

# Cross-lingual Transfer of Semantic Role Labeling Models

Mikhail Kozhevnikov and Ivan Titov

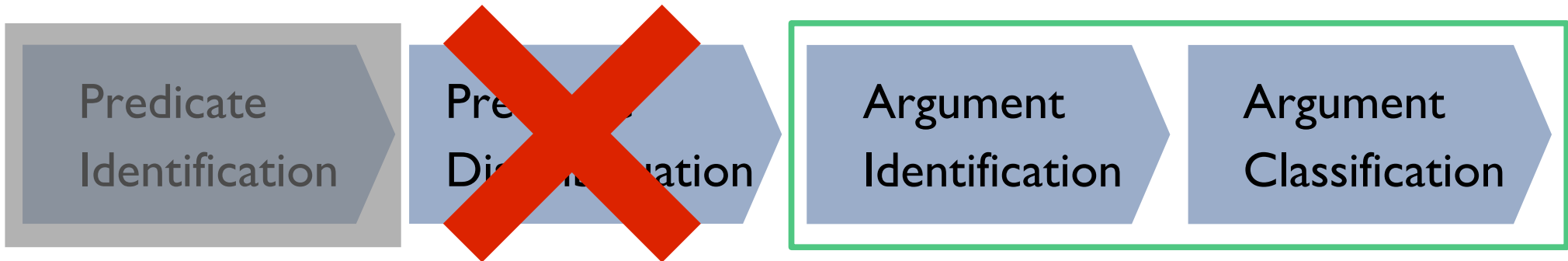
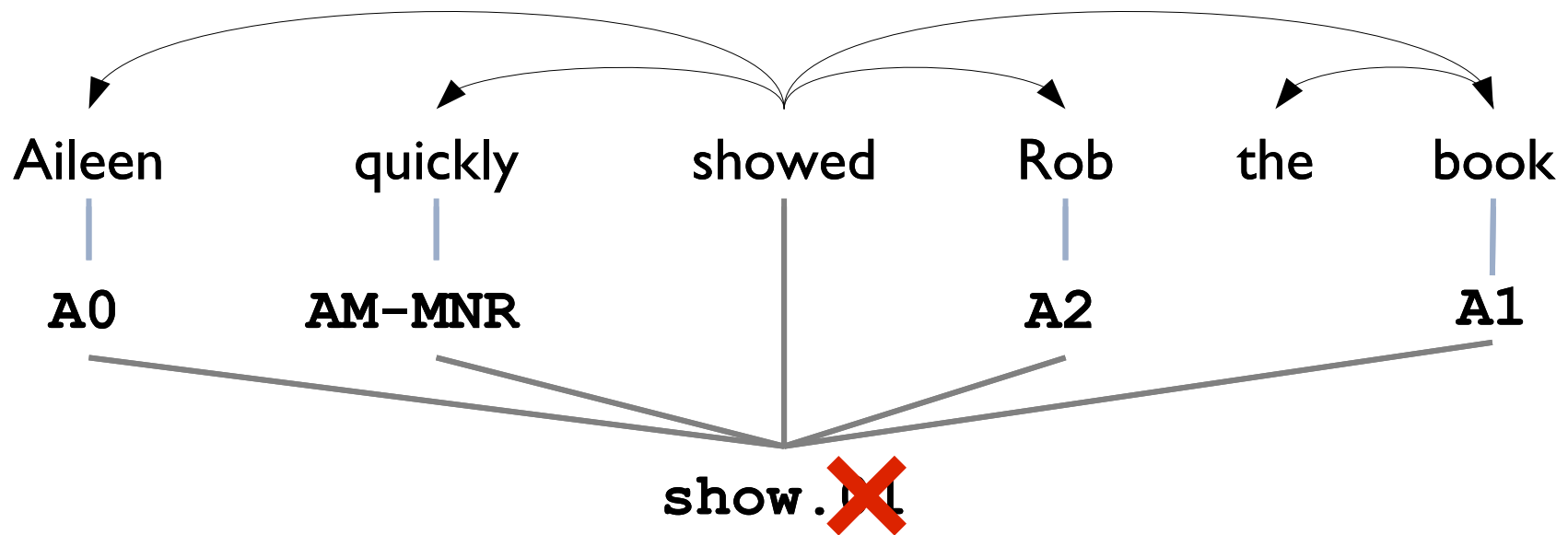


