Advancing Neural Question Generation for Formal Pragmatics

Kordula De Kuthy, Madeeswaran Kannan, Haemanth Santhi Ponnusamy, Detmar Meurers

SFB 833, Project A4, University of Tübingen

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

SFB 833

Question Generation

- Creating natural questions for a given sentence or paragraph
- Appropriately linking to the meaning of the envisaged answer phrases
- Questions in formal pragmatics:
  - Question-under-Discussion (QuD) approaches
  - link information structure and discourse structure
- Our goal:
  - Partially automate QUD annotation
  - generate all questions that can be answered by a given sentence

Our Approach

- Neural networks can be successfully trained to generate meaningful, well-formed questions (De Kuthy et al., 2020)
- Challenge for a seq2seq architecture: rare or unknown words
  - words are the basic input and output tokens
  - pretrained word embeddings with fixed vocabulary
  - the words in the question to be generated can be selected from source material
- Extend architecture with a pointer component
- Enrich the model with part-of-speech and semantic role information to improve question phrase generation.

⇒ Robust approach with systematic question type coverage

Question-Answer Data

- German QA corpus: 5.24M sentence-question-answer triples
- Sentences from German newspaper *Die Tageszeitung (taz)*
- Transformation-based question generation system (Kolditz 2015)
  - Select potential answer phrases (NPs, PPs, clauses)
  - Replace them with matching question phrases.
  - Transform syntactic representations into questions.

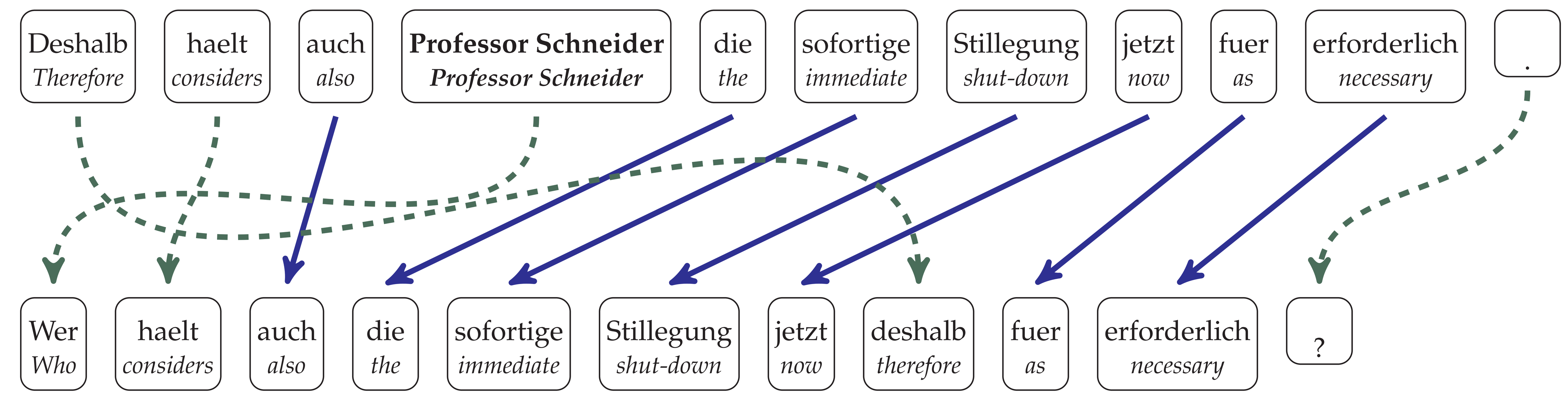
(1) A: Die Kinder essen am Sonntag Kuchen im Garten.
The children eat cake in the garden on Sunday.

Q_1: Wer isst am Sonntag Kuchen im Garten?
Who eats cake in the garden on Sunday?

Q_2: Was essen die Kinder am Sonntag im Garten?
What do the children eat in the garden on Sunday?

Q_3: Wo essen die Kinder am Sonntag Kuchen?
Where do the children eat cake on Sunday?

