Disfluent but effective? A quantitative study of disfluencies and conversational moves in team discourse

Felix Gervits[1]
Kathleen Eberhard[2]
Matthias Scheutz[1]
[1]Tufts University, Human-Robot Interaction Lab
[2]University of Notre Dame, Dept. of Psychology

Summary

Background

Grounding and disfluency in team interaction

Study

Cooperative Remote Search Task (CReST)

Findings

Effective teams make more self-repair disfluencies

Implications

Spoken dialogue systems should interpret disfluent utterances

Background

Grounding in Task-Oriented Dialogue

Teams need to communicate effectively in order to coordinate their activities and establish *common ground*

Grounding involves a mutual recognition of the shared information:

S: I’m going back into room one
D: Okay room one, like the very first starting room?
S: Yeah
D: Okay

Disfluencies Can Support Grounding

Disfluencies of all kinds are prevalent in human speech: pauses, fillers, fragments, self-repairs, etc.

Some view them merely as noise in the speech signal caused by increasing workload (Berthold & Jameson, 1999).

Others view disfluencies as serving an interpersonal function:

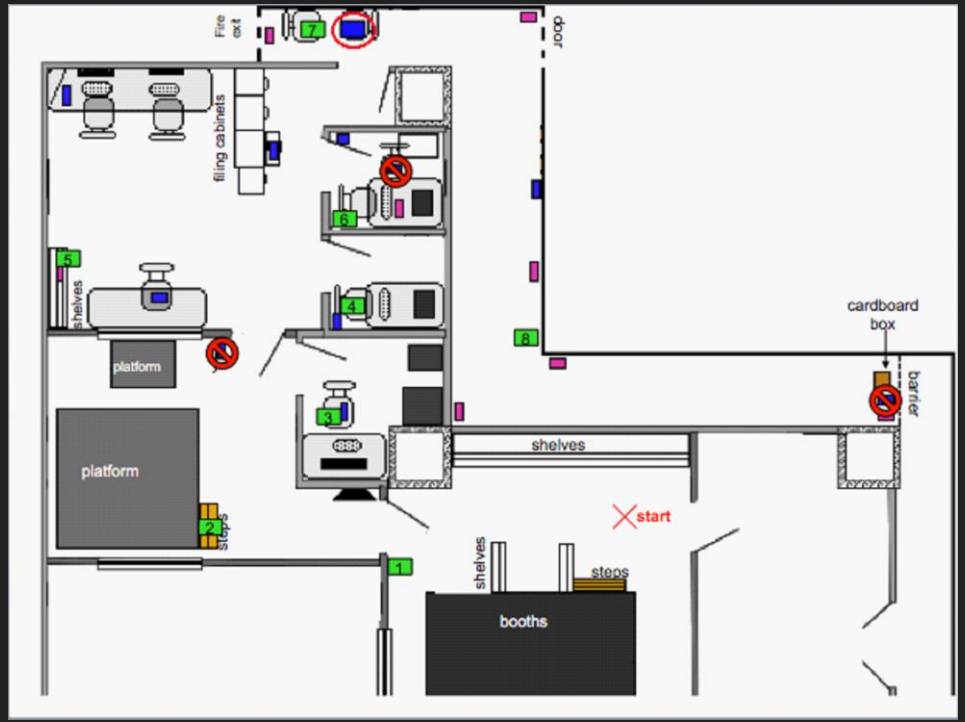
- Hold the conversational floor (Smith & Clark, 1993)
- Processing surrounding speech (Brennan & Schober, 2001)
- Resolve reference ambiguity (Arnold et al., 2007)

Unanswered Questions

1. Do these findings hold in an unscripted collaborative task?
2. Are disfluencies driven by workload or coordination?
3. How do disfluencies interact with grounding strategies?
4. Do self-repairs function as coordination devices?

