Relations and Realization in Syntax and Parsing

Columbia University
June 15, 2010

Reut Tsarfaty
The Department of Linguistics and Philology
Uppsala University
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 $\leadsto$ different realization $\leadsto$ different modeling
Statistical Parsing
Statistical Parsing

"This is easy"
Statistical Parsing

"This is easy"
Constituency-Based Statistical Parsing

```
S
  NP-SBJ  VP-PRD
    PRP    VBD  ADJP
      "This"  "is"  ADJ
          ADJ
              "easy"
```
Constituency-Based Supervised Statistical Parsing

**Model** | **Study** | **F-Score**  
---|---|---  
Treebank | Charniak 1996 | 75  
Grammar | 1996 |  
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?

And this?

And this?

And this?

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

S
  /\  
 NP  VP
  / \  /
 NNP VB NP
 |   |  |
 "John" "likes" "Mary"
Modeling Strategies

S

NP
  NP
    NNP
      "John"

VP
  VP
    VB
      "likes"

NP
  NNP
    NNP
      "Mary"
Modeling Strategies

\[
S \rightarrow NP \quad VP \\
\quad NP \quad NP \\
\quad NNP \quad VB \\
\quad "John" \quad "likes" \\
\quad NNP \\
\quad NNP \\
\quad "Mary"
\]

\[
P(NP \ VP | S) = 1 \\
P(NNP | NP) = 1 \\
P(VB NP | VP) = 1 \\
P("John" | NNP) = 0.5 \\
P("likes" | VB) = 1 \\
P("Mary" | NNP) = 0.5
\]
Modeling Strategies

⇒ P("John likes Mary") = P(NP VP|S) \times ... \times P("Mary"|NNP) = 0.25
Modeling Strategies

\[ P(\text{"Mary likes John"}) = P(\text{NP VP}|S) \times \ldots \times P(\text{"Mary"}|\text{NNP}) = 0.25 \]
Modeling Strategies

```
S
   NP
     PRP.NOM
       "He"
     VB
     NP
       "likes"
     PRP.ACC
       "her"

⇒

P(NP VP|S) 1
P(PPR.NOM|NP) 0.5
P(PPR.ACC|NP) 0.5
P(VB NP|VP) 1
P("He"|PPR.NOM) 1
P("likes"|VB) 1
P("her"|PPR.ACC) 1
```
Modeling Strategies

P(“Her likes he”) = P(NP VP | S) × ... × P(“her” | PRP.ACC) = 0.25
Example 1: Parent Encoding (Johnson 1998)

```
S
  NP@S
    PRP.NOM
      "He"
  VP@S
    VB
      "likes"
    NP@VP
      PRP.ACC
        "her"

⇒

<table>
<thead>
<tr>
<th></th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(NP@S VP@S</td>
<td>S)</td>
</tr>
<tr>
<td>P(PRP.NOM</td>
<td>NP@S)</td>
</tr>
<tr>
<td>P(PRP.ACC</td>
<td>NP@VP)</td>
</tr>
<tr>
<td>P(VB NP@VP</td>
<td>VP@S)</td>
</tr>
<tr>
<td>P(&quot;He&quot;</td>
<td>PRP.NOM)</td>
</tr>
<tr>
<td>P(&quot;likes&quot;</td>
<td>VP)</td>
</tr>
<tr>
<td>P(&quot;her&quot;</td>
<td>PRP.ACC)</td>
</tr>
</tbody>
</table>
```
Example 1: Parent Encoding (Johnson 1998)

\[
\begin{align*}
S &\quad \text{NP@S} \\
\text{PRP.NOM} &\quad \text{VP@S} \\
"He" &\quad \text{VB} \\
"likes" &\quad \text{NP@VP} \\
\text{PRP.ACC} &\quad \text{"her"}
\end{align*}
\]

\[\Rightarrow\]

