

- SHOPPING CART
- LOG IN
- MEDIA CENTER
- MAILING LIST
- CONTACT US
- HOME

[EVENTS](#)
[TRAINING](#)
[CONSULTING](#)
[NEWSLETTERS](#)
[WEBCASTS](#)
[BLOGS](#)

- [Magazine](#)
- [New Customer Management Insight](#)
- [Digital Publication](#)
- [Call Center Magazine](#)
- [Archives](#)
- [FREE Webcasts](#)
- [Podcasts](#)
- [Vendor Guide](#)
- [Multimedia Content](#)
- [Online Advice](#)
- [eNewsletter Alerts](#)
- [ICMI Forum](#)
- [ICMI Jobs](#)
- [Training](#)
- [Consulting](#)
- [Newsletters](#)
- [Publications/Tools](#)
- [Industry Research](#)
- [QUEUETIPS](#)
- [Conferences](#)
- [ICMI Global Report](#)
- [Call Center Knowledge Online](#)
- [Join ICMI](#)
- [Members Only](#)
- [Vendor Guide](#)
- [Industry Information](#)
- [About ICMI](#)
- [Download the ICMI Catalog](#)

## Multimodal Toolkits: Design, In - Speech Tags and ML, Out

**Want to write mobile data applications you can talk to? While 2.5 and 3G wireless rolls out and proves itself, check out these multimodal application development tools.**

By **Ellen Muraskin**  
 01/06/2003, 10:39 AM ET

As with any multimodal strategy, deployment rests on adding speech-awareness to the browser. IBM's first extension, including IBM TTS and ASR, is a plug-in for IE 5.5 or above on a Windows 2000 or XP platform. Qualified developers can download it from [www.alphaworks.ibm.com/tech/mmbextension](http://www.alphaworks.ibm.com/tech/mmbextension). With this, multimodal apps can be tested on PCs before the mini-browsers of pocket PCs are similarly equipped.

Igor Jablokov, project manager for IBM's Multimodal team, demonstrated such an application on his laptop, pushing the Scroll Lock button to emulate the push-to-talk button Pocket PCs will use to activate "listening" for ASR input. It's a pizza-ordering app that accepts size and various toppings via check box or ASR, and speaks aloud the same questions shown on the page. Jablokov demonstrates that it works as well on Opera's (Oslo, Norway - +47 24 16 40 00, [www.opera.com](http://www.opera.com)) lightweight, multi-device browser, with an eye toward taking multimodality to non-Windows PDAs.

Jablokov also makes the point that the X+V application will run without modification in single modalities as well: Call the app on the phone, and it will simply execute the VoiceXML modules. Browse to it, and you'll see the other half.

IBM's SDK comes with a grammar tool that automatically formats entered text in JSGF, and renders it in the International Phonetic Association's fairly self-evident and tweakable format.

► UTILITIES

- [print this article](#)
- [e-mail this article](#)

► RELATED LINKS

- [Cicero Broadens Support of Web Services and SOA](#)
- [Executive Interview: Intervoces Ken Goldberg](#)
- [Speech Makes Inroads As A Service](#)
- [Q&A: Authentication Technology for Call Center Security](#)
- [Why Mobility Will Change Telephony](#)
- [Are We Finally Ready For Remote Agents?](#)
- [Products of the Year: These Are the Sharpest Knives in the Drawer](#)
- [Aspect and Microsoft Create Speech Application](#)

### Free CallCenter Insider Newsletter

### Your Email Address

### Optional Areas of Interest

- International New s
- Advice/Tips
- Technology
- Agent Development
- IVR

**TechEncyclopedia**

**Define It**

For now, ASR and TTS run completely on the client, but IBM does intend to slim it down and enable distributed ASR and TTS for future, higher bandwidth and less expensive devices. Fat clients are preferred by the device manufacturers, who want to build more features into their products; distributed speech processing is preferred by carriers, who want to incur more network usage.

[Contact Routing:  
Your Biggest Bang  
For the Buck](#)

[The Fading Away of  
the TUI](#)

► TRAINING  
► View seminar calendar

On the platform side, IBM plans to include X+V interpretation into its WebSphere Everyplace Access mobility platform, with release aimed for the second half of 2003. The toolkit is slated to go GA second quarter. On the partner end, IBM has announcements with endpoint device and speech platform manufacturers. The latest concerns Sharp's Java-and-Linux-based Zaurus PDA, which will be added as a client to WebSphere Everyplace Access. IBM's middleware will extend multimodal browsing to the Zaurus, and will be offered alongside preexisting mobile PIM, email, and Lotus clients.

