Voice Recognition Technology: Are We Ready for It?

Staci Morrison

Boston University

New Communication Technologies

Professor Carter

December 13, 2011

Introduction

With the highly anticipated release of Apple's iPhone 4S, voice recognition technology, through the sassy personal assistant called Siri, has become the hottest new technology. Is Siri really a new technology? Voice recognition technology itself is not new.

The iPhone flurry made me curious to explore whether voice recognition technology is finding a more welcoming public today than in the past. It seems we are witnessing applications and products incorporating voice and speech recognition technology across various markets more frequently today. If so, why is voice tech arriving now? Are we approaching the age of voice technology? This paper will focus on the consumer market of emerging voice recognition technology and why or why not we can expect wide adoption of more voice technology forthcoming.

Brief History of Voice Technology

Voice technology in a form most people would recognize today has existed for nearly half a decade. In fact, the early voice technology existed before there was a market or consumer application for it[1]. Bell Laboratory researcher Homer Dudley demonstrated the Voice Operated Recorder, or voder, and vocoder, or voice encoder, at the 1939 World's Fair. The voder machine was "designed to test compression schemes for the secure transmission of voice signals over copper phone lines," and the vocoder synthesized speech. In World War II the vocoder became invaluable,

---

[1] "VoicesofDragon." (2011, October 24). The evolution of voice technology.

2

"scrambling the transoceanic conversations between Winston Churchill and Franklin Delano Roosevelt" but both technologies otherwise were left under the radar of consumer demand or use[2].

Aside from the musical voice synthesizing applications of the vocoder in the 1950s-1960s, general, consumer-oriented use of voice technology did not surface until the late 1970s. Entrepreneur Gordon Matthews invented the voice mail system, called Voice Message Exchange, or VMX in 1979. This system for organizing messages digitally became the foundation of today's voicemail[3]. 1983 saw the introduction of voice-activated typewriters, with the unveiling of a model by IBM and the prediction of widespread use by 1990[4].

Spoken Dialog Systems also found a place in general consumer use. Spoken dialog systems include the friendly robotic voices who answer calls to businesses, directing customers to "press 1 for English," etc. Although an abundance of such database assistants and other spoken, voice recognition systems have been in use for years, for the most part they are inefficient and limited[5].

Through these examples alone, we can see a presence of voice technology since its unveiling in 1939, yet without publicity and immediate popularity of Apple's Siri personal assistant. What are the reasons Siri found popularity in a technology previously plagued by awkwardness and consumer malaise? Let's begin by looking at the industry behind Siri.

———————————

[2] Apple. (2009). A Brief Vocoder History. Logic Express 9.

[3] Paul, L.R. (2005). *Made in the twentieth century: a guide to contemporary collectibles.*

[4] InfoWorld. (1980, October 27). Typewriters Learn to Listen. InfoWorld, 2(19), 22.

[5] Huerta, J. and Pieraccini, R. (2005, September 3). Where do we go from here? Research and commercial spoken dialog systems.

Reasons for Voice Technology in Mobile Market

The mobile phone market is growing exponentially. In the 50th Mobile Intelligence Report, Millenial Media and Gartner reported a 2010 global mobile advertising revenue of $1.6 million worldwide.

Mobile Advertising Revenue by Region, Worldwide
In Millions of Dollars
CHART A

	2010	2011	2015
North America	304.3	701.7	5,791.4
Western Europe	257.1	569.3	5,131.9
Asia/Pacific and Japan	868.8	1,628.5	6,925.0
Rest of the World	196.9	410.4	2,761.7
Total	$1,627.1	$3,309.9	$20,610.0

Source: Gartner, "Forecast: Mobile Advertising, Worldwide, 2008-2015"
(as cited in press release, June 16, 2011)

By the end of 2011 the worldwide revenue is projected to double and eventually jump to $20.6 million by 2015. Furthermore, the Apple iPhone has been the top selling phone for the past 2.5 years. Each of the top 20 most popular phones in the second quarter of 2011 were smartphones. As such, there is huge incentive to innovate on a mobile platform. Smartphones are becoming ubiquitous, these advertising revenue numbers for the third screen indicates critical mass[6].