\[
\begin{align*}
P(\text{NP@S VP@S}|S) &\quad 1 \\
P(\text{PRP.NOM}|\text{NP@S}) &\quad 1 \\
P(\text{PRP.ACC}|\text{NP@VP}) &\quad 1 \\
P(\text{VB NP@VP}|\text{VP@S}) &\quad 1 \\
P(\text{"He"}|\text{PRP.NOM}) &\quad 1 \\
P(\text{"likes"}|\text{VP}) &\quad 1 \\
P(\text{"her"}|\text{PRP.ACC}) &\quad 1
\end{align*}
\]
Example 2: Head-Driven Processes (Collins 1999)

```
PRP.NOM       VP
  "He"       VB      NP
            "likes" PRP.ACC
                  "her"

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)

\[
\begin{align*}
S & \rightarrow <\text{VB}> \\
& \rightarrow L,\Delta_L, <\text{VB}> \\
& \rightarrow \text{PRP.NOM} \rightarrow \text{He} \\
& \rightarrow H,\Delta_0, <\text{VB}> \\
& \rightarrow \text{VP} \\
& \rightarrow <\text{VB}> \\
& \rightarrow H,\Delta_0, V <\text{VB}> \\
& \rightarrow R,\Delta_R, V <\text{VB}> \\
& \rightarrow \text{VB} \rightarrow \text{likes} \\
& \rightarrow \text{PRP.ACC} \rightarrow \text{her}
\end{align*}
\]

\[\Rightarrow\]

\[
\begin{align*}
P(<\text{VB}>|S) & \quad 1 \\
P(L\Delta_L, H\Delta_0|<\text{VB}>, S) & \quad 1 \\
P(\text{PRP.NOM}|L,\Delta_L,<\text{VB}>, S) & \quad 1 \\
P(\text{VP}|H,\Delta_0,<\text{VB}>, S) & \quad 1 \\
P(<\text{VB}>|\text{VP}) & \quad 1 \\
P(\text{PRP.ACC}|R,\Delta_R,<\text{VB}>, S) & \quad 1 \\
P(\text{VB}|H,\Delta_0,<\text{VB}>, S) & \quad 1 \\
P(\text{"He"}|\text{PRP.NOM}) & \quad 1 \\
P(\text{"likes"}|\text{VB}) & \quad 1 \\
P(\text{"her"}|\text{PRP.ACC}) & \quad 1
\end{align*}
\]
Head-Driven Processes (Collins 1999)

⇒

\[
\begin{align*}
P(<\text{VB}>|S) & \quad 1 \\
P(L\Delta_{L_1},H\Delta_0| <\text{VB}>,S) & \quad 1 \\
P(\text{PRP.NOM}|L\Delta_{L_1},<\text{VB}>,S) & \quad 1 \\
P(VP|H,\Delta_0,<\text{VB}>,S) & \quad 1 \\
P(<\text{VB}>|VP) & \quad 1 \\
P(\text{PRP.ACC}|R,\Delta_{R_1},<\text{VB}>,S) & \quad 1 \\
P(VB|H,\Delta_0,<\text{VB}>,S) & \quad 1 \\
P(\text{PRP.NOM}) & \quad 1 \\
P(\text{VB}) & \quad 1 \\
P(\text{PRP.ACC}) & \quad 1 \\
\end{align*}
\]

Think: X-Bar Syntax!
Works amazingly well for English
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

```
S
   /\   /
  NP   VP
   /\   /
  PRP.NOM  VB  NP
  |     |     |
  "He" "likes" 'her'
```
Nonconfigurationality as Misalignment

Predicate-Argument

‘SBJ’ did ‘PRD’ to ‘OBJ’

Syntactic Configuration

