

# Conversation Transcript as Utterances (i.e. Sentences)

## Dialogue Acts

Aim: statistically analyse conversations as a string of microscopic dialogue acts. These represent the **social functions** of utterances.

Tokenisation,  
Chunking

Training Data

Embeddings

Model

Classified Chunks

Recombined Chunks

### Example Applications:

- Is one person driving the conversation?
- How opinionated are the statements?
- How do participants respond to questions?
- Can a set of conversation types be identified from statistically analysing dialogue acts?
- Do certain dialogue act statistics correspond to more popular conversations (e.g. for talk shows, podcasts)

## Topics

Aim: analyse conversations macroscopically as a trajectory through topic space. This represents the **semantic meanings** of the conversation.

Tokenisation

### Keyword Extraction

Named Entity  
Recognition

Part Of Speech  
Tagging

Remove Wrong Keywords

Determine Keyword  
Ranges

Graph of Embedded  
Keyphrases for Topic  
Extraction

Embeddings

### Example Applications:

- How do topics evolve in time?
- What level of expertise do the speakers demonstrate?
- What breadth of topics is addressed?
- Are certain topics avoided?
- Summarisation applications for conversations, e.g. for recommender systems, search engines or insights