

LAW-XVI, MARSEILLE, JUNE 24 2022

DESIDERATA FOR THE ANNOTATION OF INFORMATION STRUCTURE IN COMPLEX SENTENCES

Hannah Booth

MOTIVATION

- Past two decades: a surge of interest in **information structure (IS)** and a move towards annotating natural language for IS
- Various annotation schemes and guidelines have been developed (e.g. Buráňová et al. 2000; Calhoun et al. 2005; Paggio 2006; Götze et al. 2007; Bohnet et al. 2013; Riester et al. 2018)
- Some have acknowledged **complex sentences as an annotation challenge for IS** (Bohnet et al. 2013; Cook & Bildhauer 2013; Stede & Mamprin 2016)
- But existing schemes either gloss over the issue, or propose only crude guidelines

MOTIVATION

- Theoretical work has highlighted the **special status of dependent clauses wrt. IS** (e.g. Hooper & Thompson 1973; Bybee 2002; Bianchi & Frascarelli 2010; Lahousse 2010; Matić et al. 2014; Lahousse 2022)
- Neglect of this issue can result in **inaccurate or biased annotations**, or **unannotated data**
- This holds back **theoretical** research on the interface of information structure with e.g. morphosyntax and prosody
- And also inevitably impacts **computational** research and NLP downstream tasks which make use of IS-annotated data

THIS TALK

- Outlining **desiderata for the annotation of information structure in complex sentences** as a foundation for future approaches
 - annotation format
 - annotation process
- Informed by issues highlighted by **previous annotation schemes** (e.g. Cook & Bildhauer 2011, 2013; Bohnet et al. 2013)
- Underpinned by the latest **theoretical insights** in this area (e.g. Erteschik-Shir 2007; Bianchi & Frascarelli 2010; Lahousse 2010; Ebert et al. 2014; Matic' et al. 2014; Lahousse 2022)
- Test case: IS annotation of historical Middle Low German (c.1250-1650) texts
- Overall: showcasing the benefits which can result from a free exchange between annotation approaches and theoretical research

TERMINOLOGY

- I mainly discuss topic and focus
 - ABOUTNESS-TOPIC: entity/proposition about which information is provided (Reinhart 1981)
 - (1) Q: What did Mary give to Harry?
A: **Mary** gave a shirt to Harry.
 - FRAME-TOPIC: frame within which the main clause predication is interpreted (cf. Krifka 2007)
 - (2) Q: How is John?
A: **Last weekend** he was well.
 - INFORMATIONAL-FOCUS: new information which is most relevant to discourse (cf. Vallduví 1992)
 - (3) Q: What did Mary buy?
A: Mary bought **a cactus**.
 - CONTRASTIVE-FOCUS: element/proposition which evokes alternatives (cf. Neeleman et al. 2009)
 - (4) Q: What do you want to drink, tea or coffee?
A: I want **tea**.

THEORETICAL INSIGHTS



THE CHALLENGE OF INFORMATION STRUCTURE

- The IS properties of complex sentences is an **understudied domain** (Matić et al. 2014)
- And from the literature there is, it can be difficult to establish a general consensus on even essential questions
- This presents a challenge for annotation, where the ideal is schemes which are as theoretically neutral as possible and compatible with diverse approaches
- What is there consensus on?
 - What **crosslinguistic generalisations** can be established?
 - Can the IS status of dependent clauses **be predicted from other linguistic properties** (e.g. clause class, clause ordering)?

THE DOMAIN(S) OF INFORMATION STRUCTURE

- **Traditional view:** domain of IS = utterance (Mathesius 1975; Vallduví 1992; Steedman 2000; Komagata 2003)
 - ⇒ a complex sentence has IS articulation at the matrix level only
- **More recent consensus:** IS can operate within a single utterance at different levels (Koktová 1996; Partee 1996; Hajicová et al. 1998; Erteschik-Shir 2007; Matic' et al. 2014)
 - ⇒ **a dependent clause is a (potential) IS domain**

THE DOMAIN(S) OF INFORMATION STRUCTURE

- For dependent clauses, a distinction between **external and internal IS**

- external IS: status of a dependent clause in the overall utterance

(5) It was only [after I arrived home] that I saw them.
FOCUS

- internal IS: status of individual constituents within a dependent clause

(6) I believe [that this book Mary gave to Paul].
TOPIC

- This can yield recursion (Partee 1996)

(7) **[What convinced Susan that [our arrest]_{TOPIC} was caused by Harry]_{TOPIC}** was a rumour that someone had witnessed Harry's confession.

