Syntactic Transfer Using a Bilingual Lexicon

Greg Durrett, Adam Pauls, and Dan Klein
UC Berkeley
Mosambik hofft auf Handel mit anderen Mitgliedern
Mosambik hofft auf Handel mit anderen Mitgliedern

Mozambique hope on trade with other members
Parsing a New Language

Mosambik hofft auf Handel mit anderen

Mozambique hope on trade with other members

[Berg-Kirkpatrick and Klein (2010), Petrov et al. (2011)]
Parsing a New Language

NOUN  VERB  PREP  NOUN  PREP  ADJ  NOUN
Mosambik  hofft  auf  Handel  mit  anderen  Mitgliedern
Mozambique  hope  on  trade  with  other  members

[Berg-Kirkpatrick and Klein (2010), Petrov et al. (2011)]
[McDonald et al. (2011), Cohen et al. (2011)]
Parsing a New Language

Mozambique  hope  on  trade
Mosambik  hofft  auf  Handel
Mozambique  hope  on  trade
Mosambik  hofft  auf  Handel

[Berg-Kirkpatrick and Klein (2010), Petrov et al. (2011)]
[McDonald et al. (2011), Cohen et al. (2011)]
Parsing a New Language

Mosambik
Mozambique

hofft
hope

auf
on

Handel
trade

mit
with

anderen
other

Mitgliedern
members

[Berg-Kirkpatrick and Klein (2010), Petrov et al. (2011)]

[McDonald et al. (2011), Cohen et al. (2011)]
Parsing a New Language

NOUN Mosambik Mozambique
VERB hofft hope
PREP auf on
NOUN Handel trade
PREP mit with
ADJ anderen other
NOUN Mitgliedern members

[Berg-Kirkpatrick and Klein (2010), Petrov et al. (2011)]
[McDonald et al. (2011), Cohen et al. (2011)]
[Berg-Kirkpatrick and Klein (2010), Petrov et al. (2011)]

[Berg-Kirkpatrick and Klein (2010), Petrov et al. (2011)]
Parsing a New Language

Mozambique hopes on trade with other members.

[McDonald et al. (2011), Cohen et al. (2011)]

[Berg-Kirkpatrick and Klein (2010), Petrov et al. (2011)]
[Berg-Kirkpatrick and Klein (2010), Petrov et al. (2011)]

[McDonald et al. (2011), Cohen et al. (2011)]

Mosambik hofft auf Handel mit anderen Mitgliedern

Mozambique hope on trade with other members
Mozambique hopes for trade with other members.

[Hwa et al. (2005), Ganchev et al. (2009), McDonald et al. (2011)]
Token-level Projection

Mozambique hopes for trade with other members

Assumes access to parallel data

[Hwa et al. (2005), Ganchev et al. (2009), McDonald et al. (2011)]
Mozambique hopes for trade with other members.

- Assumes access to parallel data
- Asymmetric train and test conditions

[Hwa et al. (2005), Ganchev et al. (2009), McDonald et al. (2011)]
Mosambik hofft auf Handel mit anderen Mitgliedern
Type-level Projection

- NOUN: Mosambik
- VERB: hofft
- PREP: auf
- NOUN: Handel
- PREP: mit
- ADJ: anderen
- NOUN: Mitgliedern

DE–EN dict

- Handel: trade, deal
- hofft: hope
- mit: with
Type-level Projection

Mosambik hofft auf Handel mit anderen Mitgliedern

DE–EN dict

Handel  trade, deal
hofft  hope
mit  with
Type-level Projection

DE–EN dict

Handel  trade, deal
hofft  hope
mit  with

EN trees

... trade with those regions ...
... a deal with China ...

