WebWOZ: A Wizard of Oz Prototyping Platform

Wizard of Oz (WOZ) is an evaluation method frequently used in dialogue design and user experience research. Its name stems from the famous novel ‘The Wonderful Wizard of Oz’ (Baum, 1900) in which a formidable wizards plays all kind of tricks to deceive his opponents and make them believe in his power. Like the wizard in the novel, a person running a WOZ experiment cheats, by telling test participants that they would interact with a computer system. In practice, however, they interact with a human who mostly sits in a different room and uses a 'wizard interface' as a communication channel. By doing so it is possible to explore ideas and concepts for technology before investing a considerable amount of money and resources building them. The WOZ method has been around for more than 40 years. Its first application was reported by Erdmann & Neal (1971) who tested the concept of a self-service airline ticket kiosk before Gould et al. (1983) used to explore their idea of ‘The Listing Typewriter’. From the very beginning WOZ was mainly used to explore speech-based interaction. Even though over time different exploration areas including multi-modal interaction (e.g. Salber & Coutaz, 1993) were added, the primary WOZ scenario is still around spoken dialogue. In order to run a WOZ experiment one needs an interaction interface for test participants and an additional interface that is operated by the ‘wizard’. Both interfaces are connected so that the wizard can influence the output on a participant’s screen. In addition it is mostly the case that a participant’s voice is transmitted via a separate audio channel.

Even though the principle setup of WOZ experiments is consistent, most of the interfaces built are used for one specific experiment only. The goal of WebWOZ is to deal with this issue and provide a tool for designers and researchers that lets them explore interactions beyond the frontier of one specific experiment. Rather they are able to easily create different experiments in order to test a variety of scenarios. In addition it is meant to extend the classic WOZ paradigm by offering a plug-able architecture that allows for the incorporation of language technology components as part of an experimental setting. The following will lead you through the set-up of a typical WOZ experiment and describe the different features of WebWOZ and how they are used.

Creating a New Experiment

After you are logged in to WebWOZ as a wizard you will see a drop-down menu that holds all your experiments as well as a button that lets you organise them. If you are logged in with a new wizard account the drop-down menu will be empty. *Organise Experiments* allows you to *Add New a Experiment* as well as to *Edit* and *Delete* existing ones (c.f. Illustration 1). To add a new experiment you need to focus on the bottom part of the dialogue and enter an *Experiment Name* and a *User Name* as well as a *Password* for the primary test participant you want to interact with. When running the experiment the test participant then needs to go to the WebWOZ client interface and log in with this user name and password. Additional test participants can be added after the experiment is created.

The check-box *Include Free-text tab* allows you to add a chat-like interaction channel to the experiment that offers more flexibility for the wizard. In order to create a realistic experiment set-up a wizard is typically obliged to choose from a pre-defined set of utterances. In certain settings, however, it can be desirable to allow for more flexibility. Therefore, by adding the free-text option, the wizard is not bound to the pre-defined utterances but is rather able to interact freely.

After clicking *Add Experiment* the new experiment is shown under *Your Existing Experiments*. Here if you click on *Edit* you are able to change the experiment name as well as the free-text option and add additional test participants. Via *Delete* you can delete an experiment. If you close the organise experiments dialogue you will find the newly created experiment(s) in the drop-down menu.

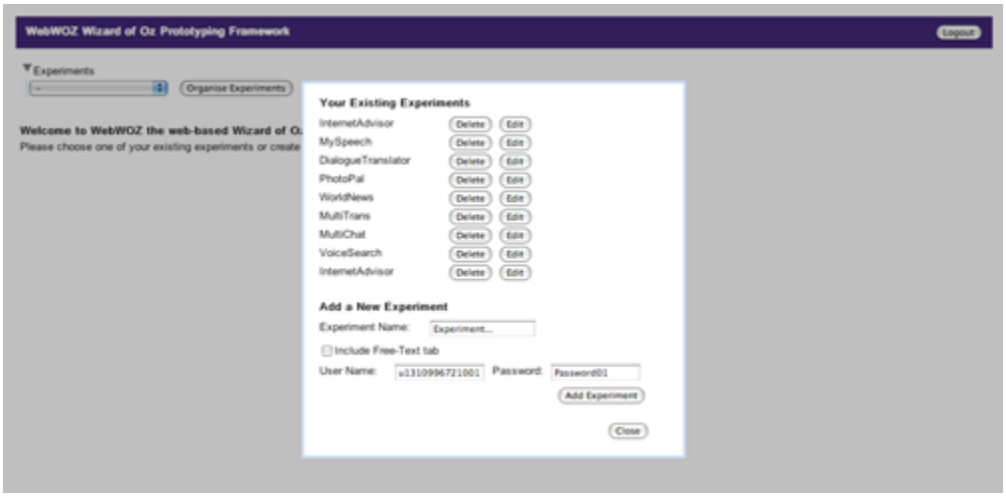

Illustration 1: Add and Edit WebWOZ Experiments

The Wizard Screen

After you choose a newly created experiment from the drop-down menu you arrive at the default wizard interface (c.f. Illustration 2). The interface is separated into different areas. On the very top there are two menus. One that allows you to switch to other *Experiments* as well as edit them and a second holding experiment related *Settings*. Both are flip-down menus highlighted by the little arrow next to hem. They appear if you click on them and disappear if you click again. Right below the menus you find a button *Enter Edit Mode* and a drop-down menu that lets you choose with which test participant you are interacting. Next to the drop-down menu you are informed about the test participant's status i.e. whether he or she is logged in, and and a button *Show Report* which holds the log file for the complete interaction with the selected test participant.

