1. Introduction

In the past few years, there has been an increasing interest in research on developing systems for efficient recording of and access to multimedia meeting data. This work often results in videos of meetings, transcripts, electronic copies of documents referenced, as well as annotations of various kinds on this data. In order to exploit this work, users need to have an interface that allows them to retrieve and browse the multimedia meeting data easily and efficiently.

When developing a language-interface for an interactive system in a new domain, the Wizard of Oz (WOz) methodology is a very useful tool. The user interacts with what they believe to be a fully automated system, when in fact another person, a 'wizard' is simulating the missing or incomplete modules, typically the speech recognition, natural language understanding and dialogue management modules.

2. Archivus

In our work we have developed a multimodal (voice, keyboard, mouse/pen) meeting browser, Archivus, whose purpose is to allow users to access multimedia meeting data in a way that is most natural to them. Since this is a new domain of interaction, users can be encouraged to try out and consistently use novel input modalities such as voice, including more complex natural language, and that in particular in this domain, such multimodal interaction can help the user find information more efficiently.

The WOz methodology is usually applied to unimodal (voice-only or keyboard-only) systems. In our case however, we have been developing a complex multimodal system and extended the methodology accordingly. We found that when the WOz methodology is extended to multimodal systems, the number of variables that have to be considered and controlled for in the experiment increases substantially.

Pilot experiments strongly suggested the use of two wizards – one supervising the user's input (Input Wizard) and the other controlling the natural language output of the system (Output Wizard). The role of the Input Wizard is to assure that the user's input (in any modality combination) is correctly conveyed to the system in the form of sets of semantic pairs. A semantic pair (SP) is a qualified piece of information that the dialogue system is able to understand. The role of the Output Wizard is to monitor and if necessary change the default prompts that are generated by the system. Changes are made for example in cases when they contribute to smoother dialogue flow, i.e. better explain the dialogue situation to the user or make the response more conversational.

3. Demonstration

In the demonstration we will show a methodology and the accompanying software that we have developed to help developers of complex language-enabled multimodal systems develop the natural language processing and dialogue management modules necessary. For the purposes of the demonstration, we will show how this methodology and software have been used in the process of developing the Archivus system, but the presented will also explain how both might be applied to other systems and domain under certain conditions.

The presenter will show a live demonstration of the Archivus system, where the audience will be able to see both the front end interface and the two back end Wizard interfaces. The presenter can then describe the functionality of the interfaces in detail, and answer any questions that the audience might have, showing concrete examples of interactions and natural language capabilities and their consequences on both the back and front-ends of the system.

If the audience is interested, a short video may be presented which shows the real conditions during the experiments with the system, including the equipment that cannot be brought to the demonstration session.