



SPEECH-SOFT
SOLUTIONS

Understanding Your IVR Options or “How to Open the Front Door to Your Customer Relationship”

Introduction

Your IVR project finally got green-lighted. The good news – there have never been more options for IVR technology: Touch-tone Only, Press or Say Digits, Directed Speech Dialog (with touch tone fall back), and Natural Language Understanding. The bad news – now you are faced with the daunting task of figuring out which one to use.

This guide is intended to help you understand the top IVR options being deployed today, from the simplest to most complex and yes, generally that means from the cheapest to most expensive as well.

IVR Basics

Before we get into the different modes, let’s review some basics of automated speech recognition (ASR) which drives all the innovations in the IVR space.

- There are two models of speech recognition: one used by dictation software that trains the recognizer (computer) to understand a single voice speaking many words. The second model limits the number of recognized words so it can handle a wide variety of speakers and accents. IVRs use the second type of recognition.
- Grammars are used to define the words that callers speak at any given point in the call flow and are an integral part of a Voice User Interface design and deployment.
- The IVR consists of a series of prompts that are played to guide callers through their tasks.
- Once an application is deployed, audio files are collected and transcribed to see if callers used that the designers did not anticipate (and therefore, are not in the grammars). This process of evaluating the grammar performance is called Tuning.

Now on to your IVR options.

Touch-tone-only

The grand-daddy of telephone automated self-service, Touch-tone-only is still the most common IVR modality. During its heyday in the 1990s, touch-tone allowed businesses to extend their business hours

to 24/7 for simple tasks. This won approval, if begrudging acceptance, from their customers. The shift to cordless phones (with handset keyboards) and then smart phones (with impossibly small keyboards) has made this modality more problematic. Callers now must take the phone away from their ear in order to respond, causing them to fall out of step with the prompting. This results in errors and opt outs. And, let's face it, there is no way that this technology feels fresh and cutting edge in the 21st century. All that being said, however, touch-tone still has its purpose. It is excellent at collecting a few items of information (credit card activations), collecting information that callers readily know (phone numbers) or for automating short tasks.



Touch-tone Benefits

- Familiar interface
- Unaffected by noisy environments or side conversations
- Easiest to code, design, test and support and, hence, the cheapest to deploy
- Quick to market
- Great for very familiar or small amounts of input from the caller

To report an outage, press 1

To change your service, press 2.

Please enter the 10 digit phone number associated with your account.

Touch-tone Challenges

- Easy to get lost in the call flow because numbers entered in menus have no real meaning.
- Can't use in a hands-free environment.
- Challenging for cordless phones handsets and smart phone keyboards.
- Can't support complicated tasks (alpha-numeric collections, address changes)
- Feels inhuman.
- Long inputs can be problematic.

Press or Say Digits

Press or Say digits have crept into use over the past few years. Callers are prompted to either press or say a specific number to select a menu option. They can also use speech to provide multi-digit information such as account numbers, dollar amounts, and dates. Press or Say Digits leverages a single digit grammar that is used throughout the many parts of the call flow to collect long digit strings (such as account and telephone numbers) and digit dates. Although Press or Say Digits does allow for hands-free interactions and solves the pesky tiny cell phone keyboard issue, it is not without its own problems. Single digits are notoriously difficult to recognize and do not perform as well as speech commands. In

addition, because these apps sound so much like Touch-tone-only, callers forget that they need to limit side conversations resulting in prompt interruptions, errors and opt outs. Recommended uses for Press or Say Digits are the same as Touch-tone-only: collecting a few items of information (credit card activations), collecting information that callers readily know (phone numbers) or automating short tasks.



To report an outage, press or say 1

To change your service, press or say 2.

To sign up as a new customer, press or say 3.

Press or Say Digits Benefits

- Simpler to code, design, test and support; hence, cheaper than other speech options
- Quick to market

Press or Say Digits Challenges

- Like Touch-tone, there are no words to engage callers and help them track where they are in their self-service journey
- Single digits recognition challenges
- Confusing modality where callers forget to limit their side conversations
- Callers overwhelmingly tend to use touch-tone (DTMF) in these applications, begging the question of why pay for speech licenses at all?
- Can't support complicated tasks (alpha-numeric collections, address changes)
- Feels inhuman

Directed Speech Dialog

The goal of a directed speech application is to guide callers by asking a series of questions, just like your best agents would do. The applications rely on grammars that are built for each collection (limiting what callers can say to a predetermined set of words or phrases). The design of the grammars and call flows requires a deep understanding of the business terminology, the caller's mental maps of the transactions they are calling to perform and knowledge of what does and doesn't work with speech recognition engines. When well designed, callers performing a task can forget that they are speaking to a computer and may even say thank you at the end. Directed Speech Dialog gives the caller speech commands to navigate a menu. Touch-tone is generally used as a fallback for error handling when callers have bad

reception or are in a noisy environment. The best Directed Speech IVRs are like plumbing, you only notice them when something goes wrong. Although they have been around since the early 2000s, speech apps are still considered cool. Directed Speech Dialog is good for complicated tasks (booking a train ticket) or tasks where speech has a real advantage (change of address).



