

Toward Voice Query Clarification

Motivation

How to transfer the concept of query suggestions to a voice-only retrieval setting?

User-centric study

14 participants from a university background
 13 tasks each: find specific information by means of a tailored Amazon Alexa skill. Example task:

Scenario:

You want to surprise your Irish partner with an Irish cocktail called B-52, but you don't know how to make it.

Interaction start:

Alexa. Find! How to do a B-52?

Response methods

Direct Answer for one meaning

Concatenate Answer for each possible meaning

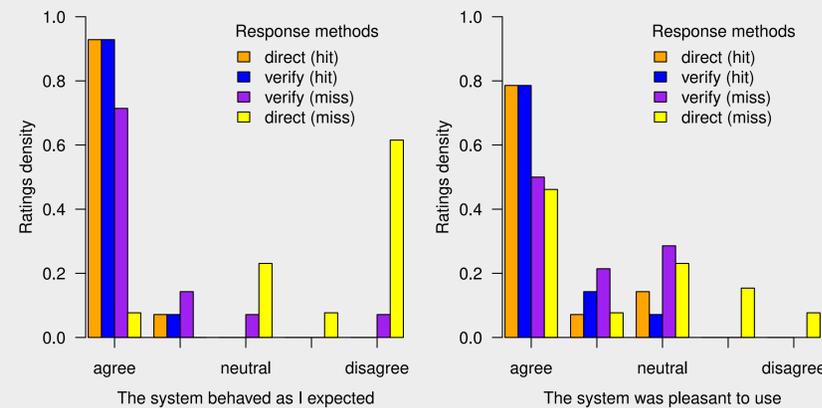
X-meanings Give 1–5 word descriptions for 3 or 5 meanings; user can choose one or ask for more

3-long-meanings Same with 8–16 words

Verify Ask if one specific meaning is meant; if no, continue like 3-meanings

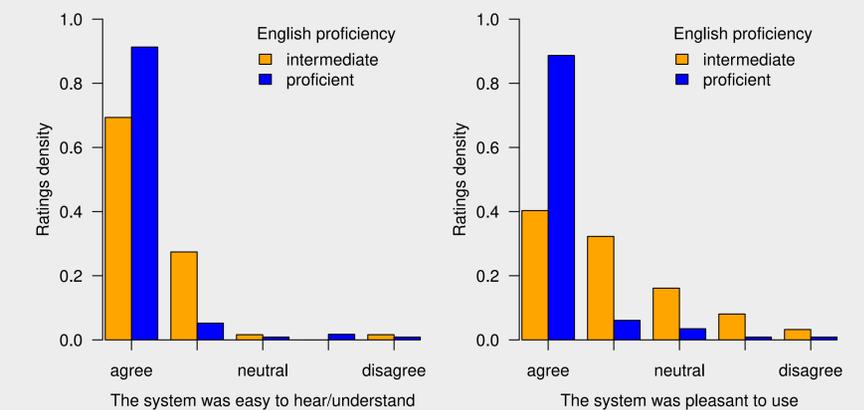
3-categories First give 3 meaning categories (like Wikipedia disambiguation pages); then continue like 3-meanings within selected category

How much does user satisfaction decrease when asked for clarification?



Alexa asking back does not decrease satisfaction

How does user experience vary for users with different background?



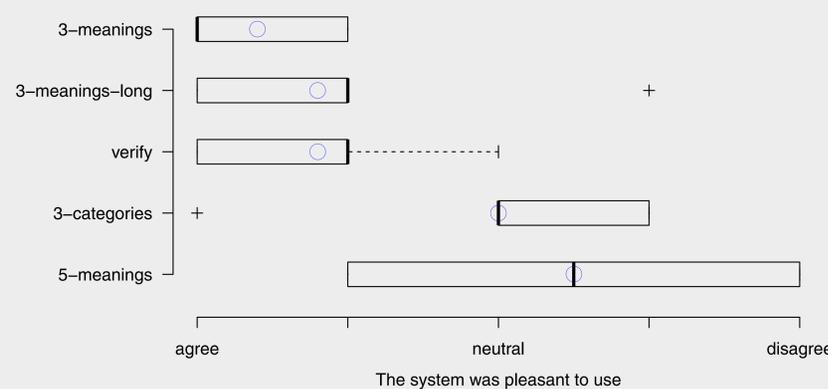
English proficiency strongly affects the experience

How should the different clarification options be presented?

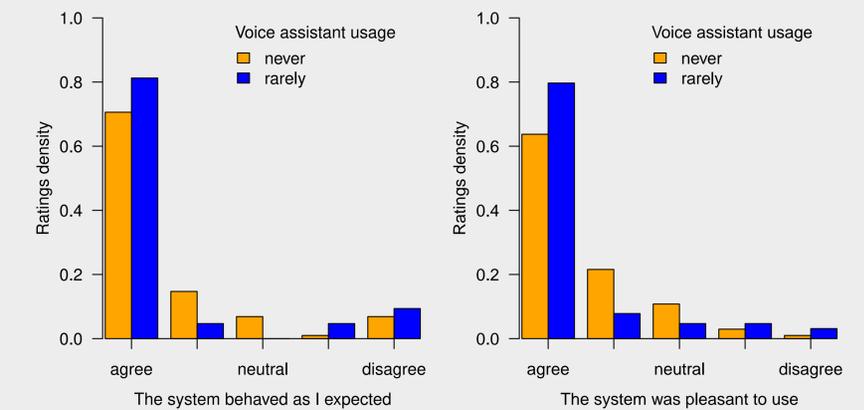
Clear differences by English proficiency

Proficient All methods rated very pleasant

Intermediate 3-meanings most pleasant:



Experimental protocol: participants directly say the intended meaning after 3 meanings are suggested



Prior voice assistant usage less so (still $p < 0.05$)

Conclusions

- Systems should ask for clarification if needed
- Systems should present 3 clarification choices
- Systems should allow users to clarify themselves
- Consider the language proficiency for evaluation