Archivus: A user performance analysis with speech, keyboard and mouse as interaction modalities

Marita Ailomaa
Artificial Intelligence Laboratory (LIA)
Ecole Polytechnique Fédérale de Lausanne (EPFL)
CH-1015 Lausanne, Switzerland

Agnes Lisowska
ISSCO/TIM/ETI,
University of Geneva
40 Blvd du Pont-D’Arve, 1211 Geneva, Switzerland

Introduction
Our poster presents part of an analysis of the results of a user experiment where the task was to answer questions about what happened in meetings using Archivus, a multimodal (mouse, keyboard and speech) dialogue system with a graphical user interface. The analysis presented here focuses on users’ performance on the tasks, and more specifically on the impact that the availability of different interaction modalities and the use of command-and-control actions made in natural language had on user performance.

The experiment
The experiment involved users trying to answer a series of true/false and short answer questions about meetings (Which two movies does Agnes suggest showing?, One of the meetings took place in Geneva.). Users were given 20 minutes to work through an ordered set of questions. On average, they were able to finish between 9 and 11 questions during that time. Each user did the experiment using one of the following modality combinations: mouse-only (M), mouse-keyboard (MK), mouse-voice (MV), voice-only (V), or mouse-voice-keyboard (MVK). It total, we had 8 users per modality combination.

It is important to note that the Archivus system is flexibly multimodal, which means that it allows users to perform the exact same actions using language (voice or keyboard) as they can with the mouse, e.g. “go to the next page”. This has particular relevance in the case of manipulating the graphical interface. However, the language modalities have the advantage of allowing users to skip steps that would be necessary with the mouse, particularly when specifying queries to the system, e.g. “Questions by Denis”.

Results
We compared task performance by adding up the scores of the solved questions in each modality condition. The result can be seen in Figure 1. Overall, MK and V users solved more questions and were more accurate than M, MV and MVK users.

Next we analyzed the relationship between task performance and use of language commands for manipulating the graphical interface, and we found tendencies showing that language commands decreased user performance (or had no impact on it). The users that scored highest (MK) made fewer language commands than the other users (see Figure 2).

For Q4 (Where was the Design meeting held?) MV users performed worse than other users, and they made much more language commands than the other groups. For Q6 on the other hand (Which two participants brought powerpoint presentations to the movie club meeting?), V users performed the worst. This question required a higher degree of
command-and-control actions, and V users may have had problems navigating through the GUI and finding the relevant items.

Figure 1. Task performance grouped by modality condition

Figure 2. Total number of language commands made for 10 different questions, grouped by modality condition

To conclude, the results indicate that natural language increases user performance in answering questions when using the Archivus system - if the users don’t (have to) make command-and-control actions. In future analyses we plan to focus on how natural language queries increase task performance.