Implementing a voice recognition technology on a smartphone such as iPhone, that has an established customer base and exceptional customer loyalty ensures

---

[6] "Justin." (2011, August 2011). Millennial Media Compares 2 Years of Mobile Advertising Trends In 50th S.M.A.R.T Report. Mobile Marketing Watch.

access to a huge pool of users. Hardwiring the Siri personal assistant into the iPhone 4S likewise builds eliminates the extra step of requiring users to download an application.

Inherent Voice Recognition Design

Mobile phones were designed for voice input. Relative to computers or tablets, mobile phones have a sophisticated built-in microphone, placed directly next to the mouth when the phone is used, which minimizes ambient noise sound degradation and maximizes voice clarity.

But the advancement of mobile phone technology progresses in line with Moore's Law[7], fitting more functions and processing power into a smaller device. Smaller devices, coupled with the increased use of smartphones to do more than make voice calls has has led to an increasingly unnatural experience of tapping input into the phone one finger poke at a time.

A shrinking phone interface still reliant on tactile input has led to an influx of users falling victim to 'fat finger syndrome.' This term was coined several years ago in reference to "the occasional tendency of stressed traders working in fast-moving electronic financial markets to press the wrong button on their keyboard and, in the process, lose their employer a mint[8]." Today, as more communication moves to personal devices, fat finger syndrome extends to everyone who types on a smartphone. Websites like www.fatfinger.com have popped up to document errors

---

[7] Intel. (2003). Moore's Law and Intel Innovation.

[8] McCurry, J. (2005, December 5).Too fat, too fast. The £1.6bn finger. *The Guardian.*

caused by the awkward experience of quickly inputting text on keyboards and consumer devices.

Likewise, the Autocorrect feature on many phones, most popularly, the iPhone, emphasizes problems that arise when users use a small touchscreen to communicate regularly. The website, DamnYouAutoCorrect.com is updated with submissions of texting faux pas daily, so many, in fact, that a book collection of the most embarrassing mis-texts was released earlier this year.[9]

Capitalizing on the existing voice recognition features of mobile phones is a natural evolution of the medium. Users are already familiar speaking to mobile phones.
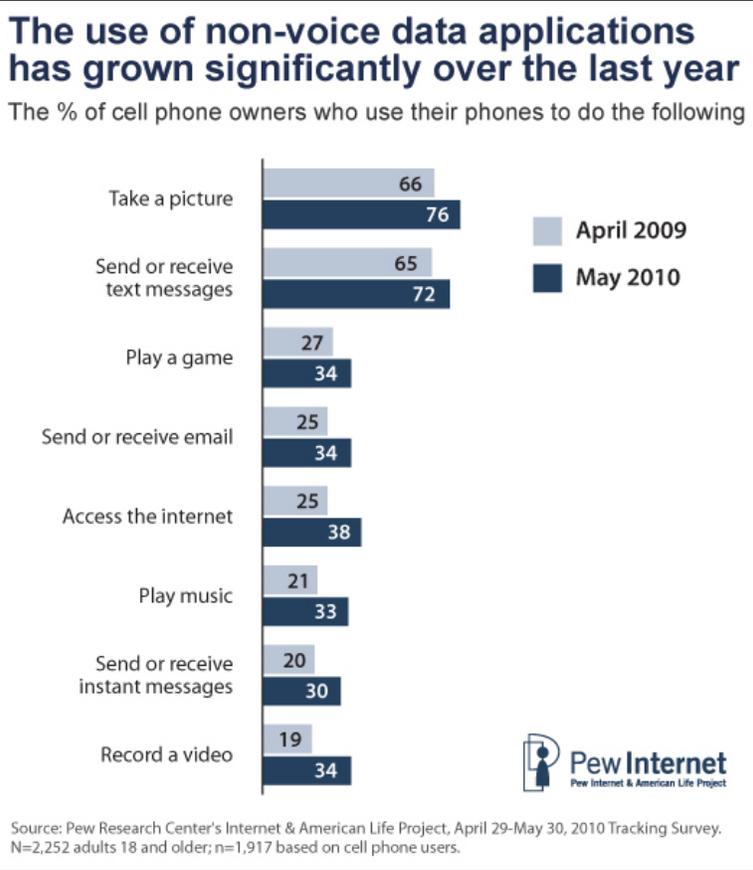
Increased Hands-free Productivity

Pew Research Center has found what we all have already noticed: people are using their cellphones for more than making voice calls.