Like Kirusa, IBM promises interoperability with preexisting VoiceXML platform makers for the voice component. Chief announcement here is with Nuance. IBM's own voice server should be rechristened by presstime as WebSphere Voice Application Access.

## Lobby7

We ran out of time and space to cover Lobby7 (Boston, MA - 617-275-7100, [www.lobby7.com](http://www.lobby7.com)) in last month's article, but they've been around as long as Kirusa, formed in June 2000 by MIT grads as a wireless services company. The name is an MIT insider reference; Lobby 7 is the infamous location where first-year students assembled on arrival.

Lobby7 makes a multimodal application server, now up to V. 2, called x|mode, and an authoring tool with pre-written modules. Their concentration is simultaneous multimodality on 2.5 and 3G networks, and they have some carrier customers, not yet announced.

X|mode runs on J2EE web application servers and comes shipped with the Tomcat web server. It works with VoiceXML and XHTML markup, but Hugo Barra, Vice President, says that they're preparing for SALT compatibility as well.

This seems to us a pure marketing decision, for two reasons: for one, Barra himself notes that they see their developer as one who is well-versed in designing speech applications, as voice is the more difficult part of the UI design. SALT, in contrast, is being pitched to web designers to add speech accessibility to their sites. Secondly, as Barra states, Lobby7 is committed to as network-centric an architecture as possible, keeping multimodal clients thin.

Barra points out that the design of a multimodal application is a "combinatorially explosive task," illustrating with a simple voice-in, screen-out address input task. This may require one or two street address fields,

states to be input by voice or drop-down menu, and other permutations of voice or stylus entry and screen or voice output. For now, this complexity precludes an app gen simplified to the drag-and-drop level.

Lobby7's authoring tool resembles a text editor, with "deep knowledge of tags." But chunks of XML code that come with x|mode do generalize a lot of tasks that only need customization for particular apps and data sets. Map access, for example (tap stylus at a point and say "zoom"), is also among the included functions. The x|mode platform works with third-party VoiceXML browsers; they've also been using both Nuance and SpeechWorks core speech engines.

### **VoiceXML Development Tools**

Back on the mono-modal side, many VoiceXML development choices remain. Most of these mostly online IDEs were started by VoiceXML interpreter platform companies in an effort to grow a developer community. Most are free for the registration, but don't count on much tech support; folks still employed at voice platform companies are busy finding and feeding their paying customers or IVR tenants. In any case, make sure they're up to VoiceXML 2.0 before you mount the learning curve.

VoiceXML development tools fall into two types, according to SpeechWorks' Rob Kassel. Some, like that of VoiceGenie, IBM and Nuance Web Server, are syntax-savvy text editors that can anticipate keystrokes in entering code, format with color and indentation, and perform some syntax checking. Other tools, like Audium's and General Magic's, hide the code under drag-and-drop GUIs and fill-in boxes.

Core speech technology vendor Nuance (Menlo Park, CA - 650-847-7839, [www.nuance.com](http://www.nuance.com)) was first to demonstrate a GUI overlay to VoiceXML, over three years ago. Though the voice web server component, Nuance Web Server, did catch on in the nascent marketplace, the V-Builder app gen didn't. Nuance is rolling V-Builder up inside a new offering, Nuance Voice Platform (NVP), for 2003.

NVP will also come with an app gen geared more to the traditional IVR base. Although this tool will produce VoiceXML code, says Marcello Typrin, Product Marketing Manager, it will be a drag-and-drop GUI and will come from a third party, as yet unannounced. (Could that be General Magic, of blessed memory? Their intellectual property has to go somewhere ...) The new platform also adds access to a more granular level of application management, and runs on leading web application servers like BEA and Tomcat, enabling pre-packaged interfaces to broad range of CTI servers, CRM packages, and databases. Nuance is additionally preparing prepackaged applications for the platform, the first to be available in Q12003. Finally, conspicuous by its absence from the SALT Forum, Nuance is cooperating closely with IBM in its X+V approach to multimodality.

							
---	--	--	--	--	---	---	---