

Using Amanda Portal's Speech-Enabled Auto-Attendant

April 28, 2003

1 Introduction

There are six steps which you should take to configure Amanda Portal to use its LumenVox-powered speech-based auto attendant:

1. Install the LumenVox software. This third-party software package performs the speech recognition function and tells Amanda what it believes that a caller has said. The installation procedure involves supplying an activation key which will enable a set maximum number of speech recognition ports. Installation of the LumenVox software is discussed in *Installing Amanda Portal*, so it will not be discussed further in this document.
2. When doing the initial configuration of Amanda using its Configuration Wizard, specify either Dialogic or GlobalCall drivers for your telephone ports—Brooktrout (Rhetorex) boards are not currently supported. The telephony boards that you use must have Dialogic's Continuous Speech Processing ("CSP") feature enabled, which may involve loading an alternate firmware load. Also, be sure to specify that the LumenVox option is selected. If you will be using text-to-speech, also enable the engine you have chosen for this function.
3. Add two items to Amanda's configuration to define the new application.
4. (Optional) Define and record the greetings that the system will play to callers.
5. Configure mailbox 990, and optionally mailbox 991 as well, to use the speech recognition application instead of the default DTMF-based one.
6. Define the name(s) to be associated with mailboxes. If necessary, specify the phonetic spelling of these names.

2 How the Auto-Attendant Application Works

Dialogic boards with the CSP feature are able to listen for your speech at the same time they're playing a greeting, message, or prompt to you. In order to do so, they have to cancel the echo of what is being played, or they would hear the playback as sound coming from the caller. As soon as they hear some sound other than the playback's echo, they can stop playing and record what the caller is saying. This process, of the caller speaking while a prompt is being played and stopping the playback, is called *barge-in*. That small portion of what the caller said in the time before the playback stops is not lost. The driver automatically saves it in a special buffer so that the entire utterance of the caller is available to be analyzed by the speech recognition engine.

Amanda must be prepared for barge-in to happen. When Amanda is playing back something normal, such as a prompt or greeting, Amanda is not programmed to stop the playback when you speak, though playback will usually stop (depending on the context) if you dial a DTMF digit.

So when Amanda is acting as a speech-recognition-based auto-attendant, the system plays prompts while waiting for barge-in to occur. When Amanda detects that you have said something, playback stops and recording starts. When you finish speaking, Amanda then passes what was recorded to the speech recognition engine. It uses a *grammar* to know what words you may have said, and it makes its best guess based on analyzing the recording. The speech engine also assigns a *confidence* value to its analysis: this is a number from 0 to 1000 indicating how sure the engine is that its analysis is correct.

If nothing was recorded, Amanda will just repeat its prompt. If you said something, but the speech recognition engine does not think that what you said is in the grammar, then Amanda will say "I'm sorry, I don't recognize that name," and then repeat the prompt. Sometimes, if you say something not in the grammar, the engine will mistakenly believe that it has recognized what you said as something else. Fortunately, the confidence score in these cases will be very low, so Amanda will be alerted to confirm the engine's understanding.

On the other hand, if the confidence in what the speech recognition engine understood is high, and the name which the engine recognized identifies a single mailbox, then Amanda will immediately transfer the call to the mailbox associated with that name.

Otherwise, if the confidence is medium or the name recognized does not uniquely identify a single mailbox, then for each mailbox, Amanda will ask the caller to say whether the mailbox is the correct one. If the caller says Yes, then the call is transferred. Otherwise, if there are more mailboxes, Amanda will ask about the next one.

By default, when Amanda asks the caller whether a particular mailbox is correct, it plays a pair of prompts with the mailbox's identification spoken inbetween, and then waits for a Yes or No response. If greetings 3 and 4 are recorded for the current box, then they will be substituted for these two system prompts. This allows per-mailbox

override of the spoken language for interacting with the caller.

The speech-recognition auto-attendant application also recognizes DTMF input. If a caller starts to dial, then the system behaves the same as non-speech-based Amanda systems.

