



Home > Topics > Developer & Web Services > Opinions > **Speech Server**  
**Launch Re-asks Old Questions**

ADVERTISEMENT

## Developer/Web Services



### Speech Server Launch Re-asks Old Questions

Coffee

By [Peter Coffee](#)

March 31, 2004



**TalkBack:**  
Sound off on  
this article

Is speech recognition a stupid computer trick, or a much-needed feature that finally works? That's the question that developers need to answer in the wake of last week's **launch of Microsoft's Speech Server 2004**.

The question must quickly be narrowed down into more specific terms. Speech recognition **can mean almost anything**. At one extreme, it might be only the speaker-independent recognition of a tiny vocabulary of individual words: "Press or say the number '1,'" as the voice-response systems on call-handling systems tell us. At the other extreme is the still unrealized goal of a system that can carry on a conversation with anyone who can be understood by another human being, like the fictional **HAL 9000**. (See the MIT Press book, "**Hal's Legacy**").

Customers appear to be growing in their acceptance, and even approval, of speech recognition technology at their point of first contact with an enterprise. Given the alternatives of long waiting times for a live attendant, or cumbersome keypad interaction, **more than two-thirds of customers** now appear to find speech recognition both convenient and effective.

The most common criticism of speech recognition as an **enterprise productivity tool** is a matter of scale. One person demonstrating speech-command systems to co-workers may be impressive; several dozen people in a crowded office bay, all talking to their computers, may be oppressive or even intolerable. Fortunately, the most **recent research** suggests that it may not be necessary to speak out loud—not even at the volume required by a headset microphone—to use speech as a command interface.

Workers are also spending more of their time, thanks to wireless

networks, away from their desks and closer to the problems that they're solving, whether that means being on the road or on the factory floor. Using voice input in these environments may be not only convenient, but also **mandatory** as a matter of safety. If background noise is overwhelming, **lip-reading algorithms** may make the difference. Location information, derived from GPS or other technologies, can also provide **valuable clues** to what a user might be intending to say.

Even so, I met last week with representatives of **Applied Voice & Speech Technologies Inc.**, who made a worthwhile point: being able to recognize speech accurately is not the same thing as having a well-designed speech-driven application, any more than having accurate keyboard or mouse-based input means having a good user interface design.

Microsoft's Speech Server, like Microsoft's Windows, certainly gives developers a rich environment in which to ply their craft, and I'm certain that we'll see Microsoft offering **industry-leading tools** toward that end. Open industry standards will also play an important role (the W3C now defines a **syntax** for representing grammars), but I'm certain that Microsoft's own applications will set a high bar that independent developers will be challenged to jump—and that alternative speech-oriented platform providers, like **Apple**, **Opera** and **IBM**, will have to face in the marketplace.

And that reminds me of a **comment I made** concerning Microsoft's proposal to settle its U.S. antitrust suit, back in 2001: "What about the proposed clause that defines a personal computer as one that has a keyboard? Did no one involved in this settlement ever see a Tablet PC demonstration? Look, Ma, no keys!... Three years from now, how many people will use voice-response systems in their cars as their mobile gateways to the Internet?"

That clause is still there in the **Final Judgment** of the case: "VI. Definitions... Q. 'Personal Computer' means any computer configured so that its primary purpose is for use by one person at a time, that uses a video display and keyboard... Servers, television set top boxes, handheld computers, game consoles, telephones, pagers, and personal digital assistants are examples of products that are not Personal Computers within the meaning of this definition."

Those issues associated with the Microsoft antitrust litigation come bubbling to the surface with last week's **European Union action**: for whatever it's worth, my past analyses of key documents and decisions in that case are now linked from a **single page** with some introductory comments as part of eWEEK's ongoing **special report** on the company's legal trials.

I wish this sort of thing weren't important to our technology choices, but it looks as if speech-based devices may compel us to ask and answer some of these questions all over again.

***Speak to me about your platform hopes and fears at [peter\\_coffee@ziffdavis.com](mailto:peter_coffee@ziffdavis.com).***



TALKBACK

[Sign In To Talk Back!](#) | [Register](#)