ACC |NP@VP)	1
P(VB NP@VP|VP@S)	1
P(”He”|PRP.NOM)	1
P(”likes”|VP)	1
P(”her”|PRP.ACC)	1

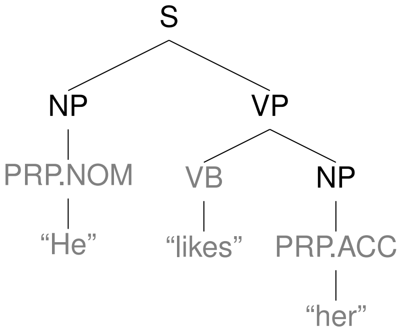
Example 1: Parent Encoding (Johnson 1998)



⇒

P(NP@S VP@S|S)	1
P(PRP.NOM |NP@S)	1
P(PRP.ACC |NP@VP)	1
P(VB NP@VP|VP@S)	1
P(”He”|PRP.NOM)	1
P(”likes”|VP)	1
P(”her”|PRP.ACC)	1

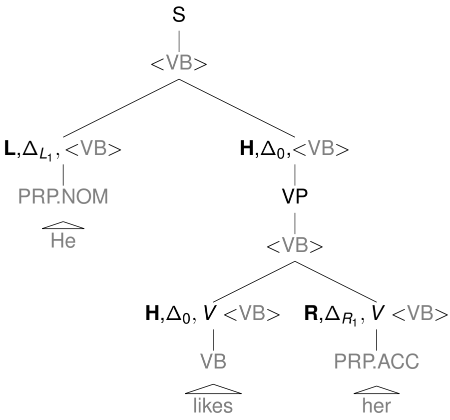
Example 2: Head-Driven Processes (Collins 1999)



$\Rightarrow$

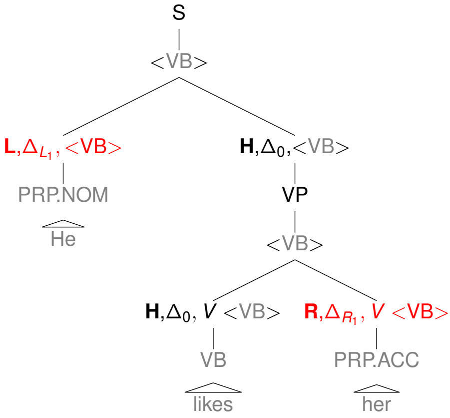
P(NP VP|S)	1
P(PRP.NOM|NP)	0.5
P(PRP.ACC|NP)	0.5
P(VB NP|VP)	1
P("He"|PRP.NOM)	1
P("likes"|VB)	1
P("her"|PRP.ACC)	1

Head-Driven Processes (Collins 1999)



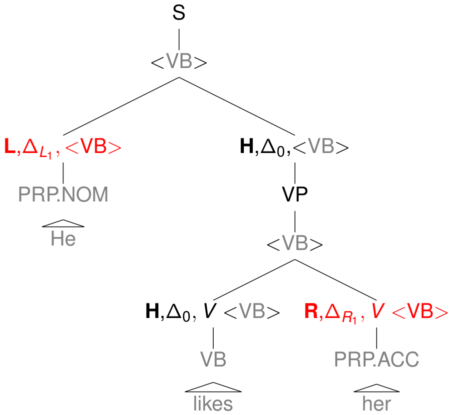
$\Rightarrow$	
$P(<VB>\|S)$	1
$P(\mathbf{L}\Delta_{L_1},\mathbf{H}\Delta_0\| <VB>,S)$	1
$P(PRP.NOM\|\mathbf{L},\Delta_{L_1},<VB>,S)$	1
$P(VP\|\mathbf{H},\Delta_0,<VB>,S)$	1
$P(<VB>\|VP)$	1
$P(PRP.ACC\|\mathbf{R},\Delta_{R_1},<VB>,S)$	1
$P(VB\|\mathbf{H},\Delta_0,<VB>,S)$	1
P("He"|PRP.NOM)	1
P("likes"|VB)	1
P("her"|PRP.ACC)	1

