MORPHOLOGY, SYNTAX AND WHAT’S IN BETWEEN
Towards Paradigmatic Syntax

Reut Tsarfaty
The Department of Linguistics and Philology
Uppsala University

MACHINE TRANSLATION AND MORPHOLOGICALLY RICH LANGUAGES
University of Haifa
January 23-27, 2011
Syntax, MT and MRLs
Syntax, MT and MRLs

Some Brief Historical Overview
Syntax, MT and MRLs

Some Brief Historical Overview

- What Kevin said

Morphology and Syntax Interact!
No, pre/post-processing won’t work
Syntax, MT and MRLs

Some Brief Historical Overview

- What Kevin said
- What Yoav said
- What all linguists say all the time
Syntax, MT and MRLs

Some Brief Historical Overview

- What Kevin said
- What Yoav said
- What all linguists say all the time

⇝ Morphology and Syntax Interact!
Syntax, MT and MRLs

Some Brief Historical Overview

- What Kevin said
- What Yoav said
- What all linguists say all the time

⇝ Morphology and Syntax Interact!
No, pre/post-processing won’t work
Today

An Answer to Kevin (Sunday):
How to model a morphosyntactic tree

A Question to Jason (Thursday):
How to learn bilingual paradigms
The Task:
Statistical Parsing

The Challenge:
Complex Form-Function Correspondence

The Method:
Following the footsteps of Morphology

The Proposal:
A Relational-Realizational Approach

⇒ A Stepping Stone
Towards computational typology
Part 1: The Task

Statistical Parsing
Statistical Parsing

"This is easy"
Statistical Parsing

"This is easy"
"This" is "easy"
Statistical Parsing

S

NP-SBJ  VP-PRD
NP-SBJ  VP-PRD

PRP

NNP

“This”

VB

VB

ADJP

ADJP

ADJ

ADJ

“easy”
Works Great for English

S
  NP-SBJ
    PRP
    “This”
  VP-PRD
    VB
    “is”
    ADJP
    ADJ
    “easy”

<table>
<thead>
<tr>
<th>Model</th>
<th>Study</th>
<th>F-Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treebank Grammar</td>
<td>Charniak 1996</td>
<td>75</td>
</tr>
<tr>
<td>Head-Driven</td>
<td>Collins 1997</td>
<td>88.6</td>
</tr>
<tr>
<td>Discriminative Reranking</td>
<td>Collins 2000</td>
<td>89.7</td>
</tr>
<tr>
<td>Discriminative Reranking</td>
<td>Johnson &amp; Charniak 2005</td>
<td>91.0</td>
</tr>
<tr>
<td>Self-Training</td>
<td>McClosky 2006</td>
<td>92.1</td>
</tr>
<tr>
<td>State-Splits</td>
<td>Petrov et al 2007</td>
<td>90.1</td>
</tr>
<tr>
<td>Forest Reranking</td>
<td>Liang Huang 2008</td>
<td>91.7</td>
</tr>
</tbody>
</table>
Not As Well For Other Languages

And what about this?

 컴퓨터에서 사용할 수 있습니다。

And this?

إغلي الماء قبل استعماله.

And this?

иш להרתיע את המים
לפני השיכוץ.

<table>
<thead>
<tr>
<th>Language</th>
<th>Parser</th>
<th>F-Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>German</td>
<td>Rafferty &amp; Manning 2008</td>
<td>79.2</td>
</tr>
<tr>
<td>Czech</td>
<td>Collins et al. 1999</td>
<td>79.3</td>
</tr>
<tr>
<td>Chinese</td>
<td>Levy &amp; Manning 2003</td>
<td>78.8</td>
</tr>
<tr>
<td>Arabic</td>
<td>Maamouri, Bies &amp; Kulick 2008</td>
<td>78.1</td>
</tr>
<tr>
<td>Hebrew</td>
<td>Tsarfaty &amp; Sima’an 2007</td>
<td>74.4</td>
</tr>
</tbody>
</table>
Not So Well For MRLs