NOUN  VERB  PREP  NOUN  PREP  ADJ  NOUN
Mosambik  hofft  auf  Handel  mit  anderen  Mitgliedern

trade, deal  with
Type-level Projection

Feature: “Goodness in English”
Value: high

DE–EN dict
- Handel: trade, deal
- hofft: hope
- mit: with

EN trees
- trade, deal with those regions
- a deal with China
Type-level Projection

NOUN Mosambik
VERB hofft
PREP auf
NOUN Handel
PREP mit
ADJ anderen
NOUN Mitgliedern

DE–EN dict

Handel trade, deal
hofft hope
mit with
Type-level Projection

DE–EN dict

Handel  trade, deal
hofft  hope
mit  with

NOUN Mosambik  VERB hofft  PREP auf  NOUN Handel  PREP mit  ADJ anderen  NOUN Mitgliedern

hope

trade, deal

hope

with

with
Type-level Projection

NOUN Mosambik
VERB hofft
PREP auf
NOUN Handel
PREP mit
ADJ anderen
NOUN Mitgliedern

DE–EN dict

Handel trade, deal
hofft hope
mit with

EN trees

no examples
Type-level Projection

Feature: “Goodness in English”
Value: low

DE–EN dict

Handel  trade, deal
hofft  hope
mit  with

EN trees

no examples
Model

- Edge-factored discriminative parser, based on MSTParser (McDonald et al. 2005)
- DELEX features: universal POS (McDonald et al. 2011)
- PROJ features: “Goodness in English”
Mozambique hopes on trade with other members.
Mozambique hopes to trade with other members.
Mosambik hofft auf Handel mit anderen Mitgliedern

Mozambique hope on trade with other members
Model: **DELEX** Features

**ATTACH**

<table>
<thead>
<tr>
<th>Feature</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOUN → PREP</td>
<td>1</td>
</tr>
</tbody>
</table>

**Mosambik hofft auf Handel mit anderen**

Mozambique hope on trade with other members
Model: **DELEX** Features

**ATTACH**

<table>
<thead>
<tr>
<th>Feature</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOUN → PREP</td>
<td>1</td>
</tr>
</tbody>
</table>

+ indicators including direction and distance

Mosambik hofft auf Handel mit anderen Mitgliedern
Mozambique hope on trade with other members
Mosambik hofft auf Handel mit anderen Mitgliedern

Mozambique hope on trade with other members
Model: \textbf{PROJ} Features

\[ \text{ATTACH} \]

<table>
<thead>
<tr>
<th>Feature</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Handel → mit</td>
<td>1</td>
</tr>
</tbody>
</table>

Mozambique    hope    on    trade    with    other    members

Mosambik    hofft    auf    Handel    mit    anderen    Mitgliedern
Model: **Proj** Features

**ATTACH**

- Insufficient target-language supervised data to use

<table>
<thead>
<tr>
<th>Feature</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Handel</td>
<td>1</td>
</tr>
</tbody>
</table>

- **Mosambik** hofft auf **Handel** mit **anderen** **Mitgliedern**
- **Mozambique** hope on **trade** with **other** **members**
Model: **PROJ** Features

**ATTACH**

<table>
<thead>
<tr>
<th>Feature</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Handel → mit</td>
<td></td>
</tr>
</tbody>
</table>

Mosambik hofft auf Handel mit anderen Mitgliedern

Mozambique hope on trade with other members
Model: **PROJ** Features

### ATTACH

<table>
<thead>
<tr>
<th>Feature</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score(Handel → mit)</td>
<td></td>
</tr>
</tbody>
</table>

- NOUN: Mosambik, Mozambique
- VERB: hofft, hope
- PREP: auf, on
- NOUN: Handel, trade
- PREP: mit, with
- ADJ: anderen, other
- NOUN: Mitgliedern, members
Model: **PROJ** Features

**ATTACH**

<table>
<thead>
<tr>
<th>Feature</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score(Handel $\rightarrow$ mit)</td>
<td></td>
</tr>
</tbody>
</table>

