

Understanding and Modelling Pronouns in Translation: Resources, Methods, Challenges and Insights

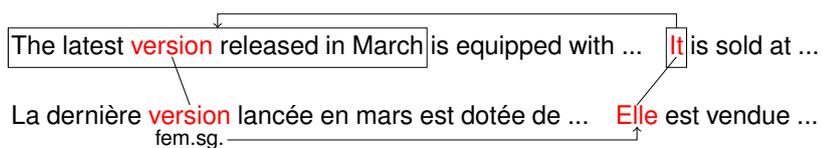
Christian Hardmeier

Uppsala universitet/IT University of Copenhagen

2022-12-01

1

Anaphoric Pronouns: The Prototypical Case



2

Common Assumptions

The antecedent of a referring pronoun is another noun phrase in the text.

Counterexamples:

Pleonastic pronouns:

But I think **it's** a tragedy when one of them doesn't see the other.

Non-nominal reference:

There's so much more information about you,
and **that's** an important thing [...]

- ▶ Evaluation of different pronoun functions
- ▶ Annotation of non-nominal coreference

4

ParCorFull

A *multilingual parallel corpus* with **rich annotations** of coreference.

- ▶ Predecessor: **ParCor**
(Guillou, **Hardmeier**, Smith and Tiedemann, 2014)
 - ▶ English, German and French
 - ▶ 11 TED talks, 8 EU Bookshop docs
 - ▶ Annotations of *pronouns and their direct antecedents*
- ▶ **ParCorFull 1.0**
(Lapshinova-Koltunski, **Hardmeier** and Krielke, 2018)
 - ▶ English and German
 - ▶ 20 TED talks, 25 news articles
 - ▶ Annotation of *nominal and non-nominal coreference*
- ▶ **ParCorFull 2.0**
(Lapshinova-Koltunski, Ferreira, Lartaud and **Hardmeier**, 2022)
 - ▶ English, German, French and Portuguese

5

Coreference annotation in ParCorFull

- ▶ Anaphoric noun phrases (including split antecedents, but not singletons)
 - ▶ Nouns with modifiers, personal and demonstrative pronouns, etc. *[the new report] – [the report] – [it]*
 - ▶ Comparative reference
same, more, less, other, ...
 - ▶ Indefinite pronouns
anyone, someone, ...
 - ▶ Substitution and ellipsis
- ▶ Extratextual reference (to slides, props, etc.)
- ▶ Temporal and local adverbs
[in the 1920s] – [then]; [in the garden] – [there]
- ▶ Event reference

6

Event reference

- ▶ Reference to events (expressed by verb phrases), parts of the discourse, etc.
- ▶ In the original ParCor, this was a catch-all category.

[Democracy is in trouble], no question about [that], and [it] comes in part from a deep dilemma. . .

. . . our mission is [to organize the world's information and make it universally accessible]. And people always say, is [that] really what you guys are still doing?

[And I thought, why can't we do that today]? And [that]'s how this project got going.

7

Translating a pronoun requires generating a matching pronoun in the target language.

Counterexample:

But the thing about tryptamines is **they** cannot be taken orally because **they**'re denatured by an enzyme [. . .] in the human gut [. . .]

Par contre les tryptamines ne peuvent pas [*tryptamines cannot*] être consommées par voie orale étant dénaturé[e]s [*being denatured*] par une enzyme [. . .] dans l'intestin de l'homme [. . .]

- ▶ Recognising and categorising non-literal translation patterns

8

Non-Literal Translation Patterns

- ▶ Matching referring expressions across languages (Lapshinova-Koltunski and **Hardmeier**, DiscoMT 2017; Šoštarić, **Hardmeier** and Szymne, WMT 2018)
 - ▶ in manually annotated data (ParCorFull)
 - ▶ in large unannotated corpora
- ▶ Matching coreference annotations across languages (Lapshinova-Koltunski, Loáiciga, **Hardmeier** and Krielke, CRAC 2019)
- ▶ Methodology:
 - ▶ Automatic word alignment (GIZA++, efmara).
 - ▶ Matching chains.
 - ▶ Finding mismatches in chains (e.g., unaligned words).

9

Explicitation and implicitation

- ▶ Different referring expressions because of content differences.
- ▶ One language has more information than the other.

the French banking giant **Société Générale**, the owner of the local Komerční banka (Commerce Bank)

le géant français **Société Générale**, propriétaire de la banque tchèque Komerční banka.