UNIVERSITÄT  
DES  
SAARLANDES



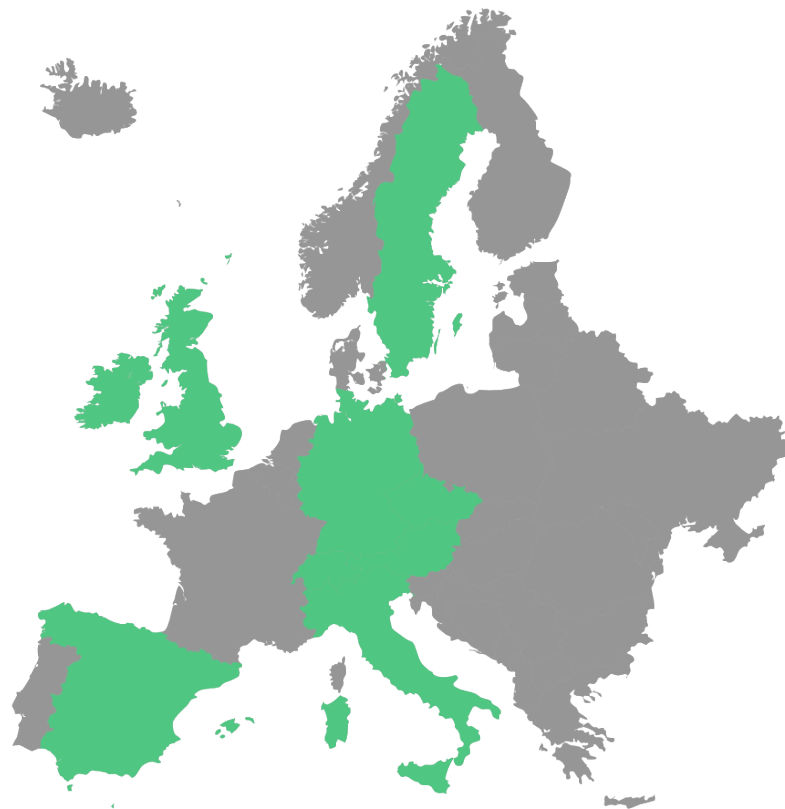
# Semantic Role Labeling

Dependency-based, like in CoNLL 2009 ST



## The Low-resource Setting

- ▶ Training requires large amounts of annotated data
- ▶ Even large corpora face coverage problems
- ▶ Very little or no data for many new languages



# Unsupervised SRL

[Grenager and Manning, 2006]

[Titov and Klementiev, 2012]

[Lang and Lapata, 2010]

Aileen quickly showed Rob the book  
c205 c087 c949 c013

The film was shown last night  
c013 c410

Show me the map  
c949 c013

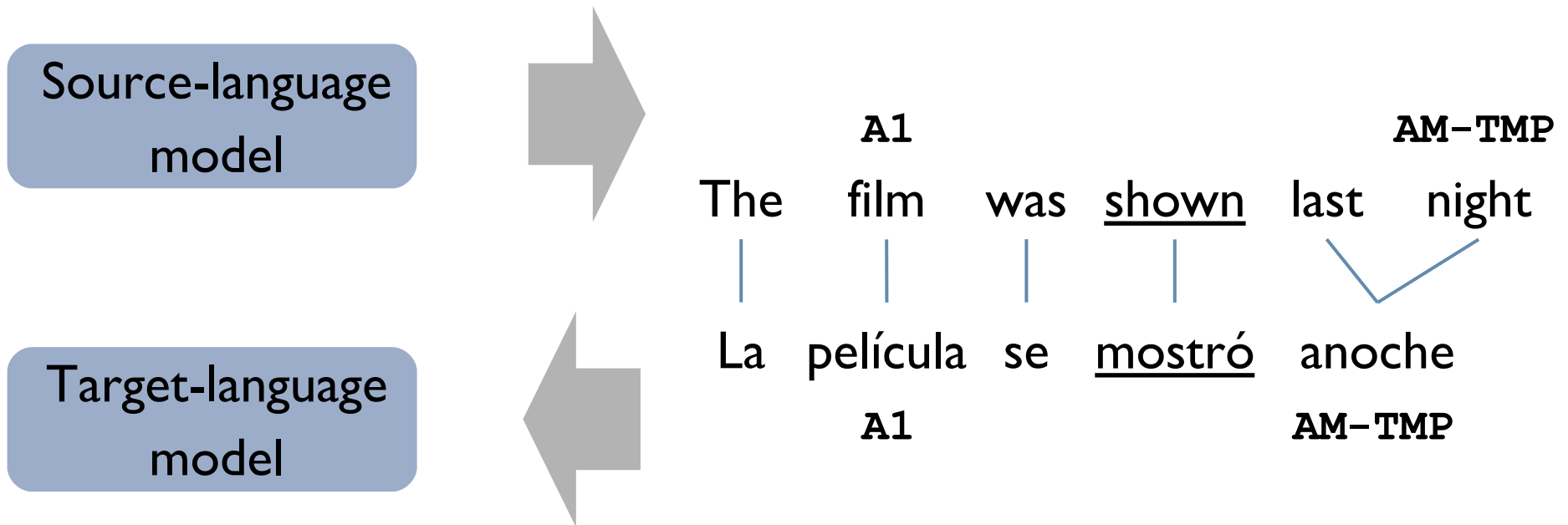
# Cross-lingual Approaches: Projection

[Pado and Lapata, 2009]

