

# Jelte van Waterschoot

Human Media Interaction  
University of Twente  
Drienerlolaan 5  
7522 NB Enschede  
The Netherlands

`j.b.vanwaterschoot@utwente.nl`  
`people.utwente.nl/j.b.vanwaterschoot`

## 1 Research Interests

My research interests lie mainly in the area of **discourse structuring**. I have an interest in **topic management**, where a virtual agent can regulate the introductions, continuations and shifts of **topics** in their conversation in order to have a **personalized, coherent** and **natural** dialogue (Rats, 1996). By personalized I mean that the agent adapts the dialogue to the user. **Coherence** is as a close link in meaning between textual units (Jurafsky and Martin, 2014). Specifically, I look at possible **break-downs** of coherence, for example when the agent misunderstands the user too often or the agent says something that is (mostly) irrelevant to the current discourse phase. With natural I mean that the conversation should mimic how humans converse with each other, including dynamic **turn-taking** and smooth **topic transitions**, especially during **interruptions**.

### 1.1 Turn-taking and Interruption Management

My research takes place in the context of the ARIA-VALUSPA<sup>1</sup> project, in which we use an agent called Alice, from Alice in Wonderland<sup>2</sup>. She can process multimodal input, such as speech and visual information. We are working on developing a dialogue manager (DM) for Alice that should be able to sustain a personalized, natural and coherent conversation (Valstar et al., 2016).

In particular for the natural aspect of the conversation we work on dynamic turn-taking, which will be based on the work of Loch (2011). His work gives us the possibility to both deal with floor and turns in one model that is easy to understand and provides us the means to have adaptive turn-taking, for example for designing a (very) dominant agent that often interrupts the user and tries to keep the turn during the interaction or perhaps a more patient (submissive) agent that refrains from interrupting the user too much.

Furthermore, an important aspect in our dialogue manager is the possibility to deal with interruptions of the conversation, using the dynamic turn taking mentioned

<sup>1</sup>[www.aria-agent.eu](http://www.aria-agent.eu)

<sup>2</sup><https://github.com/ARIA-VALUSPA/ARIA-System>

in the previous paragraph. For interruptions we look at both interruptions of the agent by the user and vice versa. In the former case, we want the agent to respond appropriately (replan behavior, ignore the interruption or hold the behavior) and in the latter we want to build an interruption manager that knows when and how to interrupt (Cafaro, Glas and Pelachaud, 2016).

### 1.2 Topic Management

Firstly, I'm interested in how topic management influences the **naturalness** of the dialogue. **Topics** can be defined as the 'aboutness' of the dialogue (Rats, 1996). The idea is that every utterance has a topic and the agent simulates understanding of the topics by responding with an utterance that is topic related, or by using a topic transition strategy to start talking about another topic. Glas has found that topic management can result in a more engaging conversation for the user, which is related to the naturalness and coherence of the dialogue (Glas, 2015).

What I am particularly interested in is designing data-driven **transition strategies**. How can the agent introduce a topic gracefully or steer towards a topic that it wants to discuss? To find out, I want to perform a Wizard-of-Oz corpus study, where a wizard representing the agent will talk freely with a participant. Both the wizard and participant get assigned topics they should discuss (which can be related or non-related topics). In particular I will look at how the wizard performs topic management during the conversation, because her (or his) actions are representative of what the autonomous agent's actions should be. I will annotate the corpus with topic transitions, cues for topic transitions and manually craft rules that can be put in a topic management module for the agent.

Furthermore, I consider that **topic recognition** is necessary in topic management, because the agent can make topic transitions more personalized and natural by recognizing the user's topics of interest. Currently, topic recognition in dialogues often works on the utterance level (select the noun phrase in the subject or object position), but it could be extended determine a topic of a dialogue segment (Rats, 1996). To learn how an agent should perform segment-level topic recognition, I will have observers in-

dicating topics in the dialogues of the corpus.

In a final experiment I want to have users talking to a version of Alice that contains topic management and is able to recognize higher-level topics and using human-like topic transition strategies. To evaluate these dialogues, observers will rate the recorded conversations and use for example metrics like human-likeness and competence (Glas, 2015).

## 2 Future of Spoken Dialog Research

Often dialogue experts are necessary for domain experts to design dialogues. I expect that in the near future a toolkit will be available, similar to the VHToolkit or our ARIA platform, which is easy enough to use for domain experts to design their own domain-related dialogues without the need for dialogue experts.

Moreover, robots and agents used in noisy environments, such as schools and museums, have a hard time interacting appropriately with the user, especially on the dialogue level with spoken language. How can we overcome this limitation?

I think it will take some decades, or maybe even more, to have collected enough multi-modal data for using only machine learning for designing dialogue systems. I expect that contemporary chatbots will then evolve into embodied agents using more multimodal interaction.

## 3 Suggestions for Discussion

Here I provide some of the topics I would like to discuss at the event:

- How to **measure** the naturalness of a conversation? This could include quantitative metrics such as number of repairs of the agent, topic distance between utterances. A qualitative measure is letting observers judge the naturalness of the conversation via a questionnaire.
- What are good **interruption strategies** for an agent? What **information** do you need for these strategies? We know how users perceive the agent with cooperative and disruptive interruptions, but how should the agent respond to interruptions of the user? How can it decide to replan behavior or just ignore it?
- What are good **tools** to **build dialogues**? How much data do you need for designing domain-specific dialogues? Should we stick to easy-to-author dialogue trees? Or are there good possibilities with machine learning?
- How do we determine the **topic** of a dialogue (segment)? Which methods already exist and which information from the dialogue should we use?

## 4 Acknowledgements

ARIA-VALUSPA is funded by European Union Horizon 2020 research and innovation program, grant agreement No 645378.

## References

- Michel Valstar, Tobias Baur, Angelo Cafaro, Alexandru Ghitulescu, Blais Potard and Johannes Wagner. 2016. *Ask Alice: An Artificial Retrieval of Information Agent*. Proceedings of the 18th ACM International Conference on Multimodal Interaction
- Mieke M. M. Rats. 1996. *Topic Management in Information Dialogues, Translation and Compiling*. ITK Dissertation Series, Tilburg, NL.
- Nadine Glas. 2015. *Engagement driven Topic Selection for an Information-Giving Agent*. Workshop on the Semantics and Pragmatics of Dialogue, SemDial 2015, Gothenburg, Sweden.
- Angelo Cafaro, Nadine Glas and Catherine Pelachaud. 2016. *The Effects of Interrupting Behavior on Interpersonal Attitude and Engagement in Dyadic Interactions*. Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems.
- Frieder Loch. 2011. *Computational Models for Turn Management using Statecharts*. Master's Thesis. University of Twente.
- Dan Jurafsky and James H. Martin. 2014. *Speech and language processing*. Vol 3. London: Pearson.

## Biographical Sketch



Jelte van Waterschoot has a bachelor degree in Artificial Intelligence from the Radboud University Nijmegen and a master degree in Artificial Intelligence from Utrecht University. Furthermore he has completed an internship at TNO<sup>3</sup> involving dynamic task distribution between human firefighters and robots during fire rescue missions. He started his PhD in the beginning of 2016 on the ARIA-VALUSPA project, working on dialogue management for a virtual agent. His supervisors are prof. dr. D.K.J. Heylen and dr. M. Theune. His main expertise lies in artificial intelligence and linguistics. He wants to develop intelligent agents that can use common-sense reasoning for problem solving, creativity and conversations. Furthermore, he is interested in participation within the university, as active member of an advising committee for the faculty.

<sup>3</sup>Dutch Organization for Applied Scientific Research