```
S
  NP
  |  VP
  |   PRP.NOM
  |    “He”
  VB
    “likes”
  NP
    NN.ACC
        ‘her’
```
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)
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
(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”
Recap:

**CONFIGURATIONAL** ——— **NONCONFIGURATIONAL**

1:1 ———— m:n

Vietnamese > English > German > Hebrew > Warlpiri
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

Morpheme-Based Morphology (Bloomfield, 1933)
- 'kids'
- 'kid'
- KID
- 's'
- plural
- 'oxen'
- 'ox'
- OX
- 'en'
- plural
- 'men'
- e
- plural
- 'sheep'
- 'sheep'
- ∅
- plural

Morphological Exponence
▶ Simple Exponence (e.g., 's' in 'cats')
▶ Cumulative Exponence (e.g., 's' in 'eats')
▶ Extended Exponence (e.g., 'i','ren' in 'children')
Morpheme-Based Morphology (Bloomfield, 1933)

'kids', 'oxen', 'men', 'sheep'

'kid', 's', 'ox', 'en', 'm..n', 'e', 'sheep', ∅

KID plural OX plural MAN plural SHEEP plural
Modeling Morphology

Morpheme-Based Morphology (Bloomfield, 1933)

- ‘kid’ ‘s’ ‘ox’ ‘en’ ‘m..n’ e ‘sheep’ ∅
- KID plural OX plural MAN plural SHEEP plural

Morphological Exponence

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

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

Morphological Exponence
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

<table>
<thead>
<tr>
<th></th>
<th><strong>LEXICAL</strong></th>
<th><strong>INFERENTIAL</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>INCREMENTAL</strong></td>
<td>Item &amp; Arrangement</td>
<td>Item &amp; Processes</td>
</tr>
<tr>
<td></td>
<td>(Bloomfield 1933)</td>
<td>(Hocket 1954)</td>
</tr>
<tr>
<td></td>
<td>(Lieber 1992)</td>
<td>(Steele 1995)</td>
</tr>
<tr>
<td><strong>REALIZATIONAL</strong></td>
<td>Distributed Morphology</td>
<td>(Extended) Word &amp; Paradigm</td>
</tr>
<tr>
<td></td>
<td>(Halle and Marantz 1993)</td>
<td>(Matthews 1972), (Anderson 1992)</td>
</tr>
<tr>
<td></td>
<td>Lexical Phonology</td>
<td>(Stump 2001), (Blevins 2006)</td>
</tr>
</tbody>
</table>

*Table: A Taxonomy of Models for Morphology (Stump 2001)*
## Word-and-Paradigm Morphology

### Paradigmatic Organization

<table>
<thead>
<tr>
<th>/EAT/</th>
<th>1Sing</th>
<th>2Sing</th>
<th>3Sing</th>
<th>1Pl</th>
<th>2Pl</th>
<th>3Pl</th>
</tr>
</thead>
<tbody>
<tr>
<td>Past</td>
<td>1SingPast</td>
<td>2SingPast</td>
<td>3SingPast</td>
<td>1PlPast</td>
<td>2PlPast</td>
<td>3PlPast</td>
</tr>
<tr>
<td>Present</td>
<td>1SingPres</td>
<td>2SingPres</td>
<td>3SingPres</td>
<td>1PlPres</td>
<td>2PlPres</td>
<td>3PlPres</td>
</tr>
<tr>
<td>Perfect</td>
<td>1SingPerf</td>
<td>2SingPerf</td>
<td>3SingPerf</td>
<td>1PlPerf</td>
<td>2PlPerf</td>
<td>3PlPerf</td>
</tr>
</tbody>
</table>

### Realization Rules

- /EAT/ +1SingPast → ‘ate’
- /EAT/ +3SingPast → ‘ate’
- /EAT/ +1SingPres → ‘eats’
- /EAT/ +3SingPres → ‘eat’
The Proposal (I): “Lifting” the Terminology

Morphological Exponence

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

Morphosyntactic Exponence

- Simple (e.g., SBJ \(\rightsimeq\) nominative)
- Cumulative (e.g., SBJ, PRD, OBJ \(\rightsimeq\) clitics)