Why?
Homework Assignment

Read:


Answer:

- What are the three main challenges in parsing of MRLs?
- How are these going to affect your MT system?
English Pronouns
English Pronouns

P(“He likes her”) = P(NP VP|S) × ... × P(“her”|PRP.ACC) = 0.25
English Pronouns

\[
P(\text{“Her likes he”}) = P(\text{NP VP|S}) \times \ldots \times P(\text{“her”|PRP.ACC}) = 0.25
\]
Example 1: Parent Encoding (Johnson 1998)

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

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

\[
\begin{align*}
S &\rightarrow <\text{VB}> \\
L, \Delta_{L_1}, <\text{VB}> &\rightarrow \text{PRP.NOM} \\
H, \Delta_0, <\text{VB}> &\rightarrow \text{He} \quad \text{VB} \\
R, \Delta_{R_1}, <\text{VB}> &\rightarrow \text{likes} \quad \text{PRP.ACC} \\
\end{align*}
\]

\[
\Rightarrow \\
P(<\text{VB}>|S) = 1 \\
P(L\Delta_{L_1}, H\Delta_0| <\text{VB}>, S) = 1 \\
P(\text{PRP.NOM}|L, \Delta_{L_1}, <\text{VB}>, S) = 1 \\
P(<\text{VB}>|\text{VP}) = 1 \\
P(<\text{VB}>|\text{VP}) = 1 \\
P(\text{PRP.ACC}|R, \Delta_{R_1}, <\text{VB}>, S) = 1 \\
P(<\text{VB}>|\text{VP}) = 1 \\
P(\text{He}|\text{PRP.NOM}) = 1 \\
P(\text{"likes"}|\text{VB}) = 1 \\
P(\text{"her"}|\text{PRP.ACC}) = 1
\]
Example 2: Head-Driven Processes (Collins 1999)

\[
\begin{align*}
S &= \langle VB \rangle \\
&\downarrow_{\text{PRP.NOM}} \\
L, \Delta_{L_1}, \langle VB \rangle &\quad \text{He} \\
&\downarrow_{\text{He}} \\
H, \Delta_0, \langle VB \rangle &\quad \text{likes} \\
&\downarrow_{\text{PRP.ACC}} \\
R, \Delta_{R_1}, \langle VB \rangle &\quad \text{her} \\
&\downarrow \\
H, \Delta_0, \langle VB \rangle &\quad \text{likes} \\
&\downarrow_{\text{PRP.ACC}} \\
\Rightarrow &\quad P(\langle VB \rangle | S) \\
&\quad P(L \Delta_{L_1}, H \Delta_0 | \langle VB \rangle, S) \\
&\quad P(\text{PRP.NOM} | L, \Delta_{L_1}, \langle VB \rangle, S) \\
&\quad P(\text{VP} | H, \Delta_0, \langle VB \rangle, S) \\
&\quad P(\langle VB \rangle | \text{VP}) \\
&\quad P(\text{PRP.ACC} | R, \Delta_{R_1}, \langle VB \rangle, S) \\
&\quad P(\text{VB} | H, \Delta_0, \langle VB \rangle, S) \\
&\quad P(\langle VB \rangle | S) \\
&\quad P(\text{He”} | \text{PRP.NOM}) \\
&\quad P(\text{”likes”} | \text{VB}) \\
&\quad P(\text{”her”} | \text{PRP.ACC})
\end{align*}
\]
What if we take a different treebank

The Modern Hebrew Treebank

S

NP
dani
Dani.MS

VB
natan
gave.3MS

ADVP
etmol
yesterday

NP
et hamatana
ACC DEF-present

PP
ledina
DAT Dina

S

NP
et hamatana
ACC DEF-present

VB
natan
gave.3MS

ADVP
etmol
yesterday

NP
dani
Dani.MS

PP
ledina
DAT Dina
Example 1:

Parent Encoding
Example 2:

The Head-Driven (HD) Approach

[Diagram showing syntactic tree structures with labeled nodes and arcs representing sentence structure.]
So What Is Going On?