(8) What convinced Susan that our arrest was caused by Harry was **[a rumour that someone had [witnessed Harry's confession.]_{FOCUS}]_{FOCUS}**

DESIDERATA – ANNOTATION FORMAT

DESIDERATA – ANNOTATION FORMAT

- **Six key aspects** which any format should be able to handle in order to achieve a theoretically sound and practically sensible IS annotation across complex sentences
 - 1 Multiplicity
 - 2 Recursion
 - 3 Discontinuity
 - 4 Supra-clausality
 - 5 Uncertainty
 - 6 Meta-annotation
- Some of these result from the overall nature of IS, others from complex sentences specifically
- All add further weight to the importance of **multidimensional, standoff formats**, as already employed for IS (Stede & Mamprin 2016; Celano 2019)

MULTIPLICITY

- Any annotation scheme needs to be able to handle **multiple, potentially cross-cutting IS articulations within a single clause/sentence**
- Topic/comment and focus/background typically crosscut (Dahl 1974; Vallduví 1992)

(9) Q: What does John drink?

A₁: John drinks beer
 TOPIC COMMENT

A₂: John drinks beer
 BACKGROUND FOCUS

RECURSION

- Many previous annotation schemes encode IS via flat spans, but **recursive IS as presented by complex sentences requires some level of hierarchisation**

- (10) **[What convinced Susan that [our arrest]_{TOPIC} was caused by Harry]_{TOPIC}** was a rumour that someone had witnessed Harry's confession.
- (11) What convinced Susan that our arrest was caused by Harry was **[a rumour that someone had [witnessed Harry's confession.]_{FOCUS}]_{FOCUS}**

DISCONTINUITY

- Many languages exhibit **discontinuous IS fields**, i.e. when a single IS status is assigned to multiple non-adjacent segments

- **Discontinuous focus** (German, Gussenhoven 1999)

(12) *What happened to the child?*

Karl hat dem Kind **einen Füller** geschenkt

Karl has the child a fountain-pen given

'Karl gave the child a fountain pen'

- **Discontinuous topic** (Serbian, Milićev & Milićević 2012)

(13) **Marija** sutra, **profesorica latinskog**, odlazi u penziju.

Mary tomorrow professor of-Latin goes to retirement

'Mary, professor of Latin, retires tomorrow.'

- Discontinuity has been treated in syntactic annotation (e.g. Boyd 2007; Maier & Lichte 2011) but the issue has been **less addressed wrt. IS**

SUPRA-CLAUSALITY

- The annotation of complex sentences in particular presents the need to encode **supra-clausal IS fields**, i.e. IS fields which span across clause boundaries

(14) [What convinced Susan that [our arrest]_{TOPIC} was caused by Harry]_{TOPIC} was a rumour that someone had witnessed Harry's confession.

- Where IS annotation is combined with syntactic annotation, the format must allow for IS annotations to **cross-cut syntactic clause boundaries**
- IS annotation cannot simply be parasitic on syntactic annotation
⇒ it must have **sufficient autonomy**

UNCERTAINTY

- Any IS annotation scheme should also be able to encode some level of **uncertainty** in contexts where:
 - a clear-cut identification of IS domains is not possible
 - classification of IS articulations cannot be made
- This is particularly critical for IS annotation across complex sentences, where theoretical knowledge is still underdeveloped
- Uncertainty annotation has growing interest (e.g. Barteld et al. 2014; Merten & Seemann 2018; Andresen et al. 2020; Beck et al. 2020) but **has not been fully exploited for IS**