Head-Driven Processes (Collins 1999)



$\Rightarrow$	
$P(\langle VB\rangle \mid S)$	1
$P(\mathbf{L}\Delta_{L_1},\mathbf{H}\Delta_0 \mid \langle VB\rangle, S)$	1
$P(PRP.NOM \mid \mathbf{L},\Delta_{L_1},\langle VB\rangle,S)$	1
$P(VP \mid \mathbf{H},\Delta_0,\langle VB\rangle,S)$	1
$P(\langle VB\rangle \mid VP)$	1
$P(PRP.ACC \mid \mathbf{R},\Delta_{R_1},\langle VB\rangle,S)$	1
$P(VB \mid \mathbf{H},\Delta_0,\langle VB\rangle,S)$	1
P("He"|PRP.NOM)	1
P("likes"|VB)	1
P("her"|PRP.ACC)	1

Head-Driven Processes (Collins 1999)



$\Rightarrow$	
$P(<VB> \mid S)$	1
$P(\mathbf{L}\Delta_{L_1}, \mathbf{H}\Delta_0 \mid <VB>, S)$	1
$P(PRP.NOM \mid \mathbf{L}, \Delta_{L_1}, <VB>, S)$	1
$P(VP \mid \mathbf{H}, \Delta_0, <VB>, S)$	1
$P(<VB> \mid VP)$	1
$P(PRP.ACC \mid \mathbf{R}, \Delta_{R_1}, <VB>, S)$	1
$P(VB \mid \mathbf{H}, \Delta_0, <VB>, S)$	1
$P(\text{"He"} \mid PRP.NOM)$	1
$P(\text{"likes"} \mid VB)$	1
$P(\text{"her"} \mid PRP.ACC)$	1

Think: X-Bar Syntax!

Works amazingly well for English

Modeling Strategies

The Setup

- We are given a treebank and (often) a formal device
- We can learn different models reflecting different theories

The Question

- How can we learn a model that captures the best theory, as it is reflected in the treebank data?

The Data

The Data

Typological Dimensions of Variation

Basic Word-Order Typology

(Greenberg 1966, Mithun 1992)

The Data

Basic Word-Order Typology

Word-Order Type

The order in which a Subject, a Verb and an Object appear in a canonical, neutral, unmarked sentence (SVO, VSO, VOS, etc) (Greenberg 1963)

Word-Order Freeness

The order is pragmatically determined (Mithun 1992)

RIGID ———————————— FREE

The Data

Basic Word-Order Typology

Word-Order Type

The order in which a Subject, a Verb and an Object appear in a canonical, neutral, unmarked sentence (SVO, VSO, VOS, etc) (Greenberg 1963)

Word-Order Freeness

The order is pragmatically determined (Mithun 1992)

RIGID ——————————— FREE

Vietnames ——————————— Warlpiri

The Data

Typological Dimensions of Variation

Basic Word-Order Typology
(Greenberg 1966, Mithun 1992)

Morphological Typology
(Sapir 1921, Greenberg 1954)

The Data

Morphological Typology

Morphological Synthesis

Morpheme-to-word ratio:

ISOLATING ——————— POLYSYNTHETIC

Vietnamese ——————— Yu'pic

The Data

Morphological Typology

Morphological Synthesis

Morpheme-to-word ratio:

ISOLATING ——— POLYSYNTHETIC
Vietnamese ——— Yu’pic

Morphological Fusion

Ease of segmentation:

AGGLUTINATIVE ——— FUSIONAL
Turkish ——— Latin

The Data

Typological Dimensions of Variation

Basic Word-Order Typology
(Greenberg 1966, Mithun 1992)

Morphological Typology
(Sapir 1921, Greenberg 1954)

Nonconfigurationality
(Hale 1983, Austin and Bresnan 1996)