Below you find an area called *Sent Utterances* that displays a history of everything that is sent to a test participant throughout an active session. In case the experiment makes use of participant input this area additionally holds either text-input coming directly from the participant or the results of a used speech recognition module.

Next to the history there is an area dedicated to *Frequently Used Utterances*. Here the designer can add utterances that are used relatively often throughout an experiment and therefore can hardly be added to one dialogue stage only. In addition the area holds the *Processing...* button that can be used to notify a test participant that input is processed, which gives the wizard some time to search for the right response utterance or formulate an appropriate one.

The area below the history is dedicated to the running dialogue and therefore the main interaction area for the wizard. Tabs are used to allow for structuring the dialogue. By default there is one activated tab called *Stored Utterances* and two deactivated tabs, one called *Wizard Correction* and a second one called *N-best List*. The latter two will be activated in cases where the wizard’s task is to enhance the output of a language technology component rather than mimicking its functionality for which they take on a special role. The first tab, however, represents a standard tab that is used to organize a dialogue by representing a certain dialogue stage. It holds pre-defined utterances dedicated to this stage. Additional tabs can be added so that it is possible to group utterances based on their occurrence throughout the dialogue. Doing

so helps to easy the task of the wizard and decrease search time. In addition the designer can define instructions for a tab, a.k.a. a dialogue stage, which is meant to help the wizard in being consistent and following a certain protocol when using utterances.

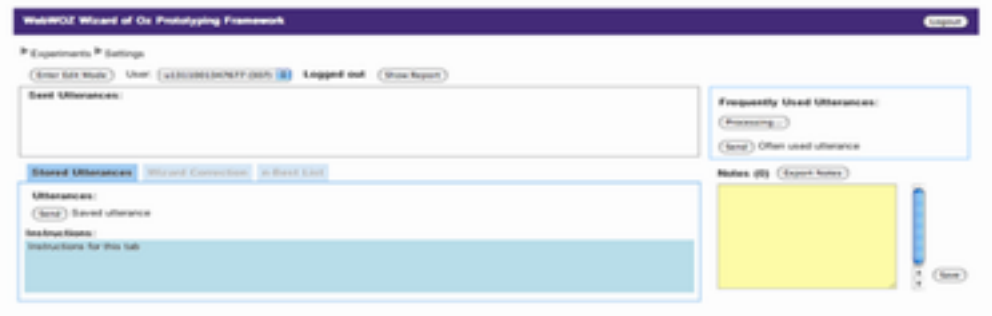

Illustration 2: Default Wizard Interface

Finally the bottom right corner of the interface allows for taking *Notes*. Every note is time stamped and added below, accessible via scrolling. In addition it is saved to a log file, which allows for keeping track of incidents happening during an experiment. The log file can be exported by clicking on *Export Notes*.

Add and Edit Utterances

In a usual WOZ experiment a wizard communicates with test participant by choosing from a pre-defined set of utterances. In order to add those utterances to an experiment please click on *Enter Edit Mode*. In edit mode an additional row of buttons is visible that lets you *Add*, *Edit* and *Delete Utterances* as well as *Tabs* (c.f. Illustration 3). When handling utterances WebWOZ distinguishes between two types. First ‘regular utterances’ for dialogue related interactions with a test participant. Second, for cases where the wizard mimics some sort of knowledge base and therefore needs to send utterances based on a certain set of domain data, WebWOZ integrates the concept of ‘domain utterances’. Domain utterances have the advantage that they are displayed in a separate area on top of the standard dialogue flow wherefore they are accessible without switching dialogue stages. In addition it is possible to add filter tags to them so that a wizard easily finds the right response by manipulating those filter options. The use of domain utterances will be discussed in more detail later in this document. For the moment, however, the focus lies on regular dialogue related utterances.

