Gyst Blog Post

Does Improving CX Really Save You Money? Here's How One Contact Center Proved It Out.

By Daniel O'Sullivan, CEO - Gyst Technologies

Does Improving CX Really Save You Money? Here's How One Contact Center Proved It Out.

Daniel O'Sullivan
Founder & CEO @ Gyst Technologies | Speech User Interfaces

82 articles ✓ Following

Note: This is an in-depth article. If you are primarily interested in the outcome of the trial, you can skip ahead to the **Conclusion** section below.

Introduction

While user-personalization has had demonstrated success for web-based interactions, it has yet to be fully leveraged in the voice channel.

Advances in speech technology such as Conversational AI, effective ASR, NLU, and Intelligent Voice Assistants are all excellent enabling technologies. These however, are

By Daniel O'Sullivan, CEO - Gyst Technologies

Does Improving CX Really Save You Money? Here's How One Contact Center Proved It Out.

Daniel O'Sullivan
Founder & CEO @ Gyst Technologies | Speech User Interfaces

82 articles ✓ Following

Note: This is an in-depth article. If you are primarily interested in the outcome of the trial, you can skip ahead to the **Conclusion** section below.

Introduction

While user-personalization has had demonstrated success for web-based interactions, it has yet to be fully leveraged in the voice channel.

Advances in speech technology such as Conversational AI, effective ASR, NLU, and Intelligent Voice Assistants are all excellent enabling technologies. These however, are

only part of the solution. Used in conjunction with a well-designed voice user interface that incorporates application specific grammar tuning, intentional pauses, unambiguous prompting and other best practices in VUI design, they represent a significant improvement over earlier technologies.

An important design consideration not addressed by this tool set however, is the individuality, navigating skills, and calling environment of the users actually engaging with the technology.

To the extent that a voice application can monitor and adjust the user experience to suit the exhibited behavior of a particular user during a voice session, a proportionate number of sessions can end with better outcomes for both the caller *and* the contact center. Empirical data gathered during a production trial at a customer site quantifies the extent of these benefits. The results are enumerated later in this article.

What makes a conversation productive?

Human conversation is a dynamic and highly individualized process. Research shows that the average English-speaking rate varies widely from 130 - 200 Words Per Minute. This wide WPM range applies to 90% of the English-speaking population.

On the listening side of the equation, additional research shows that:

- For complex material, a rate of 130 - 145 WPM may be required
- For material of average complexity, 145 - 175 WPM can be optimal
- For simple material, many listeners can accommodate over 175 WPM

Listeners can be lost to boredom, overwhelmed by complexity, or fully engaged in a conversation based on the material and the speaker's ability to deliver that material at the optimal cadence for each listener.

Good communicators are aware of this fact and continuously monitor their audience for signs of engagement, interest, and boredom. They continually adjust their speaking rate, message content and emphasis to get the message across effectively and efficiently. They make these adjustments in an instinctive, fluid, and natural way, thereby quickly "tuning in" to establish optimal harmony with the listener, keeping them fully engaged in the dialog.

Focusing on the Customer

Every user of a voice application brings with them their own unique set of cognitive, aural, verbal, and hand-eye coordination (as used in DTMF keypad entry) abilities.

Familiarity with the call flow of the voice application also varies widely from one individual to the next. A person engaging with a voice application they use frequently and know well will anticipate subsequent voice prompts and make very few input errors in order to achieve their goal.

Someone less familiar with the application will need more time to cognitively digest the instructions and options presented to them. They are also more likely to induce timeouts and input errors, or simply select the wrong path in the call flow for their particular inquiry.

Add to this the customers constantly changing calling environment variables such as background noise, poor mobile phone signals, and caller distraction, and it's easy to see why every call to a voice application is truly a unique interaction; an interaction that does not lend itself well to a solution that does not factor this in.

This is one of the principal reasons human operators are so good at answering any type of customer service inquiry - they can handle the dynamics of human conversation intuitively and with ease. The really good agents can handle the customer with a good deal of empathy for them since they know when someone is stressed out, calling from a noisy environment, when their mobile signal is fading in and out, when the kids are distracting them….the list goes on.

Callers know an agent will understand this as they try to self-serve in the voice channel. And they will opt for an agent as soon as they perceive the self-service option will not resolve their problem.

Applying the principle to voice self-service

Most applications running on Voicebots, IVR systems, and other voice channel technologies today are "static" and make no adjustments for the real-time, exhibited behavior of individual callers. As a result, all callers are handled in the same way regardless of their knowledge, cognitive abilities, navigation skills, and willingness to use voice self-service.

Specifically, all audio prompts and messages are delivered at the same WPM rate regardless of a callers skill and exhibited behavior while navigating their way through the call flow.

Without "tuning in" to a callers behavior during the call, real efficiencies in the Contact Center are lost. As shown in Table 1 below, this can have significant consequences in terms of Customer Service and the costs associated with operating Contact Centers.