The survey asked cellphone owners about usage of data applications including text messaging, taking and sending photos, accessing the internet, playing music and playing games or recording a video. Results of the research show that "compared with a similar point in 2009, cell owners ages 30-49 are significantly more likely to use a range of mobile data applications on a handheld device."[10]

---

[9] Madison, J. (2011, July 27). Frequently Asked Questions. Damn You Autocorrect!

[10] Smith, A. (2010, July 7). Mobile Access 2010.

The use of non-voice data applications has grown significantly over the last year

The % of cell phone owners who use their phones to do the following

Activity	April 2009	May 2010
Take a picture	66	76
Send or receive text messages	65	72
Play a game	27	34
Send or receive email	25	34
Access the internet	25	38
Play music	21	33
Send or receive instant messages	20	30
Record a video	19	34

Pew **Internet**
Pew Internet & American Life Project

Source: Pew Research Center's Internet & American Life Project, April 29-May 30, 2010 Tracking Survey.
N=2,252 adults 18 and older; n=1,917 based on cell phone users.

Across all areas of non-voice data applications, mobile phone owners are using their mobile phones more frequently, building demand for staying in contact even when performing other tasks such as driving. Further studies show that implementing a voice-based control for secondary tasks like "radio tuning, phone dialing, and more complex tasks involving a sequence of interactions with an in-vehicle computer system" decreased driver distraction[11].

Evidence of Future Adoption

---

[11] Harbluk, J.L. , Noy, I. Y. and Ranney, T.A. (2005). Voice Technology on Test Track Driving Performance: Implications for Driver Distraction.

According to PC World, such "voice-controlled contact and number dialing" has been standard feature on phones before the prevalence of smartphones. But "despite widespread availability, voice control never gained traction because the effort required to get it to work right wasn't worth it for most people. Voice control always required specific phrasing that sounded more like a command than natural speech. "Enunciating each word and number is a lot harder to do on a regular basis than to simply say, "Call mom"[12]

Siri solves the problem of being able to say, "Call mom." Perhaps the main differentiator of Siri to other smartphone voice technology is its ability to recognize naturally spoken language and adapt to it. Prior to its purchase by Apple, Siri was a spinoff of a technology project called Cognitive Assistant that Learns and Organizes, or CALO.[13] CALO is a project of Defense Advanced Research Projects Agency (DARPA) designed to be "an enduring personalized cognitive assistant" expected to "generate new algorithms and tools, and to yield new technology of significant value to the military." What does this mean for iPhone 4S users? Behind Siri is advanced artificial intelligence that was previously exclusive to the military.

Also, access to cloud storage significant improves the function of voice technology. Ilya Bukshteyn, senior director of sales and marketing of Microsoft's Tellme, notes that moving data processing the cloud "not only add[s] horsepower to the processing and provide[s] access to infinite databases of information; but it also

---

[12] deAgonia, M. (2011, November 27). After a Month, Siri Finds Her Voice. *PC World*.

[13] Mark, W., and Perrault, R.C. (2011). Cognitive Assistant that Learns and Organizes. In SRI International's Artificial Intelligence Center.

open[s] up user data that the scientists needed to improve applications[14]." Essentially, the cloud means faster processing for users and better feedback for engineers to improve recognition.

Potential Barriers to Mobile Voice Technology Adoption

Comparative to other general consumer spaces, the mobile industry shows the greatest potential for widespread adoption of Siri-like voice recognition technology in the near future. However, Apple has been clear to iterate this is a technology that is still in beta mode[15], with much progress to be made.

For example, the Siri can only understand English (United States, United Kingdom, Australia), French and German[16], and Apple notes, "since every language has its own accents and dialects, the accuracy rate will be higher for native speakers." Despite beta mode but to only understand native speakers of three languages eliminates a gigantic portion of the world. And what of the Americans or English iPhone lovers who speak English as a second language but neither French or German?

Added to the issue of localization, mobile applications like Siri are only effective when used within their data network. Drawing on the network connection to the cloud, once the application is outside the network, GPS and Web-based functionality falters.

_____

[14] Smith, N. ((2011, November 1). Siri Effect: Why Natural Voice Recognition Changes Everything. *Laptop Magazine*.

[15] Apple. (2011). iPhone: Siri.