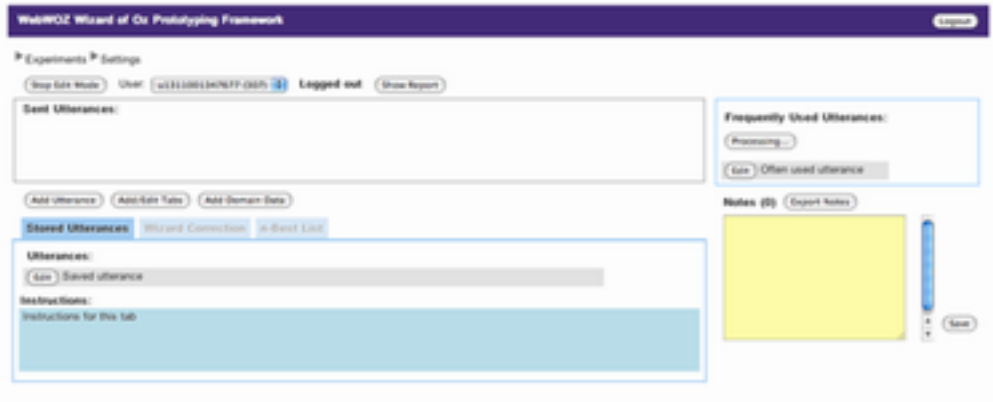

Illustration 3: WebWOZ Edit Mode

In order to add a regular dialogue utterance please click on *Add Utterance*. The following dialogue allows you to add a new utterance (c.f. Illustration 4). *Short Name / Label* lets you create a short description of the utterance. *Utterance* holds the actual utterance. With *Link to Audio File* and *Link to Video File* it is possible to connect the utterance with a prepared audio and video file, respectively. The files need to be located on an accessible server and the given link needs to be absolute (e.g. http:///link-to-file/filename.mp3). In terms of file formats it needs to be said that currently its support depends greatly on the web browser that is used on the participant's side. However, since WebWOZ is based on HTML 5, whose standardization is ongoing, the support of different media file types should be solved in the near future.

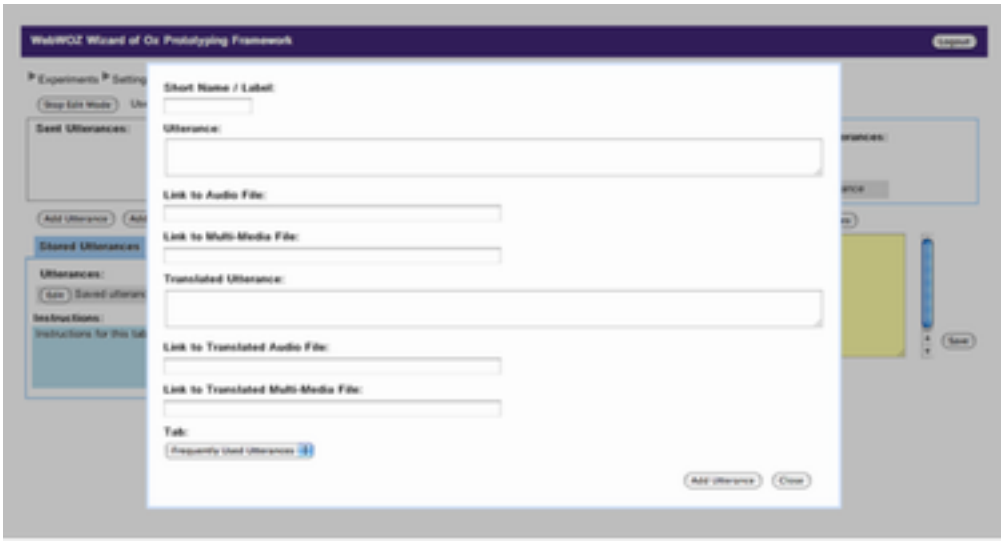

Illustration 4: Add Utterance

In case the designed experiment is based purely on text or makes use of an on-the-fly text-to-speech synthesis the fields holding the links to audio and video files can remain empty. In order to support multi-lingual experiments WebWOZ further allows you to specify a translation for your utterance including its connection to media files. Also here, if no translation is needed or an on-the-fly machine translation service is used, those fields can remain empty. Finally, the tab to which the utterance should be added, can be selected from a drop-down menu at the bottom of the dialogue. With a new experiment only two options are available. On the one hand *Frequently Used Utterances*, which describes the area next to the history, and on the other hand *Stored Utterances* standing for the first tab in the dialogue area. The following section will show you how you can add additional tabs or edit existing ones, so that it is possible to give more structure to the dialogue. For the moment you can just select one of the two existing options and click on *Add Utterance*, so that the utterance is added to the selected tab. In order to edit an utterance you need to click on the *Edit* button next to it. It will open a dialogue which lets you change all it's attributes as well as allows you to *Delete* it. In order to use an utterance, i.e. send it to a test participant, you need to exit edit mode by clicking on *Stop Edit Mode*.

Add and Edit Tabs

In order to organize utterances WebWOZ allows you add tabs. A new tab can be created entering the edit mode and clicking on *Add / Edit Tabs*. The following dialogue lets you edit and delete existing tabs as well as add new tabs (c.f. Illustration 5). In terms of editing you are able to change the name of the tab and add some instructions for the wizard, which will be displayed in the bottom. If the instructions field is empty, the box usually holding instructions is hidden, giving more space to utterances. With exception of the first tab all other tabs that are created can also be deleted, given they do not contain any utterances.