Query

- **NOUN**: Mosambik, Mozambique
- **VERB**: hofft, hope
- **PREP**: auf, on
- **NOUN**: Handel, trade
- **PREP**: mit, with
- **ADJ**: anderen, other
- **NOUN**: Mitgliedern, members
Model: **PROJ** Features

**ATTACH**

<table>
<thead>
<tr>
<th>Feature</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goodness in English</td>
<td>Score($Handel \rightarrow mit$)</td>
</tr>
</tbody>
</table>

---

**Query**

- **NOUN** Mosambik, Mozambique
- **VERB** hofft, hope
- **PREP** auf, on
- **NOUN** Handel, trade
- **PREP** mit, with
- **ADJ** anderen, other
- **NOUN** Mitgliedern, members
Query Score Computation

\[
\text{Score}(\text{Handel} \rightarrow \text{PREP}) = \log p(\text{PREP}|\text{Handel})
\]
Query Score Computation

\[
\text{Score}(\text{Handel} \rightarrow \text{PREP}) = \log p(\text{PREP} | \text{Handel})
\]

\[
= \log \sum_{w_s} p(\text{PREP} | w_s) \ p(w_s | \text{Handel})
\]
Query Score Computation

\[
\text{Score}(\text{Handel} \rightarrow \text{PREP}) = \log p(\text{PREP}|\text{Handel})
\]

\[
= \log \sum_{w_s} p(\text{PREP}|w_s) \cdot p(w_s|\text{Handel})
\]

| \(w_s\) | \(p(w_s|\text{Handel})\) |
|--------|-----------------|
| trade  | 0.5              |
| deal   | 0.5              |
Query Score Computation

Score(\textit{Handel} \rightarrow \text{PREP}) = \log p(\text{PREP} | \textit{Handel})

= \log \sum_{w_s} p(\text{PREP} | w_s) \ p(w_s | \textit{Handel})

\(w_s = \text{trade}\)

| \(w_s\)  | \(p(w_s | \textit{Handel})\) |
|---------|------------------|
| \textit{trade}  | 0.5   |
| \textit{deal}   | 0.5   |
Query Score Computation

\[ \text{Score}(\text{Handel} \rightarrow \text{PREP}) = \log p(\text{PREP} | \text{Handel}) \]

\[ = \log \sum_{w_s} p(\text{PREP} | w_s) \ p(w_s | \text{Handel}) \]

\[ w_s = \text{trade} \]

| \(x\) | \(p(x | \text{trade})\) |
|------|---------------------|
| PREP | 0.5                 |
| ADJ  | 0.5                 |

| \(w_s\) | \(p(w_s | \text{Handel})\) |
|--------|---------------------|
| trade  | 0.5                 |
| deal   | 0.5                 |
Feature Specialization

Feature: Goodness in English
Value: Score\(\text{hofft} \rightarrow \text{PREP}\)

Feature: Goodness in English
Value: Score\(\text{Handel} \rightarrow \text{PREP}\)

Mosambik
Mozambique
hofft
hope
auf
on
Handel
trade
mit
with
andreren
other
Mitgliedern
members
Feature: Goodness (VERB link) in English
Value: Score(hofft → PREP)

Feature: Goodness in English
Value: Score(Handel → PREP)
Feature: Goodness (VERB link) in English
Value: Score(hofft → PREP)

Feature: Goodness (NOUN link) in English
Value: Score(Handel → PREP)

Mosambik
Mozambique
hofft
hope
auf
on
Handel
trade
mit
with
anderen
other
Mitgliedern
members
Feature Specialization

Feature: Goodness (VERB link) in English
Value: Score(hofft → PREP)

Feature: Goodness (NOUN link) in English
Value: Score(Handel → PREP)

Similar to selective sharing of Naseem et al. (2012)
Experiments

- Sources of bilingual dictionaries
- Training
- Results
Handel

See also handel

Contents  [hide]

1 German
   1.1 Etymology
   1.2 Noun
      1.2.1 Related terms

German

Etymology
From handeln, later specialized to its current meaning.

