

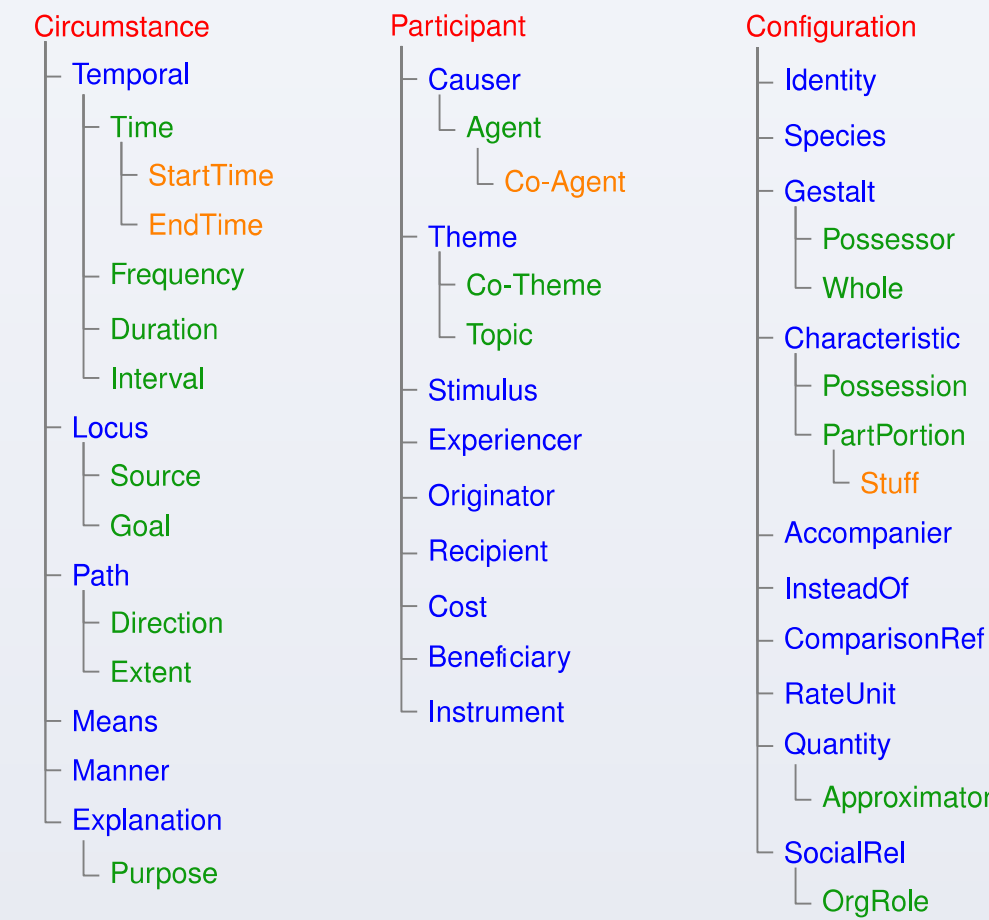
Preposition Supersenses in German-English Parallel Data

Jakob Prange¹ and Nathan Schneider^{1,2}

¹Department of Computer Science, ²Department of Linguistics, Georgetown University

INTRODUCTION

- Prepositions
 - Encode semantic relations and are polysemous (Tyler and Evans, 2003; Durrell and Brée, 1993; Müller et al., 2010)
 - Hard to translate (Hashemi and Hwa, 2014)
- SNACS hierarchy (Schneider et al., 2018)
 - 50 semantic labels for preposition
 - Based on and so far exclusively applied to English
- Case and possessives
 - SNACS framework relies on lexical items as markables
 - Does not support the annotation of morphological case markers
 - We include possessive pronouns (Blodgett and Schneider, 2018), but exclude genitive case marking on nouns, which in German is realized inflectionally, rather than through clitics
 - Case (*dative*, *accusative*) disambiguates locative prepositions in German



Das Bild hängt **über** dem Sofa . Er hängt das Bild **über** das Sofa .
The picture hangs **over** the.DAT sofa.DAT . He hangs the picture **over** the.ACC sofa.ACC .
The picture is above the sofa. He puts the picture above the sofa.

Das Flugzeug fliegt **über** dem Meer . Das Flugzeug fliegt **übers** Meer .
The airplane flies **over** the.DAT sea.DAT . The airplane flies **over-the**.ACC sea.ACC .
The airplane flies above the sea. The plane flies across the sea.