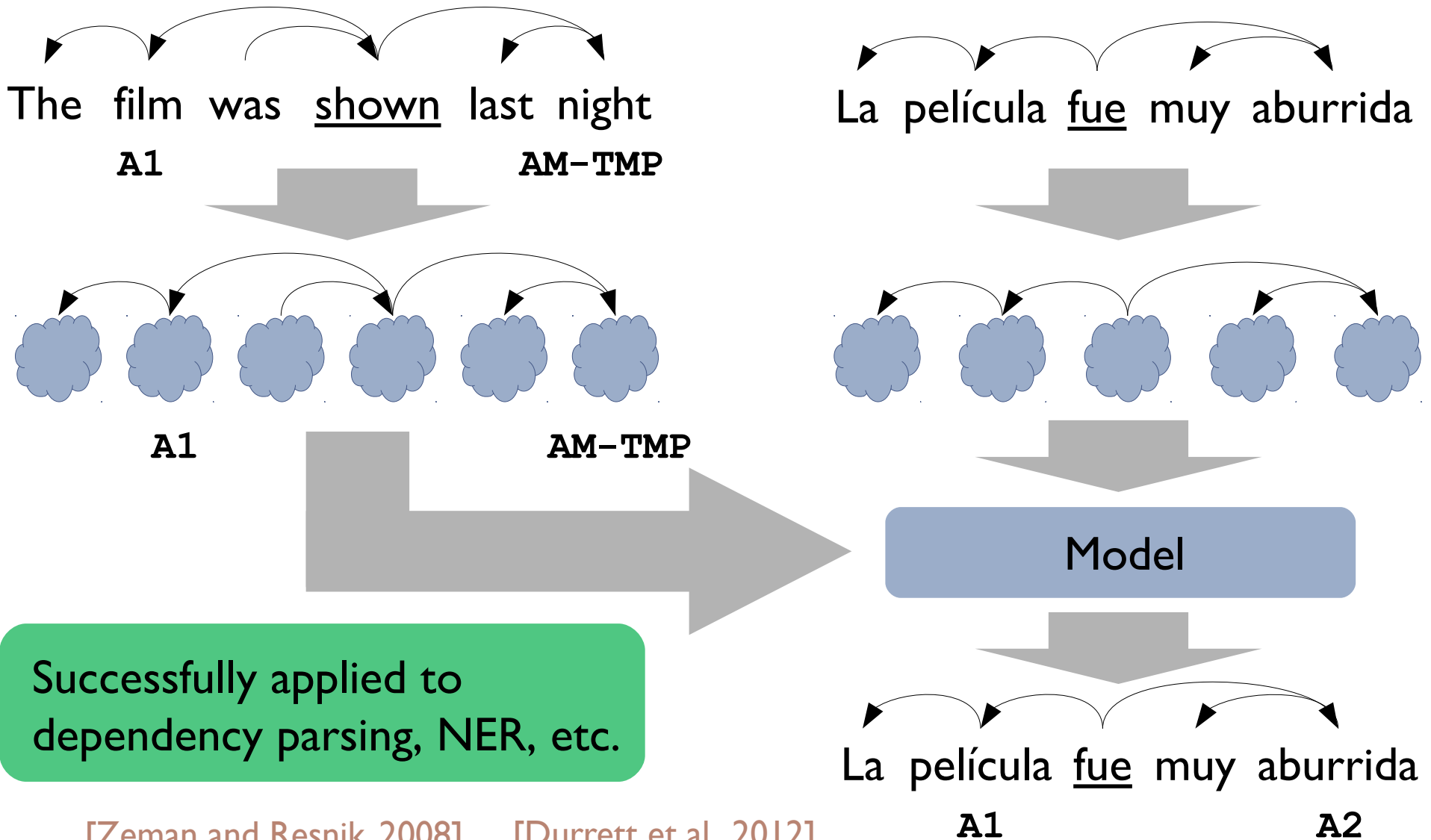
[Annesi and Basili, 2010]

[van der Plas et al., 2011]

- ▶ Run source-language model on the source side
- ▶ Propagate annotations through word-alignment links
- ▶ Train a target-language model on the output



# Cross-lingual Approaches: Model Transfer



Successfully applied to  
dependency parsing, NER, etc.

[Zeman and Resnik, 2008]

[Durrett et al., 2012]

[Søgaard, 2011]

[McDonald et al., 2011]

# Overview

## ▶ Purpose

- ▶ Create a simple model
- ▶ Compare against the alternatives in low-resource setting
- ▶ Figure out which features are useful

## ▶ Model

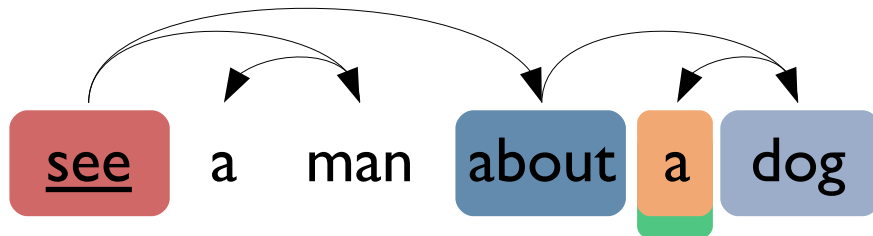
- ▶ Independent linear classifiers for each argument
- ▶ No feature selection, no second-order features

# Outline

- ▶ Motivation
- ▶ Cross-lingual Approaches
- ▶ Shared Feature Representation
- ▶ Evaluation and Baselines
- ▶ Results
- ▶ Conclusions and Outlook

