



Universiteit Utrecht

u^b

^b
UNIVERSITÄT
BERN

Assessing explicitation and implication phenomena in translation using large parallel corpora

—
Jet Hoek Sandrine Zufferey
j.hoek@uu.nl

Joint work with Jacqueline Evers-Vermeul and Ted J.M. Sanders

Vilnius, Lithuania
12 October 2017



Introduction

Coherence relations connect discourse segments

They can, but need not be marked explicitly with a connective or cue phrase

Some relations are easy to convey implicitly

Teen kills younger brother ~~because~~ he thought he deleted his Pokémon
- snopes.com

It's summer, ~~so~~ let the tabloid body shaming begin.
- boingboing.net

while other relations become very hard to reconstruct without explicit marking

5 ways to show you care ~~even though~~ you forgot about Valentine's Day
- 12news.com

If you are a young couple, Toronto Island wants you
- thestar.com

→ Supported by analyses on discourse-annotated corpora
(PDTB – Asr & Demberg 2012, RST – Das & Taboada 2013)

Implicit relations and expectedness

Differences in the linguistic marking of coherence relations have been explained through the notion of ***expectations***

Assumption:

Expected relations are easier to convey implicitly than relations that are not expected

If readers try to establish the simplest possible discourse relation (Traxler et al. 1997), cognitively complex relations should not be expected.

Hypothesis: Cognitively simple relations are more expected than relations that are cognitively more complex

Implicit relations and expectedness

Differences in the linguistic marking of coherence relations have been explained through the notion of ***expectations***

Assumption:

Expected relations are easier to convey implicitly than relations that are not expected

If readers try to establish the simplest possible discourse relation (Traxler et al. 1997), cognitively complex relations should not be expected.

Hypothesis: Cognitively simple relations are more expected than relations that are cognitively more complex

Cross-linguistically!

Cognitive complexity

- Logic
- Language Acquisition
- Language processing
- Mental Space Theory (Fauconnier, 1985)

Coherence relations and cognitive complexity

Cognitive Categories of Coherence relations – CCR

(Sanders, Spooren, & Noordman 1992)

Polarity

Positive: I like him because he always says what he thinks. **P, Q**

Negative: I like him although he always says what he thinks. **P, not-Q**

Hypothesis:

positive > negative

Order

Basic: Because her flight was cancelled, Susan missed the meeting. **P→Q**

Non-basic: Susan missed the meeting because her flight was cancelled. **Q←P**

Hypothesis:

Basic > non-basic

Coherence relations and cognitive complexity

Source of Coherence

- Objective:** The building is falling apart because its foundation was damaged in the storm.
- Subjective:** Something must have come up, because he is never late.
- Speech act:** Since you won't have time tonight, why not do your homework now?

Hypothesis:

Objective ? Speech act > Subjective

Coherence relations and cognitive complexity

Basic operation

Additive: Scott doesn't want to brush his teeth. He also doesn't want to go to bed.

Causal: He is annoying because he is always ridiculously on time.

Conditional: If he ever wants to become a millionaire, he should get off the couch.

Conditional relations are more complex than causal and additive relations

Coherence relations and cognitive complexity

Basic operation

Additive: Scott doesn't want to brush his teeth. He also doesn't want to go to bed.

Causal: He is annoying because he is always ridiculously on time.

Conditional: If he ever wants to become a millionaire, he should get off the couch.

Causal relations are more complex than additive relations?

Acquisition: Bloom et al. (1980), Evers-Vermeul & Sanders (2009)

BUT: once acquired, causal rels are processed faster than additive rels

→ Paradox of causal complexity (Sanders 2005)

→ Only for positive causal relations

→ Only a subset of positive causal relations?

Hypothesis:

Positive additive ? Positive causal > negative additive > negative causal > conditional

Parallel corpus study

Europarl Direct (Koehn 2005; Cartoni, Zufferey, & Meyer 2013)

1916 English source text relations, annotated using CCR

Translations into: Dutch French
 German Spanish

Connectives: *Also* *If*
 Although *In addition*
 Because *So*
 But *Unless*

→How are the relations expressed in the target language?

Implicitation and implicitness

Connectives are very volatile items in translation; they can be added, rephrased or removed (Halverson 2004, Zufferey & Cartoni 2014)

But this variability should be limited by each relation type's potential to remain implicit.

Implicitation and explicitation in translation

- Linguistic difference between languages in a language pair

L1 has feature X, L2 does not

- Implicitation of X from L1 into L2
- Explicitation of X from L1 into L2

- 'Random'

Relation Y is easy to convey implicitly

- Implicitation of Y from L1 into L2 AND from L2 into L1
- Explicitation of Y from L1 into L2 AND from L2 into L1

Implicitation and explicitation in translation

- Linguistic difference between languages in a language pair

L1 has feature X, L2 does not

- Implicitation of X from L1 into L2
- Explicitation of X from L1 into L2

Implicitation hypothesis
(cf. Blum-Kulka 1986)
Asymmetry hypothesis
(cf. Klaudy & Károly, 2005)

- 'Random'

Relation Y is easy to convey implicitly

- Implicitation of Y from L1 into L2 AND from L2 into L1
- Explicitation of Y from L1 into L2 AND from L2 into L1

Implicitation and explicitation in translation

- Linguistic difference between languages in a language pair

L1 has feature X, L2 does not

- Implicitation of X from L1 into L2
- Explicitation of X from L1 into L2

- 'Random'

Relation Y is easy to convey implicitly

- Implicitation of Y from L1 into L2 AND from L2 into L1
- Explicitation of Y from L1 into L2 AND from L2 into L1

The types of relations that are most often implicitated are also the ones most often explicitated in translation

(see also Hoek, Evers-Vermeul, & Sanders 2015)

Translations of coherence relations

Explicit: Because John won the race, he very is happy. }

Paraphrase: John's victory made him very happy. }

Explicit

ExplicitUNDSP: John won the race and was very happy. }

Syntax: John, who won the race, is happy. }

Implicit: John is happy. \emptyset He won the race. }

Implicit

Other:

ParaphraseCONSTR: If we want to stop climate change, we have to...
→ To stop climate change, we have to...

→The relation in the target text has to correspond to the relation in the source text



Results

Logistic regression model

Target language:

Dutch, German, French > Spanish

Order:

Basic > non-basic

Source of Coherence:

Speech act > objective, subjective

Polarity * Basic operation:

pos. causal, pos. additive, **neg. additive** > neg. causal > pos. conditional, neg. conditional

Discussion and conclusions

- Many differences in the marking of coherence relations can be explained in terms of ***cognitive complexity***

Cognitively simple relations are easier to convey implicitly than relations that are cognitively more complex

- This principle seems to hold across languages
 - Test for other, unrelated languages
- Translation corpora can be used to research translation phenomena, but also to investigate mono-lingual (non-translation) phenomena
 - Especially useful when researching **meaning** (cf. Noël 2003)



Universiteit Utrecht

Thank you!



Jet Hoek

j.hoek@uu.nl

Full paper to appear in *Journal of Pragmatics*