An Observation

- Configurational languages:
  - Tree configurations approximate grammatical functions
- Less-configurational languages:
  - Not.
So What Is Going On?

An Observation

- Configurational languages:
  - Tree configurations approximate grammatical functions

- Less-configurational languages:
  - Not.
  - No regularities to statistically learn from data

A Question

What kind of form-function correspondence patterns our parser needs to learn from the data?
Part 2: The Challenge

Modeling Form-Function Correspondence
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

Predicate-Argument Relations

‘SBJ’ did ‘PRD’ to ‘OBJ’

Syntactic Configuration

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

Predicate-Argument Relations

‘SBJ’ did ‘PRD’ to ‘OBJ’

Configurational Languages

```
S
   NP
   VP
   PRP.NOM
     “He”
   VB
     “likes”
   NP
     NN.ACC
       ‘her’
```
Nonconfigurationality as Misalignment

Predicate-Argument Relations

‘SBJ’ did ‘PRD’ to ‘OBJ’

Less-Configurational Languages
Morphosyntactic Exponence in Hebrew

Word-Order

(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)
Exponence Relations in 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
Exponence Relations in 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
Exponence Relations in 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]
Exponence Relations in Hebrew (1:many)

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
Exponence Relations in 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 Morphosyntactic Exponence

Recap:

Configurational ——— Nonconfigurational
1:1 ———————————— many : many

- Exponence relations relate grammatical functions to the formal means that realize them in the syntactic structure
- Configurationality is a special case of a 1:1 mapping between grammatical functions to configurational positions

Question:
How can we model and statistically learn generally complex, many-to-many, form-function correspondence in syntax?
Part 3: The Proposal

Following the footsteps of morphology
Modeling Morphology (i): Terminology

Morpheme-Based Morphology (Bloomfield, 1933)

- ‘kids’, ‘oxen’, ‘men’, ‘sheep’
- KID plural, OX plural, MAN plural, SHEEP plural

Morphological Exponence (Matthews 1991)

- Simple Exponence (1:1)
- Cumulative Exponence (many:1)
- Extended/Distributed Exponence (1:many)
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
- **REALIZATIONAL**: property-bundles are pre-condition for spell-out
## Modeling Morphology (III): 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 (Bloomfield 1933) (Lieber 1992)</td>
<td>Item &amp; Processes (Hocket 1954) (Steele 1995)</td>
</tr>
<tr>
<td><strong>REALIZATIONAL</strong></td>
<td>Distributed Morphology (Halle and Marantz 1993) Lexical Phonology</td>
<td>(Extended) Word &amp; Paradigm (Matthews 1972), (Anderson 1992) (Stump 2001), (Blevins 2006)</td>
</tr>
</tbody>
</table>

**Table:** A Taxonomy of Models for Morphology (Stump 2001)
The Strategy (IV): (Extended) Word-and-Paradigm

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

Morphological Exponence

- Simple (1:1, e.g., PL \(\leadsto\) ‘s’ in ‘cats’)
- Cumulative (many:1, e.g., 3RD, SG \(\leadsto\) ‘s’ in ‘eats’)
- Distributed/Extended (1:many PL \(\leadsto\) ‘i’, ‘ren’ in ‘children’)

\(\text{Morpho)Syntactic Exponence: Relations} \leadsto \text{Positions}

- Simple (1:1, e.g., SBJ \(\leadsto\) nominative)
- Cumulative (many:1, e.g., PRD, OBJ \(\leadsto\) clitics)
- Distributed/Extended (1:many SBJ \(\leadsto\) agreement)
The Proposal (I): “Lifting” the Terminology