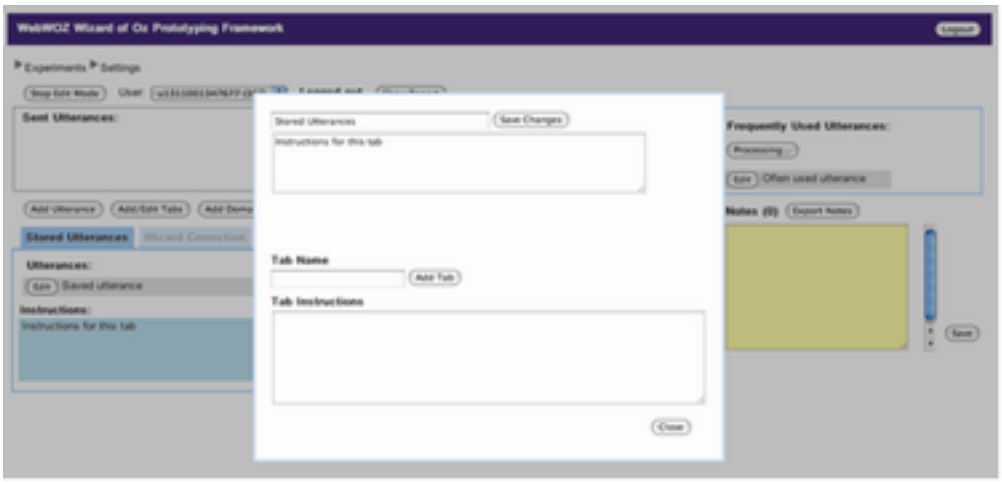

Illustration 5: Add Tabs

If you want to delete a tab that contains utterances, you first need to delete those utterances or move them to a different tab. In order to add a tab enter a *Tab Name* and optional *Tab Instructions* and click on *Add Tab*. If you Close the dialogue you will see that the tab was added to the dialogue structure. Now utterances can be added to this tab or moved to it from others.

Frequently Used Utterances

Frequently used utterances are utterances that a wizard needs to use on an irregular basis and which therefore are hard to add to only one specific dialogue tab. They are shown in a separate area on the screen (c.f. Illustration 6). Typically those utterances include ways of error recovery like “Sorry, but I did not understand you. Could you please repeat?” or gab fillers like “Please hold!”. Those utterances are treated like general utterances and can therefore be added in edit mode using the *Add Utterance* button and edited using the *Edit* button next to them. In addition the area for frequently used utterances holds a button called *Processing...*. A wizard can use this button to notify a participant that the computer is processing the last request. On the participants screen this will display the utterance Processing... followed by an additional dot every second to signalise an ongoing working process. Doing so gives the wizard some time to search for the right utterance and at the same time shows the test participant that the system is responsive. As soon as the wizard sends the next utterance the Processing... is reset so that the next time it is used it starts again with three dots at the end.

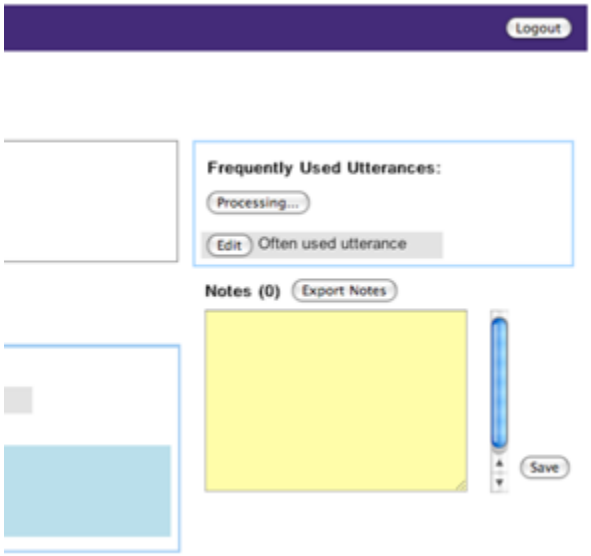

Illustration 6: Frequently Used Utterances

Taking Notes

WOZ is an evaluation method and therefore one often wants to make notes throughout an experiment and later revisit them in order to better understand a given situation. In order to support this process WebWOZ offers a notes functionality that allows the wizard to write a note within the wizard interface. Notes are time stamped and saved to a database so that they can be exported later on (c.f. Illustration 7).

Illustration 7: Taking Notes

Settings

The settings are accessible via the flip-down menu on top of the wizard screen. Here it is possible to configure the interaction pipeline between the wizard and the test participants (c.f. Illustration 8).