[16] Apple. (2011). iPhone: Siri Frequently Asked Questions.

Evidence of Online & Social Media Voice Technology

A growing number of companies are integrating voice control and speech recognition features into everyday communication software.

Noteworthy innovators include:

- Dragon (www.nuance.com/dragon), a leader in natural language recognition, with applications in the general consumer space, medical industry and business.

- Microsoft TellMe (www.microsoft.com/en-us/tellme), software integrating mobile, desktop, online voice control functions. Microsoft has implemented speech recognition applications into its Windows operating system for years and is also building voice into its Xbox media consoles.

- Yap (www.yapme.com), recently acquired by Amazon, this voice-to-text speech recognition company offers voicemail transcription[17].

- Zypr (www.zypr.net), a free Web service of Pioneer that offers voice access to Facebook, mapping tools, calendar, etc.

- Angel (www.angel.com), social media company that recently unveiled "Voice For Twitter" and "Voice For Facebook," services enabling users to speak a status update or Tweet and Angel will convert it to text and publish it online[18].

---

[17] Boulton, C. (2011, September 11). Amazon Buys Yap, Joining Apple, Google in Speech Recognition. *eWeek*.

[18] Voice for Facebook, 2011.

Google is also making strides in the voice recognition technology, but with a more holistic approach to understanding natural language. Interviewing Google engineers and Princeton linguist and translator David Bellos, *Slate Magazine* writer Jeremy Kingsley found that Google is learning human context through the data gathered in online searches. For example, Google infers a search for "hot dogs" to mean an inquiry about the food rather than the animal.

"Google was the first to really put this idea to use, and it marked a significant step from computers reading strict syntax to reading the force of meaning with context-sensitive intelligence," explains Kingsley. "Today, the algorithm has an understanding of language something like a 10-year-old's, but its rate of improvement is fast exceeding human language-learning development[19]."

Computational linguistic engineers at Google are also working on projects such as the "'Poetic' Statistical Machine Translation," an effort to teach computers to translate and infer the meaning of poetry through the interpretation of rhyme and meter[20]. Voice is vital to achieving the Holy Grail of artificial intelligence: "actually teaching a computer to act like a human being[21]," says Chewy Trewhella, new business development manager at Google.

It seems logical to assume that these learned human semantics will soon serve as the "cloud-based brain" for improved voice recognition technology applications to come.

---

[19] Kingsley, J. (2011, October 31). Google Translate. Slate Magazine

[20] Genzel, D., Och, F., Uszkoreit, J. (2010, October 10). "Poetic" Statistical Machine Translation: Rhyme and Meter.

[21] Graham, F. (2011, February 10). Speak and spell: Can voice technology give your app the edge? *BBC News*.

Voice Technology Benefits for Web-based Services

Search engines like Google are amassing large amounts of data by the minute. From this data human context and semantics are inferred, creating a database of phrases, word pairings, idioms. Because the mobile device side of voice recognition technology fetches both translating functions and data from the cloud, it makes sense to have that database constantly updated. Secondarily, housing the database of voice translating algorithms and data in the cloud reduces the amount of bandwidth necessary to use a voice application on a mobile device.

The online publication, *Ars Tecnica*, performed an in-house test of bandwidth usage by Siri in one month. Journalists used the personal assist to perform various functions ranging from five Siri inquires a day up to 15. On average, it was found that regular use of Siri will use about 30MB of data per month. For reference, one hour of online streaming music requires about 16-18MB[22].

Role of Social Media in Voice Technology Adoption

One theory supporting a move to widespread adoption of voice technology is sociologic. Zeynep Tufekci, sociologist and Assistant Professor at University North Carolina Chapel Hill, argues that social media such as Twitter and Facebook incorporates etiquette and behavior of both written and oral communication. In response to the claim the social media and networking sites are deteriorating the

---

[22] Cheng, J. (2011, November 12). "Siri, how much data do you gobble up in a month?" Ars investigates.

quality of human communication and genuine relationship building[23], Tufekci cites the rise in communal communicating as evidence of a shift toward orality.

"What we are seeing with social media is the public sphere, hitherto dominated by written culture, has been more opened up to oral psychodynamics." She further explains, "the dual nature of Twitter, allows us to comment and respond and converse with others in real-time making it more similar to talking than it is publishing[24]." The relevance of this move toward comfort with orality that, as she stated, our society is traditionally based on written communication. In the technology realm, we have seen our advancements support the leaning toward a printed or linear word: book, typewriter, word processor, fax machine, Internet, email, text messaging, now social networking.