- Construal analysis
 - Each preposition token is labeled for two slots:
 - The *scene role* occupied by the PP in the specific context, which we assume to be largely language-independent and often preserved in translation
 - The *function* prototypically contributed by the preposition type, which may be directly dependent on the lexical choice, a verb's preferred argument structure, or other language-specific parameters



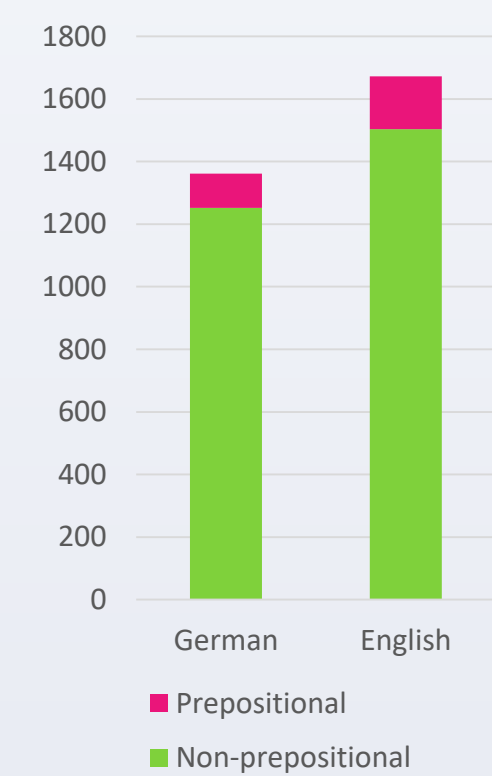
OBJECTIVES

- Corpus linguistics
 - Disambiguate prepositional semantics in German and align with parallel English annotations
 - Measure distributions of prepositions, meanings, and cross-linguistic overlap
- NLP
 - Create a parallel gold-standard corpus of preposition and case semantics
 - Use it to train WSD, alignment, MT, IR systems

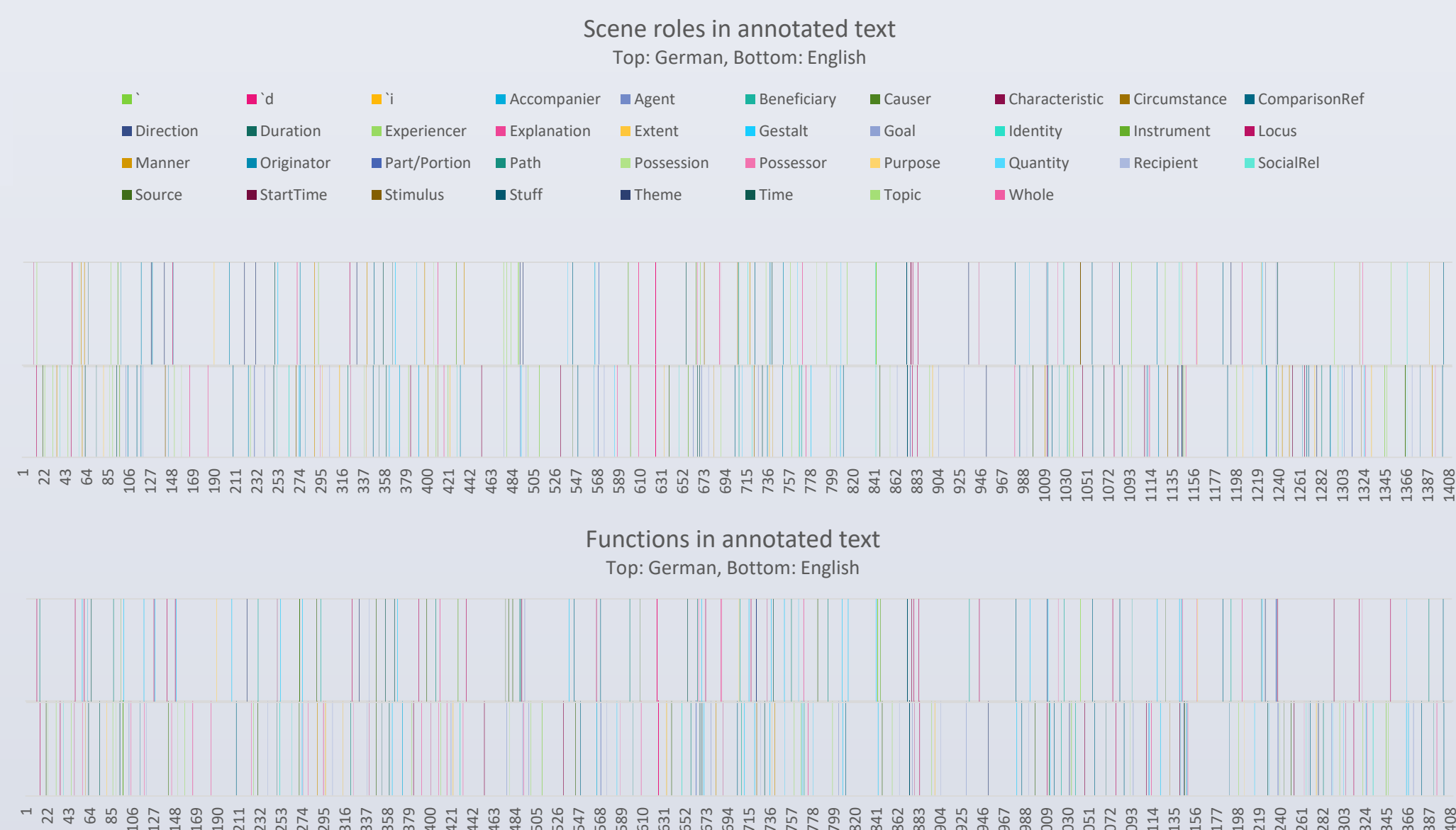
ANNOTATION

- One linguistically trained annotator
- All prepositional expressions and possessive pronouns
- Chapters 1 and 4 of the German translation of Saint-Exupéry's *Le Petit Prince*.
- Manual alignment with the SNACS-annotated parallel English text (Schneider et al., 2018)
- 1,361 tokens in German and 1,672 tokens in English

