The frequency and distribution of *um* and *uh* in acquisition

Daisy Leigh • Stanford University • ddleigh@stanford.edu

## Background

Children encounter production delays in conversation - especially turn-initially, when the demands of planning and producing an utterance slow down turn-timing. 1,2,3

Adult speakers use pragmatically meaningful delay markers (DMs) like *um* and *uh* to hold the floor when they encounter production problems. 4

The developmental trajectories of *um* and *uh* might be different: children need turn-initial DMs less as they get better at turn-timing, and turn-medial DMs more as they learn the pragmatics of conversation.

## Study 1: ‘Shem’ Case study

### Methodology

- **turn-initial**
  - um we should invite the mouses to play
  - *Shem,* English-speaking monolingual child
  - corpus data from 2.2 to 3.2
  - every DM coded for position in the turn
- **turn-medial**
  - I think um let’s play...
  - *Shem,* English-speaking monolingual child
  - every child-produced utterance coded for type

### Results

- DMs more frequent in question-responses than child-initiated turns ($\chi^2 (1, N = 305) = 36.57, p < 0.001$)
- DMs more frequent in response to wh- vs. polar questions ($\chi^2 (1, N = 153) = 36.38, p < 0.001$)
- Turn-initial DMs significantly more likely in question-responses, turn-medial DMs in child-initiated turns ($\chi^2 (1, N = 305) = 27.15, p < 0.001$)

### Predictions

- Do these initial predictions hold?
  - More complex child-directed questions (why, how) increased at the same time, suggesting that CDS becomes more complex as children’s linguistic and turn-timing skills develop - as observed by others 5

## Study 2: ‘Providence’ Corpus study

### Methodology

- **5 children (3 girls, 2 boys) from the Providence corpus**
- **Data from regular recordings between 1;4 and 3;4**
- **Every DM coded for turn-position and turn type, every child-produced utterance coded for turn-type**

### Pattern 1

- Huge variation in the frequency of DM production: children who used the most showed similar patterns to Shem
- Lexisosyntactic competence (quantified by IPSyn, VOCD, MLU and DSS) did not directly relate to DM frequency
- Exposure to DMs in child-directed speech also did not fully account for the variation
- All children increased turn-initial DM production over time, but displayed one of two broad patterns:
  - **Pattern 1:** DMs used early to hold the floor while lexisosyntactic competence develops
  - **Pattern 2:** DM production increases in line with conversational experience; turn-medial DMs acquired after turn-initial
  - Together, this suggests that some children were still acquiring the pragmatic meaning of *um* and *uh* - only using them to hold the floor mid-utterance after having acquired them to mitigate turn-timing delays

### Pattern 2

- Children start using DMs as early as (1;8), but with varying frequency
- This variation is a product of input, lexisosyntactic competence, sensitivity to pragmatic information, and a desire to hold the floor
- Children begin using DMs turn-initially when turn-timing pressures are particularly acute, and then begin producing turn-medial DMs to manage delays in more complex utterances

### References


Many thanks to Eve Clark, Rob Podesva, Meghan Summer and the Stanford Linguistics community for their invaluable discussion, guidance and support.