



KNOWLEDGE DISCOVERY

Cross-layer language models for conversational speech

Philipp Gabler, Name2, Name3, ...

MOTIVATION

Context: improvement of language models (LMs) for conversational Austrian German

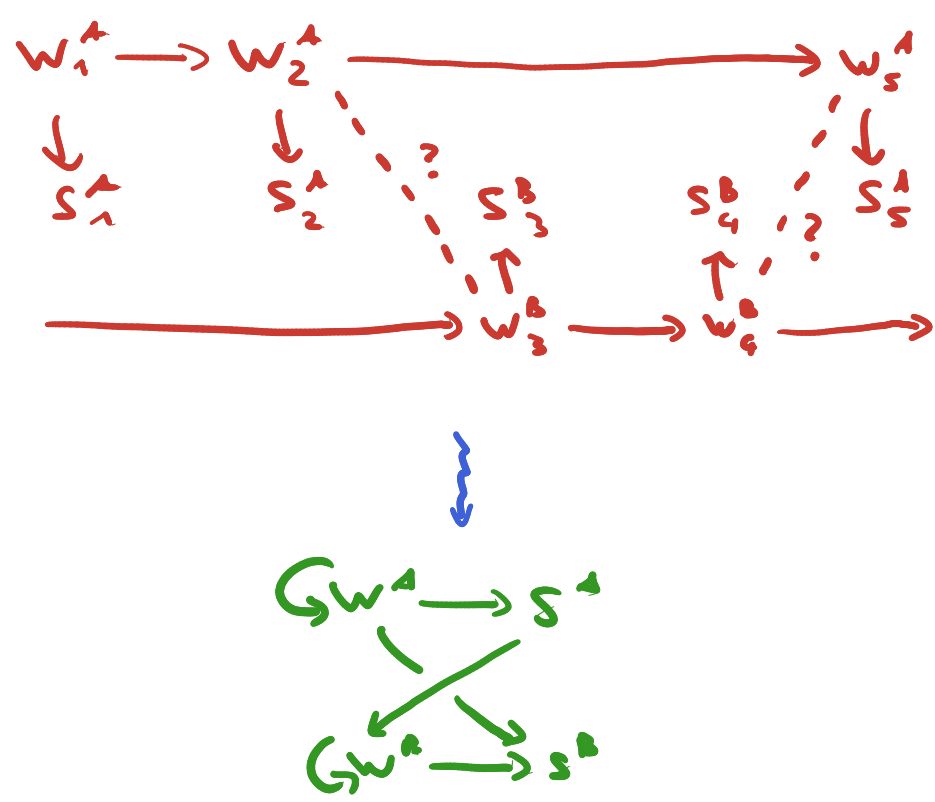
- Improve ASR for conversational speech (CS)
- Help increase knowledge about production and perception of CS
- Increase knowledge & resources for Austrian German

Research directions:

- LMs for conversational language & interaction
- Long-range dependencies, local contexts, cross-layer features, high-level semantics
- Generative/causal models for linguistic data

BACKGROUND

MODELS TAKING MULTI-SPEAKER INTERACTION SERIOUSLY?

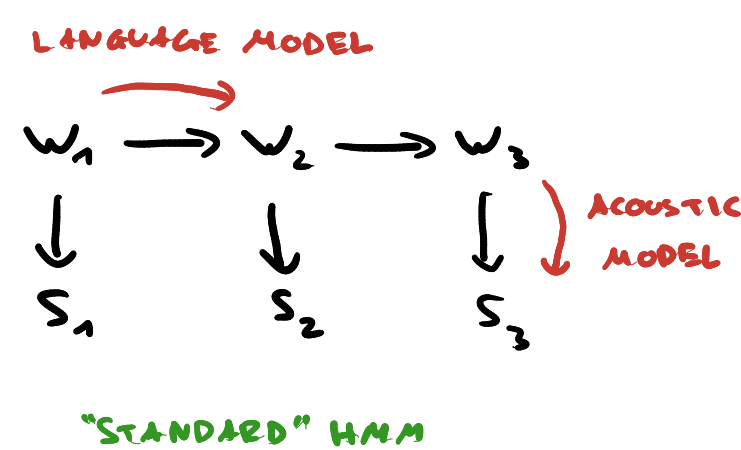
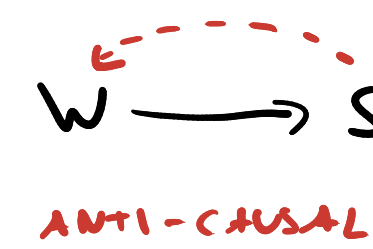


APPROACHES TAKING INTO ACCOUNT DIFFERENCES IN CAUSAL STRUCTURE?