- Distributed/Extended (e.g., OBJ \(\rightsimeq\) DOM, FS)
The Proposal (I): “Lifting” the Terminology

Morphological Exponence: Properties $\leadsto$ Words

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

Morphosyntactic Exponence: Relations $\leadsto$ Configurations

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

<table>
<thead>
<tr>
<th></th>
<th>CONFIGURATIONAL</th>
<th>RELATIONAL</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>INCREMENTAL</strong></td>
<td>Head-Driven Parsing</td>
<td>Dependency Parsing</td>
</tr>
<tr>
<td><strong>REALIZATIONAL</strong></td>
<td>Stochastic TAG, CCG</td>
<td>Relational-Realizational</td>
</tr>
</tbody>
</table>

**Table:** A Taxonomy of Statistical Parsing Frameworks (Tsarfaty 2010)
## The Proposal (IV): Relational-Realizational Modeling

<table>
<thead>
<tr>
<th>S⟨PRED⟩ FEATS</th>
<th>Affirmative</th>
<th>Interrogative</th>
<th>Imperative</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARG-ST</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>intransitive</td>
<td>S_{affirm}+{SBJ,PRD}</td>
<td>S_{inter}+{SBJ,PRD}</td>
<td>S_{imper}+{SBJ,PRD}</td>
</tr>
<tr>
<td>transitive</td>
<td>S_{affirm}+{SBJ,PRD,OBJ}</td>
<td>S_{inter}+{SBJ,PRD,OBJ}</td>
<td>S_{imper}+{SBJ,PRD,OBJ}</td>
</tr>
<tr>
<td>ditransitive</td>
<td>S_{affirm}+{SBJ,PRD,OBJ,COM}</td>
<td>S_{inter}+{SBJ,PRD,OBJ,COM}</td>
<td>S_{imper}+{SBJ,PRD,OBJ,COM}</td>
</tr>
</tbody>
</table>

**Figure:** Paradigmatic Organization

\[
\begin{align*}
S_{affirm}+\{SBJ,PRD,OBJ,COM\} \\
\langle\text{Dani, natan, et hamatana, ledina}\rangle \\
\langle\text{Dani gave ACC-the-present to-Dina}\rangle
\end{align*}
\]

**Figure:** Form-Function Separation

\[
\begin{align*}
S_{affirm}+\{SBJ,PRD,OBJ,COM\} \\
\langle\text{et hamatana, natan, Dani, ledina}\rangle \\
\langle\text{ACC-the-present gave Dani to-Dina}\rangle
\end{align*}
\]
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

S

NP-SBJ
- dani
  - Dani

VB-PRD
- natan
gave

ADVP
- etmol
- yesterday

NP_{Def+Acc}^-
OBJ
- et hamatana
- Acc
- Def-present

PP-COM
- ledina
to Dina
Tsarfaty et al. (2008, 2009)

Relational-Realizational (RR) Parsing

\[ \{SBJ, PRD, OBJ, COM\} @ S \]

- NP: dani Dani
- VB: natan gave
- ADVP: etmol yesterday
- NP + Def + Acc: et hamatana Acc Def-present
- PP: ledina to Dina
Tsarfaty et al. (2008, 2009)

Relational-Realizational (RR) Parsing

\[
S \rightarrow \{\text{SBJ}, \text{PRD}, \text{OBJ}, \text{COM}\} \@ S
\]

\[
\begin{align*}
\text{SBJ} @ S & \quad \text{PRD} @ S & \quad \text{PRD:OBJ} @ S & \quad \text{OBJ} @ S & \quad \text{COM} @ S \\
\text{NP} & \quad \text{VB} & \quad \text{ADVP} & \quad \text{NP}_{\text{+Def+Acc}} & \quad \text{PP} \\
\text{dani} & \quad \text{natan} & \quad \text{etmol} & \quad \text{et hamatana} & \quad \text{ledina}
\end{align*}
\]

\[
\text{Dani gave etmol yesterday et hamatana to Dina}
\]
Tsarfaty et al. (2008, 2009)

Relational-Realizational (RR) Parsing