Nonconfigurationality as Misalignment

Nonconfigurationality as Misalignment

Predicate-Argument Relations

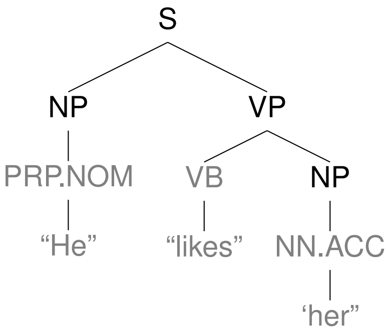
‘SBJ’ did ‘PRD’ to ‘OBJ’

Nonconfigurationality as Misalignment

Predicate-Argument Relations

‘SBJ’ did ‘PRD’ to ‘OBJ’

Syntactic Configuration

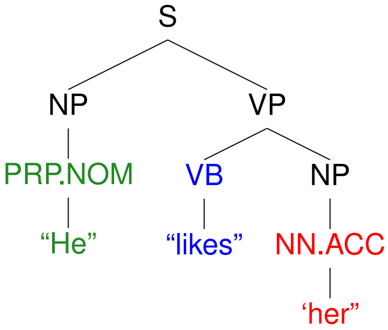


Nonconfigurationality as Misalignment

Predicate-Argument

'SBJ' did 'PRD' to 'OBJ'

Syntactic Configuration



Understanding Nonconfigurationality

Word-Order in Modern Hebrew

(1) a. dani natan et hamatana ledina
Dani gave ACC the-present to-Dina
"Dani gave the present to Dina" (SVO)

b. et hamatana natan dani ledina
ACC the-present gave Dani to-Dina
"Dani gave the present to Dina" (OVS)

c. natan dani et hamatana ledina
gave Dani ACC the-present to-Dina
"Dani gave the present to Dina" (VSO)

d. ledina natan dani et hamatana
to-dina gave Dani ACC the-present
"Dani gave the present to Dina" (VSO)

Argument Marking in Modern Hebrew (1:1)

Case-Assigning Prepositions

(2)

a. dani natan et hamatana ledina
Dani gave ACC DEF-present DAT-Dina

b. et hamatana natan dani ledina
ACC DEF-present gave Dani DAT-Dina

c. natan dani et hamatana ledina
gave Dani ACC DEF-present DAT-Dina

d. ledina natan dani et hamatana
DAT-dina gave Dani ACC DEF-present

Argument Marking in Modern Hebrew (1:many)

Differential Object-Marking

(3) a.

dani	natan	et	hamatana	ledina
Dani	gave	ACC	DEF-present	to-Dina

b.

et	hamatana	natan	dani	ledina
ACC	DEF-present	gave	Dani	to-Dina

c.

natan	dani	et	hamatana	ledina
gave	Dani	ACC	DEF-present	to-Dina

d.

ledina	natan	dani	et	hamatana
to-dina	gave	Dani	ACC	DEF-present

Argument Marking in Modern Hebrew (1:many)

Feature Spreading (Danon, 2007)

(4)

a.

dani	natan	[et	matnat	yom	hahuledet]	ledina	
Dani	gave	[ACC	present	day	DEF-birth]	to-Dina	

b.

[et	matnat	yom	hahuledet]	natan	dani	ledina	
[ACC	present	day	DEF-birth]	gave	Dani	to-Dina	

c.

natan	dani	[et	matnat	yom	hahuledet]	ledina	
gave	Dani	[ACC	present	day	DEF-birth]	to-Dina	

d.

ledina	natan	dani	[et	matnat	yom	hahuledet]	
to-dina	gave	Dani	[ACC	present	day	DEF-birth]	

Argument Marking in Modern Hebrew (1:m)

Agreement

(5)

a.

dani	natan	et	hamatana	ledina
Dani.MS	gave.3MS	ACC	DEF-present	DAT-Dina

b.

et	hamatana	natan	dani	ledina
ACC	DEF-present	gave.3MS	Dani.MS	DAT-Dina

c.

natan	dani	et	hamatana	ledina
gave.MS	Dani.3MS	ACC	DEF-present	DAT-Dina

d.

ledina	natan	dani	et	hamatana
DAT-dina	gave.3MS	Dani.MS	ACC	DEF-present

Argument Marking Modern Hebrew (many:1)

Clitics and Null Anaphors

(6) a.