A 2010 study by *American Behavioral Scientist* found that our interactions with social media may be supporting Tufekci's theory. The relationships people form online tend to be a reflection of the relationships formed offline, integrating a conversational element to online-based networking. "Internet use has become normalized, with more people spending more time engaging in various activities via the Internet everyday and with the boundaries between online and offline ever blurring[25]."

In a society that has places equal, is not preferable value on an email over a phone call, a move in the reverse direction is notable. In terms of adopting a voice

---

[23] Keller, B. (2011, May 18). The Twitter Trap.

[24] Tufekci, Z. (2011, May 19). Why Twitter's Oral Culture Irritates Bill Keller (and why this is an important issue).

[25] Wang, H and Wellman, B. (2010). Social Connectivity in America: Changes in Adult Friendship Network Size From 2002 to 2007.*American Behavioral Scientist*.

recognition technology, comfort with an oral-based communication in place of, or even supplementing, text-based communication tech such as Twitter indicates a familiarity with technology-mediated communication that was not present in the recent past.

This becomes of particular importance when considering the anxiety factor of new technology adoption paired with the unnaturalness of speaking to a machine.

Potential Barriers to Web-based Voice Communication

Perhaps the greatest barrier is reliance on an Internet connection. Because the data processing side of voice technology is performed in the cloud, there is no functionality without access to the cloud. No internet, no use whatsoever.

As mentioned, voice technology is the basic function of smartphone technology. Computers, laptops, tablets are derivatives of the desktop computer, with focus and function reliant on typed input through keypads or touchscreen. Emphasis on tactile input and visual output creates a user experience dependent upon keeping the device distant from his or her face.

This creates two major problems for voice technology: first, the microphone on the device is farther from the speaker, allowing for increased ambient noise, distortion and otherwise lowered voice clarity. Second, given non-mobile devices include higher quality output speakers than mobile phones, a voice command response may broadcast beyond the user, creating issues of privacy and social etiquette.

The New York Times recently reported on the unease of many people witnessing others speaking to their iPhones in public. "When talking to their cellphones, people sometimes start sounding like machines themselves," it noted. Yet, despite the public nuisance, people could not help but eavesdrop on the conversations between people and their Siris[26]. The paradigm of public computing would require a significant alteration.

Imagine this problem exacerbated by the normal distance between users and their computers in a setting such as a classroom or office. According to Google Research Scientist Vincent Vanhoucke, "there is a huge divide between mobile voice technology and desktop. The mic is further from the voice, the mic is less sophisticated, there is more ambient noise." Referring to Google's voice control software, Vanhouke explains, "new algorithms had to be written that accounted for increased ambient noises and decreased voice clarity. It's only compounded by the thousands of languages and approximately 230 billion words Google Search by Voice will eventually have to deal with[27]." Until then, users may have to speak more clearly and perhaps more loudly to their computers and still face the risk that software may not understand their commands.

There is also the issue of how humanlike users need a voice to be in order to encourage natural language input. Dr. Thomas Hempel, an engineer for Siemens, explains in his 2008 publication, *Usability of speech dialog systems: listening to the target audience*,

---

[26] Wingfield, N. (2011, December 2) Oh, for the Good Old Days of Rude Cellphone Gabbers. *The New York Times*.

[27] Parr, B. (2011, July 7). How Google Is Leading the Way to a Voice-Activated Future.

"if a persona design of a speech system is very realistic (humanlike) some users behave in a way that decreases the dialog success. It is known that users who do not respect the limitations of the system decrease performance. In our studies, some users overestimated the speech dialog system and expected it to behave like a real person. The recognition problems increased and the users became disappointed. if this assumption in evidenced one could conclude that it is a disadvantage if a speech dialog system is designed to be too human-like[28]."

Dr. Hempel's research is supportive of the "uncanny valley" hypothesis in the robotic field, which claims that when human replicas are made to appear too much like a human being, human observers expressed a severe drop in likeability, becoming repulsed by the humanoid[29]. Both Hempel and robotics researchers agree that up to the point of being "too human-like" human users of artificial intelligence (voice technology included) experience positive feelings of empathy and likeability for the technology.

