Seminar Computational Sociolinguistics (CSL) — Part 2

Overview of Seminar Topics and Literature

Henning Wachsmuth

henningw@upb.de

PADERBORN UNIVERSITY

Assignment of seminar topics

- **This talk**
  - Overview of all 20 possible seminar topics, 4 each by the 5 advisors
  - For each topic, two articles are given that provide the basis of the topic
- **Concept behind**
  - Each seminar participant will be assigned one topic
  - The two articles (+ *at least* one further relevant article to be found by you) should be discussed in the talk and in the article
  - You can choose topic preferences, we then assign topics
- **Your task**
  - Inform yourself about the topics and articles in this presentation
  - Choose 3 topics with preferences
  - Until Sunday, April 25, 23:59 GTM+2. Send e-mail with preferences

Choosing preferences: eMail and subsequent process

- **Your eMail**
  - Recipient. henningw@upb.de
  - Subject. ”[csl] Topic preferences“
  - Content. Your name and matriculation number + 3 topic preferences
  - Example. On the right, you see how the content of your eMail could look like

Name:
Timon Gurcke

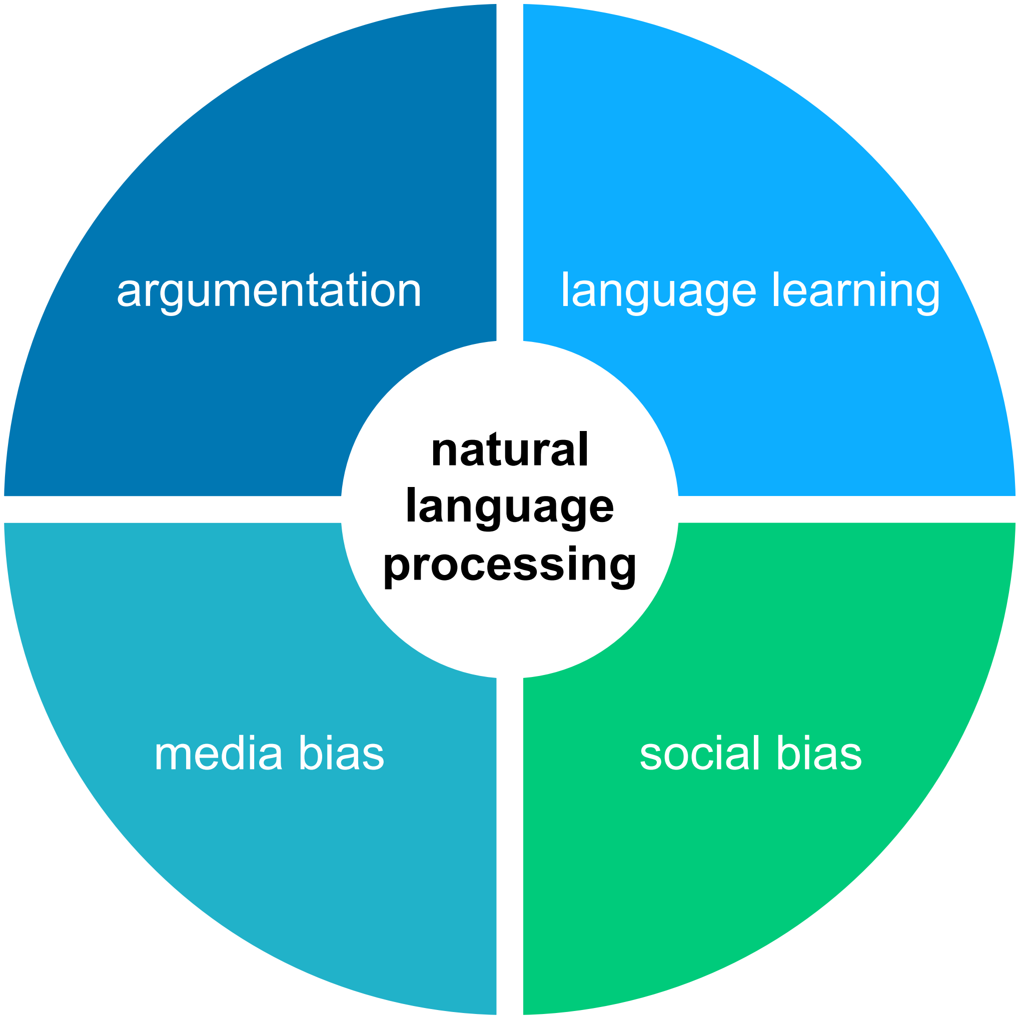
Matriculation number:
1234567

Topic preferences:
1) e. Modeling Learner Argumentation
2) b. Portraying Social Groups
3) j. Negating Claims