10

Accommodation of language differences

- ▶ Differences in grammatical systems.
This can often be analysed as obligatory explicitation.
EN: Those are **things** ∅ you have in common with your parents and with your children.
DE: [**Dinge**], [**die**] Sie mit Ihren Eltern und Kindern gemein haben.
- ▶ Differences in linguistic preferences
EN: A reaction to the medication the clinic gave me for my depression left me suicidal.
DE: Die Medikamente, die sie mir in der Ambulanz gegen meine Depressionen gaben, führten bei mir zu Selbstmordgedanken.

11

Different interpretations of corresponding referring expressions

We can create [a global parliament of mayors]. [That]'s an idea.

[We can create a global parliament of mayors]. [That]'s an idea.

EN: Think what happens [when we collect all of that data and we can put it together in order to find patterns we wouldn't see before]₁. [This]₁, I would suggest, perhaps [it]₁ will take a while, but [this]₁ will drive . Fabulous, lots of people talk about .

DE: Was passiert, [wenn wir all diese Daten sammeln und wir sie zusammenfügen können, um Muster zu erkennen, die wir nicht vorher sehen konnten]₁. Vielleicht dauert [dies]₁ ja noch eine Weile, aber [es]₁ wird eine Revolution in der Medizin. Fabelhaft — sehr viele Leute sprechen [darüber]₁.

12

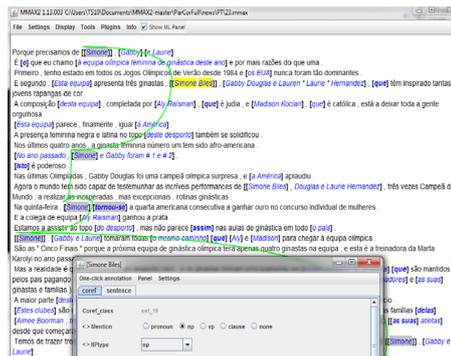
Annotation errors

- ▶ Annotation errors
- ▶ Word alignment errors
 - ▶ Statistical word alignment is a linguistically ill-defined task.
- ▶ Inconsistent interpretation of annotation guidelines across languages

13

Visualisation challenges

- ▶ In each language, we only see
 - ▶ one chain and
 - ▶ the properties of one markable at a time.
- ▶ Very easy to miss inconsistencies even in one language.
- ▶ Word alignment is not shown.



14

Lessons learnt

- ▶ Parallel corpus with rich coreference annotation is a very valuable resource.
- ▶ Difficult to achieve consistent annotation quality, especially over long time.
- ▶ What would we need?
 - ▶ Better corpus visualisation/navigation.
 - ▶ Word alignment with proper linguistic foundation.
 - ▶ Resources for *continuous* quality checks and annotator (re)training.

15

Referring pronouns agree in gender and number with their antecedent.

Counterexample:

Notional concord:

So I think Deep Mind, what's really amazing about Deep Mind is that it can actually – they're learning things in this unsupervised way.

▶ Studying linguistic preferences across languages and genres

16

Understanding Translation

How do the production and interpretation of referring expressions vary across languages?

Human production study:

Referring back to organisational named entities

Last week, Intel announced the shutdown of the factory.
In the press release, _____

FC Barcelona won the World Cup three times since 2009.
Next year, _____

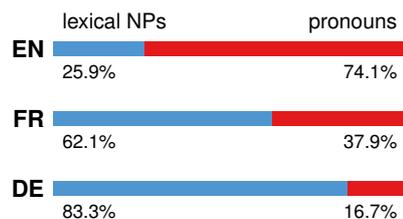
AC/DC achieved international success in 1976.
In the next forty years, _____

Ongoing work with Luca Bevacqua, Sharid Loáiciga and Hannah Rohde

17

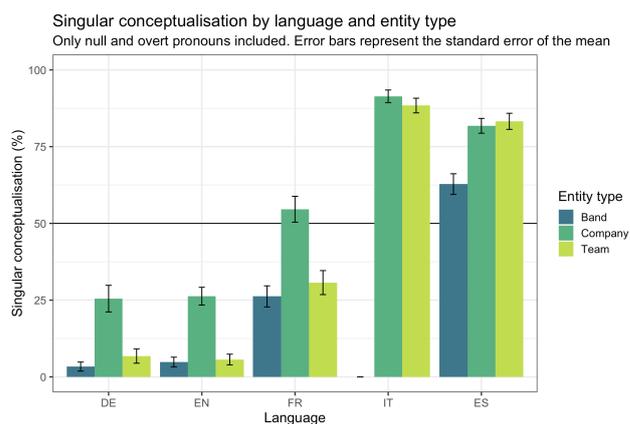
Named Entity Reference: Results

FC Barcelona won the World Cup three times since 2009.
Next year, FC Barcelona/the club/it...



18

Singular vs. plural conceptualisation



19

Two Studies

Study 1: Constructed stimuli

- ▶ Prompt sentences constructed by the authors.
- ▶ Four types of named entities: Companies, publishers, sport teams and music bands.

