

THE FUTURE
IS NOW –
YOU CAN TALK TO
YOUR MACHINES
AND THEY LISTEN!



KEY LIMETM
INTERACTIVE

THE FUTURE
IS NOW

You can talk to your machines and they listen!

HOW KEY LIME INTERACTIVE (KLI) IS HELPING
REFINE VOICE ACTIVATED EXPERIENCES

BY EUGENIO SANTIAGO

VP User Research

What is happening today

We are in the early stages of another tech boom. Controlling devices and products using your voice is here in a big way. While speech recognition has been around since 1952, the recent advancements in deep learning and big data have resulted in the design and deployment of various speech recognition systems worldwide. And all the big boys are involved: Google, Amazon, Microsoft, IBM, Apple, and Baidu to name a few.

The applications of voice recognition are growing too. You see it integrated in various industries such as automotive (in-car systems), health care (documentation), military, education, and now in your daily life with devices like the Amazon Echo and Google Home entering households.

Approach

At KLI, we hang our hat on implementing user research in innovative ways. In that vein, we have developed a robust methodology for companies looking to enable voice control for their future products. This methodology is aimed at optimizing the user experience by making it as seamless and frustration-free for your customers to interact with your product via voice control.

It is a multi-phased approach consisting of a quantitative study to collect a large sampling of voice commands, iterative tests to refine the experience, and a final qualitative test (in-person) to validate the performance of the system based on the previous learnings.

This approach has successfully proven to be cost

effective and rigorous while providing actionable recommendations.

What we've learned

In deploying this research approach, we learned a great deal. Chief among them was how voluminous the verbal requests are for even the simplest of commands. This learning amplified the importance of how large the voice command database needs to be for your customers to interact with the system as naturally as possible.

Our initial phase of research focuses on this specifically. We sourced a user sample from a geographically diverse area with the intent of collecting variations in commands and vocalizations which vary in terms of accent, pronunciation, pitch, cadence speed, etc. This data collection method ensures we adequately catalog a sufficient number of utterances; we learned the possible commands users will speak is far greater than the initial list you come up with.

While a flexible language taxonomy that accounts for a range of users is important, there are other elements of the interaction that play a significant role in how well the user experience is perceived. While observing users interact with the product, we identified a few more nuances about language that were not initially apparent – the problem that arises with discontinuous speech.

Each of these systems has an activation phrase followed by a window of time for the user to provide the desired command. If the user does not complete the command within that timeframe, regardless of how inclusive your command database is, the system will likely fail.

This presents an interesting business issue to solve for when you think about in-home systems-like Amazon's Alexa, that doesn't exist when dealing with say the voice activation system in your vehicle. When a command fails while speaking to Alexa, who does the user blame? Is Alexa at fault or did the skill fail?

For those unfamiliar with a 'skill', it is defined by Amazon as the following: a capability that enables customers to interact with devices in a more intuitive way using voice. Examples of Alexa skills include the ability to play music, answer general questions, set an alarm, and more. In use, a person would give a command such as "Alexa, ask airport security for the wait time at LGA", then the Airport Security skill would check the TSA's public API for the last reported wait times and then verbally respond back to the user.

As developers of these skills you need to have a strong understanding of the entire user interaction, what you can control, what you cannot, and how both can potentially impact how users feel about *your* brand.

While the product went through the grinder a bit during user testing and took some lumps, when it came time to run our validation study, it was amazing to see how comfortable most participants became with the system in a limited amount of time and how genuinely excited they were about the prospect of controlling the product using their voice.

Right away users started to offer up suggestions for how they wanted to communicate with the product and what they wished it could do (the version they tested had limited functionality but functionality that would roll out for generation one). It was crystal clear, although the product had come a long way since the start of development and had reached a point the design team, software team, and company overall could be proud of, the need to further refine and continuously update the product would be needed to delight customers.

What's next

As the Internet of Things (IoT) continues to expand and pull in items that were previously disconnected, voice controlled devices will become the norm. If you are working on voice controlled devices currently or are moving closer to that reality, remember the following

"As developers of these skills you need to have a strong understanding of the entire user interaction, what you can control, what you cannot, and how both can potentially impact how users feel about your brand."

things:

1. You need a flexible language taxonomy that accounts for a range of users
2. The variations in command utterances are far greater than you expect
3. Conducting user testing will uncover findings that your software team alone will not find
 - It won't slow you down if you incorporate it into your schedule
 - It won't cost you money but rather save you money by identifying issues before your users are exposed to them
4. [For skill creators] Your command database isn't the only thing to be mindful of. The quirks of interacting with systems like the Amazon's Alexa, the Google Assistant, Siri, etc. are also part of the experience
5. The bar to delight customers rises as rapidly as the technology advances. Continue to refine your experience and test iteratively

To learn more about Key Lime Interactive and how you can partner with us to optimize your voice activated experiences, contact info@keylimeinteractive.com CJM can serve as a powerful tool for your team and help you see that vision come to fruition.



**Optimize the Experience.
Inform Design.**

[@keylimeinteractive.com](https://www.keylimeinteractive.com)
305.809.0555