- **Subsequent process**
  - We will assign topics based on preferences, special reasons, and randomly
  - If you don‘t send your e-mail in time, we will assign any free topic to you
  - The final schedule will be decided based on the topic assignment
    You can get a rough idea of the schedule from the ordering on the next slides.
  - Topic assignment and schedule will be announced until the next session

Overview of topics

Recap: CSL-related areas of the CSS group



Topics for the seminar talks and articles

- **Understanding Sociolinguistic Variables**

a.	Aligning AI with Human Values	Max
b.	Portraying Social Groups	Max
c.	Analyzing Cultural Differences	Henning
d.	Understanding Idiomatic Expressions	Mei-Hua

- **Analyzing and Improving Argumentation**

e.	Modeling Learner Argumentation	Mei-Hua
f.	Assessing Learner Argumentation	Mei-Hua
g.	Supporting Learner Argumentation	Mei-Hua
h.	Modeling Audience of Argumentation	Henning
i.	Detecting Fallacious Argumentation	Henning
j.	Negating Claims	Milad
k.	Controlling Claim Generation	Milad

More on next slide…

Topics for the seminar talks and articles

- **Analyzing and Adjusting Media Bias**

  l. Understanding News Quality — Wei-Fan
  m. Analyzing Misinformation — Wei-Fan
  n. Understanding Moral Foundations — Milad
  o. Mining Moral Foundations — Milad
  p. Analyzing Media Bias — Wei-Fan
  q. Adjusting Media Bias — Wei-Fan

- **Assessing Social Bias and Impact**

  r. Analyzing Social Bias in Representation — Henning
  s. Detecting Social Bias in Generation — Max
  t. Understanding Social Impact — Max

Topics and Literature

Topics supervised by Henning 1

c. Analyzing Cultural Differences

- Given literature:

  Hovy and Purschke (2018). Dirk Hovy and Christoph Purschke. Capturing Regional Variation with Distributed Place Representations and Geographic Retrofitting. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, pages 4383–4394, 2018.

  Tian et al. (2020). Yufei Tian, Tuhin Chakrabarty, Fred Morstatter, and Nanyun Peng. Identifying Cultural Differences through Multi-Lingual Wikipedia. arXiv:2004.04938, 2020.

h. Detecting Fallacious Argumentation

- Given literature:

  Habernal et al. (2018). Ivan Habernal, Henning Wachsmuth, Iryna Gurevych, and Benno Stein. Before Name-calling: Dynamics and Triggers of Ad Hominem Fallacies in Web Argumentation. In Proceedings of the 16th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pages 386–396, 2018.

  Delobelle et al. (2019). Pieter Delobelle, Murilo Cunha, Eric Massip Cano, Jeroen Peperkamp, Bettina Berendt. Computational Ad Hominem Detection. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop, pages 203–209, 2019.

Topics supervised by Henning 2

i. Modeling Audience of Argumentation

- Given literature:

El Baff et al. (2018). Roxanne El Baff, Henning Wachsmuth, Khalid Al-Khatib, and Benno Stein. Challenge or Empower: Revisiting Argumentation Quality in a News Editorial Corpus. In Proceedings of the 22nd Conference on Computational Natural Language Learning, pages 454–464, 2018.

Lukin et al. (2017). Stephanie Lukin, Pranav Anand, Marilyn Walker and Steve Whittaker. Argument Strength is in the Eye of the Beholder: Audience Effects in Persuasion. In Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers, pages 741–752, 2017.

r. Analyzing Social Bias in Representation

- Given literature:

Lauscher et al. (2020). Anne Lauscher, Rafik Takieddin, Simone Paolo Ponzetto, and Goran Glavaš. AraWEAT: Multidimensional Analysis of Biases in Arabic Word Embeddings. In Proceedings of the Fifth Arabic Natural Language Processing Workshop, pages 192–199, 2020.

Spliethöver and Wachsmuth (2020). Maximilian Spliethöver and Henning Wachsmuth. Argument from Old Man’s View: Assessing Social Bias in Argumentation. In Proceedings of the 7th Workshop on Argument Mining, pages 76–87, 2020.

Topics supervised by Milad 1

j. Negating Claims

- Given literature:

Hidey and McKeown (2019). Christopher Hidey and Kathy McKeown. Fixed That for You: Generating Contrastive Claims with Semantic Edits. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), pages 1756–1767, 2019.

Bilu et al. (2015). Yonatan Bilu, Daniel Hershcovich, and Noam Slonim. Automatic Claim Negation: Why, How and When. In Proceedings of the 2nd Workshop on Argumentation Mining, pages 84–93, 2015.

k. Controlling Claim Generation

- Given literature:

Schiller et al. (2020). Benjamin Schiller, Johannes Daxenberger, and Iryna Gurevych. Aspect-Controlled Neural Argument Generation. arXiv:2005.00084, 2020.