Where are you travelling?
And where are you travelling from?
You're calling about a loan, is that correct?

Directed Speech Dialog Benefits

- Feels human
- Truly hands free; no problems with small keyboards and handsets
- Handles collecting names, addresses and alpha-numeric account or member ids

Direct Speech Dialog Challenges

- Requires periodic tuning (updating grammars) to optimize performance
- Noisy environments and side conversations can trip up the recognition and disrupt the experience
- Slower to market
- More expensive to design, build, deploy

Natural Language Understanding

Natural Language Understanding (NLU) is today's most sophisticated speech recognition solution: the IVR can actually understand answers to open-ended question like "How can I help you?". If you have a large number of nested menus and you find that your callers aren't finding their way through to important options, then NLU may be a viable solution for you. Placement of the NLU at either the top of the call or after authentication depends on what makes most sense for your caller population, the tasks you are automating and your business rules. Remember, NLUs are merely an enhancement for the main menu in Directed Speech Dialog IVRs. They are not used throughout an application (a common misconception).

NLU uses machine learning to train the speech recognizer to understand that many different phrases (utterances) can mean the same thing. This is done by grouping them into semantic categories which can be leveraged by the application. But before you get all Star Trek on me, understand that it is humans who do the categorization which makes NLUs: you guessed it, very labor intensive and expensive. They

require a sizable data collection where all the different ways your callers talk about a task are gathered. NLU also require close collaboration between the IVR designer and the business experts to define appropriate categories as well as additional collections to handle vague phrases.

NLU collections were all the rage after Apple's Siri hit the stores and trendy companies had to prove their cool factor by deploying these Siri-like menus in their IVRs. The hype didn't always live up to the reality (not unlike Siri itself) and the promises for improved effectiveness were not always achieved. The high cost of deploying NLUs – and of keeping them up to date with the latest products – has led some prominent companies to abandon them for more traditional Directed Speech menus.

Still, you should consider them if you have complicated menus or you just want to use the latest and greatest technology because your brand demands it. NLUs work well for banking's straightforward terminology but less so for tech support or health care queries where callers stumble over trying to describe their reason for calling.



In a few words, please tell me why you are calling today.

NLU Benefits

- Eliminates the need for callers to navigate through multiple menus
- Most human
- Considered cutting edge

NLU Challenges

- Labor intensive and costly to design and deploy
- Not readily adaptable to new products
- Requires ongoing tuning after deployment

Conclusion

In this double-jeopardy era of raised customer expectations and heightened customer empowerment through social media and online reviews, your customer care technology decisions have greater import than ever before – and none more so than the front door to your customer relationship—yes, your IVR.

As you evaluate the various options, please be aware that you can start small and build to a more complex solution. You don't have to decide on a single model for all your IVR transactions. In fact, many companies use an assortment of IVR modalities. For example, touchtone-only for credit card activations and NLU main menus for telephone banking across all accounts.

Speech-Soft Solutions teams of elite professionals have been helping companies design and deploy cost effective, customer impacting IVR speech applications in broad vertical markets since 2003. We work

closely with your team to tailor applications to your customer vision and internal business requirements. We leverage our experience and best practices to quickly design and deploy speech applications that improve customer experience and reduce operating cost. From customer interaction and Omni-channel strategies to contact center integration and voice biometrics, Speech-Soft is a great partner to work with on your new or evolving customer care interactions.

Speech-Soft Solutions, LLC is a global systems integrator, providing multi-channel strategies for Contact Center, speech-enabled Interactive Voice Response (IVR), Computer Telephony Integration (CTI) solutions, Omni channel architectures, business intelligence and professional services using state-of-the-art technologies and industry-leading business partners, to meet the needs of virtually any industry.

Our industry expertise in Cisco, eGain, Avaya, Nuance, VXML, CCXML, IP/VOIP and CTI technologies will help you keep your customer connected, while your business can focus on growth and increase its competitive advantage.



SPEECH-SOFT
SOLUTIONS