Privacy Concerns

An integration of voice technology, especially on a personal device, comes with its own set of privacy concerns. Though not the focus of this research, privacy is worth noting as a barrier to large scale adoption. Would companies like Apple have

---

[28] Hempel, T. (2008) Usability of speech dialog systems: listening to the target audience. p.70.

[29] Brown, M. (2011, July 19). Exploring the uncanny valley of how brains react to humanoids. Wired UK.

ownership of every voice clip we speak to Siri? Considering the amount of personal data companies like Google, Apple, Amazon already maintain on their users, coupling this data with a personal voice seems one step from recreating an entire user digitally. Would voice clips be stored like text-based data? Can that be anonymized?

If the cloud-reliant Amazon Silk is any indication[30], there will be some push back on implementation of cloud-reliant voice software as well.

Future of Voice Technology

Everett Rogers developed a theory of diffusion to explain the process of exposing users to new innovations before widespread adoption. Rogers' theory is dependent on five levels of end user interaction: relative advantage, compatibility, complexity, trialability and observability[31].

Relative advantage

Does voice recognition technology - in the general consumer market - offer a relative advantage of existing, widespread technology? The answer is both yes and no, depending on the specific application of the voice technology. As discussed, the application for voice technology when applied to activities where it can minimize distraction, such as while driving, is an advantage. Assuming the vehicle has a built-in voice control system and easily accessible activation button, drivers agree to the relative advantage over hand controlled technology.

---

[30] Mello, J.P. (2011, October 15). Congress Targets Amazon's SIlk Browser. *PC World*.

[31] Grant, A.E. and Meadows, J.H. (2010). Communication Technology Update and Fundamentals, (12 Ed.). pp. 48-49.

However, when applied to social media, online functions and virtual personal assisting, users must initiate the voice technology with a button pressed by the hand or with another touch-sensitive feature. In this case, the relative part of relative advantage becomes subjective: it may be relatively easier to continue using touch to operate the device due to proximity of the fingers to the touchscreen. Or, it may not because of the inherent unnaturalness of typing.

Compatibility

According the voice recognition/voice control applications available in the consumer market today, we can see the technology is generally compatible with existing consumer technology such as email, Web browsing, calendars, etc.

Complexity

Is the technology very difficult or complex for end users? Here we run into another possible yes and no answer. Siri again is an example of a more user friendly, therefore, less complex, voice technology. The personal assistant is built in to the iPhone 4S so users can begin to use the feature out of the box. With other applications, both in the mobile market and personal computer/device space, the user must first seek out the voice technology of choice, download and/or purchase it. Then, there may be a process for initiating the software on the machine, activating specific functions and training it to personalized commands.

To install and configure the speech recognition feature of Microsoft XP, for example, requires several, multiple-step processes, the explanation for which spans thirteen pages of text[32].

Trialability

Users must be able to use a new technology on a trial basis before deciding to adopt it. Although much of the software and mobile applications available today are not especially intuitive, many do offer a free or trial version. Usually these versions, such as the Dragon Go! variant of the Dragon Dictation mobile application, lack the sophistication and usefulness of the paid application, but nonetheless provide a glimpse of the technology.

Likewise, Siri is a beta software, arguably a trial version of voice technology yet to be released by competitors Google, Amazon, Dragon, etc.

Observability

Are others using the new technology? Early adopters have been using Dragon for years and other had downloaded Siri before it was commandeered by Apple. In niche groups, usually those of early technology adopters or users who used voice technology as a speech replacement, the consumer application has been observable. Again, on this point I refer to Siri. Apple exposed voice technology to a wide audience, encompassing both early adopters and early majority adopters[33]. In doing this, more

---

[32] Microsoft. (2011, September 23). Microsoft Support: How to install and configure speech recognition in Windows XP.

[33] Grant, A.E. and Meadows, J.H. (2010). Communication Technology Update and Fundamentals, (12 Ed.). pp. 49-50.

of the general public is able to observe the technology in use more frequently, encouraging others to use it as well.