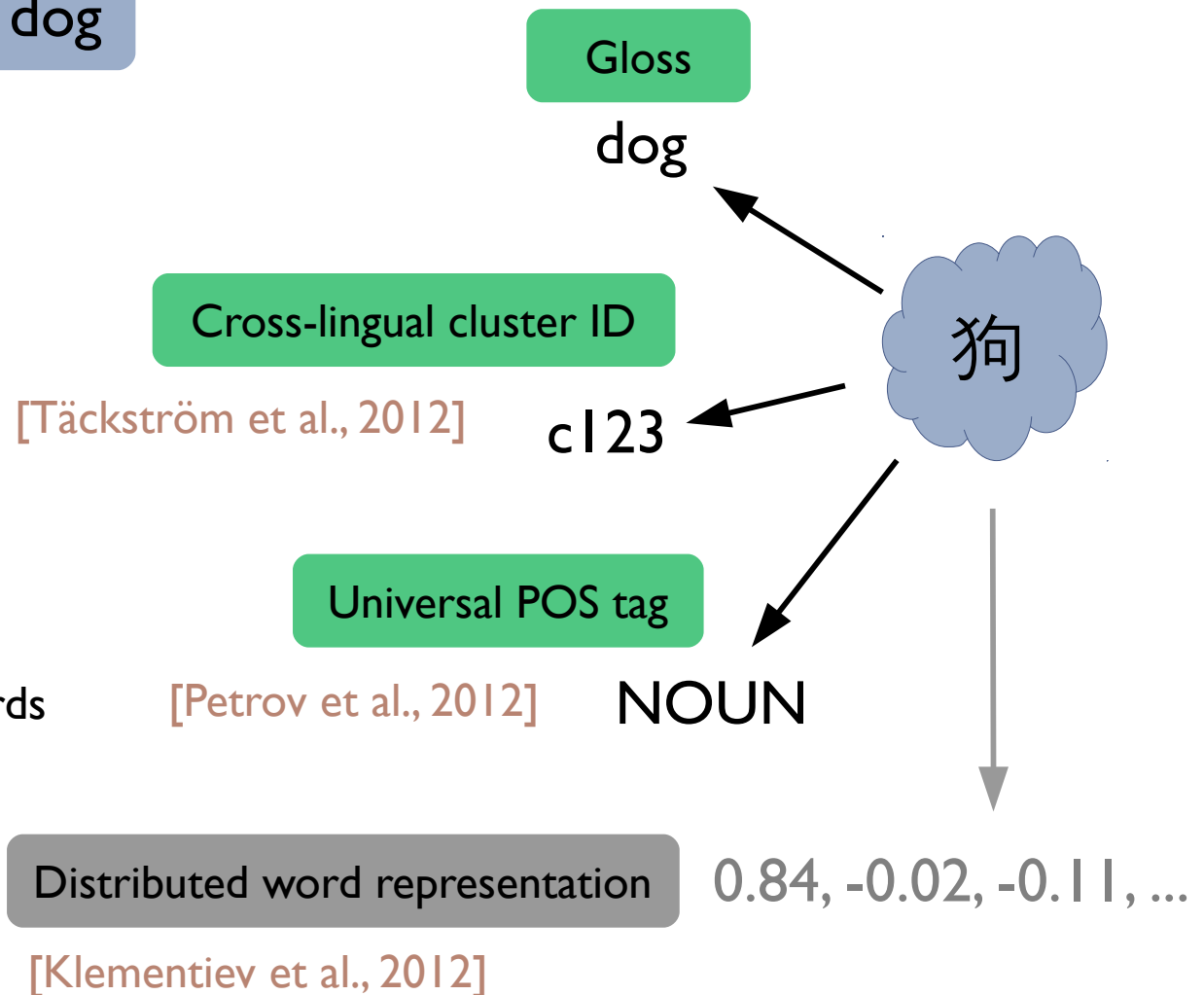


# Shared Feature Representation



Features of an argument instance include attributes of

- Argument word
- Predicate word
- Parent
- Children
- Siblings
- Preceding and following words



# Dependency features

## ▶ (Unlabeled) dependency structure

▶ Gold-standard dependencies stripped of dependency relations

▶ Direct transfer [McDonald et al., 2011]

## ▶ Dependency relations

▶ Currently only PCEDT

▶ Need more homogeneous treebanks

[Zeman et al., 2012] [McDonald et al., 2013]

Pair	UAS
En-Zh	35%
Zh-En	42%
En-Cz	36%
Cz-En	39%
En-Fr*	67%

En-Fr: evaluation against predicted, not gold syntax

### ▶ English-Czech

- ▶ Prague Czech-English Dependency Treebank 2.0 [Hajič et al., 2012]
- ▶ Similar dependency relations on the two sides

### ▶ English-French

- ▶ Annotation projection from English (PropBank)
- ▶ 1000 sentences manually corrected [van der Plas et al., 2011]