Study 2: Corpus stimuli

- ▶ Prompt sentences were extracted from OntoNotes and simplified.
- ▶ Continuations were constructed to increase chances of eliciting a reference to the named entity.
- ▶ Generally longer and more complex than the constructed stimuli.
- ▶ Unrelated filler items also based on corpus data.

20

Generating prompts from corpora

Original:

In the final trading, the House was insistent on setting aside \$500 million to carry out base closings ordered to begin in fiscal 1990.

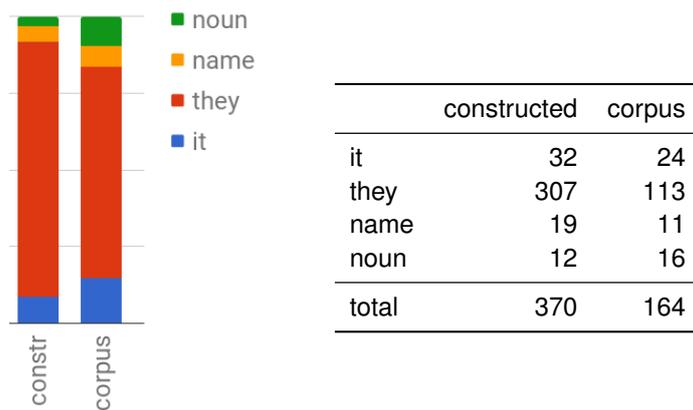
Prompt:

The House showed insistence on setting aside \$500 million to carry out base closings ordered to begin in fiscal 1990.

In an amended piece of legislation, _____

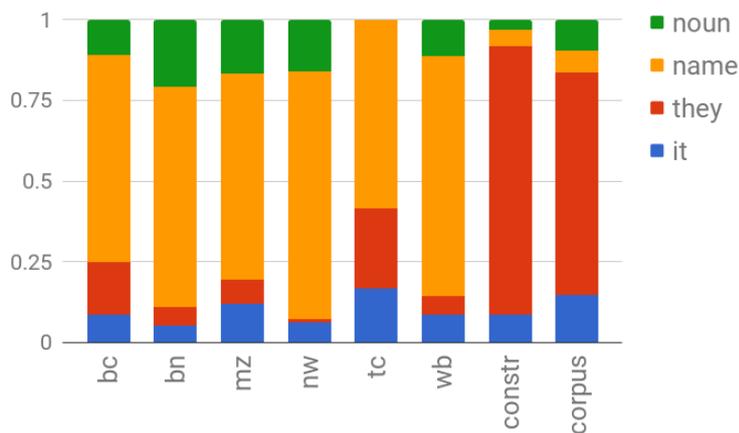
21

Continuation Studies: Results



22

Results including Corpus Study on OntoNotes



23

Conclusions

- ▶ Good pronoun translation is far more complex than enforcing gender agreement.
- ▶ Referring expression use differs significantly across languages. Good translation should respect target language conventions.
- ▶ Genre, register and modality also have strong effects.
- ▶ Annotation and exploration is made difficult by the lack of tools.
- ▶ ParCorFull 2.0 covers 4 European languages and can be used to study these phenomena or construct test suites.

24

Modelling Cross-Lingual Coreference

Work done with Gongbo Tang
(now at Beijing Language and Culture University)



25

Latent Anaphora Resolution for Cross-Lingual Pronoun Prediction

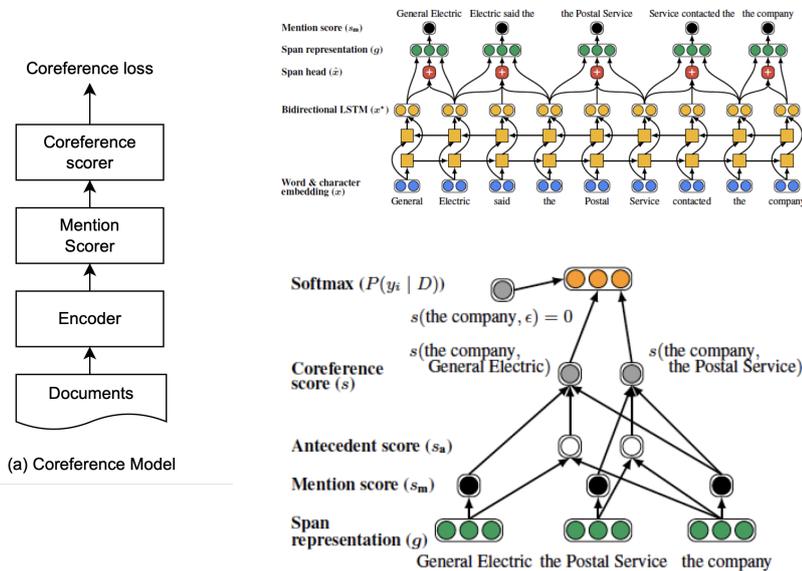
Christian Hardmeier Jörg Tiedemann Joakim Nivre