Table 1
The impact of poor CX on the Customer and the Contact Center

IVR/Voicebot Problem	Customer Impact	Contact Center Impact
Conversation moves too fast	Customer makes input errors and gets frustrated	Zero-outs, hangups, venting to agents
Conversation moves too slow	Customer believes an agent will be faster	More calls to agents
Excessive timeouts and reprompts	Longer calls and increased customer frustration	Zero-outs, hangups, venting to agents
Low goal completion	Customer wasted their time and is annoyed	Repeat use of self-service is discouraged
Long handle times	Customer believes an agent will be faster	Longer calls and more calls to agents
Low engagement rates	Customer wants to self-serve but can't	Agents handle mundane/ repetitive questions
No performance analysis	Pain and friction points further annoy customers	Blind to where customers have trouble

Trial results on a client application

In coordination with our client, a large and well-known healthcare insurance provider, we conducted a trial to determine what effect dynamically adjusting the audio playback rate of voice prompts in their IVR would have on voice self-service performance.

Existing voice prompts were speed adjusted in direct relation to individual caller skills. Caller skill here refers to how well a particular user navigates a Conversation Turn in the call flow, as compared to thousands of samples taken earlier from the calling population.

A Conversation Turn is defined as a request for user input that is spoken by the voice system, followed by the response (spoken or keyed in via DTMF) given by that user. Each no-input and no-match retry prompt and response encountered by the user counts as an additional Conversation Turn. For a more detailed explanation of how this process works, there is a short video available [here](#).

The client's voice application handles inquiries for medical insurance claims, benefits, member coverage, and general information. It serves primarily members (generally novice users) and providers (generally expert users). During the trial, audio playback speed adjustment levels of 100, 110, 114, 117, and 119 percent were used. A playback level of 100 indicates the normal playback rate of the audio, 110 represents 110 percent of normal, and so forth.

As Table 2 below shows, for the standard (unadjusted audio) phone calls, the mean number of Conversation Turns was 31.41% (428,820/1,365,172), with a 95% binomial confidence interval ranging from 31.33% to 31.49%).

For the speed adjusted ports, the mean number of Conversation Turns was 22.95% (19,826/86,405), with a 95% adjusted-Wald binomial confidence interval ranging from 22.67% to 23.23%. Because the binomial confidence intervals did not overlap, the difference was statistically significant ($p < .05$).

This shows that the mean number of Conversation Turns for speed adjusted and standard calls were, respectively, 4.4 and 3.2. This represents a **36.90%** increase in caller engagement in voice self-service when speed adjusted audio is used.

Table 2

Changes in voice self-service engagement

	Standard	Speed Adjusted
Calls Answered	428,820	19,826
Conversation Turns	1,365,172	86,405
Conversation Turns Per Call	3.2	4.4
Caller Engagement		**+36.90%**
95% CI Upper Limit	31.49%	23.23%
Observed Percentage	31.41%	22.95%
95% CI Lower Limit	31.33%	22.67%

Referring now to Table 3 below, with standard audio, the rate of first-time errors was 21.75% (296,898 errors divided by 1,365,172 opportunities for error, with a 95% binomial confidence interval ranging from 21.68% to 21.82%).

With the speed adjusted audio, the first-time error rate was 17.74% (15,324/86,405, with a 95% binomial confidence interval ranging from 17.48% to 17.99%). This led to an absolute reduction in first-time errors (combined no-input and no-match events) of 4.1% (relative reduction of 18.5%). Because the binomial confidence intervals did not overlap, the difference was statistically significant ($p < .05$).

Table 3
First try error/reprompt messages

	Standard	Speed Adjusted	Difference
First Try Errors	296,898	15,324	
Conversation Turns	1,365,172	86,405	
95% CI Upper Limit	21.82%	17.99%	-4.34%
Observed Percentage	21.75%	17.74%	**-4.01%**
95% CI Lower Limit	21.68%	17.48%	-3.69%
CI Range	0.14%	0.51%	

Table 4 below shows how using speed adjusted audio impacted overall error rates in the IVR. While the greatest reduction in errors, 4.01% absolute or **18.45%** relative, occurred at the first try level, some errors were actually introduced at levels 2 and 3 during the trial.

While second and third try errors are of course, undesirable, reducing first try errors has the advantage of preventing downstream errors and keeping the Conversation Turn completely clean. This has special significance in terms of keeping the caller moving through the call flow and is especially effective at encouraging marginal users to persist.

Notably, the overall input error rate decreased by 3.59% absolute or **12.51%** relative for speed adjusted callers during the trial. This translates to those callers experiencing 52,112 fewer input errors. It also means they did not have to listen and respond to 52,112 additional error/retry messages. In addition to keeping these calls "cleaner", this also helped to keep them shorter.