Morphological Exponence: Properties $\leadsto$ Words

- Simple (1:1, e.g., $\text{PL} \leadsto \text{‘s’ in ‘cats’}$)
- Cumulative (many:1, e.g., $\text{3RD,SG} \leadsto \text{‘s’ in ‘eats’}$)
- Distributed/Extended (1:many $\text{PL} \leadsto \text{‘i’, ‘ren’ in ‘children’}$)

(Morpho)Syntactic Exponence: Relations $\leadsto$ Positions

- Simple (1:1, e.g., $\text{SBJ} \leadsto \text{nominative}$)
- Cumulative (many:1, e.g., $\text{PRD,OBJ} \leadsto \text{clitics}$)
- Distributed/Extended (1:many, e.g., $\text{SBJ} \leadsto \text{agreement}$)
The Proposal (II): Modeling Assumptions

**CONFIGURATIONAL vs. RELATIONAL Approaches**

- **CONFIGURATIONAL:**
  configurations are primary, relations are derived

- **RELATIONAL:**
  relations are primary, configurations are computed

**INCREMENTAL vs. REALIZATIONAL Approaches**

- **INCREMENTAL:**
  constructive operations,
  individually/incrementally define/add relations

- **REALIZATIONAL:**
  interpretive operations,
  sets of relations are precondition to realization
### Table: A Taxonomy of Generative Syntactic Frameworks

<table>
<thead>
<tr>
<th></th>
<th>CONFIGURATIONAL</th>
<th>RELATIONAL</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>INCREMENTAL</strong></td>
<td>X-Bar Theory</td>
<td>Dependency Grammar</td>
</tr>
<tr>
<td><strong>REALIZATIONAL</strong></td>
<td>Tree Adjoining Grammar</td>
<td></td>
</tr>
</tbody>
</table>
The Proposal (IV): Relational-Realizational Modeling

The \textit{Relational} Assumption

- Paradigms organize the syntactic domain
- Cells in paradigms define sets of relations
- Sets of relations are realized in different configurations

The \textit{Realizational} Assumption

- Sets of relations (Arg-St) are primitives
- Rules interpret sets of relations as surface forms
- Rules can refer to multiple relations and span clauses

Realization in Syntax is Recursive!
Realization of a cells refers to function cells in other paradigms
The Proposal (IV): Relational-Realizational Modeling

<table>
<thead>
<tr>
<th>ARG-ST</th>
<th>AFFIRMATIVE</th>
<th>INTERROGATIVE</th>
<th>IMPERATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>intransitive</td>
<td>$S_{affirm} {\text{SBJ,PRD}}$</td>
<td>$S_{inter} {\text{SBJ,PRD}}$</td>
<td>$S_{imper} {\text{SBJ,PRD}}$</td>
</tr>
<tr>
<td>transitive</td>
<td>$S_{affirm} {\text{SBJ,PRD,OBJ}}$</td>
<td>$S_{inter} {\text{SBJ,PRD,OBJ}}$</td>
<td>$S_{imper} {\text{SBJ,PRD,OBJ}}$</td>
</tr>
<tr>
<td>ditransitive</td>
<td>$S_{affirm} {\text{SBJ,PRD,OBJ,COM}}$</td>
<td>$S_{inter} {\text{SBJ,PRD,OBJ,COM}}$</td>
<td>$S_{imper} {\text{SBJ,PRD,OBJ,COM}}$</td>
</tr>
</tbody>
</table>