META-ANNOTATION

- IS annotation schemes should also allow for some form of **meta-annotation**, i.e. information about a given IS annotation, which explains the choices made
- This is particularly relevant in the context of IS, which lacks consensus on key definitions
- Judgements are often less clear-cut and more subjective than at other linguistic levels, even with a carefully operationalised set of diagnostics
- Meta-annotation can promote the **usability** of the resource for theoretical studies, making the decision behind the annotation **more transparent** (as used already in some resources Laprun et al. 2002; Romary et al. 2010)

DESIDERATA – ANNOTATION PROCESS

DESIDERATA – ANNOTATION PROCESS

- Manual IS annotation based on pragmatic context-based judgements alone is a relatively **subjective and time-intensive process**, especially for complex sentences
- To explore: to what extent can the IS of dependent clauses be **predicted from pre-existing annotations for morphosyntactic features**?

DESIDERATA – ANNOTATION PROCESS

- Each dependent clause is annotated in **two separate stages**:
 - 1 **external IS** (IS status in overall utterance)
 - 2 **internal IS** (IS of individual constituents within the clause)
- The two tasks are largely independent of each other and informed by different considerations
 - Identification of external IS for a dependent clause does not necessarily imply it has internal IS
- The approaches to the annotation process were tested in a **pilot IS annotation of dependent clauses** in a Middle Low German text from the Corpus of Historical Low German (CHLG, Booth et al. 2020)
 - *Chronica Nova*, Dietric Engelhus (historical chronicle, 1435 CE)
 - 709 clauses annotated as a dependent clauses (Penn treebank-style)
- Pilot annotation conducted using Annotald tool (Beck et al. 2015)

GENERALISATIONS: EXTERNAL IS

- **Crosslinguistic syntax-IS correspondences which can be identified for external IS**
 - *D* occurs before host clause \approx **TOPIC**
(Marchese 1977; Lehmann 1984; Thompson 1985; Chafe 1984; Lehmann 1988; Diessel 2001)
 - *D* is conditional clause \approx **TOPIC** (Haiman 1978; Schiffrin 1992; Ebert et al. 2014)
 - *D* is clefted \approx **FOCUS** (Matić et al. 2014)
 - *D* is nonfactive complement (under e.g. 'say', 'think') \approx **FOCUS** (Dehé & Wichmann 2010)
 - *D* is factive complement (under e.g. 'regret', 'discover') \approx **BACKGROUND**, i.e. neither topic nor focus (Matić et al. 2014)
 - *D* is restrictive relative clause \approx **FOCUS** (Schachter 1973)
 - *D* is non-restrictive relative clause \approx **BACKGROUND** (Umbach 2006; Song 2014)

Figure: Hand-crafted rule-based algorithm for assigning external IS to dependent clauses

```
case D is before host clause
      external IS := TOPIC
case D is not before host clause
      if D is conditional clause then
        external IS := TOPIC
      elif D is clefted then
        external IS := FOCUS
      elif D is complement clause then
        if D is nonfactive then
          external IS := FOCUS
        else
          external IS := BACKGROUND
      elif D is relative clause then
        if D is RRC then
          external IS := FOCUS
        else
          external IS := BACKGROUND
      else
        external IS := BACKGROUND
```

TEST CASE – EXTERNAL IS

- **Reference set:** all dependent clauses were manually annotated for external IS on the basis of contextual pragmatic judgements alone (irrespective of other linguistic features)
- **Test set:** a fresh round of manual annotation was then performed relying exclusively on the rule-based algorithm and the output compared against the first round
- **Overall accuracy:** 81.6%

	P	R	F
TOPIC	.849	.753	.798
FOCUS	.860	.636	.731
BACKGROUND	.704	.884	.783

Table: Per tag performance of hand-crafted rule-based algorithm for annotation of external IS

- **The exploitation of pre-annotated morphosyntactic/lexical features can play a useful role in informing the annotation of dependent clauses for external IS**