The next step of Diffusion Theory, and final step for most new communication technology is the decision making process by a potential user. This step is comprised of knowledge, persuasion, acceptance or rejection, implementation and confirmation. Again using Siri as an example, a potential user must be aware of the personal assistant technology. Then, he or she must form positive attitude toward Siri, leading to acceptance of the iPhone 4S. Finally, if the user is please with this choice, confirmation has been reached and the innovation is adopted.

Feedback via blogs, media reporting and reported sales show that users are in the range of acceptance and implementation for Siri[34]. This is a positive sign for the consumer voice technology space in general, as a move into confirmation of Siri indicates a move toward majority adoption.

Conclusion

If Rogers theory is accurate, greater voice technology is poised to infiltrate consumer technology very soon. The technology today is far more sophisticated than voice recognition in the past, plus the power and nearly limitless storage capacity of cloud computing allows for huge databases of ever-growing vocabulary.

It will be interesting to observe what technologies will build off the framework of voice technology once it reaches critical mass. *The Boston Globe*[35] just reported

---

[34] Apple. (2011, October 10). iPhone 4S Pre-Orders Top One Million in First 24 Hours.

[35] Stelter, S and Wingfield, N. (2011, December 5). Xbox to take on cable box duties. *The Boston Globe*.

that Microsoft will be including voice controls into its Xbox system, which will also integrate mainstream television programming, controlled by voice and gesture. Together with the Internet buzz that the next generation Kinect controller will be able to read lips and understand facial expressions[36], we may see voice technology as a gateway to completely touch-less computing.

---

[36] Yin-Poole, W. (2011, November 28). Kinect 2 so accurate It can lip read. Eurogamer.

References

Apple. (2009). *A Brief Vocoder History.* Logic Express 9. Retrieved from http://documentation.apple.com/en/logicexpress/instruments/index.html#chapter=7%26section=10%26tasks=true.

Apple. (2011). iPhone: Siri. Retrieved from http://www.apple.com/iphone/features/siri.html.

Apple. (2011). Siri Frequently Asked Questions. Retrieved from http://www.apple.com/iphone/features/siri-faq.html.

Apple. (2011, October 10). iPhone 4S Pre-Orders Top One Million in First 24 Hours. Retrieved from http://www.apple.com/pr/library/2011/10/10iPhone-4S-Pre-Orders-Top-One-Million-in-First-24-Hours.html.

Boulton, C. (2011, September 11). Amazon Buys Yap, Joining Apple, Google in Speech Recognition. *eWeek.* Retrieved from http://www.eweek.com/c/a/Mobile-and-Wireless/Amazon-Buys-Yap-Joining-Apple-Google-in-Speech-Recognition-552139/.

Brown, M. (2011, July 19). Exploring the uncanny valley of how brains react to humanoids. *Wired UK.* Retrieved from http://www.wired.co.uk/news/archive/2011-07/19/uncanny-valley-tested.

Cheng, J. (2011, November 12). "Siri, how much data do you gobble up in a month?" Ars investigates. Ars Tecnica. Retrieved from http://arstechnica.com/apple/news/2011/11/how-data-heavy-is-siri-on-an-iphone-4s-ars-investigates.ars.

deAgonia, M. (2011, November 27). After a Month, Siri Finds Her Voice. *PC World.* Retrieved from http://www.pcworld.com/article/244840/after_a_month_siri_finds_her_voice.html#tk.rss_news.

Genzel, D., Och, F., Uszkoreit, J. (2010, October 10). "Poetic" Statistical Machine Translation: Rhyme and Meter. Retrieved from http://static.googleusercontent.com/external_content/untrusted_dlcp/research.google.com/en/us/pubs/archive/36745.pdf.

Graham, F. (2011, February 10). Speak and spell: Can voice technology give your app the edge? *BBC News.* Retrieved from http://www.bbc.co.uk/news/business-12418373.

Grant, A.E. and Meadows, J.H. (2010). *Communication Technology Update and Fundamentals, (12 Ed.).* Burlington, MA: Elsevier.

Harbluk, J.L. , Noy, I. Y. and Ranney, T.A. (2005). Voice Technology on Test Track Driving Performance: Implications for Driver Distraction. Human Factors: The Journal

of the Human Factors and Ergonomics Society. 47(439) [Abstract]. DOI: 10.1518/0018720054679515.

Hempel, T. (2008). *Usability of speech dialog systems: listening to the target audience*. Erlangen, Germany: Springer.

