Theory of mind in language, minds, and machines: a multidisciplinary approach

Dijk, B.M.A. van

Acknowledgements

I remember calling my partner Willeke, right after a job interview in Max van Duijn's office and mentioning I was sure that we had a match. My feeling was right, as from that moment we shared many unforgettable experiences that contributed to this dissertation in direct and indirect ways: from visiting chaotic classrooms to record children's stories to cycling London's empty roads in the middle of the night. Max, thank you for the freedom to develop my own line of research within our collaboration, for always being supportive, and for being open to talk about anything. We discussed many things larger than research for which I am grateful. The most remarkable moments were perhaps the ones we were not together: you presented our story corpus twice exactly when Willeke was giving labour to my two children. Not many supervisors can say the same thing.

Marco Spruit, perhaps you felt at first a bit puzzled about my topic in relation to your own work, but our collaboration turned out great. In fact, we are now collaborating in the medical domain - as your field of expertise - with the tools and insights obtained during my time as PhD candidate, which I consider a good plot twist. Thank you for always being straight to the point, for showing how in academia one can properly unwind, and for staying composed when reviewers did not seem to appreciate our work.

Tom Kouwenhoven, starting from our Leiden coffee strolls, you were my fellow office mate in Huygens 126. You helped me solve numerous issues with my code and always provided a constructive critical eye on ongoing projects. I enjoyed working on our joint papers and always felt motivated going to the office knowing I would see you. You have become a good friend, illustrated by the fact that you cooked for hours in my kitchen while not even being able to enjoy the food yourself.

Barend Beekhuizen, thank you for your hospitality at the University of Toronto, for the insightful discussions we had about children's language use, and for the op-

portunity to present my work. Toronto was a milestone in my time as PhD candidate that was beyond my imagination.

Ageliki Nicolopoulou, it was very special to meet the author of many papers seminal to my own work. Thank you for making our symposium a success; I feel privileged to have met you and to have discussed children's captivating story worlds.

Thanks to many research assistants for help with collecting and labelling data: Iris Jansen, Isabelle Blok, Li Kloostra, Lola Vandame, Nikita Ham, Werner de Valk, and Yasemin Tunbul. Special thanks to Werner for the great drawings accompanying our experiments as the Creative Intelligence Lab's (CIL) artist-in-residence, and to Li Kloostra for proofreading various papers.

Thanks to the CIL members for making this PhD journey a fun one, including (but not limited to) Dan Xu, Danica Mast, Giulio Barbero, Maarten Lamers, Marcello Gómez-Maureira, Marianne Bossema, Max Peeperkorn, Peter van der Putten, Ramira van der Meulen, Rob Saunders, Ross Towns, Tessa Verhoef, Tom Breedveld, and Zane Kripe. I was also happy to be part of Marco's Translational Data Science lab, and thank Alireza Shojaifar, Armel Lefebvre, Chaïm van Toledo, Hielke Muizelaar, Injy Sarhan, Jim Achterberg, Marcel Haas, Max van Haastrecht, and Samar Samir. I feel lucky to still collaborate with many of you. Special thanks to Max van Haastrecht as co-organiser of the PhD seminar and the fun social events we organised together. Peter Dekker, thanks for the fruitful exchanges we had about living the PhD life. Suzan Verberne, thank you for your advice at various stages of my project and the opportunity to present my work in your Text Mining and Retrieval lab. I also thank the Media Technology MSc programme staff for the opportunity to develop my teaching skills during my PhD.

Thanks dear Albert van Dijk and Marianne van Dijk - who we miss dearly - for your love and support. The unique Van Dijk thing I share with Daan and Tobias van Dijk I will always value. Thanks also to Leonor Villa Acosta and Cristhian Pinzon Villa for sharing many precious moments. Pieter Verduijn, having your artistic rendition of this dissertation's contents as the cover is an honour. Theresa Montenarello, thanks for helping with digitalising Pieter's work while still being in your baby bubble. Thanks also to Teun and Corrie Verduijn for caring for my children many times while I was working on this dissertation.