An Example: Copy vs.Generate Decisions



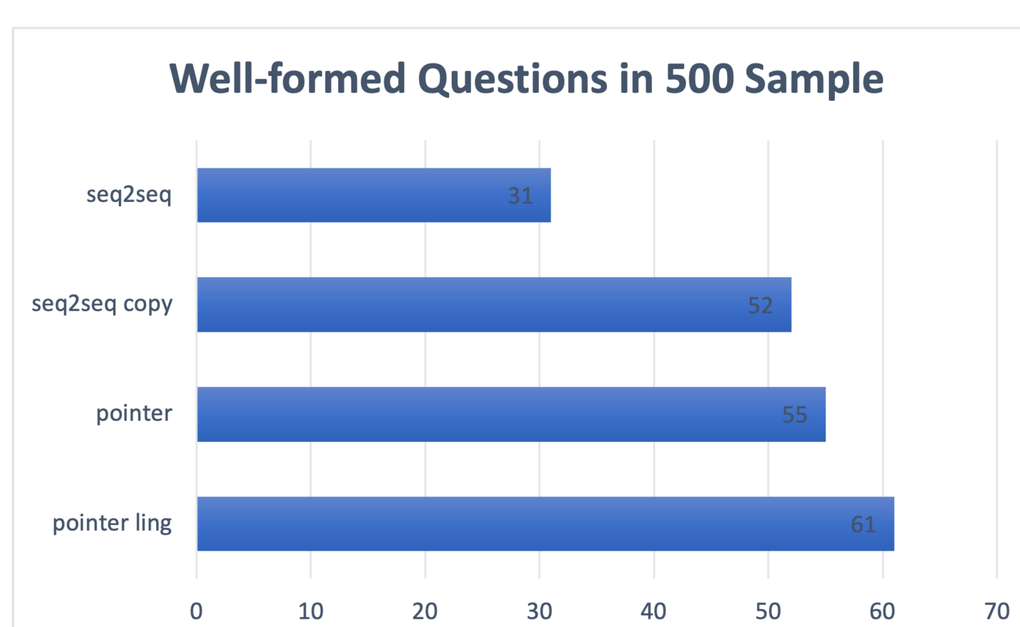
Neural Question Generation

- Basic Sequence-to-sequence architecture (Sutskever et al., 2014):
  - encoder network learns source sequence representation
  - decoder network generates target sequence
  - attention mechanism aligns source and target sequences
- Our Pointer Model:
  - Maxout pointer mechanism with gated self-attention (Zhao et al., 2018)
  - Input sequences: surface tokens, span of answer phrase, parts of speech (POS), semantic role labels (SRL)
  - Copy score: attention scores computed between encoder and decoder hidden state
  - Copy & generation module compete for predicted question

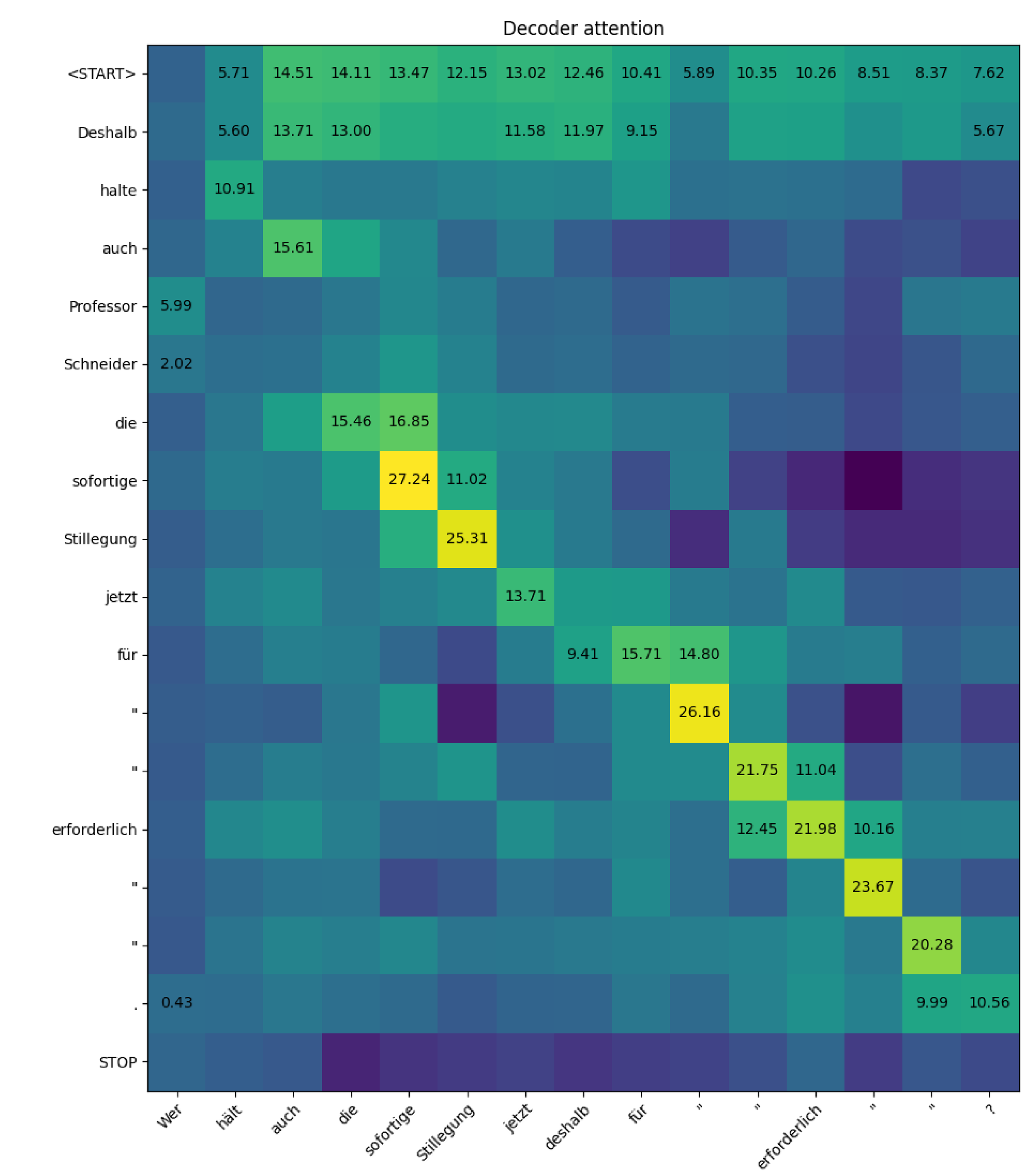
Quantitative Evaluation

- High BLEU score ⇒ high similarity between neural and gold Q

Model	Size	Features	BLEU
Seq2seq	500k	Word, Ans, POS	71.25
Seq2seq + Copy	500k	Word, Ans, POS	84.24
Pointer	500k	Word, Ans	89.40
Pointer	500k	Word, Ans, POS, SRL	**91.45**



Softmaxed Attention Weight



Distribution of Error Types