Dear Apple: Go Big with Siri and Nuance in iOS 5

Gary Morgenthaler | 6/4/11

As someone who served on the boards of **Siri** and Nuance—two companies whose technologies are rumored to be deeply integrated into Apple's upcoming iOS 5—I might have a unique take on what's possible at the upcoming Apple World Wide Developer Conference (WWDC) in San Francisco June 6-10. I throw my speculations into the ring with so many others, informed not by any privileged knowledge (the product team is appropriately tight-lipped), but by my long-time involvement with speech recognition and artificial intelligence (AI). With Siri and Nuance, Apple can leapfrog Google's Android in the short-term. Or, they can do something "Apple-esque" and fundamentally change the rules of the game and reclaim leadership for good.

From a standing start three years ago, Android has managed to overtake iOS as the leading smartphone platform. Today, Google has 35 percent (versus Apple's 26 percent) of the U.S. smartphone market, according to a **comScore** report for April 2011. How Apple uses Siri and Nuance in their next move will have important repercussions for the company as well as the developer, technology, and user communities worldwide.

What's Rumored

Of all the rumors coming out of the rumor mill, the following is most pertinent for this conversation: It is widely reported that Apple is working on a licensing agreement with Nuance to embed speech recognition capabilities deeply into iOS 5. It is also speculated that Siri, a natural language AI technology acquired by Apple in April 2010, will figure prominently in iOS 5. In addition, Apple may be working on a "world phone" or "iPhone 4S or 5″ which would be compatible with both GSM and CDMA—which means it would work on almost all of the world's cellular networks.

What's Possible

If these rumors are true, Apple is in an advantageous position. But it remains to be seen have far they will push their advantage. For example, Apple can either integrate Nuance and Siri a little—or a lot. I say, go big. Don't use the incredible power of these two best-in-class technologies to manage Apple-only applications. Sure, it would be fun to say, "Open iTunes. Play *Born this Way* by Lady Gaga," instead of typing it out. But it would be far more transformative to open up the API to allow all of Apple's 100,000-plus registered developers to dream up ways to use voice recognition and natural language AI in their own third-party apps.

Die-hard Apple fans fear a repeat of history in which Apple chooses to control innovation. That didn't work out so well for Apple back in the 1980s and 1990s when Microsoft ate their lunch by opening their hardware platform and encouraging a proliferation and domination of Windows-compatible systems. I don't believe that will happen again; but it is possible.

A quick primer on what Nuance and Siri can do will shed light on why I am so fascinated by their achievements. Both born out of the prestigious **SRI International** technology R&D center in Menlo Park, CA, they serve different parts of the spectrum in which your verbal intent is translated into action.

For example, you say, "Book me a table for two at **Il Fornaio** in Palo Alto at 6 p.m. tonight."

Nuance would perform speech recognition and parse out every sound in that phrase. It would map sounds into syllables and syllables into words. This is a non-trivial task and Nuance does it extremely well (so well the company has a market cap of more than $6 billion today). It would then present those sounds as transcriptions. But it ends there. It does not understand those transcriptions.

Siri is far better at understanding. It's a gifted natural language AI technology that knows the myriad of ways people express intent. Siri knows that by "book" you don't mean a paperback novel but the action to "reserve" a table. It adds a deep layer of intelligence on top of Nuance and helps turn your spoken intent into action. It can go right to **OpenTable** and reserve a table for you. With one verbal command, you can skip the thumb typing and avoid the three Web screens to complete your task.

What's Apple-esque

It would be quintessential Apple to rewrite the rules and surprise the world—like they did with the iPad. Apple owns nearly 90 percent of the tablet market today because they dictated what a tablet should be—with its industrial design, elegant UI, treasure chest of applications and an entire supporting ecosystem.

They can also dictate what a smartphone should be. It must have deeply embedded voice recognition and artificial intelligence capabilities, a bevy of related applications, and the ability to run across the world's telecommunications networks. (Nuance, by the way, speaks 56 languages and dialects to date.) Apple should offer APIs that let developers tap directly into Nuance's technology to speech-enable more types of interactions, and into Siri's technology to translate users' voice commands into practical actions.

That's how Apple can deliver on the promise of Siri—by leveraging speech recognition and natural-language understanding to make the mobile platform more productive and capable than a keyboard-enabled, big screen system allows. In this way, Apple can give its own developer community powerful new tools, reclaim the high ground, and make it exceedingly difficult for others to follow.

Gary Morgenthaler is a partner at Morgenthaler Ventures in Menlo Park, CA.