dani	natan	et	hamatana	ledina
Dani.MS	gave.3MS	ACC	DEF-present	DAT-Dina

“Dani gave the present to Dina”

b.

natati	et	hamatana	ledina
gave.1S	ACC	DEF-present	DAT-Dina

“I gave the present to Dina”

c.

natatiha	ledina
gave.1S.ACC.3FS	DAT-Dina

“I gave it to Dina”

Language Types and Modeling Strategies

Recap:

CONFIGURATIONAL ———— NONCONFIGURATIONAL

1:1 ———————————————— m:n

Vietnamese > English > German > Hebrew > Warlpiri

Language Types and Modeling Strategies

Recap:

CONFIGURATIONAL ———— NONCONFIGURATIONAL

1:1 ———————————————— m:n

Vietnamese > English > German > Hebrew > Warlpiri

- Realization is the mapping of functions to forms
- Different Languages show different realization strategies
- Different realization strategies may require different models

Question:

How can we model generally complex form-function mappings?

Modeling Morphology

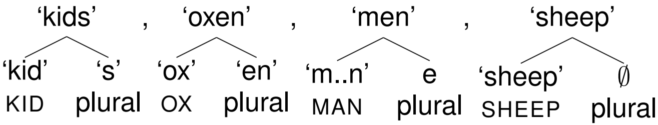
Modeling Morphology

Morpheme-Based Morphology (Bloomfield, 1933)

‘kids’		,	‘oxen’		,	‘men’		,	‘sheep’	
‘kid’	‘s’		‘ox’	‘en’		‘m..n’	e		‘sheep’	$\emptyset$
KID	plural		OX	plural		MAN	plural		SHEEP	plural

Modeling Morphology

Morpheme-Based Morphology (Bloomfield, 1933)



Morphological Exponence

- Simple Exponence (e.g., ‘s’ in ‘cats’)
- Cumulative Exponence (e.g., ‘s’ in ‘eats’)
- Extended Exponence (e.g., ‘i’,‘ren’ in ‘children’)

Modeling Morphology: Primitives and Processes

Lexical vs. Inferential Approaches

- Lexical:
  morphemes are primary, properties stored in the lexicon
- Inferential:
  properties are primary, forms are computed

Incremental vs. Realizational Approaches

- Incremental:
  morphemes/properties are accumulated incrementally
  (“monotonic” rules)
- Realizational:
  property-bundles are pre-condition for rule application
  (“spell-out” rules)

Modeling Morphology: A Taxonomy

	LEXICAL	INFERENTIAL
INCREMENTAL	Item & Arrangement (Bloomfield 1933) (Lieber 1992)	Item & Processes (Hocket 1954) (Steele 1995)
REALIZATIONAL	Distributed Morphology (Halle and Marantz 1993) Lexical Phonology	(Extended) Word & Paradigm (Matthews 1972), (Anderson 1992) (Stump 2001), (Blevins 2006)

Table: A Taxonomy of Models for Morphology (Stump 2001)

Word-and-Paradigm Morphology

Paradigmatic Organization

/EAT/	1Sing	2Sing	3Sing	1Pl	2Pl	3Pl
Past	1SingPast	2SingPast	3SingPast	1PlPast	2PlPast	3PlPast
Present	1SingPres	2SingPres	3SingPres	1PlPres	2PlPres	3PlPres
Perfect	1SingPerf	2SingPerf	3SingPerf	1PlPerf	2PlPerf	3PlPerf

Realization Rules

$$\begin{array}{c} /\mathrm{EAT}/ \\ +1\mathrm{SingPast} \\ | \\ \text{‘ate’} \end{array} , \quad \begin{array}{c} /\mathrm{EAT}/ \\ +3\mathrm{SingPast} \\ | \\ \text{‘ate’} \end{array} , \quad \begin{array}{c} /\mathrm{EAT}/ \\ +1\mathrm{SingPres} \\ | \\ \text{‘eats’} \end{array} , \quad \begin{array}{c} /\mathrm{EAT}/ \\ +3\mathrm{SingPres} \\ | \\ \text{‘eat’} \end{array}$$