Study

Cooperative Remote Search Task (CReST) Corpus



- 8 minutes of data were collected from each of 10 dyads

  (2712 utterances, and 15194 words)

- Conversational moves and disfluencies were annotated

- Team effectiveness was measured objectively based on performance

Task Description

Dialogue Event Annotation

Dialogue moves (from Carletta et al., 1997)

Initiation

- Instruct
- Explain
- Wh- Question
- Yes/No Question
- Check
- Align

Response

- Acknowledge
- Wh- Reply
- Yes/No Reply

Ready “OK” + Initiation move

Disfluencies (from HCRC Coding Manual)

Self-repairs:

Repetition - “Look- look in the box”

Substitution - “Pink- I mean blue box”

Insertion - “In the room- the nearby room”

Deletion - “We don’t have- uh let’s hurry up”

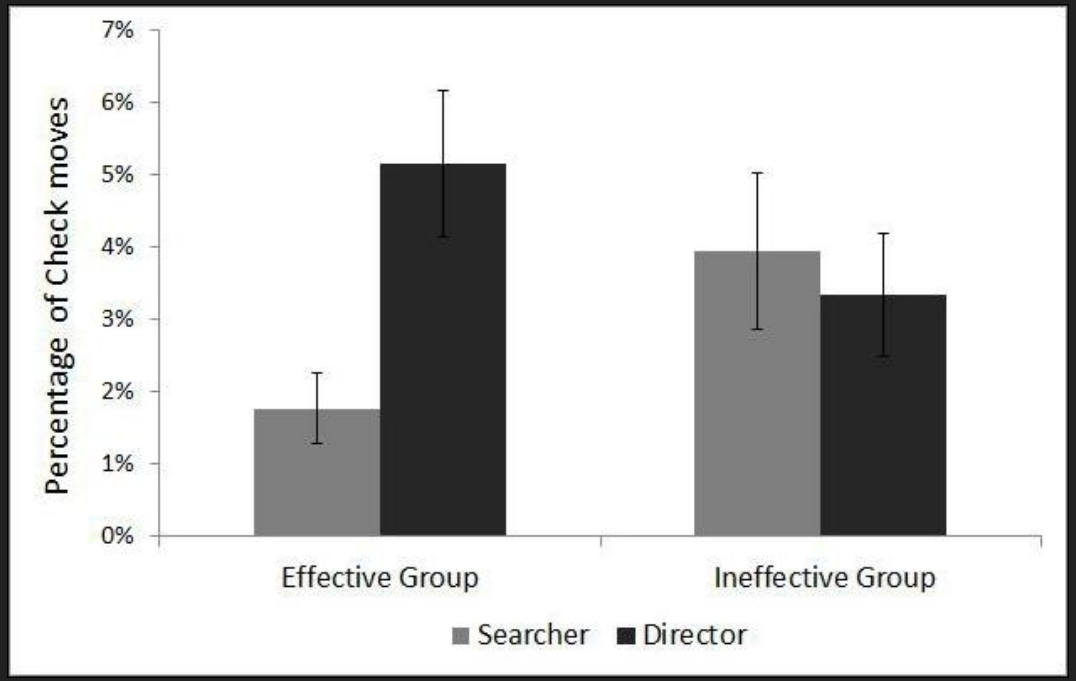
*Pauses were not included in the analysis

Findings

Check Moves

Group x Speaker interaction for *Check* moves ($F(1,32) = 7.053$, $p = .012$).