Noun
Handel m

1. deal
2. trade, trading
Handel

See also handel

Contents [hide]

1 German
   1.1 Etymology
   1.2 Noun
      1.2.1 Related terms

German

Etymology
From handeln, later specialized to its current meaning.

Noun
Handel m

1. deal
2. trade trading
Abstract

We consider the problem of using a bilingual dictionary to transfer lexico-syntactic information from a resource-rich source language to a resource-poor target language. In contrast to past work that used bitexts to transfer analyses of specific sentences at the token level, we instead use features to transfer the behavior of words at a type level. In a discriminative dependency parsing framework, our approach produces gains across a range of target languages, using two different low-resource training methodologies (one weakly supervised and one indirectly supervised) and two different dictionary sources (one manually constructed and one automatically constructed).

Table 14: Syntactic Transfer Using a Bilingual Lexicon

| $w_s$       | $p(w_s | \text{Handel})$ |
|-------------|-------------------------|
| deal        | 0.333                   |
| trade       | 0.333                   |
| trading     | 0.333                   |

Figure 1: Sentences in English and German both containing words that mean "demand." The fact that the English demand takes nouns on its left and right indicates that the German verlangen should do the same, correctly suggesting attachments to Verzicht and Gewerkschaften.
Da ist zum Beispiel der Handel mit Drogen

Trafficking in drugs is one example

Unkontrollierter Handel mit leichten Waffen

Uncontrolled trade in light weapons

Ich will mit dem Handel anfangen

I will start with trade
Da ist zum Beispiel der Handel mit Drogen

Trafficking in drugs is one example

Unkontrollierter Handel mit leichten Waffen

Uncontrolled trade in light weapons

Ich will mit dem Handel anfangen

I will start with trade

| $w_s$     | $p(w_s|\text{Handel})$ |
|-----------|------------------------|
| trade     | 0.66                   |
| trafficking | 0.33                 |
Unkontrollierter Handel mit leichten Waffen

Uncontrolled trade in light weapons

Unkontrollierter Handel mit Drogen

Traffic in drugs is one example

Ich will mit dem Handel anfangen

I will start with trade

| $w_s$       | $p(w_s|\text{Handel})$ |
|-------------|-------------------------|
| trade       | 0.768                   |
| trafficking | 0.057                   |
| commerce    | 0.05                    |
| trading     | 0.03                    |
| market      | 0.02                    |
| marketing   | 0.02                    |
| business    | 0.013                   |
| traffic     | 0.007                   |
| sales       | 0.002                   |
| deal        | 0.002                   |
| ...         | ...                     |
Small Target Training Set

- 8 target languages, CoNLL shared task treebanks
- Train on 400 trees from CoNLL training set
- Test on CoNLL test set
Small Target Training Set

- 8 target languages, CoNLL shared task treebanks
- Train on 400 trees from CoNLL training set
- Test on CoNLL test set

- Universal POS tags generated from gold tags using mapping of Petrov et al. (2011)
Small Target Training Set

- 8 target languages, CoNLL shared task treebanks
- Train on 400 trees from CoNLL training set
- Test on CoNLL test set

- Universal POS tags generated from gold tags using mapping of Petrov et al. (2011)

- Source corpus: parsed English Gigaword
Abstract

We consider the problem of using a bilingual dictionary to transfer lexico-syntactic information from a resource-rich source language to a resource-poor target language. In contrast to past work that used bitexts to transfer analyses of specific sentences at the token level, we instead use features to transfer the behavior of words at a type level. In a discriminative dependency parsing framework, our approach produces gains across a range of target languages, using two different low-resource training methodologies (one weakly supervised and one indirectly supervised) and two different dictionary sources (one manually constructed and one automatically constructed).

Figure 1: Sentences in English and German containing words that mean “demand.” The fact that the English demand takes nouns on its left and right indicates that the German verlangen should do the same, correctly suggesting attachments to Verzicht and Gewerkschaften.