Lastly I would to like express my deep gratitude to Willeke Verduijn, not my best friend or buddy but life partner, who has seen it all the last four years. Please know I learn more from you every day about love, patience, kindness, and what truly matters in life, than any degree could offer.

List of publications

Asterisks denote equal contributions.

1. Van Dijk, B.M.A. and Van Duijn, M.J. (2021). Modelling Characters' Mental Depth in Stories Told by Children Aged 4-10. In Fitch, T., Lamm, C., and Leber, H., editors, *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 43, pages 2384-2390.

2. Van Dijk, B.M.A.,* Van Duijn, M.J.,* Verberne, S., and Spruit, M.R. (2023). ChiS-Cor: A Corpus of Freely-Told Fantasy Stories by Dutch Children for Computational Linguistics and Cognitive Science. In Jiang, J., Reitter, D., and Deng, S., editors, *Proceedings of the 27th Conference on Computational Natural Language Learning*, pages 352-363. Association for Computational Linguistics.

3. Van Dijk, B.M.A., Spruit, M.R., and Van Duijn, M.J. (2023). Theory of Mind in Freely-Told Children's Narratives: A Classification Approach. In Rogers, A., Boyd-Graber, J., and Okazaki, N., editors, *Findings of the Association for Computational Linguistics*, pages 12979-12993. Association for Computational Linguistics.

4. Van Duijn, M.J., Van Dijk, B.M.A., and Spruit, M.R. (2022). Looking from the Inside: How Children Render Character's Perspectives in Freely-told Fantasy Stories. In Clark, E., Brahman F., and Iyyer, M., editors, *Proceedings of the 4th Workshop on Narrative Understanding*, pages 66-76. Association for Computational Linguistics.

5. Van Dijk, B.M.A., Van Duijn, M.J., Kloostra, L., Spruit, M.R., and Beekhuizen, B.F. (2024). Using a Language Model to Unravel Semantic Development in Children's Use of a Dutch Perception Verb. In Zock, M., Chersoni, E., Hsu, Y., and

De Deyne, S., editors, *Proceedings of the 8th Workshop on Cognitive Aspects of the Lexicon*, pages 98-106. European Language Resources Association.

6. Van Duijn, M.J.,* Van Dijk, B.M.A.,* Kouwenhoven, T.,* De Valk, W.M., Spruit, M.R., and Van Der Putten, P.W.H. (2023). Theory of Mind in Large Language Models: Examining Performance of 11 State-of-the-Art Models vs. Children Aged 7-10 on Advanced Tests. In Jiang, J., Reitter, D., and Deng, S., editors, *Proceedings of the 27th Conference on Computational Natural Language Learning*, pages 389-402. Association for Computational Linguistics.

7. Van Dijk, B.M.A., Kouwenhoven, T., Spruit, M.R., and Van Duijn, M.J. (2023). Large Language Models: The Need for Nuance in Current Debates and a Pragmatic Perspective on Understanding. In Bouamor, H., Pino, J., and Bali, K., editors, *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 12641–12654. Association for Computational Linguistics.

Curriculum Vitae

Bram van Dijk was born in Bogotá, Colombia in 1991. He obtained his bachelor's degree in Social Science in 2014 and his master's degree in History and Philosophy of Science in 2018, both from Utrecht University. After briefly working as a research assistant for various Dutch universities, he started in 2020 as a PhD candidate collaborating with dr. Max van Duijn's NWO-funded project 'A Telling Story', that was focused on unravelling Theory of Mind in children's stories. In 2023 their joint paper presenting the Dutch children's story corpus ChiSCor won the best paper award at the Conference for Computational Natural Language Learning in Singapore. In the same year, Bram was briefly a visiting researcher at the University of Toronto, collaborating with dr. Barend Beekhuizen on analysing the semantics of children's use of perception verbs. Bram completed courses in deep learning, text mining, and science communication, and co-taught the course 'Sciences and Humanities' in the Media Technology MSc programme as part of developing complementary academic skills in his PhD trajectory. Currently he is working as a postdoctoral researcher in prof. dr. Marco Spruit's Translational Data Science lab at the Leiden University Medical Center, where he focuses on the application of Computational Linguistics in the medical domain.