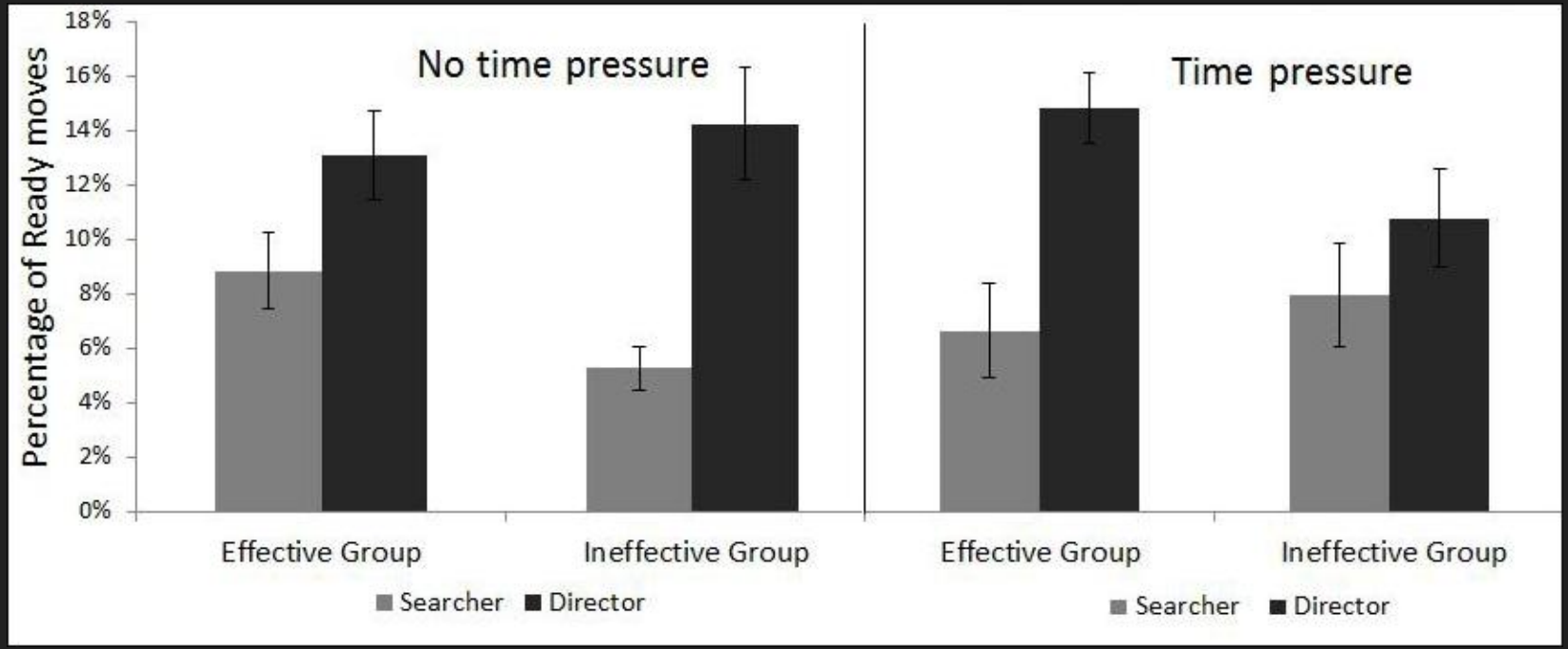
e.g., “You said the box is in the corner?”