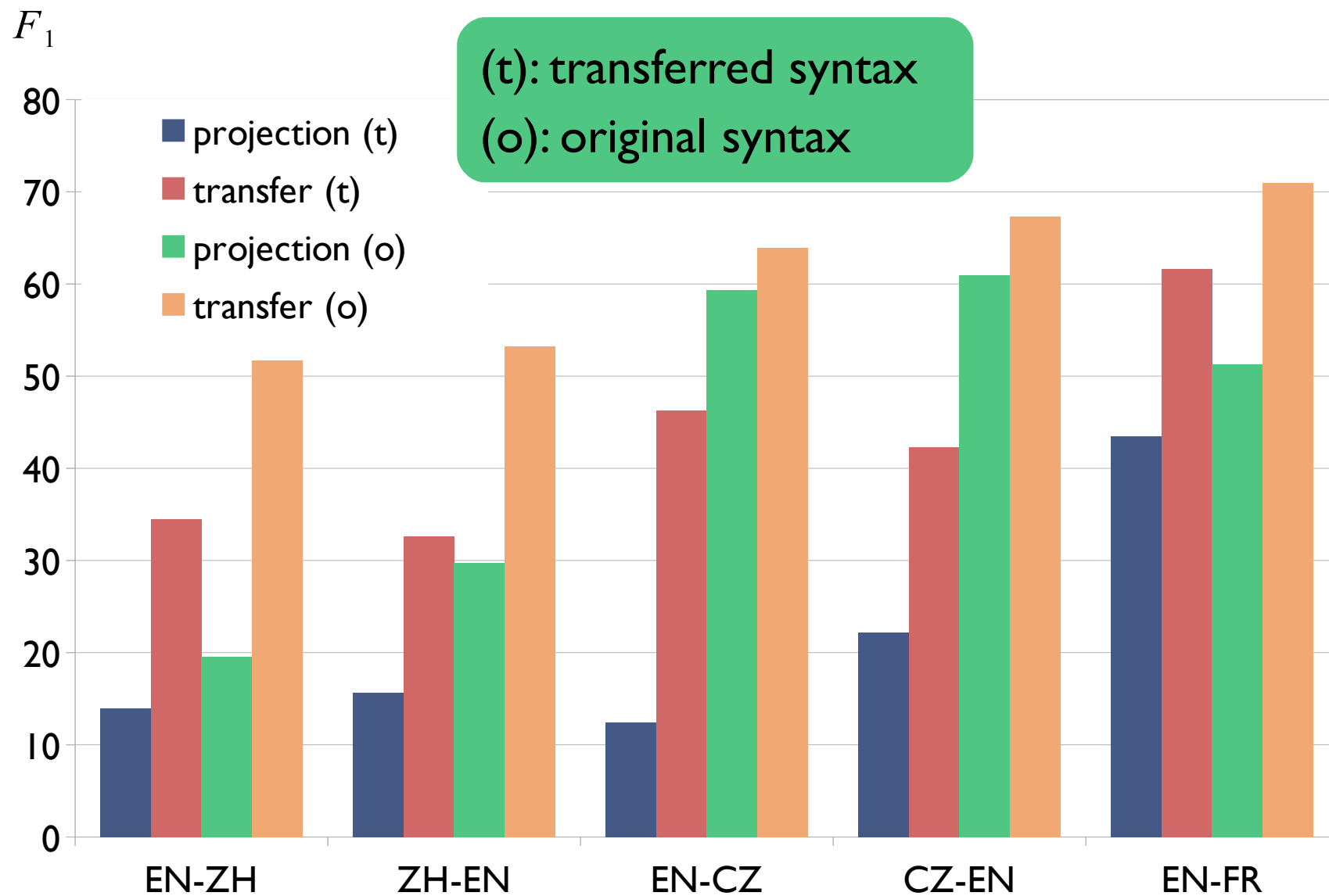
### ▶ English-Chinese

- ▶ Chinese Treebank, guidelines similar to PropBank
- ▶ Core roles only (no modifiers)

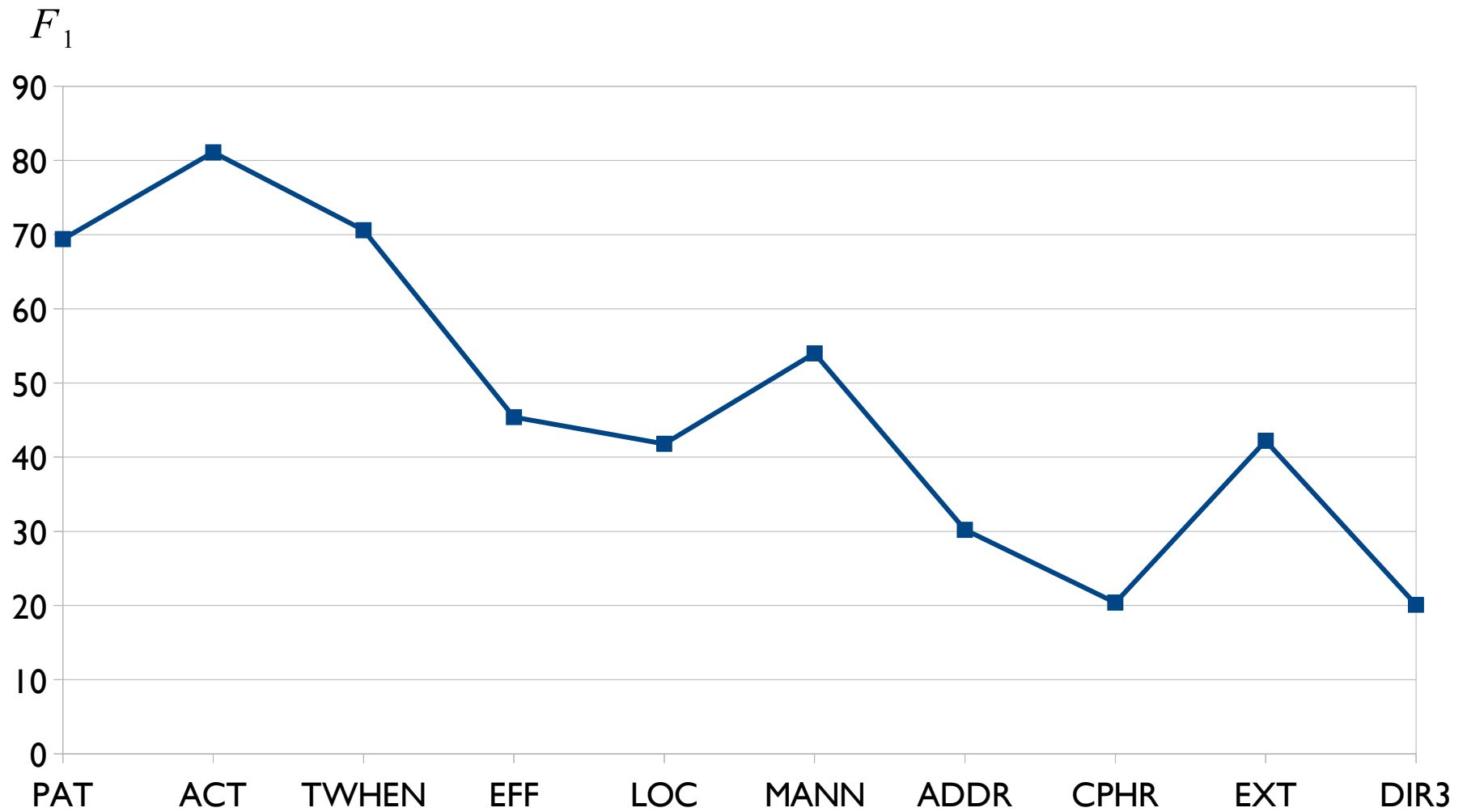
# Baselines

- ▶ **Annotation projection baseline**
  - ▶ Apply annotation projection to parallel data (except for French)
  - ▶ Train a lexicalized model on the output
  - ▶ Compare in terms of F1 score
- ▶ **Unsupervised baseline** [Titov and Klementiev, 2012]
  - ▶ Compare to an unsupervised SRL system using cluster measures
  - ▶ Classification only