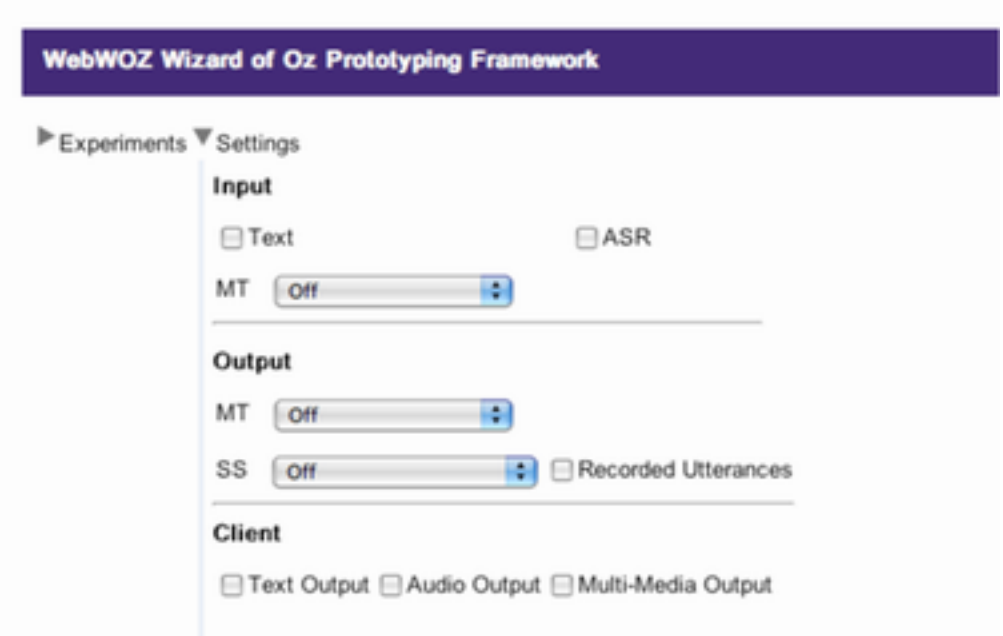

Illustration 8: Settings

Input

Input defines how the test participant interacts with the system. If the Text box is checked the participant will be able to use a chat-like input window to interact with the system. *ASR* activates speech recognition for the participant. When activated the recognition language can be defined and furthermore it is possible to activate a *Correction Mode* which is used in experiments where the wizard rather enhances the output of language technology component than replaces it as a whole (e.g. the wizard corrects the output of the speech recognizer). Correction includes *Wizard Correction* which lets the wizard amend what is recognized and *N-best List* which allows to choose from a list of possible recognition results. Depending on which option is selected, the output arrives at the dedicated tab within the dialogue structure.

Both, text input and speech recognition can be activated at the same time, even though the use of one only modality might be more likely. In cases where the experiment tries to explore a multi-lingual setting it is further possible to push the client input, whether text based or recognized by the speech recognizer, through an on-the-fly machine translation (MT) system before it arrives at the wizard. For this purpose WebWOZ incorporates two external *MT*

systems. If a MT system on the input side is selected it is possible to define source as well as target language. Also here a *Correction Mode* can be activated, which works the same ways as already explained in connection with the ASR option. However, a correction can only happen at the last stage of an input. Hence, if an MT system is used the correction mode for ASR, if set, will be deactivated. On the other hand, if a correction mode for ASR is activated, it is not possible to use MT on the input side.

Even though WebWOZ supports text as well as speech recognized input by participants, it is often the case with WOZ experiments, that advanced input modalities are evaluated. For example a speech recognizer that goes beyond the rather limited functionality offered by the one used in WebWOZ, or alternatively some sort of multi-modal interaction. In this case it is possible that neither the *Text* nor the *ASR* box is checked for which the experiment set-up needs to cater for an alternative method of transferring input from a participant to the wizard. In most cases this can be achieved via some sort of Voice-Over-IP system in combination with live video streams.

Output

The *Output* settings concern everything that goes from the wizard to the test participant. Similar to the input side it is here possible to send the utterance first through an MT system before it arrives at the participant. In addition it is possible to send it to an on-the-fly text-to-speech synthesizer (TTS). WebWOZ integrates two different text-to-speech systems. Both, however, have still experimental status for which their functionality can not be guaranteed. Even though on-the-fly MT as well as TTS is supported, in most WOZ experiments a researcher might prefer a more controlled set-up. Therefore it is possible to use pre-recorded utterances, either based on audio or based on multi-media files, both stored on an external server. So if the *Recorded Utterances* box is checked, neither TTS nor MT is available on the output side. Rather the pre-recorded files that are connected to an utterance (c.f. Add and Edit Utterances) will be played to a participant.

Client

Finally, the *Client* settings define what actually arrives at a test participant. Options consist of *Text Output*, which is basically the utterance in pure text form, *Audio Output* either from an on-

the-fly TTS or pre-recorded, and *Multi-Media Output* coming from a connected multi-media file or a special purpose web-service producing on-the-fly multi-media content. Options can be used in different combination. Note, however, that if none of the options is selected, nothing will arrive at the participant.

Domain Utterances

As already mentioned earlier in some WOZ experiments the wizard mimics a certain knowledge. For example an information system whose final recommendation depends on different parameters. In order to help the wizard run such an experiment, WebWOZ incorporates the concept of 'domain utterances'. Domain utterances are special utterances that serve the purpose of giving a recommendation. In order to arrive at a final recommendation it is possible to add different filters and values to domain utterances. Important here is that all the domain utterances of an experiment share the same filters. Therefore, filter values need to be set for all of them.

Domain utterances can be created in edit mode by clicking on *Add Domain Data.* The following dialogue is similar to the one discussed for adding regular utterances. One difference, however, is that it does not allow for defining a specific tab the utterance should go into. That is because domain utterances are situated in a separate area on top of the actual dialogue structure. After adding a domain utterance this additional area appears. Also, an additional button to add filters is now visible (c.f. Illustration 9). The new area consists of several tabs which are active or inactive depending on the experiments settings. In addition to domain data it also holds filters for domain utterances and the option to use free text to interact with a test subject. The free text option can be activated in the general experiment settings via *Organise Experiments* which can be found in the *Experiments* flip-down menu. Filters can be added through clicking on the *Add Filter* button. If non of these elements is used within an experiment, the whole area disappears.