Figure: Paradigmatic Organization

Figure: Realization Rules
Realization Rules
Realization Rules

A PCFG Approximation

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
Realization Rules

Form-Function Separation

S

{SBJ, PRD, OBJ, COM}@S

NP

dani
Dan

VB

natan
gave

ADVP

etmol
yesterday

NP_{Def+Acc}

et hamatana
Acc Def-present

PP

ledina
to Dina
Realization Rules

Morphological and Syntactic Realization

S

{SBJ,PRD,OBJ,COM}@S

SBJ@S

NP

dani
Dani

PRD@S

VB

natan
gave

PRD:OBJ@S

ADVP

etmol
yesterday

OBJ@S

NP+Def+Acc

et hamatana
Acc Def-present

COM@S

PP

ledina
to Dina
Realization Rules: Economy and Generalization

S
{PRD,SBJ,OBJ,COM}@S

SBJ@S
NP
dani
Dani

PRD@S
VB
natan
gave

PRD:OBJ@S
ADVP
etmol
eyesterday

OBJ@S
NP+Def+Acc
et hamatana
Acc Def-present

COM@S
PP-COM
ledina
to Dina

S
{PRD,SBJ,OBJ,COM}@S

OBJ@S
NP+Def+Acc
et hamatana
Acc Def-present

PRD@S
VB
natan
gave

PRD:OBJ@S
ADVP
etmol
eyesterday

SBJ@S
NP
dani
Dani

COM@S
PP-COM
ledina
to Dina
The Generative Model

Projection:

\[ P \{g_{r_i}\}_{i=1}^n \@ P \]

Configuration:

\[ \{g_{r_i}\}_{i=1}^n \@ P \]

\[ g_{r_1} \@ P \quad g_{r_1} : g_{r_2} \@ P \quad \ldots \quad g_{r_n} \@ P \]

Realization:

\[ g_{r_1} \@ P \quad g_{r_1} : g_{r_2} \@ P \quad \ldots \quad g_{r_n} \@ P \]

\[ C_1 \quad \ldots C_{1:2i}\ldots \quad C_n \]
The Probabilistic Model

The RR Probabilities:

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

Projection \[ P_p(\{g_{ri}\}_{i=1}^n | P) \times \]

Configuration \[ P_c(\langle gr_0 : gr_1, g_1, \ldots \rangle | \{g_{ri}\}_{i=1}^n, P) \times \]

Realization \[ \prod_{i=1}^n P_{r_1}(C_i | gr_i, P) \times \]

\[ P_{r_2}(\langle C_{01}, \ldots, C_{0m_0} \rangle | gr_0 : gr_1, P) \times \]

\[ \prod_{i=1}^n P_{r_2}(\langle C_{i1}, \ldots, C_{imi} \rangle | gr_i : gr_{i+1}, P) \]

The RR Parser:

\[ \pi^* = \arg\max_{\pi} P(\pi) = \arg\max_{\pi} \prod_{r \in \pi} P_{RR}(r) \]
Part IV: Applications

- Cross-Linguistic Parsing
- Cross-Linguistic Variation
Application I: Parsing Modern Hebrew

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/Gender

Estimation
Relative Frequency + Simple Unknown Words Smoothing

Parsing
Exhaustive Viterbi Parsing (using BitPar, Schmid 2004)

Evaluation
PARSEVAL (i) Overall, and (ii) Per Category Evaluation
# A Taxonomy of PCFG-based Parsers

<table>
<thead>
<tr>
<th></th>
<th><strong>CONFIGURATIONAL</strong></th>
<th><strong>RELATIONAL</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>INCREMENTAL</strong></td>
<td>Head-Driven Parsing</td>
<td>(Collins 1999)</td>
</tr>
</tbody>
</table>

| **REALIZATIONAL** | **RELATIONAL-REALIZATIONAL** | (Tsarfaty et al. 2009) |

**Table:** A Taxonomy of PCFG-Based Parsing Frameworks
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)
<table>
<thead>
<tr>
<th>Feature</th>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ø</td>
<td>83.06</td>
<td>83.49</td>
</tr>
<tr>
<td>(5914)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>gender</td>
<td>82.18</td>
<td>83.70</td>
</tr>
<tr>
<td>(10765)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>case/def</td>
<td>79.53</td>
<td>83.66</td>
</tr>
<tr>
<td>(12700)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>gender/case/def</td>
<td>80.89</td>
<td>84.13</td>
</tr>
<tr>
<td>(13028)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Swedish Parsing Results Using Gold Standard Input