Table 14: Small Target Training Set
Abstract

We consider the problem of using a bilingual dictionary to transfer lexico-syntactic information from a resource-rich source language to a resource-poor target language. In contrast to past work that used bitexts to transfer analyses of specific sentences at the token level, we instead use features to transfer the behavior of words at a type level. In a discriminative dependency parsing framework, our approach produces gains across a range of target languages, using two different low-resource training methodologies (one weakly supervised and one indirectly supervised) and two different dictionary sources (one manually constructed and one automatically constructed).

DELEX \text{MANUAL} + PROJ (MATERIAL)

<table>
<thead>
<tr>
<th>Language</th>
<th>DELEX (MATERIAL)</th>
<th>DELEX+PROJECT (MANUAL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DE</td>
<td>324K 0.89</td>
<td>22K 0.64</td>
</tr>
<tr>
<td>EL</td>
<td>196K 0.94</td>
<td>23K 0.43</td>
</tr>
<tr>
<td>ES</td>
<td>165K 0.89</td>
<td>206K 0.74</td>
</tr>
<tr>
<td>IT</td>
<td>158K 0.91</td>
<td>78K 0.65</td>
</tr>
<tr>
<td>NL</td>
<td>251K 0.87</td>
<td>50K 0.72</td>
</tr>
<tr>
<td>PT</td>
<td>165K 0.85</td>
<td>46K 0.53</td>
</tr>
<tr>
<td>SV</td>
<td>307K 0.93</td>
<td>28K 0.60</td>
</tr>
</tbody>
</table>

Figure 1: Sentences in English and German both containing words that mean "demand." The fact that the English demand takes nouns on its left and right indicates that the German verlangen should do the same, correctly suggesting attachments to Verzicht and Gewerkschaften.
Syntactic Transfer Using a Bilingual Lexicon
Greg Durrett, Adam Pauls, and Dan Klein
Computer Science Division
University of California, Berkeley
{gdurrett,adpauls,klein}@cs.berkeley.edu

Abstract
We consider the problem of using a bilingual dictionary to transfer lexico-syntactic information from a resource-rich source language to a resource-poor target language. In contrast to past work that used bitexts to transfer analyses of specific sentences at the token level, we instead use features to transfer the behavior of words at a type level. In a discriminative dependency parsing framework, our approach produces gains across a range of target languages, using two different low-resource training methodologies (one weakly supervised and one indirectly supervised) and two different dictionary sources (one manually constructed and one automatically constructed).

DELEX
DELEX+PROJ(MANUAL)
DELEX+PROJ(AUTOMATIC)

score(handel prep) = p(handel | prep) p(was | handel) p(x | trade)

Figure 1: Sentences in English and German both containing words that mean “demand.” The fact that the English demand takes nouns on its left and right indicates that the German verlangen should do the same, correctly suggesting attachments to Verzicht and Gewerkschaften.

Table 14: Small Target Training Set
Small Target Training Set

UAS (8 languages averaged)

<table>
<thead>
<tr>
<th>Method</th>
<th>400 train trees</th>
<th>Open-class token coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>DELEX</td>
<td>75.1</td>
<td>60.8</td>
</tr>
<tr>
<td>DELEX+PROJ (MANUAL)</td>
<td>76.7 +1.6</td>
<td>89.9</td>
</tr>
<tr>
<td>DELEX+PROJ (AUTOMATIC)</td>
<td>77.4 +2.3</td>
<td></td>
</tr>
</tbody>
</table>

Abstract

We consider the problem of using a bilingual dictionary to transfer lexico-syntactic information from a resource-rich source language to a resource-poor target language. In contrast to past work that used bitexts to transfer word order and affixes, our approach produces gains across a range of target languages, using two different low-resource training methodologies (one weakly supervised and one indirectly supervised) and two different dictionary sources (one manually constructed and one automatically constructed).