Table 4

All error/reprompt messages

	Standard	Speed Adjusted	Absolute	Relative	Errors
First Try Errors	296,898 (21.75%)	15,324 (17.74%)	-4.01%	-18.45%	-58,251
Second Try Errors	74,608 (5.47%)	4,990 (5.78%)	0.31%	5.67%	4,500
Third Try Errors	20,262 (1.48%)	1,380 (1.60%)	0.11%	7.61%	1,639
Totals	391,768 (28.70%)	21,694 (25.12%)	-3.59%	-12.51%	-52,112

Table 5 shows how using speed adjusted audio impacted average handle time (AHT) in the IVR. As indicated earlier, this voice application serves primarily the clients members (insurance policy beneficiaries), and providers, generally the medical service provider or doctor's office administrative staff.

Since members call the application relatively infrequently, they tend to be less skilled at navigating the call script. Additionally, they tend to be less inclined to learn how to use the IVR. Many will opt for a human as soon as self-service becomes error prone, challenging, or simply unproductive for them.

Providers on the other hand, call the application several times per day and are generally calling for a specific, well-defined purpose such as benefits coverage or a claims inquiry. They know from past experience that the IVR is the fastest way to answer their inquiries and that dealing with an agent may actually take longer.

In summary, while it is difficult to say why some of the caller types shown in Table 5 show an increase in AHT while others show a decrease, this is likely due to the various skill levels and attitudes towards the IVR that each type has.

In the aggregate, faster audio and fewer retry messages contributed to shorter handle times in a good way here, while increased caller engagement in the IVR contributed to longer handle times, also in a good way. The distribution of AHT increases and decreases among the various caller types may vary, but the means definitely justifies the end goal here.

Table 5

Average handle time by user type

	Standard		Speed Adjusted				
	AHT	Answered	AHT	Answered	Change	Percent	Usage
Broker	69	7,200	75	284	6	8.70%	748
Employer	104	3,236	87	50	-17	-16.35%	-931
Member	125	217,894	130	11,106	5	4.00%	19,083
Provider	141	188,480	134	7,912	-7	-4.96%	-22,912
Spanish	103	1,853	76	34	-27	-26.21%	-849
Unknown	40	10,157	41	440	1	2.50%	177
Totals		428,820		19,826		-0.49%	-4,684

Conclusion

As the trial data above indicates, when a voice application has the capacity to monitor the skill and exhibited behavior of individual callers, and automatically tune the playback speed of voice prompts (WPM spoken) to suit the callers environment and capabilities, a significant number of self-service phone calls can result in better outcomes for both the caller and the contact center.

In summary, the trial results indicate:

- Engagement in the IVR increased by **36.9%**
- First try error messages were reduced by **18.45%**
- Overall, error messages were reduced by **12.51%**

Callers using speed adjusted audio had 36.9% more engagement (Conversation Turns) with the IVR than callers using standard audio. They also encountered 12.51% fewer error messages and thus, had to reenter information 12.51% fewer times.

The difference in cost between a call handled by voice self-service and a call handled by an agent can vary between $2 - $6 or more, depending on the length of the call, the knowledge and training level the agent receives, onshore/offshore sourcing, and other factors. For our calculations below, we will assume a cost differential of $4 per call between the two.

From Table 2 above, Standard calls consisted of 3.2 Conversation Turns on average. Thus, had the 36.9% increase in engagement the adjusted audio callers experienced been handled by agents, the additional cost for a contact center handling 10,000 self-service calls per day would be:

$4 x (10000 x .369)/3.2 = **$4,612** per day.

Put another way, replacing standard audio with adjusted speed audio in this particular voice application generates **$1,683,380** in annual cost savings for the contact center.

Direct cost savings aside, having customers experience fewer error messages and requests for the reentry of information, along with handling their inquiries on their first contact and freeing up agents for less mundane calls, all contribute to additional benefits in terms of improved customer service and brand image.

While voice applications tend to be similar in structure within each of the healthcare, financial, government, utilities, travel, and retail verticals, calling populations vary from region to region, as do caller demographics. Conducting a rigorous trial with A/B testing on large sample sizes like this is the best way to learn what precise benefits this technology will provide for a given voice application.

About Gyst Technologies

At Gyst Technologies, we develop advanced personalization software for Contact Center IVR systems, Voice-Enabled Virtual Assistants, Conversational AI Services, and any self-service channel that uses voice as a means of user communication. We have reduced enterprise costs and improved the customer experience on millions of user interactions to date.

Contact Us today to learn how we can do the same for your Contact Center.

Thank you for downloading this AWS and Gyst Blog Post! Carahsoft is the distributor for AWS public sector solutions available via GSA, NASPO, The Quilt and other contract vehicles.

To learn how to take the next step toward acquiring AWS's solutions, please check out the following resources and information:

For additional resources:
carah.io/AWS-Resources

For upcoming events:
carah.io/AWS-Events

For additional AWS solutions:
carah.io/AWS-Solutions

For additional public sector solutions:
carah.io/AWS.Solutions

To set up a meeting:
AWS@carahsoft.com
888-662-2724

To purchase, check out the contract vehicles available for procurement:
carah.io/AWS-Contracts