If the mailbox contains an initial greeting (the current greeting setting of the mailbox) and if the global configuration parameter `no_first_barge` is set to true, then when Amanda first answers the phone, barge-in will not be allowed until after the first greeting has been played. This allows callers who call from noisy environments to hear something before the barge-in occurs. Otherwise, if it occurred right away, the caller would not hear anything useful from Amanda and so would not know what to do.

3 Additional Configuration Requirements

After you have run the Configuration Wizard, there are two changes that you need to make to Amanda's configuration database. You can make these changes by running Amanda's Configure command from the Windows Start menu.

The first parameter you will modify is called `tcl_source_files`. In addition to the filenames listed, you should add, on the end of the list, the name `lumena.tcl`. This file defines the LumenVox-based auto-attendant application.

The other parameter to modify is called `tcl_start_files`. This setting will usually be empty initially. You should set it to `startl.tcl`.

4 Greetings

The speech-based auto-attendant application can use as many as three different greetings from each mailbox that is configured to use it (as discussed in Section 5, below).

The mailbox's current greeting is played first. It is played only once, unless greeting 7 has not been recorded. In mailbox 990, which is used to answer new incoming calls, this greeting usually says something like "Thank you for calling ABC Corporation." This greeting will usually be greeting 1 or 2 in the mailbox, and it may be changed by the auto-scheduler.

Greeting 7 prompts the caller to say something, and it is played right after the current greeting, or whenever the system is back at the top of its "loop." The default greeting 7 for mailbox 990 is recorded as "How may I direct your call?" There is no default greeting 7 for mailbox 991. If greeting 7 is not recorded in a given box which is using this application, then the current greeting will be used as the "instructions" prompt for the caller in place of greeting 7.

Greeting 6 gives the caller additional information if the caller does not say anything for several seconds. By default in mailbox 990 it says “You may say the name of the person or department you wish to reach.” If greeting 6 is not recorded, as is the default in mailbox 991, then not additional prompting will be used, and only greeting 7 (or the current greeting, if greeting 7 is not available) will be played over and over to the caller. The configuration parameter `max_prompt` is used to control the number of times that the application tries to get the caller to say a recognized name before giving up.

Because the speech recognition engine uses Mulaw recordings, and the Dialogic CSP algorithm works best when the format being played is the same as the format being recorded, you should always record your greetings which may be barged-in-upon in Mulaw format. This is the default format in which Amanda records greetings.

5 Configuring Mailboxes 990 and 991

Every Amanda Portal mailbox can be configured as to what actions it performs when a caller executes that mailbox, or, separately, what should happen when a user logs into that same mailbox. These mailbox parameters are called `CALLER_CODE` and `USER_CODE`, respectively. Their settings correspond to classes of service for the mailbox.

To enable Amanda to act as a speech-recognition auto-attendant, you must change at least mailbox 990 from using a DTMF-based behavior to using the speech-enabled behavior. As you know, by default incoming calls are directed automatically to mailbox 990 (though you can change this on a per-port basis via the `box_grt` parameter). When Amanda comes to mailbox 990, it will execute whatever is defined to be that mailbox’s `CALLER_CODE`. You can change this using the Amanda Station client program. After logging in as the administrator (mailbox 999), highlight mailbox 990, then double-click on the “Method Information” icon on the right. The first tab that will be open will be the one entitled “Caller Code” and the value shown will probably be `method_dos_caller`, which is a function which behaves the same as other Amanda (originally DOS-based) products. You must change this to `method_lumen_aa`, then click OK to save the change.

If a caller cannot be understood by the speech recognition engine, or the caller does not understand what the prompts are telling him to do, then after a few tries (`max_prompt` plus 1), this function will exit which will cause Amanda to execute the `DONE_CODE` for the mailbox. By default, this will switch processing to mailbox 991, whose default greeting says “To reach the person you are calling, enter their extension. For an employee directory, enter 411. To speak to the operator, press 0 or stay on the line.” In other words, it invites the caller to dial using traditional DTMF input. You may elect to leave mailbox 991 configured just as it is, or you may want to change its `CALLER_CODE` also to `method_lumen_aa`. Without changing the greeting/prompt, the caller will be informed that he should dial a mailbox number. In actuality, with this change, he can still speak a name associated with a mailbox (just as he could have dialed a mailbox number even when processing was on mailbox 990).

