

Contact Center Automation Cost Estimation Tools

Speech Recognition or DTMF?

Introduction

Your agent costs are sky-rocketing. To save your business, you must automate your contact center. But what's the best solution– DTMF or speech recognition? This paper explores some of the questions that may arise as you choose between the two alternatives.

Is Your Application Suited to DTMF

The answer to this question depends upon the type of information that must be collected from the caller in order to complete a transaction or inquiry. DTMF is well-suited for navigating small menus or lists, entering numeric information such as accounts or phone numbers, and responding to yes/no questions. Entry of dates, times, and alpha-numeric data can also be accommodated with DTMF, but the user- interface for entering these can be complicated and frustrating to the caller. Selection from long lists of products or services can also be implemented with DTMF. However, this generally requires either 1) use the touch-tone keypad to spell the product/service name, or 2) a long and wordy prompt that enumerates the caller's options. Finally, collection of extremely complex data such as postal addresses is virtually impossible to handle with DTMF.

Speech recognition, on the other hand provides a natural and intuitive means of gathering data of these types. Using speech recognition, long prompts can be shortened, thus reducing call duration. Callers can say the name of a product or service, rather than fumble with a keypad to spell the name.

What Can Speech Recognition Do That DTMF Can't

In the past few years, natural language capability has been added to speech recognition engines. These products can understand unconstrained caller speech. They enable the automation of applications such as help-desks or call routers in which callers' requests are fairly unpredictable. DTMF cannot handle these types of applications.

Should Existing DTMF Applications Be Converted to Speech Recognition

The answer depends mostly on how successful your existing application has been in reaching your goals for automation. If a large percentage of your calls are already automated *and* if you have a large caller base that is satisfied with the current DTMF application, there may only be limited benefit in converting to speech recognition. If on the other hand, a large percentage of callers either 1) hang-up prematurely or 2) “zero-out” to an agent, then speech recognition could improve your automation rate. This is true even if the functionality of the speech recognition application is identical to that of the original DTMF application. In fact, studies have shown that speech recognition often increases the call completion rate. And a recent studyⁱ has shown that callers are less likely to request an agent when using a speech-enabled application, than when using a DTMF application.

How Will My Callers React

Studiesⁱⁱ have indicated that callers prefer speech-recognition interfaces to DTMF. Speech recognition provides a natural and intuitive way for callers to interact with an application. Speech recognition applications can flatten long or complex DTMF menu structures, and they can collect multiple pieces of information in a single prompt, so callers spend less time on the phone. Also, callers appreciate the convenience of having service available 24 hours, 7 days a week.

What’s the Cost of Implementing Speech Recognition

Speech applications can range from a low of \$10,000 to over \$150,000. While it is impossible to determine the exact cost of a Speech Application without doing a detailed analysis, it is possible to get some budgetary numbers by answering a few simple question found in the attached is “Speech Application Customer Questionnaire”. By classifying the complexity of your application as Simple, Average or High and the size as Small, Medium or Large, you can determine the estimated cost of developing a Speech Enabled VoiceXML application from the attached “Speech Application Cost Estimation Table”.

What’s the Return on Investment

In general, the initial cost of deploying speech recognition is greater than that for DTMF. There are a few reasons for this: Hardware such as echo cancellers for barge- in must be purchased. Many speech recognition solutions require one or more speech servers in addition to the IVR system. Speech user- interfaces are more complex than those for DTMF applications, and hence, incur additional expense for design and coding. Speech recognition applications require one or more performance analysis and tuning cycles to fine-tune the applications; these cycles are not required in DTMF application. Finally, licenses for the speech recognition engine software must be purchased.

Despite these extra costs, in many cases, speech recognition applications often generate a greater ROI than DTMF applications, and can pay for themselves in less than a year. The reasons are obvious: more functions can be automated with speech recognition than with DTMF; more callers are willing to use speech recognition than DTMF; callers wait in queues for shorter lengths of time; call durations are shortened, and call completion rates are increased. As a result savings are realized in port costs and agent costs. Finally, there are less tangible benefits to be realized. Customer satisfaction increases, agents are freed to handle complex tasks requiring human interaction (such as up-selling), and agent productivity increases.

Real Soft, Inc.

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](#) or call us at +1 609 409 3636 ext 124.

¹ UI Design Update Newsletter – July 2003 from Human Factors International - Suhm, B., Bers, J., McCarthy, D., Freeman, B., Getty, D., Godfrey, K., and Peterson, P. (2002). A Comparative Study of Speech in the Call Center: Natural Language Call Routing vs. Touch-tone Menus. Paper presented at ACM SIGCHI, Minneapolis, Minnesota.

² Harris Interactive Speech Satisfaction Study, 2003.