RESULTS



- Proportion of prepositions used is similar
 - 8% in German, 10% in English
- Only 73 prepositions have a directly corresponding expression
 - English uses a lot of prepositions where German uses case to express the same meaning
 - Include case markers in SNACS!
- Construal analysis helps!
 - Of aligned prepositions:
 - 80% have the same scene role
 - 55% have the same function



EXAMPLES

Ich habe viel **bei** Erwachsenen gelebt und ich habe sie ganz **aus** der Nähe
I have much near grown-ups.DAT lived and I have them entirely from the.DAT proximity.DAT

betrachtet .
observed .
I have lived a great deal **among** grown-ups . I have seen them intimately , close **at** hand .

Erzählt ihr den Erwachsenen [...]
If-tell you.PL the.DAT grown-ups.DAT
If you were to say **to** the grown-ups [...]

Zum Glück **für** den Asteroiden B 612 zwang ein türkischer Diktator **sein** Volk **unter**
To-the fortune for the.DAT Asteroid.DAT B 612 forced a Turkish dictator his people under
Androhung der Todesstrafe , sich **nach** europäischer Art zu kleiden .
threat the.GEN death-penalty.GEN , RFL after European manner to dress .
Fortunately , however , **for** the reputation of Asteroid B-612 , a Turkish dictator made a law that
his subjects , **under** pain of death , should change **to** European costume .

CONCLUSIONS

- Preposition supersenses, especially when applied with the **construal analysis**, can help exploit semantic similarities in German and English
- These would otherwise be overshadowed by **language-specific** construals and constructions
- To increase coverage for languages morphologically richer than English, **nonlexical case markers** should be added to the current framework

ONGOING & FUTURE WORK

- More annotation needed
 - Annotation guidelines for German are in development
 - We are looking for annotators -- please reach out if interested!
- Extend the preposition supersense framework to include morphological case markers
 - Design needs to be generalizable across many languages
 - Annotation tool (UCCAApp, Abend et al., 2017) and general workflow need to be adapted
 - How to deal with interaction between prepositions and case?
- Build a system that jointly learns to assign supersenses in a multilingual environment

REFERENCES

Omri Abend, Shai Yerushalmi, and Ari Rappoport. 2017. UCCAApp: Web-application for syntactic and semantic phrase-based annotation. In Proc. Of ACL 2017, System Demonstrations, pages 109–114, Vancouver, Canada.

Austin Blodgett and Nathan Schneider. 2018. Semantic supersenses for English possessives. In Proceedings of the Eleventh International Conference on Language Resources and Evaluation, Miyazaki, Japan. ELRA.

Martin Durrell and David Brée. 1993. German temporal prepositions from an english perspective. The Semantics of Prepositions: From Mental Processing to Natural Language Processing, 3:295.

Homa B Hashemi and Rebecca Hwa. 2014. A comparison of MT errors and ESL errors. In LREC, pages 2696–2700.

Jena D Hwang, Archana Bhatia, Na-Rae Han, Tim O’Gorman, Vivek Srikumar, and Nathan Schneider. 2017. Double trouble: The problem of construal in semantic annotation of adpositions. In Proceedings of the 6th Joint Conference on Lexical and Computational Semantics (* SEM 2017), pages 178–188.

Antje Müller, Olaf Hülscher, Claudia Roch, Katja Keßelmeier, Tobias Stadfeld, Jan Strunk, and Tibor Kiss. 2010. An annotation schema for preposition senses in German. In Proceedings of the Fourth Linguistic Annotation Workshop, pages 177–181. Association for Computational Linguistics.

Nathan Schneider, Jena D Hwang, Vivek Srikumar, Jakob Prange, Austin Blodgett, Sarah R Moeller, Aviram Stern, Adi Bitan, and Omri Abend. 2018. Comprehensive supersense disambiguation of English prepositions and possessives. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, Melbourne, Australia. ACL. To appear.

Andrea Tyler and Vyvyan Evans. 2003. The semantics of English prepositions: Spatial scenes, embodied meaning, and cognition. Cambridge University Press.

Henk Van Riemsdijk. 2007. Case in spatial adpositional phrases: The dative-accusative alternation in German. Pitar Mos: A building with a view. Festschrift for Alexandra Cornilescu, pages 1–23.

Joost Zwarts. 2005. The case of prepositions: Government and compositionality in German PPs. In 21st Annual Meeting of the Israel Association for Theoretical Linguistics, Haifa.

CONTACT

Jakob Prange, Department of Computer Science, Georgetown University, jp1724@georgetown.edu