6 Defining Names for Mailboxes

The auto-attendant application builds a grammar based on the name(s) in the 411 directory that you define for each mailbox. This happens automatically when the system comes up, and each time you add, modify, or delete any mailbox's directory entry. You may assign as many names to each mailbox as you like.

It is best to use both first and last names for recognition. The longer the phrase that the speech recognition engine can match to the recording, the more sure the engine will be that it has it right. Furthermore, by using both first and last names, it is unlikely that any names will not map one-to-one to a mailbox number. If your company has both John Smith and John Jones, then if you allow callers to say simply "John," no matter how sure Amanda is that the caller said "John," Amanda will still have to ask the caller if they meant John Smith or John Jones. Also, the speed at which the speech recognition engine can analyze a recording is related to the number of words in the grammar (the number of different names you have defined). Listing a first, last, and first+last name for each mailbox triples the number of names to be recognized, compared to listing only first+last names. So if your company has a large number of employees, it becomes mandatory to list only first+last names, while smaller companies may list all the possible variations.

Using the Amanda Portal Station client, highlight a mailbox and double-click on the "General Information" icon. Be sure that each mailbox has a name filled into the Name field at the top of this dialog. If that user does not have a name and extension recording, then the text will be used via text-to-speech (if you have that option enabled) to speak the name of the mailbox. If neither is available, then Amanda can identify the mailbox only by its number, which will not be very helpful to callers.

Next, in the same dialog, use the "Add Name..." button to add one or more names to be associated with this mailbox. *Do not spell the names phonetically or use sound-alike spellings.* You should spell each name just as that person actually spells the name. The next section discusses how the speech recognition engine understands the names and how you can modify this behavior. Of course, you may also enter other words as names, such as "Accounting," "Sales," etc. Associate these names with mailboxes just as you would any other name.

7 Phonetic Spelling of Names

The LumenVox speech recognition engine has a dictionary of approximately 120,000 names which are commonly found in the United States. For each name in the dictionary, there are one or more phonetic pronunciations of that name in the international phonetic alphabet. You will find that this dictionary already contains the vast majority of the names that you will be assigning to mailboxes, so you will not need to take any

further action—just add the name to its mailbox and you’re finished. To check the pronunciation(s) for a given name, run LumenVox’s PhoneticSpeller program. Then just type in the name and the program will display the phonetic spelling(s) for that name that the speech engine will be listening for. Do not use things like ‘op per a der’ or some such, because (1) there are perfectly good (correct) alternatives in the dictionary, which will be used (punch them into PhoneticSpeller and see), and (2) adding spurious stuff hurts the recognizer. So, if the speller is okay, than use it, but for anything that isn’t within one or (at most) two consonants of being correct, try spelling it phonetically instead. In the case of vowels, unless the opposition is a full vowel versus IX or AX, spell it phonetically if the speller isn’t dead on, because vowels are critical to the recognition process. In particular, the insertion or deletion of a vowel compared to the correct pronunciation of a word or name will very likely cause the recognizer to miss that word.

If a name is not found in the dictionary, then the LumenVox software will generate its best guess as to the pronunciation of the name. Although this usually works fairly well for common English words, it does not work for names very well because names do not necessarily follow American English pronunciation rules, and because names not in the dictionary are likely to be foreign, and thus even less likely to follow the pronunciation rules.

If the pronunciation that is shown is not correct for this name, you can tell Amanda the correct pronunciation by specifying its international phonetic spelling. To do so, when you are in the dialog to add a name, you will see a button labeled “Phonics...” Pressing that button will bring up a subordinate dialog box which will allow you to spell out the name phonetically, and if your computer is equipped with a soundcard, speakers, and Microsoft’s SAPI 5.x software, then the client can speak out the indicated pronunciation so that you can verify that it is correct. The dialog includes a list of the phonemes used for American English. Double-clicking on one will add it to the phonetic spelling that is building up in the bottom of the dialog. Clicking the “Speak” button causes the client to try to use SAPI to speak the phonetic pronunciation. When you are satisfied as to the pronunciation, click OK to save the changes and close the dialog boxes. Doing so will automatically update the Amanda server to “listen for” the indicated pronunciation whenever that particular name is associated with a mailbox.