[Diagram of a tree structure with nodes labeled as follows: S, {PRD,SBJ,OBJ,COM}@S, SBJ@S, PRD@S, PRD:OBJ@S, OBJ@S, COM@S, NP, NP+Def+Acc, PP-COM, SBJ@S, PRD@S, PRD:OBJ@S, OBJ@S, COM@S, NP, NP+Def+Acc, PP-COM, et hamatana Acc Def-present, natan gave, et mol yesterday, et hamatana Acc Def-present, dani Dani, ledina to Dina]
The Model Parameters

Projection:

\[ P \]
\[ \{ gr_i \}_{i=1}^n \circ P \]

Configuration:

\[ \{ gr_i \}_{i=1}^n \circ P \]
\[ gr_1 \circ P \]
\[ gr_1 \circ gr_2 \circ P \]
\[ \ldots \]
\[ gr_n \circ P \]

Realization:

\[ gr_1 \circ P \]
\[ gr_1 \circ gr_2 \circ P \]
\[ \ldots \]
\[ gr_n \circ P \]
\[ C_1 \]
\[ \ldots C_1 : 2_i \ldots \]
\[ C_n \]
The Probabilistic Model

The RR Probabilities:

\[ P_{\text{RR}}(r) = \]

\[ = \]

\[ = \]

\[ = \]

\[ = \]

\[ = \]

\[ = \]

Projection

Configuration

Realization

\[ \pi^* = \arg\max_{\pi} P(\pi) = \arg\max_{\pi} \prod_{r \in \pi} P_{\text{RR}}(r) \]
Application I: Parsing Modern Hebrew
## A Taxonomy of PCFG-based Parsers

<table>
<thead>
<tr>
<th><strong>Incremental</strong></th>
<th><strong>Configurational</strong></th>
<th><strong>Relational</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Head-Driven Parsing</td>
<td>(Collins 1999)</td>
<td>Relational-Realizational</td>
</tr>
<tr>
<td>(Collins 1999 enhanced)</td>
<td></td>
<td>(Tsarfaty et al. 2009)</td>
</tr>
</tbody>
</table>

**Table:** A Taxonomy of PCFG-Based Parsing Frameworks
# A Taxonomy of PCFG-based Parsers

<table>
<thead>
<tr>
<th>Incremental</th>
<th>Relational</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Configurational</strong></td>
<td><strong>Relational</strong></td>
</tr>
<tr>
<td>Head-Driven Parsing</td>
<td>Dependency Parsing</td>
</tr>
<tr>
<td>(Collins 1999)</td>
<td>(Collins 1999 enhanced)</td>
</tr>
<tr>
<td>Flattened Trees</td>
<td>Relational-Realizational</td>
</tr>
<tr>
<td>(Johnson 1998)</td>
<td>(Tsarfaty et al. 2009)</td>
</tr>
</tbody>
</table>

**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

<table>
<thead>
<tr>
<th>Category</th>
<th>First Value</th>
<th>Second Value</th>
<th>Third Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP</td>
<td>77.39 / 74.32</td>
<td>77.94 / 73.75</td>
<td>78.96 / 76.11</td>
</tr>
<tr>
<td>PP</td>
<td>71.78 / 71.14</td>
<td>71.83 / 69.24</td>
<td>74.4 / 72.02</td>
</tr>
<tr>
<td>SBAR</td>
<td>55.73 / 59.71</td>
<td>53.79 / 57.49</td>
<td>57.97 / 61.67</td>
</tr>
<tr>
<td>ADVP</td>
<td>71.37 / 77.01</td>
<td>72.52 / 73.56</td>
<td>73.57 / 77.59</td>
</tr>
<tr>
<td>ADJP</td>
<td><strong>79.37 / 78.96</strong></td>
<td>78.47 / 77.14</td>
<td>78.69 / 78.18</td>
</tr>
<tr>
<td>S</td>
<td><strong>73.25 / 79.07</strong></td>
<td>71.07 / 76.49</td>
<td>72.37 / 78.33</td>
</tr>
<tr>
<td>SQ</td>
<td>36.00 / <strong>32.14</strong></td>
<td>30.77 / 14.29</td>
<td><strong>55.56 / 17.86</strong></td>
</tr>
<tr>
<td>PREDP</td>
<td>36.31 / 39.63</td>
<td><strong>44.74 / 39.63</strong></td>
<td>44.51 / <strong>46.95</strong></td>
</tr>
</tbody>
</table>
Take Home

<table>
<thead>
<tr>
<th>CONFIGURATIONAL</th>
<th>RELATIONAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>INCREMENTAL</td>
<td>Head-Driven Parsing</td>
</tr>
<tr>
<td></td>
<td>Dependency Parsing</td>
</tr>
<tr>
<td>REALIZATIONAL</td>
