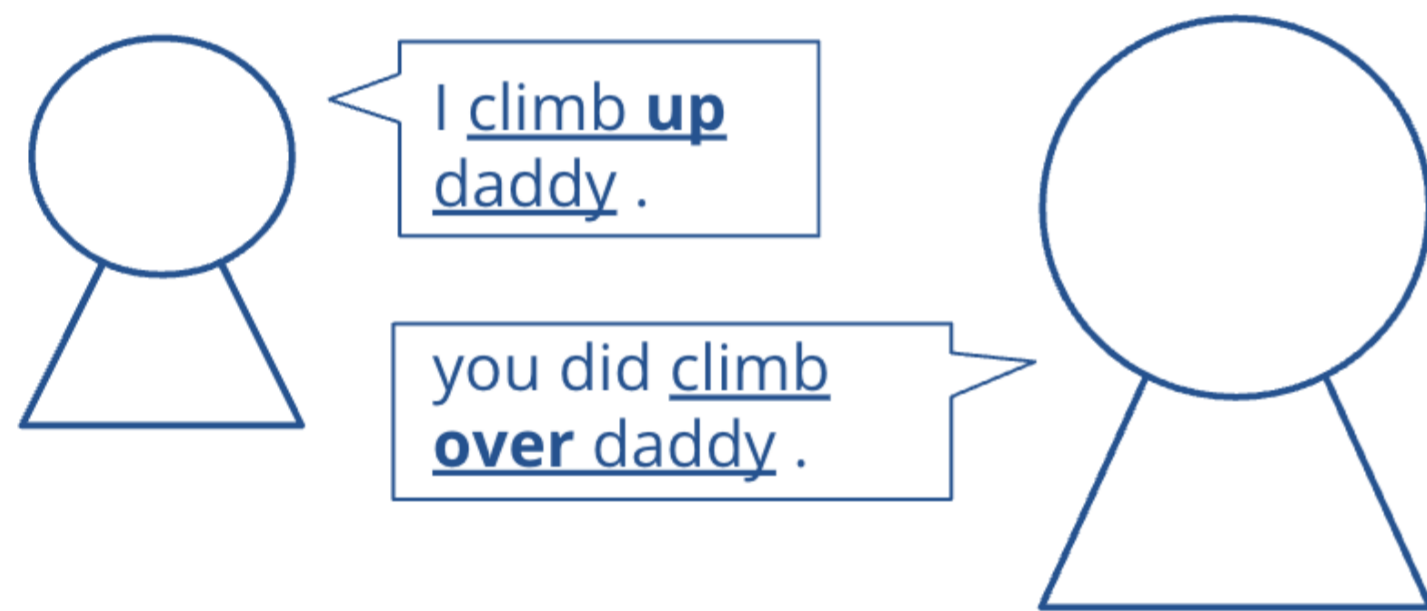


Tailored feedback in language learning: synergies in linguistics and computer science

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**.



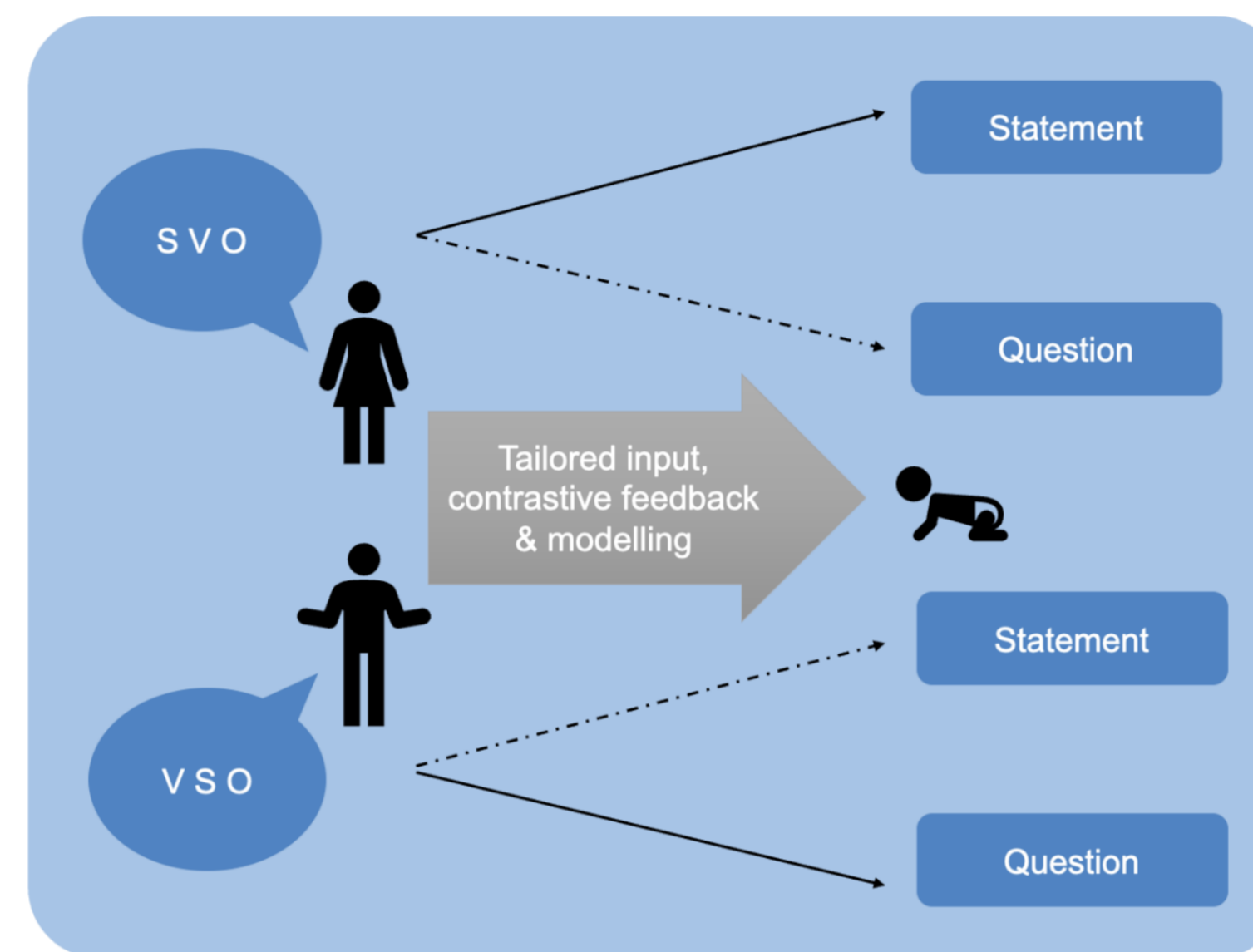
Example Corrective Feedback from Hiller & Fernández 2016