Alshomary et al. (2021). Milad Alshomary, Wei-Fan Chen, Timon Gurcke, Henning Wachsmuth. Belief-based Generation of Argumentative Claims. In Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume, pages 224–233, 2021.

Topics supervised by Milad 2

n. Understanding Moral Foundations

- Given literature:

Fulgoni et al. (2016). Dean Fulgoni, Jordan Carpenter, Lyle Ungar, and Daniel Preoţiuc-Pietro. An Empirical Exploration of Moral Foundations Theory in Partisan News Sources. In Proceedings of the Tenth International Conference on Language Resources and Evaluation, pages 3730–3736, 2016.

Graham et al. (2009). Jesse Graham, Jonathan Haidt, and Brian A Nosek. Liberals and Conservatives Rely on Different Sets of Moral Foundations. Journal of Personality and Social Psychology 96(5):1029-46, 2009.

o. Mining Moral Foundations

- Given literature:

Kaur and Sahsahara (2016). Rishemjit Kaur and Kazutoshi Sasahara. Quantifying Moral Foundations from Various Topics on Twitter Conversations. In Proceedings of the 2016 IEEE International Conference on Big Data, pages 2505-2512, 2016.

Araque et al. (2020). Oscar Araque, Lorenzo Gatti, and Kyriaki Kalimeri. Exploiting a Moral Lexicon and Embedding Similarity for Moral Foundations Prediction. Knowledge-Based Systems 191, 2020.

Topics supervised by Wei-Fan 1

l. Understanding News Quality

- Given literature:

Louis and Nenkova (2014). Annie Louis and Ani Nenkova. Verbose, Laconic or Just Right: A Simple Computational Model of Content Appropriateness under Length Constraints. In Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics, pages 636–644, 2014.

Arapakis et al. (2016). Ioannis Arapakis, Filipa Peleja, Barla Berkant, and Joao Magalhaes. Linguistic Benchmarks of Online News Article Quality. In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 1893–1902, 2016.

m. Analyzing Misinformation

- Given literature:

Kumar et al. (2016). Srijan Kumar, Robert West, and Jure Leskovec. Disinformation on the Web: Impact, Characteristics, and Detection of Wikipedia Hoaxes. In Proceedings of the 25th International Conference on World Wide Web, pages 591–602, 2016.

Volkova and Jang (2018). Svitlana Volkova and Jin Yea Jang. Misleading or Falsification? Inferring Deceptive Strategies and Types in Online News and Social Media. In Companion Proceedings of the The Web Conference 2018, pages 575–583, 2018.

Topics supervised by Wei-Fan 2

p. Analyzing Media Bias

- Given literature:

  Baly et al. (2019). Ramy Baly, Georgi Karadzhov, Abdelrhman Saleh, James Glass, Preslav Nakov. Multi-Task Ordinal Regression for Jointly Predicting the Trustworthiness and the Leading Political Ideology of News Media. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), pages 2109–2116, 2019.

  Chen et al. (2020). Wei-Fan Chen, Khalid Al Khatib, Henning Wachsmuth, and Benno Stein. Analyzing Political Bias and Unfairness in News Articles at Different Levels of Granularity. In Proceedings of the Fourth Workshop on Natural Language Processing and Computational Social Science, pages 149–154, 2020.

q. Adjusting Media Bias

- Given literature:

  Chen et al. (2018). Wei-Fan Chen, Henning Wachsmuth, Khalid Al-Khatib, and Benno Stein. Learning to Flip the Bias of News Headlines. In Proceedings of The 11th International Natural Language Generation Conference, pages 79–88, 2018.

  Madaan et al. (2021). Nishtha Madaan, Inkit Padhi, Naveen Panwar, and Diptikalyan Saha. Generate Your Counterfactuals: Towards Controlled Counterfactual Generation for Text. arXiv:2012.04698, 2021.

Topics supervised by Max 1

a. Aligning AI with Human Values

- Given literature:

Gabriel (2020). Iason Gabriel. Artificial Intelligence, Values, and Alignment. Minds and Machines 30, pages 411–437, 2020.

Irving and Askell (2019). Geoffrey Irving and Amanda Askell. AI Safety Needs Social Scientists. 10.23915/distill.00014, 2019.

b. Portraying Social Groups

- Given literature:

Park et al. (2021). Chan Young Park, Xinru Yan, Anjalie Field, Yulia Tsvetkov. Multilingual Contextual Affective Analysis of LGBT People Portrayals in Wikipedia. arXiv:2010.10820, 2021.

Sap et al. (2020). Maarten Sap, Saadia Gabriel, Lianhui Qin, Dan Jurafsky, Noah A. Smith, and Yejin Choi. Social Bias Frames: Reasoning about Social and Power Implications of Language. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pages 5477–5490, 2020.