The Proposal (I): “Lifting” the Terminology

Morphological Exponence

- Simple (e.g., PL $\rightsquigarrow$ ‘kids’)
- Cumulative (e.g. 3PER+SING $\rightsquigarrow$ ‘eats’)
- Distributed/Extended (e.g. PL $\rightsquigarrow$ ‘children’)

Morphosyntactic Exponence

- Simple (e.g., SBJ $\rightsquigarrow$ nominative)
- Cumulative (e.g., SBJ,PRD,OBJ $\rightsquigarrow$ clitics)
- Distributed/Extended (e.g., OBJ $\rightsquigarrow$ DOM, FS)

The Proposal (I): “Lifting” the Terminology

Morphological Exponence : Properties $\rightsquigarrow$ Words

- Simple (e.g., PL $\rightsquigarrow$ ‘kids’)
- Cumulative (e.g. 3PER+SING $\rightsquigarrow$ ‘eats’)
- Distributed/Extended (e.g. PL $\rightsquigarrow$ ‘children’)

Morphosyntactic Exponence : Relations $\rightsquigarrow$ Configurations

- Simple (e.g., SBJ $\rightsquigarrow$ nominative)
- Cumulative (e.g., SBJ,PRD,OBJ $\rightsquigarrow$ clitics)
- Distributed/Extended (e.g., OBJ $\rightsquigarrow$ DOM, FS)

The Proposal (II): Modeling Principles

Configurational vs. Relational Approaches

- Configurational:
  configurations are primary, relations are derived
- Relational:
  relations are primary, configurations are derived

Incremental vs. Realizational Approaches

- Incremental:
  Syntactic rules are monotonic
  (incrementally accumulate relations)
- Realizational:
  Syntactic rules define spellout
  (relations as precondition to realization)

The Proposal (III): A Taxonomy of Parsing Frameworks

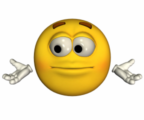
	Configurational	Relational
Incremental	Head-Driven Parsing 	Dependency Parsing
Realizational	Stochastic TAG, CCG	Relational-Realizational 

Table: A Taxonomy of Statistical Parsing Frameworks (Tsarfaty 2010)

The Proposal (IV): Relational-Realizational Modeling

S⟨PRED⟩ FEATS / ARG-ST	Affirmative	Interrogative	Imperative
intransitive	S_{affirm}+{SBJ,PRD}	S_{inter}+{SBJ,PRD}	S_{imper}+{SBJ,PRD}
transitive	S_{affirm}+{SBJ,PRD,OBJ}	S_{inter}+{SBJ,PRD,OBJ}	S_{imper}+{SBJ,PRD,OBJ}
ditransitive	$\boxed{S_{affirm}\text{+\{SBJ,PRD,OBJ,COM\}}}$	S_{inter}+{SBJ,PRD,OBJ,COM}	S_{imper}+{SBJ,PRD,OBJ,COM}

Figure: Paradigmatic Organization

$\boxed{S_{affirm}\text{+\{SBJ,PRD,OBJ,COM\}}}$

⟨	NP_{nom}	,	VB	,	$NP_{def.acc}$	,	NP_{dat}	⟩
	Dani		natan		et hamatana		ledina	
	Dani		gave		ACC-the-present		to-Dina	

$\boxed{S_{affirm}\text{+\{SBJ,PRD,OBJ,COM\}}}$

⟨	$NP_{def.acc}$	,	VB	,	NP_{nom}	,	NP_{dat}	⟩
	et hamatana		natan		Dani		ledina	
	ACC-the-present		gave		Dani		to-Dina	

Figure: Form-Function Separation

The Model

Relational-Realizational (RR) Parsing

(Tsarfaty, Sima'an and Scha 2008, 2009)