GENERALISATIONS: INTERNAL IS

- **Crosslinguistic syntax-IS correspondences which can be identified for internal IS**
(concerns whether a clause is asserted or presupposed)
 - *D* occurs before host clause \implies **no internal IS** (Komagata 2003; Lelandais & Ferré 2017)
 - *D* is central (i.e. event-structuring) adverbial clause \implies **no internal IS** (De Cat 2012; Lahousse & Borremans 2014)
 - *D* is peripheral (i.e. discourse-structuring) adverbial clause \implies **internal IS**
(De Cat 2012; Lahousse & Borremans 2014)
 - *D* is factive complement \implies **no internal IS** Hooper & Thompson (1973); Maticić et al. (2014)
 - *D* is nonfactive complement \implies **internal IS** (Hooper & Thompson 1973; Maticić et al. 2014)
 - *D* is restrictive relative clause \implies **no internal IS** (Depraetere 1996; Maticić et al. 2014)
 - *D* is non-restrictive relative clause \implies **internal IS** (Depraetere 1996; Bybee 2002)

Figure: Hand-crafted rule-based algorithm for deciding whether to assign internal IS to dependent clauses

```
case D is before host clause
      status := no internal IS
case D is not before host clause
      if D is adverbial clause then
        if D is central adverbial clause then
          status := no internal IS
        else
          status := internal IS
      elif D is complement clause then
        if D is factive then
          status := no internal IS
        else
          status := internal IS
      elif D is relative clause
        if D is RRC then
          status := no internal IS
        else
          status := internal IS
      else
        status := unknown
```

TEST CASE – INTERNAL IS

- **Reference set:** each dependent clause was manually annotated on the basis of pragmatic judgements alone for the presence/absence of internal IS
- **Test set:** a fresh round of (manual) annotation was then performed using the rule-based algorithm for presence/absence of internal IS, without paying attention to pragmatic context
- **Overall accuracy:** 88.3%
- **The exploitation of pre-annotated morphosyntactic/lexical features can also play a useful role in determining whether a dependent clause constitutes an independent IS domain, with internal IS**

CONCLUSION

CONCLUSION

- State-of-the-art theoretical research on the IS properties of complex sentences can help inform **more accurate, nuanced annotation schemes, where complex sentences have proved a particular challenge**
- **Annotation format:** specific demands of IS and complex sentences adds further weight to the importance of **multidimensional, standoff formats**
- **Annotation process:** external and internal IS merit separate tasks, which both exploit **pre-annotated non-IS features**

ACKNOWLEDGEMENTS

Many thanks to:

- the Flemish Research Foundation (FWO) for generous funding (Postdoctoral Junior Fellowship, 2021–2024, grant number 12ZL522N)



REFERENCES I

- Andresen, Melanie, Michael Vauth & Heike Zinsmeister. 2020. Modeling ambiguity with many annotators and self-assessments of annotator certainty. In *Proceedings of the 14th linguistic annotation workshop*, 48–59. Barcelona, Spain: Association for Computational Linguistics. <https://aclanthology.org/2020.law-1.5>.
- Barteld, Fabian, Sarah Ilden, Ingrid Schröder & Heike Zinsmeister. 2014. Annotating descriptively incomplete language phenomena. In *Proceedings of LAW VIII - the 8th linguistic annotation workshop*, 99–104. Dublin, Ireland: Association for Computational Linguistics and Dublin City University. doi:10.3115/v1/W14-4915. <https://www.aclweb.org/anthology/W14-4915>.
- Beck, Christin, Hannah Booth, Mennatallah El-Assady & Miriam Butt. 2020. Representation problems in linguistic annotations: Ambiguity, variation, uncertainty, error and bias. In *Proceedings of the 14th linguistic annotation workshop*, 60–73. Barcelona, Spain: Association for Computational Linguistics. <https://aclanthology.org/2020.law-1.6>.
- Beck, Jana, Aaron Ecay & Anton Karl Ingason. 2015. Annotald. version 1.3. 7.
- Bianchi, Valentina & Mara Frascarelli. 2010. Is topic a root phenomenon? *Iberia: An International Journal of Theoretical Linguistics* 2(1). 43–88.
- Bohnet, Bernd, Alicia Burga & Leo Wanner. 2013. Towards the annotation of Penn TreeBank with information structure. In Ruslan Mitkov & Jong C. Park (eds.), *Proceedings of the Sixth International Joint Conference on Natural Language Processing*, 1250–1256. Nagoya: Asian Federation of Natural Language Processing.
- Booth, Hannah, Anne Breitbarth, Aaron Ecay & Melissa Farasyn. 2020. A Penn-style treebank of Middle Low German. In *Proceedings of the 12th Language Resources and Evaluation Conference*, 766–775. Marseille, France: European Language Resources Association. <https://www.aclweb.org/anthology/2020.lrec-1.96>.