<table>
<thead>
<tr>
<th></th>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\emptyset$</td>
<td>78.65</td>
<td>77.71</td>
</tr>
<tr>
<td>(8696)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><code>gender</code></td>
<td>73.20</td>
<td>78.09</td>
</tr>
<tr>
<td>(11382)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><code>case/def</code></td>
<td>74.90</td>
<td>79.09</td>
</tr>
<tr>
<td>(11239)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><code>gender/case/def</code></td>
<td>68.97</td>
<td>77.89</td>
</tr>
<tr>
<td>(13347)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Application II: Cross-Linguistic Variation

1: Apply the model to different languages, e.g.,
   - Hebrew: a Semitic Language
   - Swedish: a Germanic Language

2: Learn the distribution of model parameters
   - RR-Projection
   - RR-Configuration
   - RR-Realization

3: Instantiate typological parameters for UG
Application II: Cross-Linguistic Variation

1: Apply the model to different languages, e.g.,
   - **Hebrew**: a Semitic Language
   - **Swedish**: a Germanic Language

2: Learn the distribution of model parameters
   - RR-Projection
   - RR-Configuration
   - RR-Realization

3: Instantiate typological parameters for UG
Parameter 1: Basic Word-Order (Greenberg 1963)

Basic Word-Order Parameter in Hebrew: 
\[ P(< \text{configuration} > | \{\text{SBJ, PRD, OBJ}\} \odot S) \]

<table>
<thead>
<tr>
<th>Probability</th>
<th>Configuration</th>
</tr>
</thead>
<tbody>
<tr>
<td>35.3%</td>
<td>SBJ PRD OBJ □</td>
</tr>
<tr>
<td>15.6%</td>
<td>SBJ PRD □ OBJ □</td>
</tr>
<tr>
<td>12.3%</td>
<td>□ PRD SBJ OBJ □</td>
</tr>
<tr>
<td>10.3%</td>
<td>SBJ □ PDR OBJ □</td>
</tr>
<tr>
<td>6.5%</td>
<td>□ SBJ PRD OBJ □</td>
</tr>
<tr>
<td>4.1%</td>
<td>SBJ □ PRD □ OBJ □</td>
</tr>
<tr>
<td>3.7%</td>
<td>□ PRD SBJ □ OBJ □</td>
</tr>
<tr>
<td>3%</td>
<td>OBJ PRD SBJ □</td>
</tr>
<tr>
<td>1.7%</td>
<td>□ SBJ PRD □ OBJ □</td>
</tr>
<tr>
<td>1.7%</td>
<td>□ PRD OBJ SBJ □</td>
</tr>
<tr>
<td>1.3%</td>
<td>SBJ □ PRD OBJ □</td>
</tr>
<tr>
<td>1%</td>
<td>□ PRD □ SBJ OBJ □</td>
</tr>
</tbody>
</table>
Basic Word-Order Parameter in Swedish:
\[ P(< configuration > | \{ \text{SBJ,PRD,OBJ} \} @ S) \]