The Model

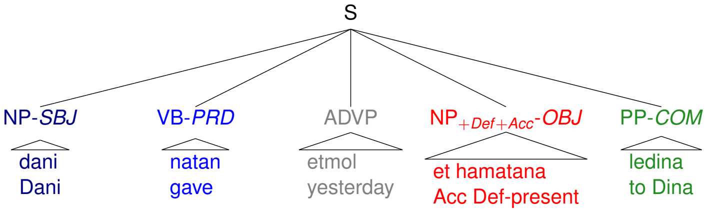
Relational-Realizational (RR) Parsing

(Tsarfaty, Sima'an and Scha 2008, 2009)

- Separate *Form* and *Function*
  - First Generate Grammatical *Relations*
  - Then Spell-out (Morpho)Syntactic Realization
- Separate Means of *Realization*
  - First Generate *Configuration*
  - Then *Morphosyntactic* Representation

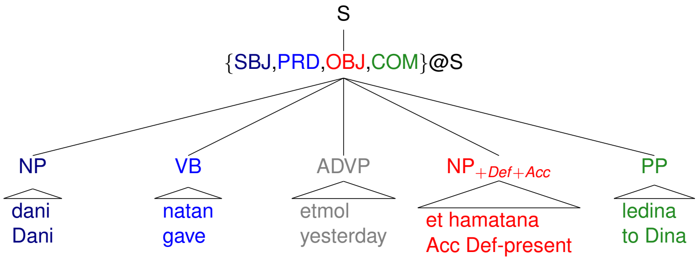
Tsarfaty et al. (2008, 2009)

Relational-Realizational (RR) Parsing



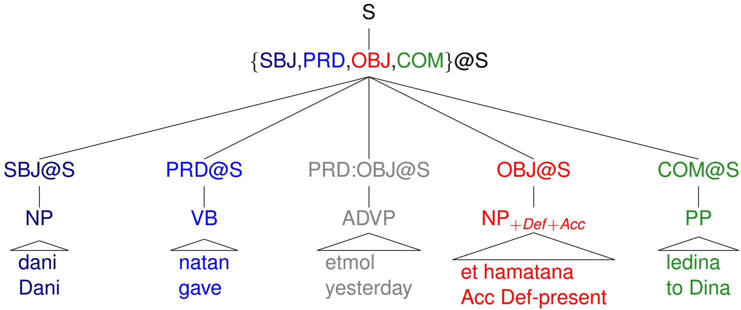
Tsarfaty et al. (2008, 2009)

Relational-Realizational (RR) Parsing



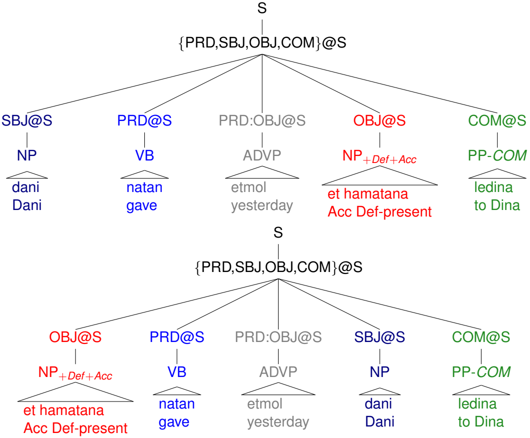
Tsarfaty et al. (2008, 2009)

Relational-Realizational (RR) Parsing



Tsarfaty et al. (2008, 2009)

Relational-Realizational (RR) Parsing



The Model Parameters

Projection:

$$P$$

$$\{gr_i\}_{i=1}^{n}@P$$

Configuration:

$$\{gr_i\}_{i=1}^{n}@P$$

$$gr_1@P \quad gr_1 : gr_2@P \quad \ldots \quad gr_n@P$$

Realization:

$$gr_1@P \quad gr_1 : gr_2@P \quad \ldots \quad gr_n@P$$

$$C_1 \quad ..C_{1:2_i}.. \quad C_n$$

The Probabilistic Model

The RR Probabilities:

