



## White Paper: Contact Center Automation

# How to Build a Speech Recognition Application

### Abstract

Using speech recognition to automate many of your call center functions can provide many benefits. However, the process of building a successful speech recognition application is complex. This paper provides a detailed description of the steps required to create a speech recognition application including: requirements definition, dialogue design, application design, application development, testing, deployment and tuning.

### Introduction

You've done your homework, and you already know the benefits that an automated speech recognition application can add to your call center:

- ✓ Increased automation – you'll free up your agents those routine tasks and let them handle those interactions that really require the human touch
- ✓ Out-of-hours coverage
- ✓ Greater customer satisfaction – you'll decrease the amount of time customers have to wait on queue for an agent
- ✓ Large return on investment – savings in agent and port costs will pay for the cost of automating in less than a year

So now what? What does it really require to build a successful speech recognition application?

### Building a Speech Recognition Application

The process of building a successful speech recognition application involves the following steps:

#### Requirements Definition

Defining your application requirements is the first step in creating a speech application. During this stage, define your business goals and identify your customer needs. Decide at a high-level what your application must do to achieve these.

- ✓ Which functions within your call-center will you automate, and which will remain with an agent?
- ✓ What age, gender, and geographical area do you expect most of your callers to come from?
- ✓ How often do you expect them to call your application?
- ✓ What are your business goals – greater automation, increased customer satisfaction, both?
- ✓ Are there any legal restrictions on the application?
- ✓ What is your expected call-volume?

These are some of the issues that must be addressed during the requirements definition stage.



## Dialog Design

Your speech application should be “user-friendly”: natural sounding, easy-to-use, and an overall enjoyable experience for the caller. Dialog design is the process used to attain these goals. The steps required are:

1. Determine the general character or “*persona*” of your automated “agent”. In other words, how will the caller perceive the application – friendly, unfriendly, helpful, unhelpful, knowledgeable, unknowledgeable?
2. Create a *call-flow* to describe the order in which events (messages, database accesses, transfers) occur in the application.
3. Pick confirmation, error recovery, and help strategies. The strategies selected should take into account the application requirements, as well as, best practices for voice user interface design.
4. Design the *prompts*, that is, the messages that callers will hear. Prompt design includes the precise wording, as well as, the placement of pauses, and the use of proper inflection to obtain the greatest naturalness and clarity of meaning.
5. Design the *grammars*. For each prompt, the grammar describes the set of all possible responses the caller can give to that prompt. More precisely, the grammar describes the set of *expected* responses from the caller (because in fact, the caller can and sometimes will respond in unexpected ways that make little sense). The job of the grammar designer is to use knowledge of caller behavior to decide what the most likely caller responses are.

## Application Design

The issues addressed in the design phase of a speech recognition application are similar to those addressed when designing any piece of software. Additionally, the application design will incorporate the dialog design created earlier.

## Application Development

There are three major components to your application that require development. They are:

- ✓ The application logic. The application might be developed in VoiceXML, a standard language developed by the World Wide Web Consortium. Or, depending upon the platform upon which your application will run, a proprietary language developed specifically for that platform might be used.
- ✓ The grammars. The language or tool used to develop the grammars will also depend upon the software platform and/or recognition software used.
- ✓ The audio prompts. Natural sounding, clear prompts require more than a tape recorder placed near voice talent while he/she reads a list of prompts. To ensure that the correct mood and meaning of each prompt is conveyed to your callers, the voice talent should be coached. This will allow the intentions of the dialog designer to be communicated to the voice talent. The context of each prompt within the application and the overall persona of the entire application are important points to include in the coaching and recording session.



## Testing

Once the individual components of the application have been developed, they are assembled and the application is tested. The testing should include:

- ✓ The application logic.
- ✓ The grammars. Check that each grammar is named correctly. Also, that each grammar contains the phrases it was designed to contain and excludes extra phrases. Also, testing should verify that the grammars return the correct values to the application. Most recognition engines include text-based tools that allow you to check grammar coverage and return values without inputting real speech.
- ✓ The recognition. This is your opportunity to eliminate gross problems with recognition. Since speech recognition technology is never 100% accurate, define test criteria that distinguish “working satisfactorily” from “not working satisfactorily” rather than “not working” from “working”. Recruit 10 to 15 callers to test each scenario. (The same callers may be used to test different scenarios).

Of course, any problems identified by testing should be corrected before the application is deployed.

## Deployment and Tuning

Once requirements have been written and the application has been designed, coded and thoroughly tested, the application is for a *limited* deployment. The purpose of this phase is two-fold: 1) to determine how well the application performs for a real caller base and 2) to fine-tune the application for optimal performance before exposing it to the entire caller population. The steps for this phase are:

1. Collect speech and logs from real callers. Typically, data from at least 1000 calls spread over several days is required for a meaningful analysis.
2. Transcribe the speech data.
3. Analyze the application performance. Some of the things frequently measured in this step include
  - Recognition accuracy – how often did the recognizer come up with the correct response
  - Call automation rate – how many calls were able to be completed without transferring to an agent
  - In-grammar rate – how often did the grammar include the caller response
  - Average number of re-prompts – how many times did you have to prompt the user for the same piece of information before it was recognized by the application
4. Identify problem areas within the application.
5. Determine the cause of problems. As an example, the analysis might show that at a particular prompt, a larger percentage of callers respond with a phrase that is not included in the grammar. That’s the problem. The cause could be a missing phrase in the grammar. It could just as likely be that the prompt is confusing.
6. *Tune* the application – that is, modify the call flow, the prompts, the grammars, specific recognizer parameters, the application parameters, or any combination of these required to improve the performance of the application.
7. Repeat steps 1 through 5, if required.

Once the application has been tuned to meet pre-defined success criteria, it is ready to be fully deployed.



## How RSI Can Help

RSI has extensive experience in developing IVR applications ranging from DTMF to Natural Language Speech Recognition using VXML. It has to its advantage an ISO 9001:2000 and CMM Level 4 Offshore Center of Excellence to deliver cost-effective solutions.

Whether you are looking for business case development, design consultation, mentoring, temporary development resources or outsourcing, our speech team equipped with state of the art infrastructure is ready to address your needs. Let us know how we may begin helping you. Please do contact us via email at [marketing@realsoftinc.com](mailto:marketing@realsoftinc.com) or call us at +1 609 409 3636.