In a discriminative dependency parsing framework, we instead use features to transfer the behavior of words at a type level. In a distributional level, we instead use features to transfer the meaning of words that mean “demand.” The fact that the English sentence “the senators demand abandonment on the reform” translates into the German sentence “die Gewerkschaften verlangen Verzicht auf die Reform,” which is correct, shows that our approach correctly suggests “the” (DE) to “die” (DT) and “demands” (VERLABEN) to “verlangen” (VM). Figure 1 shows a sample of sentences in English and German.

Table 14:

<table>
<thead>
<tr>
<th>Method</th>
<th>400 train trees</th>
<th>Open-class token coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>DELEX</td>
<td>75.1</td>
<td>60.8</td>
</tr>
<tr>
<td>DELEX+PROJ (MANUAL)</td>
<td>76.7 +1.6</td>
<td>89.9</td>
</tr>
<tr>
<td>DELEX+PROJ (AUTOMATIC)</td>
<td>77.4 +2.3</td>
<td></td>
</tr>
</tbody>
</table>

Our approach produces gains across a range of target languages, using two different low-resource training methodologies (one weakly supervised and one indirectly supervised) and two different dictionary sources (one manually constructed and one automatically constructed).
Abstract

We consider the problem of using a bilingual dictionary to transfer lexico-syntactic information from a resource-rich source language to a resource-poor target language. In contrast to past work that used bitexts to transfer analyses of specific sentences at the token level, we instead use features to transfer the behavior of words at a type level. In a discriminative dependency parsing framework, our approach produces gains across a range of target languages, using two different low-resource training methodologies (one weakly supervised and one indirectly supervised) and two different dictionary sources (one manually constructed and one automatically constructed).

Table 14:

<table>
<thead>
<tr>
<th>Language</th>
<th>Small Target Training Set</th>
</tr>
</thead>
<tbody>
<tr>
<td>DE</td>
<td>324K 0.91 22K 0.64</td>
</tr>
<tr>
<td>DE</td>
<td>320K 0.89 58K 0.55</td>
</tr>
<tr>
<td>EL</td>
<td>196K 0.94 23K 0.43</td>
</tr>
<tr>
<td>ES</td>
<td>165K 0.89 206K 0.74</td>
</tr>
<tr>
<td>IT</td>
<td>158K 0.91 78K 0.65</td>
</tr>
<tr>
<td>NL</td>
<td>251K 0.87 50K 0.72</td>
</tr>
<tr>
<td>PT</td>
<td>165K 0.85 46K 0.53</td>
</tr>
<tr>
<td>SV</td>
<td>307K 0.93 28K 0.60</td>
</tr>
</tbody>
</table>

Figure 1: Sentences in English and German both containing words that mean "demand." The fact that the English demand takes nouns on its left and right indicates that the German verlangen should do the same, correctly suggesting attachments to Verzicht and Gewerkschaften.

Score(Handel PREP) = p(PREP | Handel) p(w | Handel) p(x | trade)

The senators demand strict new ethics rules...

DT NNS VBP JJ JJ NNS NNS

Gewerkschaften verlangen Verzicht auf die Reform
NN VVFIN NN APPR

Unions demand abandonment on the reform
ART NN

Figure 1: Sentences in English and German both containing words that mean "demand." The fact that the English demand takes nouns on its left and right indicates that the German verlangen should do the same, correctly suggesting attachments to Verzicht and Gewerkschaften.