$$
\begin{array}{ll}
\multicolumn{2}{l}{\mathbf{P}_{\mathbf{RR}}(r) =} \\
\textit{Projection} & \mathbf{P}_{\mathbf{p}}(\{gr_i\}_{i=1}^{n} | P) \times \\
\textit{Configuration} & \mathbf{P}_{\mathbf{c}}(\langle gr_0 : gr_1, g_1, \ldots \rangle | \{gr_i\}_{i=1}^{n}, P) \times \\
\textit{Realization} & \prod_{i=1}^{n} \mathbf{P}_{\mathbf{r_1}}(C_i | gr_i, P) \times \\
 & \mathbf{P}_{\mathbf{r_2}}(\langle C_{0_1}, \ldots, C_{0_{m_0}} \rangle | gr_0 : gr_1, P) \times \\
 & \prod_{i=1}^{n} \mathbf{P}_{\mathbf{r_2}}(\langle C_{i_1}, \ldots, C_{i_{m_i}} \rangle | gr_i : gr_{i+1}, P)
\end{array}
$$

The RR Parser:

$$\pi^* = argmax_\pi P(\pi) = argmax_\pi \prod_{r \in \pi} \mathbf{P}_{\mathbf{RR}}(r)$$

Application I: Parsing Modern Hebrew

A Taxonomy of PCFG-based Parsers

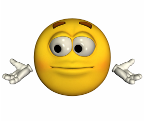
	CONFIGURATIONAL	RELATIONAL
INCREMENTAL	Head-Driven Parsing (Collins 1999) 	
REALIZATIONAL		Relational-Realizational (Tsarfaty et al. 2009) 

Table: A Taxonomy of PCFG-Based Parsing Frameworks

A Taxonomy of PCFG-based Parsers

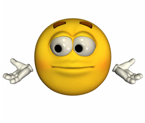
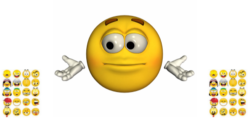
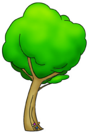
	Configurational	Relational
Incremental	Head-Driven Parsing (Collins 1999) 	Dependency Parsing (Collins 1999 enhanced) 
Realizational	Flattened Trees (Johnson 1998) 	Relational-Realizational (Tsarfaty et al. 2009) 

Table: A Taxonomy of PCFG-Based Parsing Frameworks

Case Study: Differential Object-Marking

Data

The Modern Hebrew Treebank v2, head annotated.
6500 sentences, 500/5500/500 dev/train/test split

Models

- Grammatical Functions: PRD, SBJ, OBJ, COM, CNJ
- Morphological Splits: **PoS/Def/Acc**
- Conditioning Context: Horizontal/Vertical

Estimation

Relative Frequency + Simple Unknown Words Smoothing

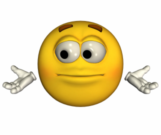
Parsing

Exhaustive Viterbi Parsing (using BitPar, Schmid 2004)

Evaluation

PARSEVAL (i) Overall, and (ii) Per Category Evaluation

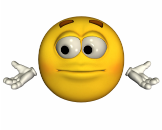
Overall Results



74.66/74.35 (7385)	73.52/74.84 (21399)	**76.32/76.51 (13618)**

Overall Results

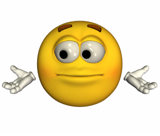
74.66/74.35
(7385)



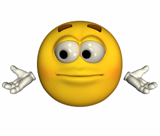
73.52/74.84
(21399)

76.32/76.51
(13618)

A Relational-Incremental Model

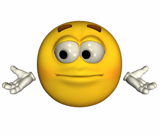


73.52/74.84
(21399)

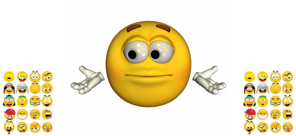


76.32/76.51
(13618)

A Relational-Incremental Model



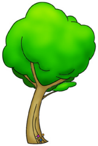
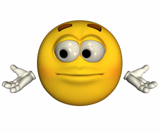
73.52/74.84
(21399)



72.84/74.62
(16460)

76.32/76.51
(13618)