Ready Moves

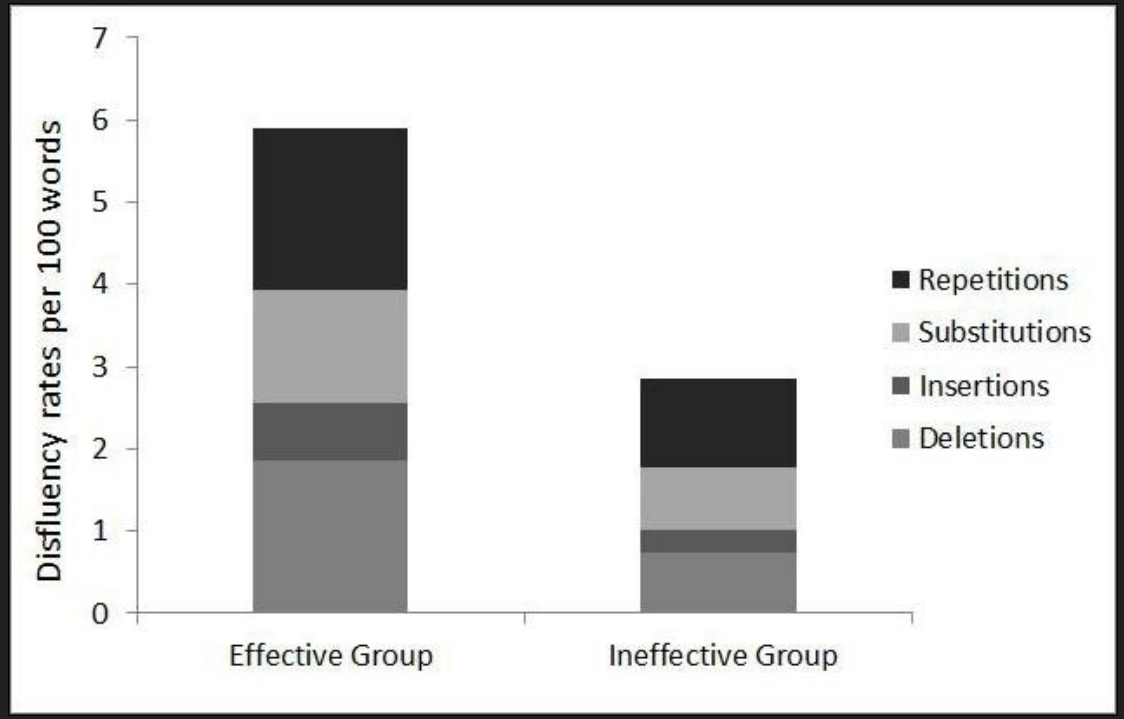
Group x Speaker x Time Pressure interaction for *Ready* moves ($F(1,32) = 4.657$, $p = .039$).

e.g., “OK, walk through the door”



Disfluencies

There was a significant effect of Group (F(4,33) = 2.787, p = .042) on rates of *Insertions* (*F*(1,36) = 4.292, *p* = .046) and *Deletions* (*F*(1,36) = 4.414, *p* = .043).



Disfluency Examples

Deletion disfluency:

D: There’s also one in the second- [pause] uh, we only have three minutes to do this, okay
S: Okay, second cubicle I got that

Substitution/Insertion disfluency:

S: Well [pause] see the two pink boxes?
D: Yes
S: On the right corner - the inside corner
D: Yes

Grounding via Disfluency

Ungrounded:

D: If you look completely straight- straight- straight [pause] like keep walking straight before you even hit the wall, there should be some shelving it looks like. Open the blue box there.

S: Wait w- where- where? Sorry {laughs}

Grounding via Disfluency

Grounded (via installments):

D: If you: turn around go out of that room
S: Okay
D: Straight in front of you should be a chair
S: Yes
D: At a table, there’s a blue box there
S: Yes

Grounding via Disfluency

Grounded (via disfluency):

S: [pause] just as I was about to turn right [pause] there’s kind of this uh stage in front of me a:nd there’s steps up to it and the box – the green box is uh right in front of that on [pause] the- on the step

D: Okay

Implications

Dialogue Systems that Handle Disfluencies

- Effective teams in our study produced twice as many self-repair disfluencies, and interpreted the information that they signalled
- Dialogue systems could benefit from using disfluent utterances
- Focus should shift from “filtering out” to interpreting disfluency function

Mechanisms Needed for Disfluency Handling

1. Identifying the type of disfluency
- requires online incremental processing for real-time prediction
2. Identifying the function of disfluency
- retrieval difficulty, workload, clarification
3. Using the disfluency to interpret speech and make predictions
- e.g., for clarification, supplement the referential description
4. Integrating with embodied capabilities
- Speak or act based on the obtained information

Conclusion

1. Disfluencies have been associated with a coordination function in previous laboratory studies.
2. We tested for the benefit of disfluencies in an unscripted, collaborative, remote search task using the CReST corpus.
3. In our corpus, self-repair rates were higher in effective teams, and were linked to more efficient grounding.
4. Future dialogue systems for situated interaction could benefit from parsing disfluencies to utilize their benefit.