Huerta, J. and Pieraccini, R. (2005, September 3). Where do we go from here? Research and commercial spoken dialog systems. http://www.isca-speech.org/archive_open/sigdial6/sgd6_001.pdf.

InfoWorld. (1980, October 27). Typewriters Learn to Listen. *InfoWorld,* 2(19), 22.

Intel. (2003). *Moore's Law and Intel Innovation.* Retrieved from http://www.intel.com/about/companyinfo/museum/exhibits/moore.htm.

"Justin." (2011, August 2011). Millennial Media Compares 2 Years of Mobile Advertising Trends In 50th S.M.A.R.T Report. Mobile Marketing Watch. Retrieved from http://www.mobilemarketingwatch.com/millennial-media-compares-2-years-of-mobile-advertising-trends-in-50th-s-m-a-r-t-report-18060/.

Keller, B. (2011, May 18). The Twitter Trap. *The New York Times*. Retrieved from http://www.nytimes.com/2011/05/22/magazine/the-twitter-trap.html?_r=1.

Kingsley, J. (2011, October 31). Google Translate. *Slate Magazine*. Retrieved from http://www.slate.com/articles/technology/technology/2011/10/google_translate_will_google_s_computers_understand_languages_be.single.html.

Madison, J. (2011, July 27). Frequently Asked Questions. Damn You Autocorrect! Retrieved from http://damnyouautocorrect.com/10785/damn-you-autocorrect-frequently-asked-questions/.

Mark, W., and Perrault, R.C. (2011). Cognitive Assistant that Learns and Organizes. In SRI International's Artificial Intelligence Center. Retrieved from http://www.ai.sri.com/project/CALO.

McCurry, J. (2005, December 5).Too fat, too fast. The £1.6bn finger. *The Guardian*. Retrieved from http://www.guardian.co.uk/business/2005/dec/09/japan.internationalnews.

Mello, J.P. (2011, October 15). Congress Targets Amazon's SIlk Browser. PC World. Retrieved from http://www.pcworld.com/article/241989/congress_targets_amazons_silk_browser.html.

Microsoft. (2011, September 23). *Microsoft Support: How to install and configure speech recognition in Windows XP.* Retrieved from http://support.microsoft.com/kb/306537.

Parr, B. (2011, July 7). How Google Is Leading the Way to a Voice-Activated Future. *Mashable*. Retrieved from http://mashable.com/2011/07/07/voice-technology-google/.

Paul, L.R. (2005). *Made in the twentieth century: a guide to contemporary collectibles*. Lanham, Maryland: Scarecrow Press.

Smith, A. (2010, July 7). Mobile Access 2010. Retrieved from http://pewinternet.org/Reports/2010/Mobile-Access-2010.aspx.

Smith, N. (2011, November 1). Siri Effect: Why Natural Voice Recognition Changes Everything. *Laptop Magazine*. Retrieved from http://blog.laptopmag.com/siri-effect-why-natural-voice-recognition-changes-everything.

Stelter, S and Wingfield, N. (2011, December 5). Xbox to take on cable box duties. *The Boston Globe*, p. B7.

Tufekci, Z. (2011, May 19). Why Twitter's Oral Culture Irritates Bill Keller (and why this is an important issue). Technosociology.org. Retrieved from http://technosociology.org/?p=431.

Voice for Facebook. (2011). Angel. Retrieved from http://www.voiceforfacebook.com/.

"VoicesofDragon." (2011, October 24). The evolution of voice technology. Retrieved from http://voicesofdragon.com/2011/10/24/the-evolution-of-voice-technology/.

Wingfield, N. (2011, December 2) Oh, for the Good Old Days of Rude Cellphone Gabbers. *The New York Times*. Retrieved from http://www.nytimes.com/2011/12/03/technology/virtual-assistants-raise-new-issues-of-phone-etiquette.html.

Wang, H and Wellman, B. (2010). Social Connectivity in America: Changes in Adult Friendship Network Size From 2002 to 2007. *American Behavioral Scientist 2010; 53; 1155-1164*. DOI: 10.1177/0002764209356247.

Yin-Poole, W. (2011, November 28). Kinect 2 so accurate it can lip read. *Eurogamer*. Retrieved from http://www.eurogamer.net/articles/2011-11-25-kinect-2-so-accurate-it-can-lip-read.