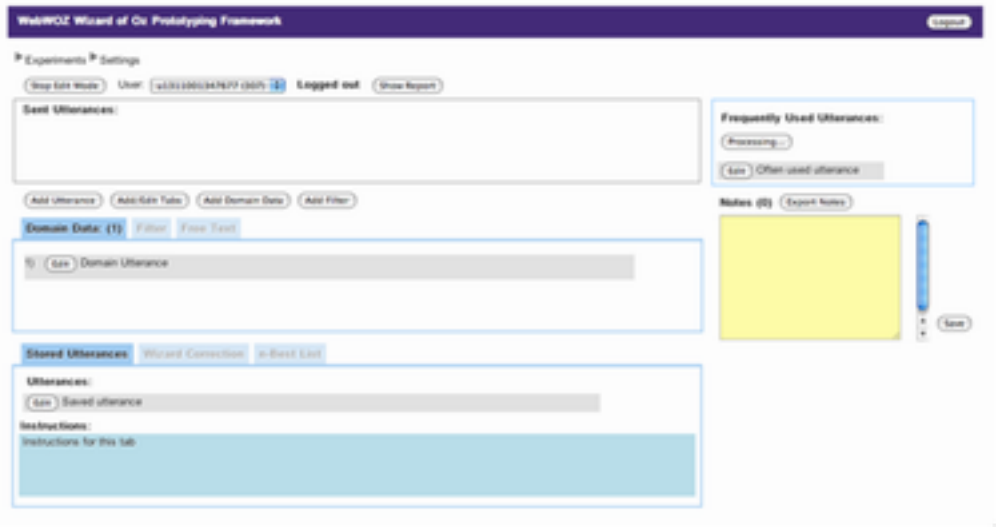

Illustration 9: Area for Domain Utterances

Adding Filters

In order to make it easier for wizards to search through domain utterances WebWOZ allows to add different filters to them. A filter can be added in edit mode by clicking on *Add Filter* (cf. Illustration 10). Note, however, that at least one domain utterance needs to be created for this button to be visible. A *Filter Name* needs to be specified as well as a *Default Value* which is added to all the existing domain utterances as well as to the ones that are created afterwards. In addition to the default value additional filter values can be specified.

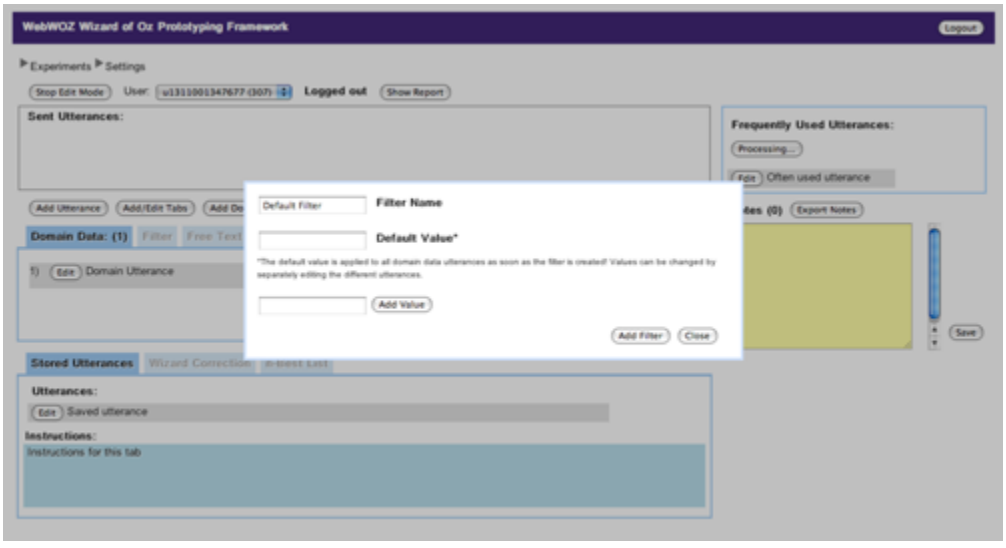

Illustration 10: Add Filters

In theory it is possible to define unlimited values as well as it is possible to define unlimited filters. However, the limited screen space might make it difficult to handle them.

After adding a filter it can be deleted as well as edited by clicking on *Edit* next to it. Also, now it is possible to set the filter values for existing domain utterances by clicking on *Edit* next them. The following dialogue will then have a drop-down menu for all the filters that were created holding the different filter values (c.f. Illustration 11).