Results Per Category


NP	77.39 / 74.32	77.94 / 73.75	**78.96** / **76.11**
PP	71.78 / 71.14	71.83 / 69.24	**74.4** / **72.02**
SBAR	55.73 / 59.71	53.79 / 57.49	**57.97** / **61.67**
ADVP	71.37 / 77.01	72.52 / 73.56	**73.57** / **77.59**
ADJP	**79.37** / **78.96**	78.47 / 77.14	78.69 / 78.18
S	**73.25** / **79.07**	71.07 / 76.49	72.37 / 78.33
SQ	36.00 / **32.14**	30.77 / 14.29	**55.56** / 17.86
PREDP	36.31 / 39.63	**44.74** / 39.63	44.51 / **46.95**

Take Home

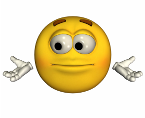
	Configurational	Relational
Incremental	Head-Driven Parsing 	Dependency Parsing
Realizational	Stochastic TAG, CCG	Relational-Realizational 

Table: A Taxonomy of Statistical Parsing Frameworks (Tsarfaty 2010)

Application II: Probabilistic Universal Grammar

Basic Word-Order Parameter:

$P(<\textit{configuration}>|\{\text{SBJ,PRD,OBJ}\}@\text{S})$

Probability	*Configuration*	tri-	bi-
0.2%	OBJ □ SUBJ PRD	OSV	OV
0.2%	PRD OBJ SBJ □	VOS	VO
0.2%	□ PRD OBJ □ SBJ □	VOS	VO
0.2 %	PRD SBJ □ OBJ □	VSO	VO
0.4 %	□ PRD □ SBJ □ OBJ □	VSO	VO
0.6 %	OBJ □ PRD SBJ □	OVS	OV
0.8 %	OBJ PRD □ SBJ □	OVS	OV
1 %	□ PRD □ SBJ OBJ □	VSO	VO
1.3%	SBJ □ PRD OBJ □	SVO	VO
1.7%	□ PRD OBJ SBJ □	VOS	VO
1.7%	□ SBJ PRD □ OBJ □	SVO	VO
3%	OBJ PRD SBJ □	OVS	OV
3.7%	□ PRD SBJ □ OBJ □	VSO	VO
4.1%	SBJ □ PRD □ OBJ □	SVO	VO
6.5%	□ SBJ PRD OBJ □	SVO	VO
10.3%	SBJ □ PDR OBJ □	SVO	VO
12.3%	□ PRD SBJ OBJ □	VSO	VO
15.6%	SBJ PRD □ OBJ □	SVO	VO
35.3%	SBJ PRD OBJ □	SVO	VO

Application II: Probabilistic Universal Grammar

Differential Object-Marking Parameter:
P($<$ *morphosyntactic representation* $>|$ OBJ@S)

Probability	Realization
5.8%	NP.DEF.ACC$\langle$*PRP*$\rangle$@*S*
6.5%	NP.DEF.ACC$\langle$*NNT*$\rangle$@*S*
6.7%	NP.DEF.ACC$\langle$*NN.DEF*$\rangle$@*S*
7.4%	NP.DEF.ACC$\langle$*NNP*$\rangle$@*S*
8.8%	NP$\langle$*NNT*$\rangle$@*S*
14.7%	NP.DEF.ACC$\langle$*NN*$\rangle$@*S*
43.5%	NP.$\langle$*NN*$\rangle$@*S*

Towards Computational Typology?

Can we Use the RR parameters to...

- Quantify Intra-Language Variation?
- Quantify Cross-Linguistic Variation?
- Learn Parameters Settings from Data?
- Quantify Nonconfigurationality?

RRRecap

- Languages are different!
  $\rightsquigarrow$Modeling strategies should accommodate differences
- Nonconfigurational languages are not configurational!
  $\rightsquigarrow$ Modeling strategies should account for misalignments
- Modeling Morphology vary in underlying assumptions
  $\rightsquigarrow$ Inferential-Realizational approaches model m:n mapping
- Modeling Morphosyntax meets similar considerations
  $\rightsquigarrow$ Relational-Realizational modeling allows for misalignments

Thank You!

Questions?

Let's Try it for Different Languages!

For more Information

Relational-Realizational Parsing

Reut Tsarfaty, University of Amsterdam

PhD Manuscript, 2010