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*

Form-function mapping of early questions



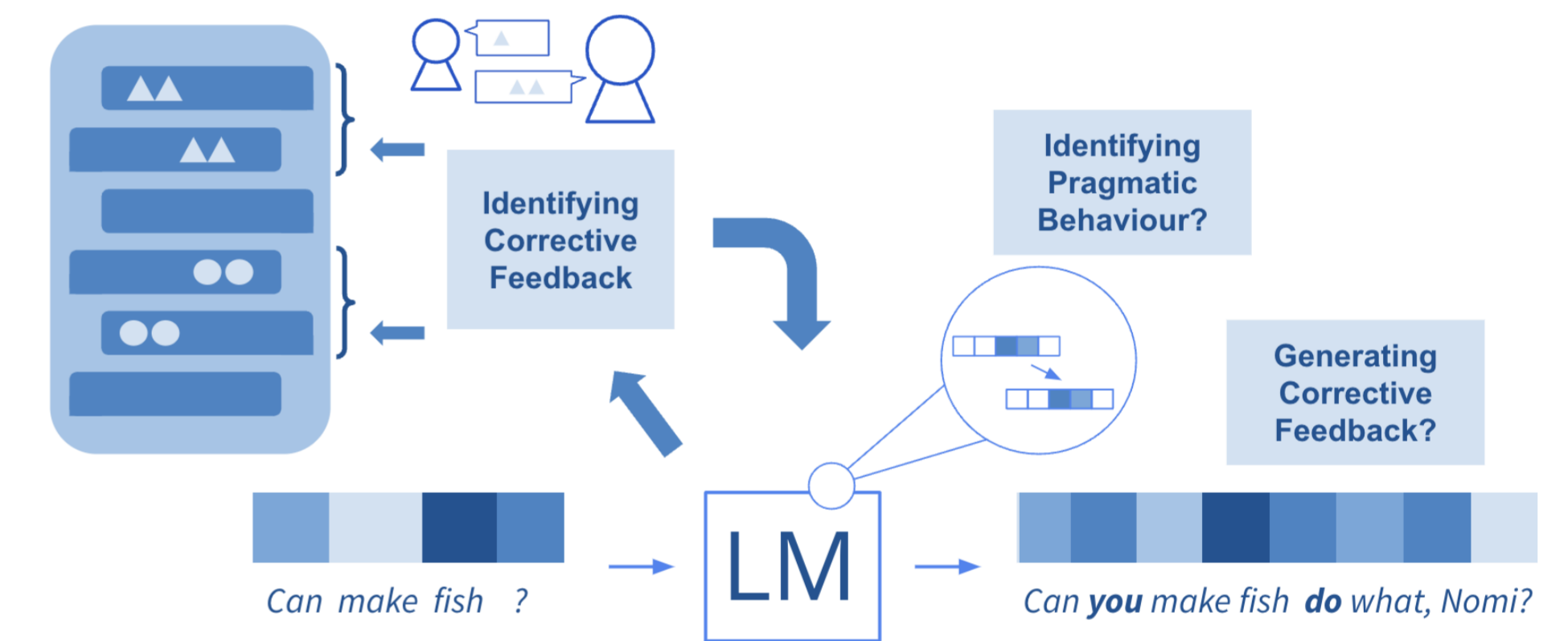
Types of Caregiver Feedback

<p>Modelling Q-A sequences @ age 1;02:29</p> <p>What does the doggie say? What does the dog say? Woof-woof.</p>	<p>Exemplifying Q-A Congruence @ age 1;10.03</p> <p>What's this? What's what? Oh, that's a sheet, honey, a sheet.</p>
<p>Feedback on Q-A dialogue @ age 1;10.10</p> <p>What's this? Pyjamas. Pyjamas, right</p>	<p>Rephrasing child Q's @ age 1;11.20</p> <p>Can make fish? Can you make fish do what, Nomi? Can make fish? Bite.</p>

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 Casillas & Frank (2017) The development of children's ability to track and predict turn structure in conversation. Hiller & Fernández (2016) A data-driven investigation of corrective feedback on subject omission errors in first language acquisition. Kramer & Rawlins (2009) Polarity particles: an ellipsis account. Marcus (1993) Negative evidence in language acquisition. Reich (2002) Question/answer congruence and the semantics of wh-phrases. Sachs (1983) Talking about the there and then: The emergence of displaced reference in parent-child discourse. Shatz (1979) How to do things by asking: Form-function pairings in mothers' questions and their relation to children's responses.