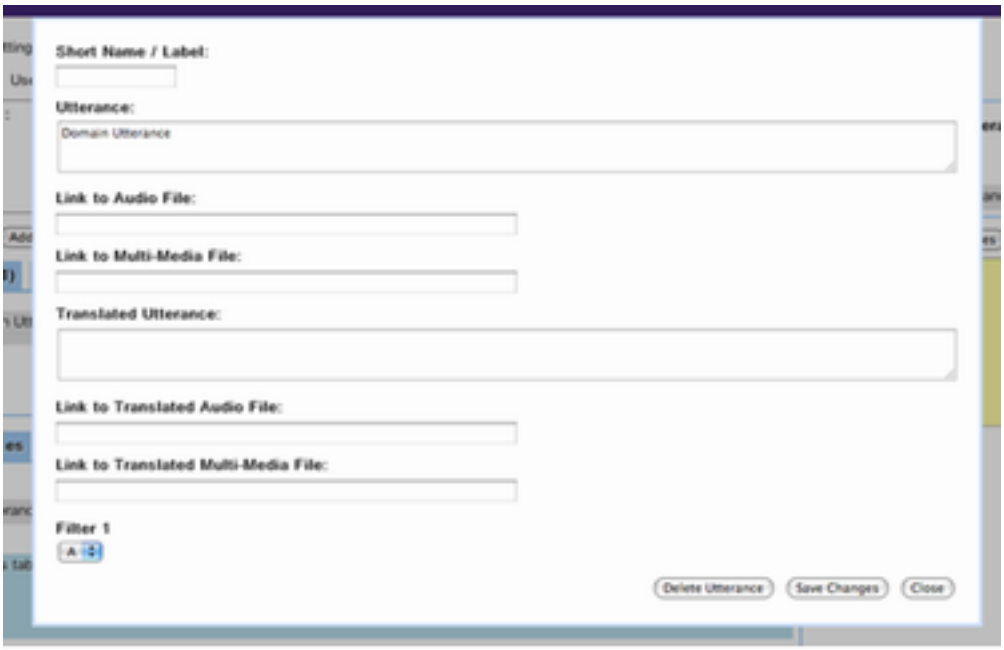

Illustration 11: Setting Filter Values for Domain Utterances

Without changes a domain utterance always uses the *Default Value* for filters. By adapting those values it is however possible to help the wizard searching for domain utterances. So is it possible to reduce the amount of domain utterances by adapting the filters (c.f. Illustration 12 & 13).

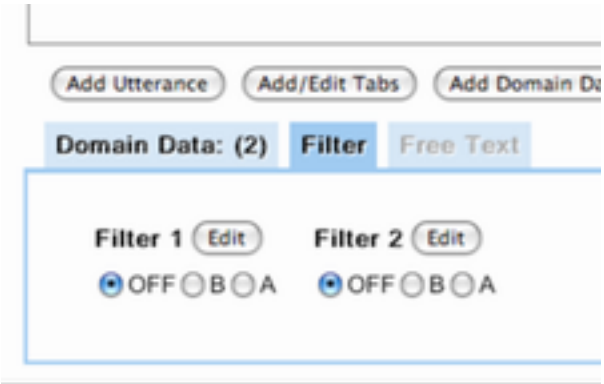

Illustration 12: No Filter Values Selected

Illustration 13: Reduced Amount of Domain Utterances through filtering

Free Text

In order have more flexibility when interacting with a test participant an experiment designer can activate the free text option. This can be done through selecting the I*nclude Free-Text tab* option in the general experiment settings. In order to so one needs to click on the *Organise Experiments* button which can be found under the *Experiments* flip-down menu. Next to any of *Your Existing Experiments* an *Edit* button can be found which leads to the dialogue for changing the general experiment settings (c.f Illustration 14 & 15). Here it is possible to change the name of the experiment, add additional test participants and adapt their login parameter, as well as setting the free text option. If the box next to Include Free-Text tab is checked and the changes saved, the dedicated tab in the wizard interface is activated (c.f. Illustration 16). Now, it is possible for a wizard to interact freely with a test participant. However, this sort of interaction is only advised for experiments that make use of on-the-fly services for TTS and MT or focus on a pure text based interaction, as no media files or translations can be prepared.