<td>Stochastic TAG, CCG</td>
</tr>
<tr>
<td></td>
<td>Relational-Realizational</td>
</tr>
</tbody>
</table>

Table: A Taxonomy of Statistical Parsing Frameworks (Tsarfaty 2010)
Application II: Probabilistic Universal Grammar

Basic Word-Order Parameter:
$P(<\text{configuration}>|\{\text{SBJ,PRD,OBJ}\}@S)$

<table>
<thead>
<tr>
<th>Probability</th>
<th>Configuration</th>
<th>tri-</th>
<th>bi-</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2%</td>
<td>OBJ □ SBJ PRD</td>
<td>OSV</td>
<td>OV</td>
</tr>
<tr>
<td>0.2%</td>
<td>PRD OBJ SBJ □</td>
<td>VOS</td>
<td>VO</td>
</tr>
<tr>
<td>0.2%</td>
<td>□ PRD OBJ □ SBJ □</td>
<td>VOS</td>
<td>VO</td>
</tr>
<tr>
<td>0.2%</td>
<td>PRD SBJ □ OBJ □</td>
<td>VSO</td>
<td>VO</td>
</tr>
<tr>
<td>0.4%</td>
<td>□ PRD □ SBJ □ OBJ □</td>
<td>VSO</td>
<td>VO</td>
</tr>
<tr>
<td>0.6%</td>
<td>OBJ □ PRD SBJ □</td>
<td>OVS</td>
<td>OV</td>
</tr>
<tr>
<td>0.8%</td>
<td>OBJ PRD □ SBJ □</td>
<td>OVS</td>
<td>OV</td>
</tr>
<tr>
<td>1%</td>
<td>□ PRD □ SBJ OBJ □</td>
<td>VSO</td>
<td>VO</td>
</tr>
<tr>
<td>1.3%</td>
<td>SBJ □ PRD OBJ □</td>
<td>SVO</td>
<td>VO</td>
</tr>
<tr>
<td>1.7%</td>
<td>□ PRD OBJ SBJ □</td>
<td>VOS</td>
<td>VO</td>
</tr>
<tr>
<td>1.7%</td>
<td>□ SBJ PRD □ OBJ □</td>
<td>SVO</td>
<td>VO</td>
</tr>
<tr>
<td>3%</td>
<td>OBJ PRD SBJ □</td>
<td>OVS</td>
<td>OV</td>
</tr>
<tr>
<td>3.7%</td>
<td>□ PRD SBJ □ OBJ □</td>
<td>VSO</td>
<td>VO</td>
</tr>
<tr>
<td>4.1%</td>
<td>SBJ □ PRD □ OBJ □</td>
<td>SVO</td>
<td>VO</td>
</tr>
<tr>
<td>6.5%</td>
<td>□ SBJ PRD OBJ □</td>
<td>SVO</td>
<td>VO</td>
</tr>
<tr>
<td>10.3%</td>
<td>SBJ □ PDR OBJ □</td>
<td>SVO</td>
<td>VO</td>
</tr>
<tr>
<td>12.3%</td>
<td>□ PRD SBJ OBJ □</td>
<td>VSO</td>
<td>VO</td>
</tr>
<tr>
<td>15.6%</td>
<td>SBJ PRD □ OBJ □</td>
<td>SVO</td>
<td>VO</td>
</tr>
<tr>
<td>35.3%</td>
<td>SBJ PRD OBJ □</td>
<td>SVO</td>
<td>VO</td>
</tr>
</tbody>
</table>
Differential Object-Marking Parameter: 
\[ P(< \text{morphosyntactic representation}> | \text{OBJ@S}) \]

<table>
<thead>
<tr>
<th>Probability</th>
<th>Realization</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.8%</td>
<td>NP.DEF.ACC\langle PRP\rangle @ S</td>
</tr>
<tr>
<td>6.5%</td>
<td>NP.DEF.ACC\langle NNT\rangle @ S</td>
</tr>
<tr>
<td>6.7%</td>
<td>NP.DEF.ACC\langle NN.DEF\rangle @ S</td>
</tr>
<tr>
<td>7.4%</td>
<td>NP.DEF.ACC\langle NNP\rangle @ S</td>
</tr>
<tr>
<td>8.8%</td>
<td>NP\langle NNT\rangle @ S</td>
</tr>
<tr>
<td>14.7%</td>
<td>NP.DEF.ACC\langle NN\rangle @ S</td>
</tr>
<tr>
<td>43.5%</td>
<td>NP.\langle NN\rangle @ S</td>
</tr>
</tbody>
</table>
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?
Languages are different!  
\(\Rightarrow\) Modeling strategies should accommodate differences

Nonconfigurational languages are not configurational!  
\(\Rightarrow\) Modeling strategies should account for misalignments

Modeling Morphology vary in underlying assumptions  
\(\Rightarrow\) Inferential-Realizational approaches model m:n mapping

Modeling Morphosyntax meets similar considerations  
\(\Rightarrow\) 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