ASSOCIATION → PREDICTION

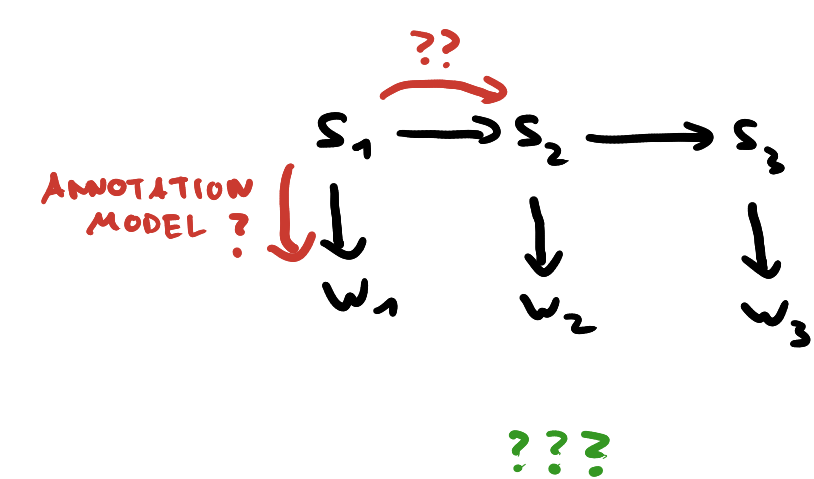
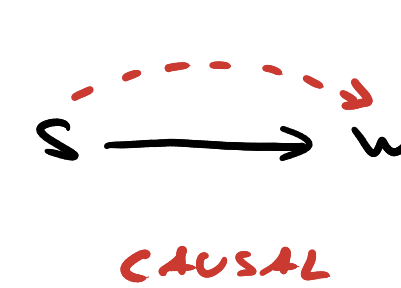
MECHANISM → INTERVENTION/COUNTERFACTUAL

READ SPEECH



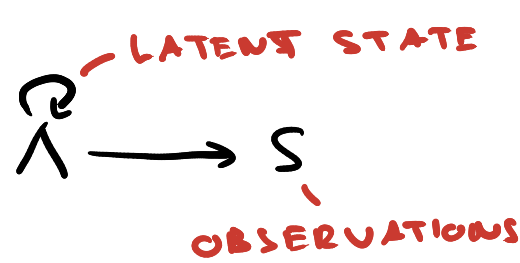
W: WRITTEN SIGNAL, S: SPOKEN SIGNAL (BOTH OBSERVED)