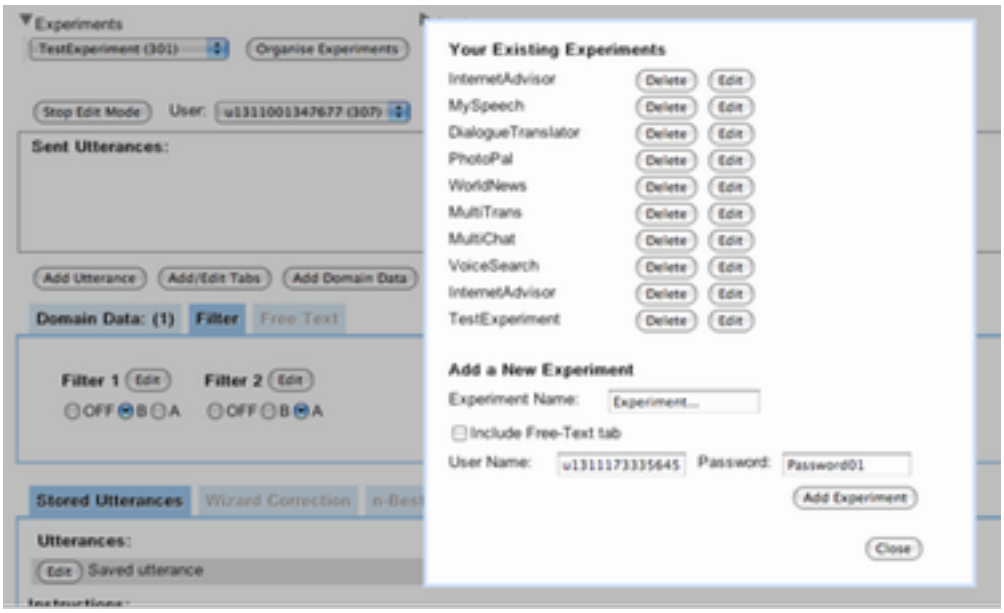

Illustration 14: Edit Experiment Settings

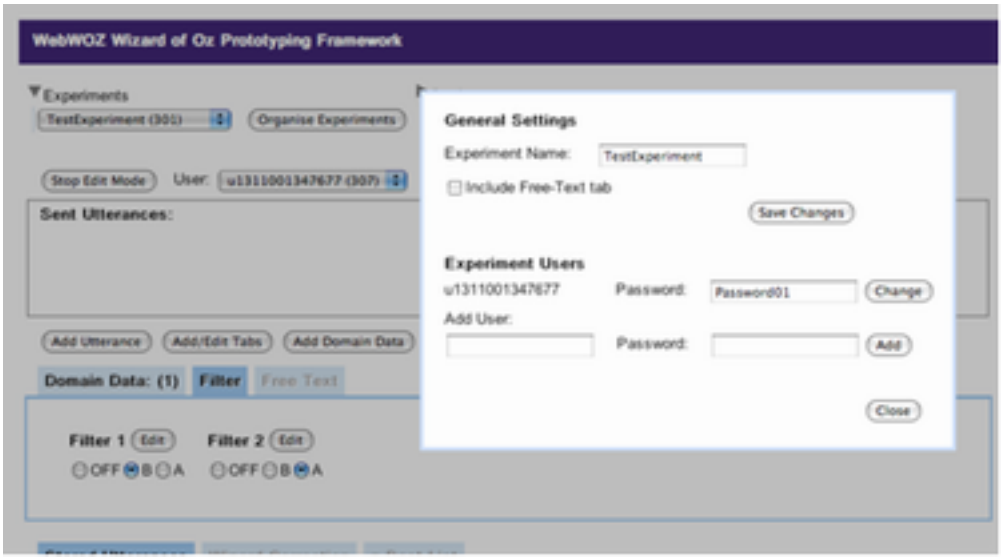

Illustration 15: General Experiment Settings

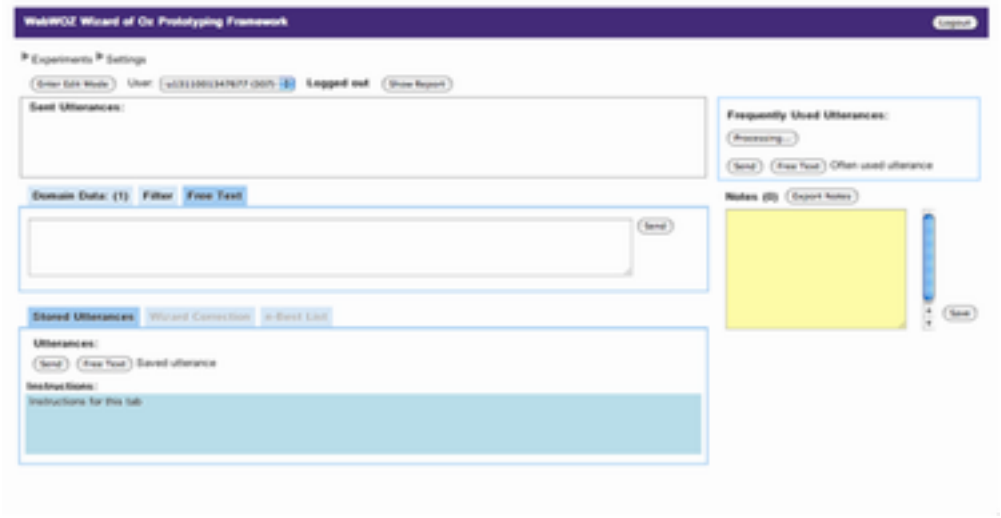

Illustration 16: Activated Free Text Interaction

With the Free Text option an additional function of WebWOZ is activated. As an experiment sometimes requires the adaption or concatenation of prepared utterances it is possible to use the *Free Text* button to send an utterance to the free text field instead of sending it directly to a test participant. There it can be amended or another utterance can be added before it is send on (c.f. Illustration 17).

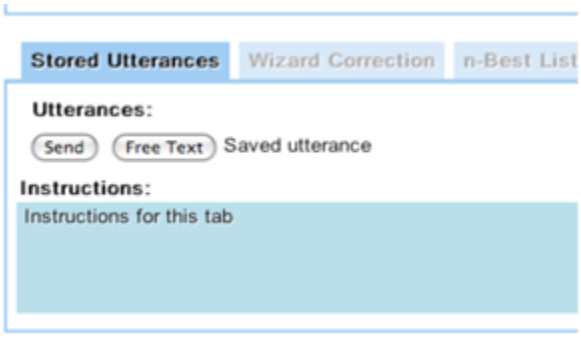

Illustration 17: Sending Utterances to the Free Text Field

References

- Baum, L. F. (1900). The Wonderful Wizard of Oz. Hill.
- Erdmann, R. L., & Neal, A. S. (1971). Laboratory vs. Field Experimentation in Human Factors: An Evaluation of an Experimental Self-service Airline Ticket Vendor, Human Factors 13, 521-531.
- Gould, J. D., Conti, J., & Hovanyecz, T. (1983). Composing Letters with a Simulated Listening Typewriter. Communications of the ACM 26(4), 295-308.