Table 14:

<table>
<thead>
<tr>
<th>Language</th>
<th>Small Target Training Set</th>
</tr>
</thead>
<tbody>
<tr>
<td>DE</td>
<td>324K 0.91 22K 0.64</td>
</tr>
<tr>
<td>DE</td>
<td>320K 0.89 58K 0.55</td>
</tr>
<tr>
<td>EL</td>
<td>196K 0.94 23K 0.43</td>
</tr>
<tr>
<td>ES</td>
<td>165K 0.89 206K 0.74</td>
</tr>
<tr>
<td>IT</td>
<td>158K 0.91 78K 0.65</td>
</tr>
<tr>
<td>NL</td>
<td>251K 0.87 50K 0.72</td>
</tr>
<tr>
<td>PT</td>
<td>165K 0.85 46K 0.53</td>
</tr>
<tr>
<td>SV</td>
<td>307K 0.93 28K 0.60</td>
</tr>
</tbody>
</table>

Figure 1: Sentences in English and German both containing words that mean "demand." The fact that the English demand takes nouns on its left and right indicates that the German verlangen should do the same, correctly suggesting attachments to Verzicht and Gewerkschaften.

Table 14:

<table>
<thead>
<tr>
<th>Language</th>
<th>Small Target Training Set</th>
</tr>
</thead>
<tbody>
<tr>
<td>DE</td>
<td>324K 0.91 22K 0.64</td>
</tr>
<tr>
<td>DE</td>
<td>320K 0.89 58K 0.55</td>
</tr>
<tr>
<td>EL</td>
<td>196K 0.94 23K 0.43</td>
</tr>
<tr>
<td>ES</td>
<td>165K 0.89 206K 0.74</td>
</tr>
<tr>
<td>IT</td>
<td>158K 0.91 78K 0.65</td>
</tr>
<tr>
<td>NL</td>
<td>251K 0.87 50K 0.72</td>
</tr>
<tr>
<td>PT</td>
<td>165K 0.85 46K 0.53</td>
</tr>
<tr>
<td>SV</td>
<td>307K 0.93 28K 0.60</td>
</tr>
</tbody>
</table>
We consider the problem of using a bilingual dictionary to transfer lexico-syntactic information from a resource-rich source language to a resource-poor target language. In contrast to past work that used bitexts to transfer analyses of specific sentences at the token level, we instead use features to transfer the behavior of words at a type level. In a discriminative dependency parsing framework, our approach produces gains across a range of target languages, using two different low-resource training methodologies (one weakly supervised and one indirectly supervised) and two different dictionary sources (one manually constructed and one automatically constructed).

Table 1: Small Target Training Set

<table>
<thead>
<tr>
<th>Language</th>
<th>Train Trees</th>
<th>Target Trees</th>
<th>UAS (8 languages averaged)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DELEX</td>
<td>100</td>
<td>200</td>
<td>71.8</td>
</tr>
<tr>
<td>DELEX+PROJ</td>
<td>400</td>
<td>200</td>
<td>73.5 (+1.7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>400</td>
<td>73.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>75.7 (+2.1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>75.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>77.4 (+2.3)</td>
</tr>
</tbody>
</table>
No Target Training Set

[McDonald et al. (2011)]
No Target Training Set

Train

- ES trees
  - DET NOUN VERB ADJ

- PT trees
  - NOUN VERB NOUN PREP DET NOUN

- SV trees
  - PRON VERB PRON ADV

Test

- DE CoNLL test set

[McDonald et al. (2011)]
No Target Training Set

Train

ES trees

PT trees

SV trees

Test

DE CoNLL test set

[McDonald et al. (2011)]
Syntactic Transfer Using a Bilingual Lexicon

Greg Durrett, Adam Pauls, and Dan Klein

Computer Science Division
University of California, Berkeley
{gdurrett,adpauls,klein}@cs.berkeley.edu

Abstract

We consider the problem of using a bilingual dictionary to transfer lexico-syntactic information from a resource-rich source language to a resource-poor target language. In contrast to past work that used bitexts to transfer analyses of specific sentences at the token level, we instead use features to transfer the behavior of words at a type level. In a discriminative dependency parsing framework, our approach produces gains across a range of target languages, using two different low-resource training methodologies (one weakly supervised and one indirectly supervised) and two different dictionary sources (one manually constructed and one automatically constructed).