REFERENCES II

- Boyd, Adriane. 2007. Discontinuity revisited: An improved conversion to context-free representations. In *Proceedings of the linguistic annotation workshop*, 41–44. Prague, Czech Republic: Association for Computational Linguistics.
<https://aclanthology.org/W07-1506>.
- Buráňová, Eva, Eva Hajičová & Petr Sgall. 2000. Tagging of very large corpora: Topic-Focus articulation. In *Proceedings of the 18th conference on computational linguistics (coling)*, 139–144. Saarbrücken, Germany, Universität des Saarlandes: Association for Computational Linguistics. doi:10.3115/990820.990841. <https://doi.org/10.3115/990820.990841>.
- Bybee, Joan. 2002. Main clauses are innovative, subordinate clauses are conservative: consequences for the nature of constructions. In Joan Bybee & Michael Noonan (eds.), *Complex sentences in grammar and discourse: Essays in honor of sandra a. thompson*, 1–17. Amsterdam: John Benjamins.
- Calhoun, Sasha, Malvina Nissim, Mark Steedman & Jason Brenier. 2005. A framework for annotating information structure in discourse. In Adam Meyers (ed.), *Frontiers in corpus annotation ii: Pie in the sky, acl2005 conference workshop*, 45–52. Ann Arbor, Michigan, June 2005.
- Celano, Giuseppe G. A. 2019. Standoff annotation for the Ancient Greek and Latin Dependency Treebank. In *Proceedings of the 3rd international conference on digital access to textual cultural heritage, brussels, belgium*, 149–153. New York: Association for Computing Machinery.
- Chafe, Wallace. 1984. How people use adverbial clauses. In *Proceedings of the tenth annual meeting of the berkeley linguistics society*, 437–449. Berkeley, CA: Berkeley Linguistics Society.

REFERENCES III

- Cook, Philippa & Felix Bildhauer. 2011. Annotating information structure: the case of topic. In Stefanie Dipper & Heike Zinsmeister (eds.), *Beyond semantics: Corpus-based investigations of pragmatic and discourse phenomena* Bochumer Linguistische Arbeitsberichte 3, 45–56. Bochum: Ruhr-Universität Bochum, Sprachwissenschaftliches Institut.
- Cook, Philippa & Felix Bildhauer. 2013. Identifying “aboutness topics”: two annotation experiments. *Dialogue & Discourse* 4(2). 118–141.
- Dahl, Östen. 1974. *Topic, comment, contextual boundedness and focus*. Hamburg: Buske.
- De Cat, Cécile. 2012. Towards an interface definition of root phenomena. In Lobke Aelbrecht, Liliane Haegeman & Rachel Nye (eds.), *Main clause phenomena: New horizons*, 135–158. Amsterdam: John Benjamins.
- Dehé, Nicole & Anne Wichmann. 2010. Sentence-initial *I think (that)* and *I believe (that)*: prosodic evidence for uses as main clause, comment clause and discourse marker. *Studies in Language* 34(1). 36–74.
- Depraetere, Ilse. 1996. Foregrounding in English relative clauses. *Linguistics* 34. 699–731.
- Diessel, Holger. 2001. The ordering distribution of main and adverbial clauses: A typological study. *Language* 77. 433–455.
- Ebert, Christian, Cornelia Ebert & Stefan Hinterwimmer. 2014. A unified analysis of conditionals as topics. *Linguistics and Philosophy* 37(5). 353–408.
- Erteschik-Shir, Nomi. 2007. *Information structure: the syntax-discourse interface*. Oxford: Oxford University Press.

