Tailored feedback in language learning: synergies in linguistics and computer science

Johannes Heim & Arabella Sinclair

{JOHANNES.HEIM, ARABELLA.SINCLAIR}@ABDN.AC.UK

Problem Statement

Caregivers provide feedback to children in dialogue via repetition, elaboration, corrective feedback, and modelling adult conversation. While previous research has primarily focused on the use and usefulness of corrective feedback to syntactic constructions (Marcus, 1993; Hiller & Fernández 2016), there is a notable gap in our knowledge about the role of situation-specific modelling of adult-like behaviour for pragmatic development.

Form-function mapping of early questions

Prior work analysing this behaviour is often based on individual case studies due to the resource intensity of analysing child language data. We exploit the powerful representations within, and classification ability of computational language models (LMs, large pretrained causal transformer models e.g. the GPT family) to recognise and analyse different types of feedback.

RQ: Can we automatically trace and identify different types of Caregiver Feedback about question-behaviour in dialogue? making use of:
  – predictable turn-taking behaviour (Casillas & Frank, 2017)
  – the lexical overlap between speaker-turns (Reich, 2011)
  – response markers (Kramer & Rawlins, 2009) - e.g. yes, no

Types of Caregiver Feedback

Examples taken from the Sachs corpus (Sachs, 1983).

Method

We plan to apply techniques to classify utterances in dialogue to identify between-speaker specific repetition patterns common to feedback giving. We then plan to use LM representations of these utterances alongside other more surface based features (e.g. lexical overlap, speaker identifier and punctuation marks) as a tool to re-analyse the dialogue data. The resulting analysis will benefit important questions in each discipline and settle long-standing controversies about language learning, including the role of caregiver feedback in learning conversational skills. This approach can be extrapolated to virtually any emerging form-function mapping with active caregiver input.

Contributions

This has implications for research in both linguistics and computational models of dialogue. Firstly, it can shed light on fundamental questions of how children exploit caregiver feedback to advance their conversational skills; secondly, analysing properties of repetition and feedback in human language can have implications for incorporating successful interaction strategies in language models used in dialogue agents.

References