# I Talk You Decode:
## Strategic Processing of Context-Sensitive Elements in Speech Interactions

## Introduction

Resolving context-sensitive elements (e.g. pronouns, indexicals) presents a core challenge in natural language understanding. Research in this area, especially coreference resolution, has developed from traditionally using rule- or feature-based methods (building on Hobbs 1978) to more recently applying statistical models and deep learning methods (Clark and Manning 2015, Lee et al. 2017, amongst many others). This project builds on current theoretical approaches to processing context-sensitive elements, but integrates the topic into a wider interdisciplinary research framework concentrating on data-driven computational pragmatics. The empirical work focuses on the interpretation of context-sensitive elements in task-oriented speech interactions. Spoken language consists of multiple communication channels of which only one uses words. Extralinguistic information (e.g. gestures, facial expressions and gaze etc.) therefore plays a crucial role to decode references of context-sensitive expressions.
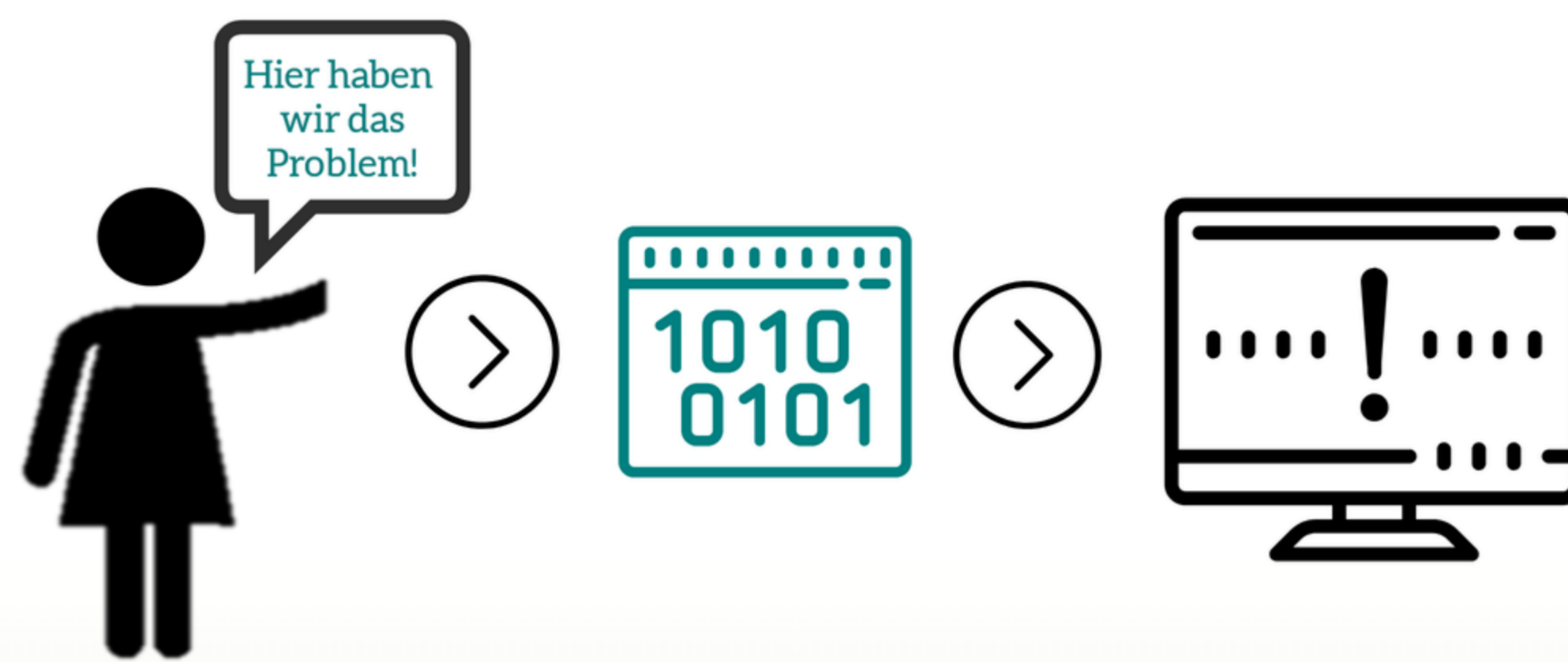


## Objectives

▶ Develop a computer-processible, **multidimensional** discourse analysis framework for context-sensitive elements

▶ Contribute an integrative theoretical approach for understanding and interpreting context and multimodality in spoken language to the field of NLP research.

## Research Questions

▶ **(Q1)** What does an automated system need for a human-like understanding of context-sensitive elements in spoken language?

▶ **(Q2)** How do linguistic and contextual constraints affect referring relations?

▶ **(Q3)** What linguistic input and multi-modal cues are offered by speakers to decode reference-shifting expressions?

## Methods

Mixture of linguistic experimental and computational methods:
**Experimental Pragmatics:** spontaneous speech experiments to investigate the spectrum of (multimodal) utterance choices by speakers
**Deep Learning:** neural network architecture to explore context sensitivity in the data

## Data

Annotated audio- and video recordings collected in task-oriented experiments performed by German native-speakers. The preprocessed data will be used to train a multi-dimensional processing algorithm. The algorithm functions as a predictive matcher for input signals and possible interpretations of context-sensitive elements.

## Study Design

Two different experiment setups that involve a situated teacher-learner scenario:

▶ **Experiment I dyads:** A single participant adopts a teacher role and explains a simple task (object manipulation) to a passive listener.

▶ **Experiment II groups:** A group of participants (2) collaboratively perform a simple task and explain their steps to a passive listener.

The role of the passive listener in both speech experiments is taken by a member of the research team whose task it is to express non-verbal signs (e.g. nodding) of following the explanation during the experiments.

### References

Clark, K., & Manning, C. D. (2015). Entity-centric coreference resolution with model stacking. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing* (Volume 1: Long Papers), 1405-1415.

Hobbs, J. R. (1978). Resolving pronoun references. In *Lingua, 44(4),* 311-338.

Lee, K., He, L., Lewis, M., Zettlemoyer, L. (2017). End-to-end neural coreference resolution. In *Proceedings of EMNLP, 9-11, 2017,* 188–197.

## Lisa Nußbaumer
*a01110284@unet.univie.ac.at*
**University of Vienna, Linguistics Department**