<table>
<thead>
<tr>
<th>Probability</th>
<th>Configuration</th>
</tr>
</thead>
<tbody>
<tr>
<td>35.5%</td>
<td>SBJ PRD OBJ □</td>
</tr>
<tr>
<td>18.9%</td>
<td>SBJ PRD □ OBJ □</td>
</tr>
<tr>
<td>13.9%</td>
<td>□ PRD SBJ PBJ □</td>
</tr>
<tr>
<td>8.1%</td>
<td>SBJ PRD OBJ</td>
</tr>
<tr>
<td>4.7%</td>
<td>□ PRD SBJ □ OBJ □</td>
</tr>
<tr>
<td>3.5%</td>
<td>OBJ PRD SBJ</td>
</tr>
<tr>
<td>2.6%</td>
<td>SBJ PRD OBJ □</td>
</tr>
<tr>
<td>1.7%</td>
<td>OBJ PRD SBJ COM □</td>
</tr>
<tr>
<td>1.6%</td>
<td>PRD SBJ OBJ □</td>
</tr>
<tr>
<td>1.6%</td>
<td>□ PRD SBJ OBJ</td>
</tr>
<tr>
<td>1%</td>
<td>□ PRD SBJ □ OBJ</td>
</tr>
</tbody>
</table>
Parameter 2: Inflectional Systems

The Object-Marking Parameter in Hebrew:
\[ P(< \text{morphosyntactic representation} > | \text{OBJ}@S) \]

<table>
<thead>
<tr>
<th>Probability</th>
<th>Realization</th>
</tr>
</thead>
<tbody>
<tr>
<td>43.5%</td>
<td>NP.\langle NN \rangle</td>
</tr>
<tr>
<td>14.7%</td>
<td>NP.DEF.ACC\langle NN \rangle</td>
</tr>
<tr>
<td>8.8%</td>
<td>NP\langle NNT \rangle</td>
</tr>
<tr>
<td>7.4%</td>
<td>NP.DEF.ACC\langle NNP \rangle</td>
</tr>
<tr>
<td>6.7%</td>
<td>NP.DEF.ACC\langle NN.DEF \rangle</td>
</tr>
<tr>
<td>6.5%</td>
<td>NP.DEF.ACC\langle NNT \rangle</td>
</tr>
<tr>
<td>5.8%</td>
<td>NP.DEF.ACC\langle PRP \rangle</td>
</tr>
</tbody>
</table>
Parameter 2: Inflectional Systems

The Object-Marking Parameter in Swedish: $P(<\textit{morphosyntactic representation}> | \text{OBJ@S})$

<table>
<thead>
<tr>
<th>Probability</th>
<th>Realization</th>
</tr>
</thead>
<tbody>
<tr>
<td>46%</td>
<td>NP.IND.NOM</td>
</tr>
<tr>
<td>20%</td>
<td>NP.DEF.NOM</td>
</tr>
<tr>
<td>13.4%</td>
<td>S</td>
</tr>
<tr>
<td>7.3%</td>
<td>NP.DEF.NOM-OBJ</td>
</tr>
<tr>
<td>4.9%</td>
<td>VP</td>
</tr>
<tr>
<td>3.6%</td>
<td>NP.IND</td>
</tr>
<tr>
<td>2.8%</td>
<td>NP.NOM</td>
</tr>
</tbody>
</table>
Towards Computational Typology and Statistical UG

We can potentially use the RR parameters to...

- Quantify Intra-Language Variation
- Quantify Cross-Linguistic Variation
- Quantify Nonconfigurationality
- Learn Probabilistic P&P
Conclusion

We presented a Relational-Realizational Architecture for Specifying and Learning Morphosyntactic Descriptions

- Simple
- Formal
- Robust
- Implementable
- Interpretable
- Explanatory

Paradigms augmented with realization rules may constitute a useful and powerful modeling strategy also for Statistical MT
Homework Assignment II

Read:
Tsarfaty, R. *Relational-Realizational Parsing*
PhD Thesis, University of Amsterdam, 2010

Implement:

- **Transfer rules** to model *morphosyntactic variation*
- **Alignment tools** to align *paradigm cells*
Thank You!

Questions?
Thank You!

Questions?

Now a postdoc at Uppsala, looking for a tenure track position :)}
Thanks to. My PhD Committee: Remko, Khalil, Joakim, Mark, Yoad, Henk, Michiel. My morphology teachers: Jim, Greg. Lots of Colleagues and co-authors. Dutch and Swedish funding agencies.