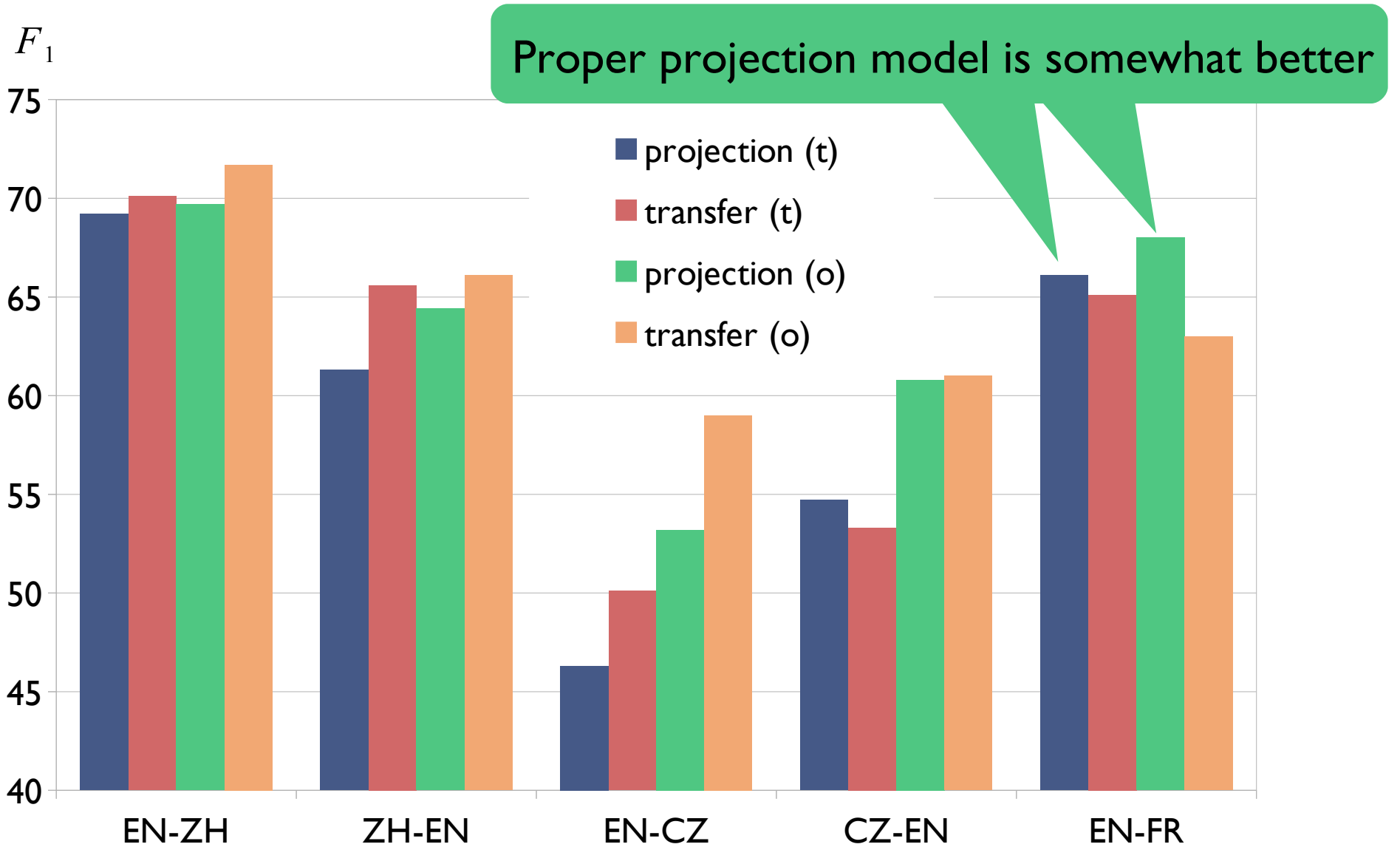
# Argument Identification



# Argument Classification: Top Ten PCEDT 2.0 Labels

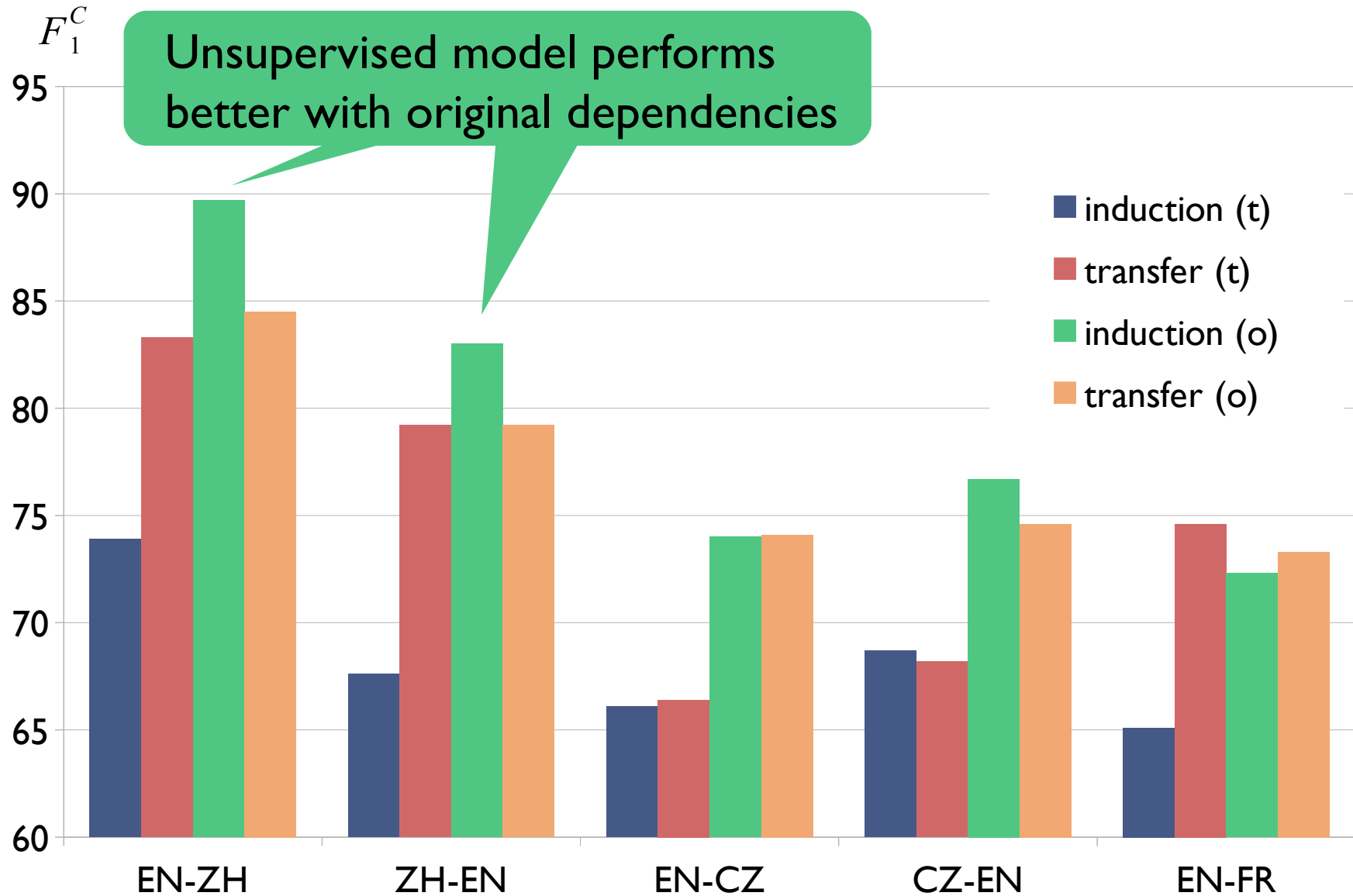


# Argument Classification: Supervised Evaluation



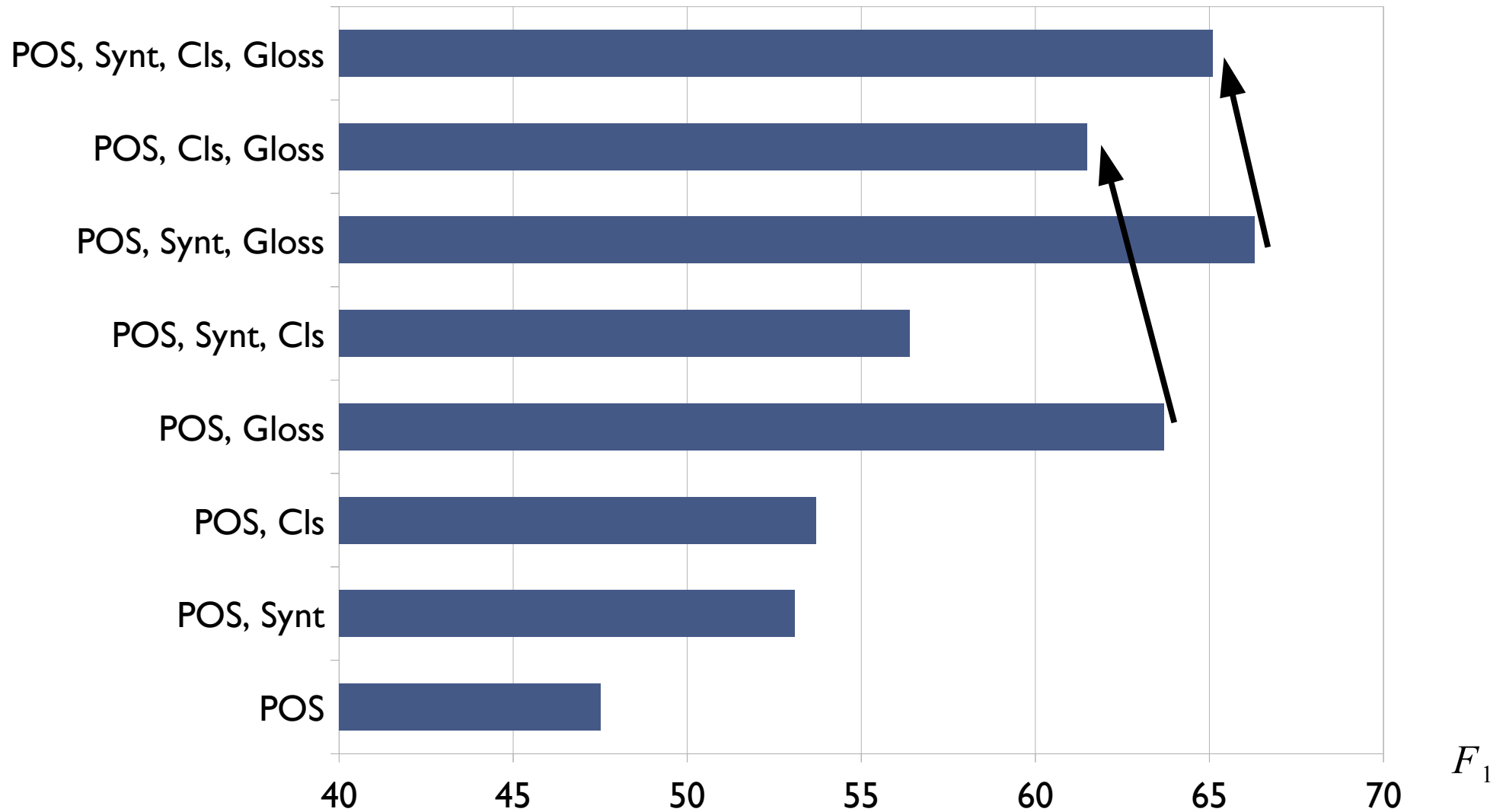
Comparable performance

# Argument Classification: Unsupervised Evaluation



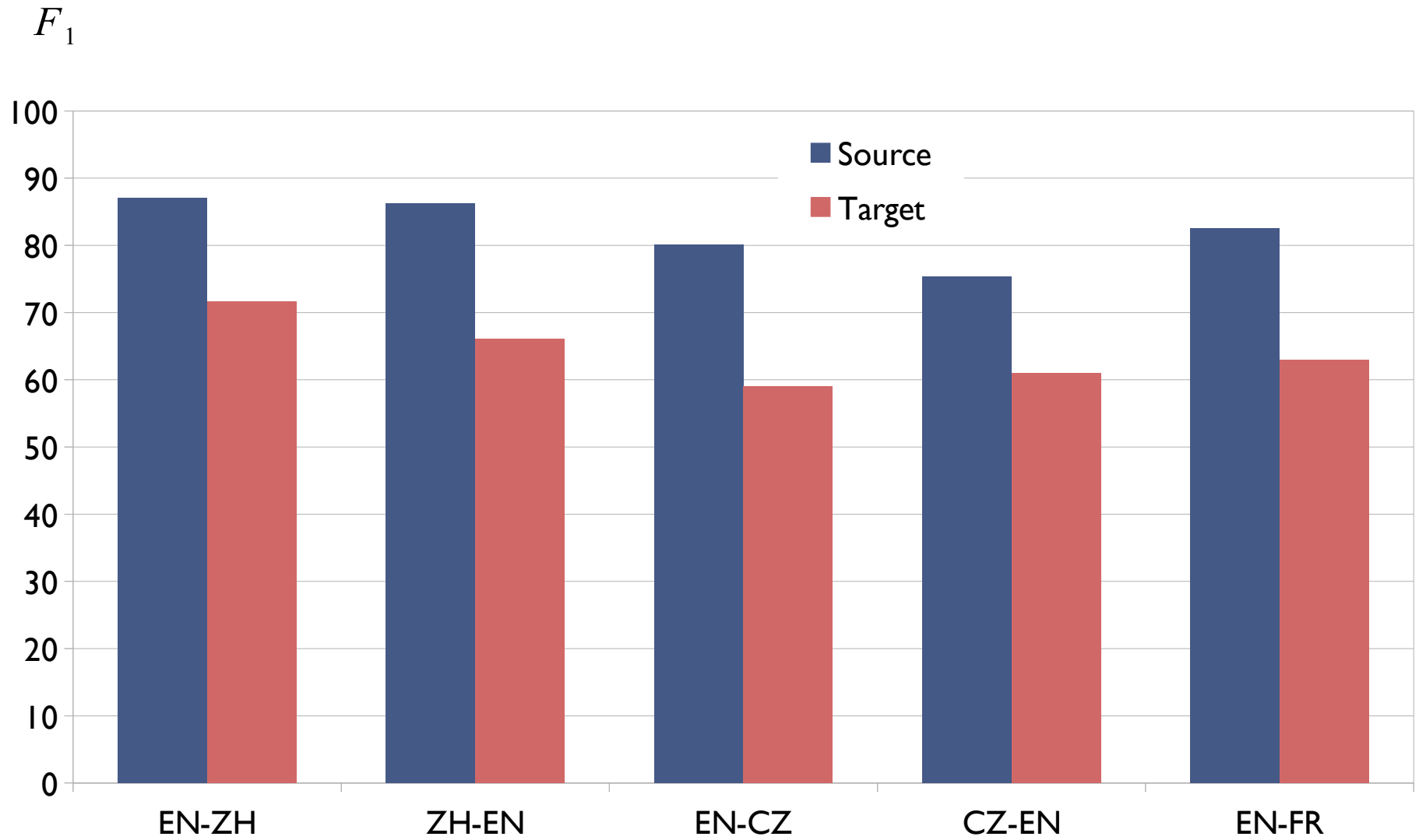


# Feature Group Contribution: En-Fr, transferred syntax



$F_1$

# Performance Drop Due to Transfer



## Conclusions

- ▶ Transfer performs comparably to annotation projection
  - ▶ Easy to implement and trivial to apply
  - ▶ Does not require high-quality parallel data
- ▶ May work better than unsupervised SRL where no accurate dependency parser is available for the target language

## Current and Future Work

- ▶ Better shared representation for dependency features
- ▶ Model inter-argument dependencies
- ▶ Multiple source languages
- ▶ Domain adaptation techniques
  - ▶ Refine using parallel data

# Acknowledgments

- ▶ Alexandre Klementiev
- ▶ Ryan McDonald
- ▶ Oscar Täckström
- ▶ Slav Petrov