REFERENCES IV

- Götze, Michael, Thomas Weskott, Cornelia Endriss, Ines Fiedler, Stefan Hinterwimmer, Svetlana Petrova, Anne Schwarz, Stavros Skopeteas & Ruben Stoel. 2007. Information structure. In Steffi Dipper, Michael Götze & Stavros Skopeteas (eds.), *Information structure in cross-linguistic corpora: Annotation guidelines for phonology, morphology, syntax, semantics, and information structure*, 147–187. Potsdam: Universitätsverlag Potsdam.
- Gussenhoven, Carlos. 1999. On the limits of focus projection in English. In Peter Bosch & Rob van der Sandt (eds.), *Focus: linguistic, cognitive, and computational perspectives*, 43–55. Cambridge: Cambridge University Press.
- Haiman, John. 1978. Conditionals are topics. *Language* 54(3). 564–589.
- Hajicová, Eva, Barbara B. H. Partee & Petr Sgall. 1998. *Topic-focus articulation, tripartite structures, and semantic content*. Dordrecht: Kluwer.
- Hooper, Joan B & Sandra A Thompson. 1973. On the applicability of root transformations. *Linguistic Inquiry* 4(4). 465–497.
- Koktová, Eva. 1996. Wh-extraction and the topic-focus articulation of the sentence. In Barbara H. Partee & Petr Sgall (eds.), *Discourse and meaning: Papers in honor of eva hajičová*, 255–271. Amsterdam: John Benjamins.
- Komagata, Nobo. 2003. Information structure in subordinate and subordinate-like clauses. *Journal of Logic, Language and Information* 12(3). 301–318.
- Krifka, Manfred. 2007. Basic notions of information structure. In Caroline Féry & Manfred Krifka (eds.), *Interdisciplinary Studies on Information Structure*, 13–56. Potsdam: Universitätsverlag.
- Lahousse, Karen. 2010. Information structure and epistemic modality in adverbial clauses in French. *Studies in Language* 34(2). 298–326.

REFERENCES V

- Lahousse, Karen. 2022. Is focus a root phenomenon? In Davide Garassino & Daniel Jacob (eds.), *When data challenges theory: unexpected and paradoxical evidence in information structure*, 148–182. Amsterdam: John Benjamins.
- Lahousse, Karen & Marijke Borremans. 2014. The distribution of functional-pragmatic types of clefts in adverbial clauses. *Linguistics* 52(3). 793–836. doi:doi:10.1515/ling-2014-0009. <https://doi.org/10.1515/ling-2014-0009>.
- Laprun, Christophe, Jonathan Fiscus, John Garofolo & Sylvain Pajot. 2002. Recent improvements to the ATLAS architecture. In *Proceedings of the second international conference on human language technology (hlt'02)*, 263–268.
- Lehmann, Christian. 1984. *Der relativsatz: Typologie seiner strukturen, theorie seiner funktionen, kompendium seiner grammatik*. Amsterdam: John Benjamins.
- Lehmann, Christian. 1988. Towards a typology of clause linkage. In John Haiman & Sandra A. Thompson (eds.), *Clause combining in grammar and discourse*, 181–225. Amsterdam: John Benjamins.
- Lelandais, Manon & Gaëlle Ferré. 2017. How are three syntactic types of subordinate clauses different in terms of informational weight? *Anglophonia* 23. <http://journals.openedition.org/anglophonia/1200>.
- Maier, Wolfgang & Timm Lichte. 2011. Characterizing discontinuity in constituent treebanks. In Philippe de Groote, Markus Egg & Laura Kallmeyer (eds.), *Formal grammar*, 167–182. Berlin, Heidelberg: Springer.
- Marchese, Lynell. 1977. Subordinate clauses as topics in Godie. In Martin Mould & Thomas J. Hinnebusch (eds.), *Papers from the 8th conference on african linguistics*, 157–164. Los Angeles, CA: Department of Linguistics, University of California.
- Mathesius, Vilém. 1975. *A functional analysis of present day English on a general linguistic basis*. The Hague: Mouton.

