Errors

What’s their impact on overall performance?

Error detection

- What's "error"? -> all miss-communication
- Distribution of error levels -> Clarks model
- Distribution is different in:
  - human-human dialogue
  - machine-human dialogue
  - Probably because the DS tasks are too simple
- It's important to measure their impact:
  - for instance, on task success
  - They do not always have an impact on user satisfaction
  - Can we learn to measure their impact from data?
  - Possible in task-oriented domains

Error prevention

- The dialogue system's prompts can trigger errors
  - Jose Luis example:

    DS: Are you looking for the telephone number of a person or a company?

    Client: Yes. (The DS expected "person" or "company")