Uppsala University
Department of Linguistics and Philology
Box 635, 751 26 Uppsala, Sweden
firstname.lastname@lingfil.uu.se

EMNLP 2013

26

Neural coreference models (Lee et al., 2017)



27

Cross-lingual coreference model

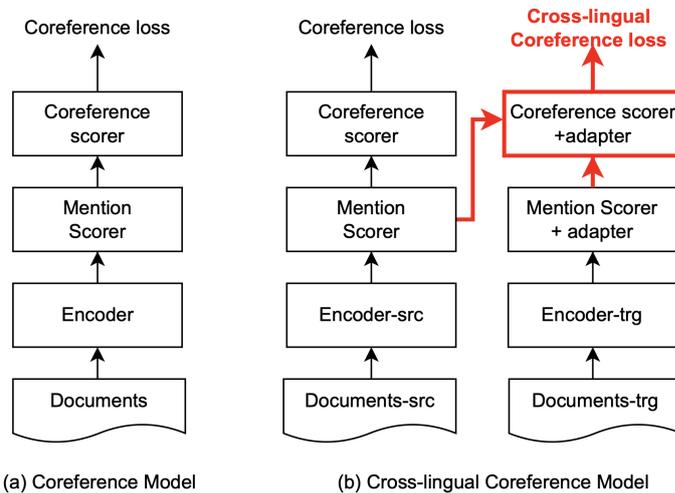
Motivation: Exploit signal from multilingual text for better coreference resolution.

- ▶ Use second “copy” of coreference system in target language.
- ▶ Initialised from pretrained system, with adapter layers.
- ▶ Model scores coreference between target-language anaphors and source-language antecedents.
- ▶ Cross-lingual coreference loss:
 - ▶ Let $S = \{s_1, \dots, s_m\}$ be the source mentions and $T = \{t_1, \dots, t_n\}$ be the target mentions.
 - ▶ The network predicts a score s_{ij} for pairs (s_i, t_j) .

$$\hat{j} = \arg \max_j s_{ij} \text{ for given } i; \quad L = \sum_{i=1}^m e^{-s_{i\hat{j}}}$$

28

Cross-lingual coreference model



29

Experimental results

OntoNotes; TL data synthetically translated with MT systems from Facebook and Helsinki

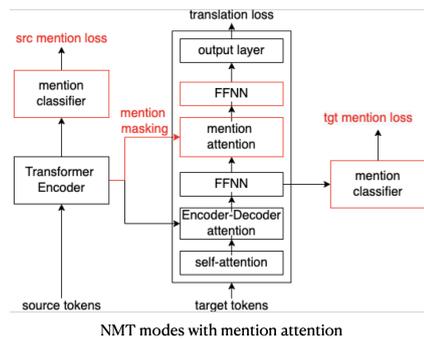
	Mention detection		Coreference	
	F	Δ	F	Δ
English	85.42	–	73.42	–
English–Arabic	86.13	0.71	74.58	1.16
English–Catalan	86.17	0.75	74.81	1.39
English–Chinese	86.02	0.60	74.53	1.11
English–Dutch	86.29	0.87	75.16	1.74
English–French	85.93	0.51	74.37	0.95
English–German	86.02	0.60	74.20	0.78
English–Italian	86.13	0.71	74.65	1.23
English–Russian	86.17	0.75	74.50	1.08
English–Spanish	86.21	0.79	74.50	1.08

30

Models with mention attention

New features:

- ▶ Mention attention module
- ▶ Mention classifiers:
Is this part of a mention?
- ▶ Mention loss
- ▶ Mention masking:
Only pass mention info
to attention module



Loss ratio: MT : M-src : M-tgt = 10 : 1 : 1

31

Experimental results

- ▶ WMT English to German (newstest2017)
- ▶ Evaluation: BLEU; APT for *it, they*

Model	BLEU	Pronouns	Ambig. pronouns
Baseline	28.01	60.1	50.4
Ours	28.23	61.2	52.2

32

Conclusions

- ▶ Cross-lingual data carries information relevant for coreference resolution.
- ▶ Effects on MT/coref performance are very consistent, but rather small.
- ▶ Significant cross-lingual variance in coreference structures for complex and non-obvious reasons.
- ▶ Annotating coreference involves potentially subjective *interpretation* – cross-lingual study exposes this.

33