REFERENCES VI

- Matić, Dejan, Rik Van Gijn & Robert D. Van Valin Jr. 2014. Information structure and reference tracking in complex sentences. In Rik van Gijn, Jeremy Hammond, Dejan Matić, Saskia van Putten & Ana Vilacy Galucio (eds.), *Information structure and reference tracking in complex sentences*, 1–42. Amsterdam: John Benjamins.
- Merten, Marie-Luis & Nina Seemann. 2018. Analyzing constructional change: Linguistic annotation and sources of uncertainty. In *Proceedings of the sixth international conference on technological ecosystems for enhancing multiculturalism TEEM'18*, 819–825. New York, NY, USA: Association for Computing Machinery. doi:10.1145/3284179.3284320. <https://doi.org/10.1145/3284179.3284320>.
- Milićev, Tanja & Nataša Milićević. 2012. Leftward movement with discontinuous appositive constructions. *Acta Linguistica Hungarica* 59(1-2). 205–220.
- Neeleman, Ad, Elena Titov, Hans Van de Koot & Reiko Vermeulen. 2009. A syntactic typology of topic, focus and contrast. In Jeroen van Craenenbroeck (ed.), *Alternatives to cartography*, 15–52. Berlin: de Gruyter.
- Paggio, Patrizia. 2006. Annotating information structure in a corpus of spoken Danish. In *Proceedings of the 5th international conference on language resources and evaluation (Irec2006)*, 1606–1609. Genova, Italy.
- Partee, Barbara H. 1996. Allegation and local accommodation. In Barbara H. Partee & Petr Sgall (eds.), *Discourse and meaning: papers in honor of eva hajicová*, 65–86. Amsterdam: John Benjamins.
- Reinhart, Tanya. 1981. Pragmatics and linguistics: An analysis of sentence topics. *Philosophica* 27(1). 53–94.

REFERENCES VII

- Riester, Arndt, Lisa Brunetti & Kordula De Kuthy. 2018. Annotation guidelines for Questions under Discussion and information structure. In Evangelia Adamou, Katharina Haude & Martine Vanhove (eds.), *Information structure in lesser-described languages: studies in prosody and syntax*, 403–443. Amsterdam: John Benjamins.
- Romary, Laurent, Amir Zeldes & Florian Zipser. 2010. [Tiger2/] documentation [Technical Report]. Inria-00593903v2.
- Schachter, Paul. 1973. Focus and relativization. *Language* 41(1). 19–46.
- Schiffrin, Deborah. 1992. Conditionals as topics in discourse. *Linguistics* 30. 165–197.
- Song, Sanghoun. 2014. Information structure of relative clauses in English: a flexible and computationally tractable model. *Language and Information* 18(2). 1–29.
- Stede, Manfred & Sara Mamprin. 2016. Information structure in the Potsdam commentary corpus: topics. In *Proceedings of the tenth international conference on language resources and evaluation (LREC'16)*, 1718–1723. Portorož, Slovenia: European Language Resources Association (ELRA). <https://aclanthology.org/L16-1271>.
- Steedman, Mark. 2000. Information structure and the syntax-phonology interface. *Linguistic Inquiry* 31(4). 649–689.
- Thompson, Sandra A. 1985. Grammar and written discourse: Initial vs. final purpose clauses in English. *Text-Interdisciplinary Journal for the Study of Discourse* 5(1-2). 55–84.
- Umbach, Carla. 2006. Non-restrictive modification and backgrounding. In *Proceedings of the ninth symposium on logic and language*, 152–159. Budapest: Hungarian Academy of Sciences.
- Vallduví, Enric. 1992. *The informational component*. New York: Garland Press.