SPONTANEOUS SPEECH



NEW APPROACH

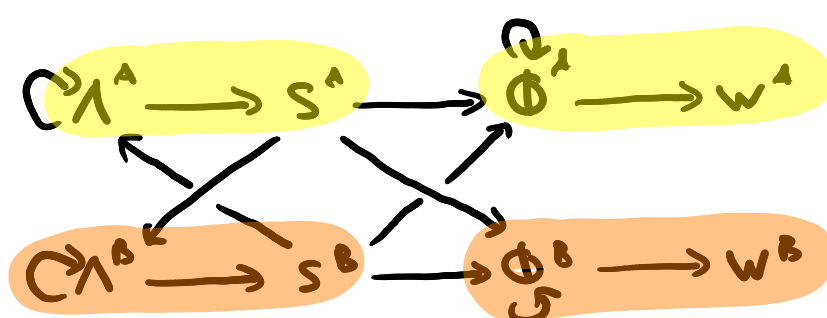
SPEAKER MODULE



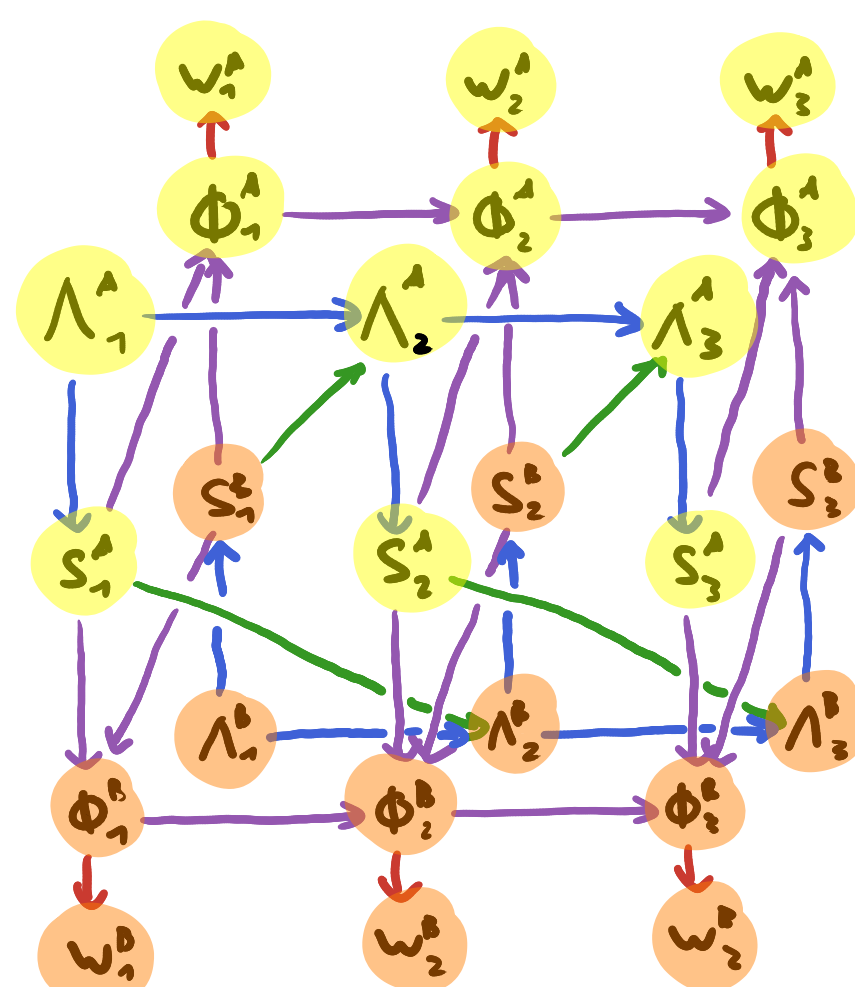
WRITER MODULE



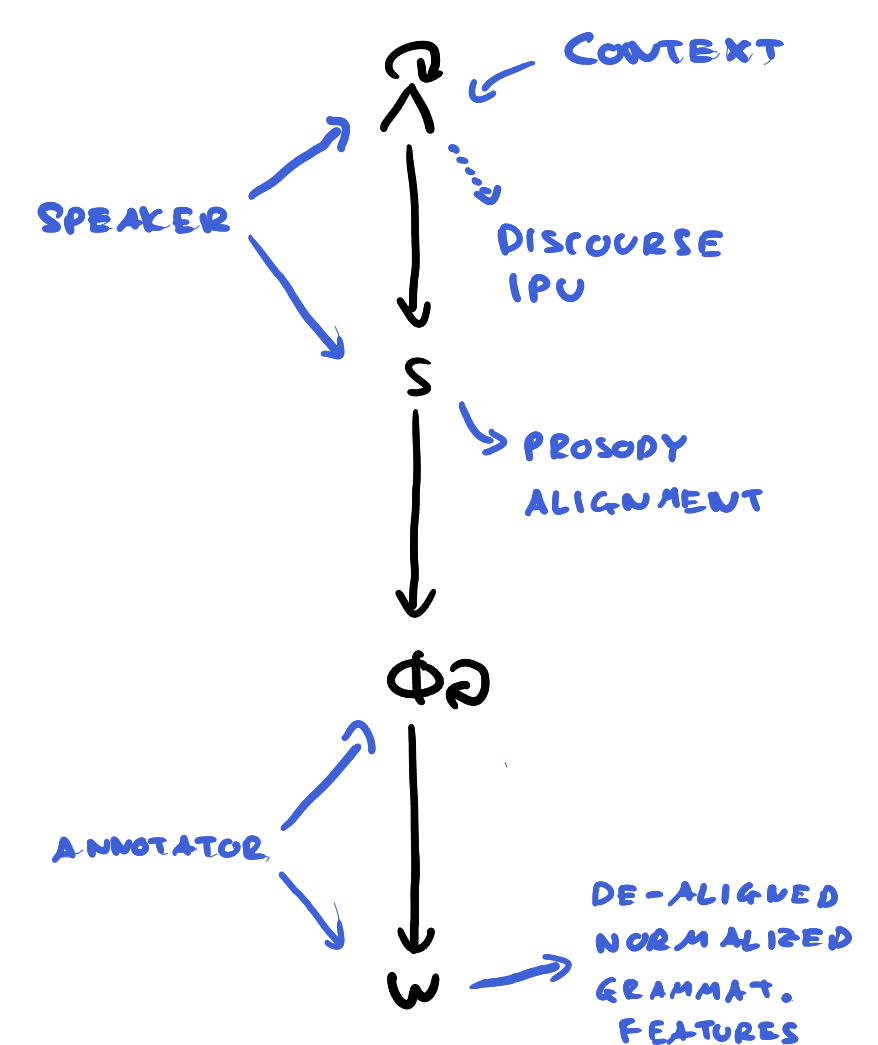
CONVERSATIONAL SPEECH



Λ, Φ : UNOBSERVED INTERNAL STATES



A B
SPEAKER PERCEPTION
SPEAKER PRODUCTION
ANNOTATOR PERCEPTION
ANNOTATOR PRODUCTION



PROJECT PLAN

Vision:

- Flexible toolbox of conditional LMs
- Clearly distinguish production & perception
- Latent & observed model for every participant
- Few concrete assumptions made

Remaining questions:

- Concrete structure (coupling)
- Representation (states, alignment, priors)
- Inference
- Simplified setting for exploration