Figure 1: Sentences in English and German both containing words that mean "demand." The fact that the English demand takes nouns on its left and right indicates that the German verlangen should do the same, correctly suggesting attachments to Verzicht and Gewerkschaften.

Table 14: No Target Training Set

UAS (8 languages averaged)

<table>
<thead>
<tr>
<th>Language</th>
<th>Delex</th>
<th>Delex+Proj</th>
</tr>
</thead>
<tbody>
<tr>
<td>DE</td>
<td>324K</td>
<td>0.91</td>
</tr>
<tr>
<td>EL</td>
<td>196K</td>
<td>0.94</td>
</tr>
<tr>
<td>ES</td>
<td>165K</td>
<td>0.89</td>
</tr>
<tr>
<td>IT</td>
<td>158K</td>
<td>0.91</td>
</tr>
<tr>
<td>NL</td>
<td>251K</td>
<td>0.87</td>
</tr>
<tr>
<td>PT</td>
<td>165K</td>
<td>0.85</td>
</tr>
<tr>
<td>SV</td>
<td>307K</td>
<td>0.93</td>
</tr>
</tbody>
</table>

DELEX DELEX+PROJ
No Target Training Set

<table>
<thead>
<tr>
<th></th>
<th>Type-level* (this work)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DELEX</td>
</tr>
<tr>
<td></td>
<td>DELEX+PROJ</td>
</tr>
<tr>
<td><strong>UAS (8 languages averaged)</strong></td>
<td></td>
</tr>
<tr>
<td>65</td>
<td>62.6</td>
</tr>
<tr>
<td>63</td>
<td></td>
</tr>
<tr>
<td>61</td>
<td>+2.0</td>
</tr>
<tr>
<td>60.6</td>
<td></td>
</tr>
<tr>
<td>59</td>
<td></td>
</tr>
<tr>
<td>57</td>
<td></td>
</tr>
<tr>
<td>55</td>
<td></td>
</tr>
</tbody>
</table>

Table 14: No Target Training Set
No Target Training Set

Type-level* (this work)

- DELEX
- DELEX+PROJ

<table>
<thead>
<tr>
<th>UAS (8 languages averaged)</th>
<th>60.6</th>
<th>62.6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+2.0</td>
<td></td>
</tr>
</tbody>
</table>

Token-level (McDonald et al. 2011)

- MULTIDIR
- MULTIPROJ

<table>
<thead>
<tr>
<th>UAS (8 languages averaged)</th>
<th>61.1</th>
<th>63.8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>+2.7</td>
</tr>
</tbody>
</table>
We consider the problem of using a bilingual dictionary to transfer lexico-syntactic information from a resource-rich source language to a resource-poor target language. In contrast to past work that used bitexts to transfer analyses of specific sentences at the token level, we instead use features to transfer the behavior of words at a type level. In a discriminative dependency parsing framework, our approach produces gains across a range of target languages, using two different low-resource training methodologies (one weakly supervised and one indirectly supervised) and two different dictionary sources (one manually constructed and one automatically constructed).

Figure 1: Sentences in English and German both containing words that mean “demand.” The fact that the English demand takes nouns on its left and right indicates that the German verlangen should do the same, correctly suggesting attachments to V erzicht and Gewerkschaften.
Related Work

- Täckström et al. (2012): type-level transfer using word clusters
Related Work

- Täckström et al. (2012): type-level transfer using word clusters

- Naseem et al. (2012): adaptation of universal POS transfer to the target language
Conclusion

- Bilingual dictionaries allow transfer of linguistic knowledge at a type level
Conclusion

- Bilingual dictionaries allow transfer of linguistic knowledge at a type level

- Type-level transfer:
  - Comparable to token-level transfer
  - Simpler to implement
  - Does not require bitext
Conclusion

- Bilingual dictionaries allow transfer of linguistic knowledge at a type level

- Type-level transfer:
  - Comparable to token-level transfer
  - Simpler to implement
  - Does not require bitext

Thank you!