Topics supervised by Max 2

s. Detecting Social Bias in Generation

- Given literature:

Sheng et al. (2019). Emily Sheng, Kai-Wei Chang, Premkumar Natarajan, Nanyun Peng. The Woman Worked as a Babysitter: On Biases in Language Generation. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing, pages 3407–3412, 2019.

Dhamala et al. (2021). Jwala Dhamala, Tony Sun, Varun Kumar, Satyapriya Krishna, Yada Pruksachatkun, Kai-Wei Chang, and Rahul Gupta. BOLD: Dataset and Metrics for Measuring Biases in Open-Ended Language Generation. IN Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, pages 862–872, 2021.

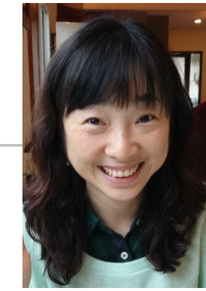
t. Understanding Social Impact

- Given literature:

Hovy and Spruit (2016). Dirk Hovy and Shannon L. Spruit. The Social Impact of Natural Language Processing. In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), pages 591–598, 2016.

Bender et al. (2021). Emily M. Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, pages 610–623, 2021.

Topics supervised by Mei-Hua 1



d. Understanding Idiomatic Expressions

- Given literature:

  Li et al. (2020). Jen-Yu Li and Thomas Gaillat. Automatic Detection of Unexpected/Erroneous Collocations in Learner Corpus. In Proceedings of the Joint Workshop on Multiword Expressions and Electronic Lexicons, pages 101–106, 2020.

  Kurfalı et al. (2020). Murathan Kurfalı and Robert Östling. Disambiguation of Potentially Idiomatic Expressions with Contextual Embeddings. In Proceedings of the Joint Workshop on Multiword Expressions and Electronic Lexicons, pages 85–94, 2020.

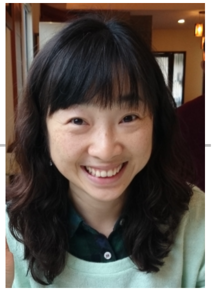
e. Modeling Learner Argumentation

- Given literature:

  Beigman Klebanov et al. (2016). Beata Beigman Klebanov, Christian Stab, Jill Burstein, Yi Song, Binod Gyawali, and Iryna Gurevych. Argumentation: Content, Structure, and Relationship with Essay Quality. In Proceedings of the Third Workshop on Argument Mining, pages 70–75, 2016.

  Wambsganss et al. (2020). Thiemo Wambsganss, Christina Niklaus, Matthias Söllner, Siegfried Handschuh, and Jan Marco Leimeister. A Corpus for Argumentative Writing Support in German. In Proceedings of the 28th International Conference on Computational Linguistics, pages 856–869, 2020.

Topics supervised by Mei-Hua 2



f. Assessing Learner Argumentation

- Given literature:

Wachsmuth et al. (2016). Henning Wachsmuth, Khalid Al-Khatib, and Benno Stein. Using Argument Mining to Assess the Argumentation Quality of Essays. In: Proceedings of the 26th International Conference on Computational Linguistics, pages 1680–1692, 2016.

Stiegelmayr et al. (2018). Andreas Stiegelmayr and Margot Mieskes. Using Argumentative Structure to Grade Persuasive Essays. In Proceedings of the International Conference of the German Society for Computational Linguistics and Language Technology GSCL 2017: Language Technologies for the Challenges of the Digital Age, pages 301-308, 2018.

g. Supporting Learner Argumentation

- Given literature:

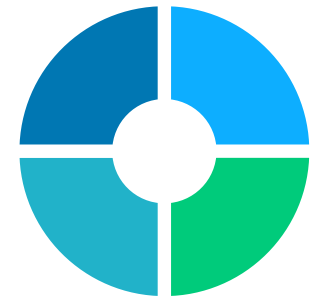
Wambsganss et al. (2020). Thiemo Wambsganss, Christina Niklaus, Matthias Cetto, Matthias Söllner, Siegfried Handschuh, and Jan Marco Leimeister. AL: An Adaptive Learning Support System for Argumentation Skills. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, pages 1–14, 2020.

Wambsganss et al. (2021). Thiemo Wambsganss, Sebastian Guggisberg, and Matthias Soellner. ArgueBot: A Conversational Agent for Adaptive Argumentation Feedback. In Proceedings of the 16th International Conference on Wirtschaftsinformatik, 2021.

Sum-up

Conclusion

- **Seminar topics**
  - 20 candidate topics related to 1–2 of the four areas
  - Each of you will be assigned one of these topics
  - Given + further literature form the basis of talk and article



- **Topic assignment**
  - You choose topic preferences, we assign topics
  - Inform yourself about the topics of the given literature
  - Send me your topic preferences by Sunday this week!

- **Next up**
  - Topic assignment will be done until the next session
  - Basics of scientific presentation in the next session
  - Talk preparation starts then