(This article first appeared in **Dr**. **Dobb's Journal**, a publication for developers. It was a side-bar article to a speech recognition theme. SALT [speech application language tags] was Microsoft's XML-based tool for encorporating their speech application recognition integration [SAPI, their speech-to-text and text-to-speech handling in Windows] to Windows-based solutions.)

SALT and VoiceXML

Speech application language tags (SALT) and VoiceXML are both XML solutions designed for telephony applications. The obvious question is, are they competing standards? The short answer is no. SALT does not intend to replace VoiceXML. VoiceXML is an excellent language for telephony-only systems. On the other hand, SALT provides features that VoiceXML currently does not. There are four main differences between the two languages.

• Multimodal support. SALT supports multimodal operations. Multimodal features allow more than one type of input, or mode. That means the device's display, graphics, buttons, stylus, or any other input/output capabilities may be used in combination with conventional speech to navigate a Web site. Multimodal capabilities become significant with proliferation of smart phones, personal digital assistants (PDAs), tablet PCs, and other mobile communication devices. VoiceXML is intended for telephony applications only and there is no clear way to extend it with multimodal capabilities.

• Programming model. SALT presents an unobtrusive and simple programming model. SALT extends HTML by only a handful of tags and adopts current Web programming and execution models. For HTML, SALT uses existing script and event models that are supported by visual markup browsers. As a result, developers have a rapid learning curve because they use their existing skills and knowledge. In contrast, VoiceXML uses over 40 tags and creates an entirely new markup language, which developers must learn.

• Programming logic. SALT adheres fully to the declarative principles of XML and uses standard XML and Web programming mechanisms. SALT is event driven so that program flow is convenient and encapsulated into objects. VoiceXML is an awkward use of XML principles. It introduces <if> and <goto> program logic, elements otherwise foreign to XML and HTML programming. The fixed flow and system-driven programming logic of VoiceXML constrains users and ultimately limits the company's telephony system overall. Its mixture of markup and script makes the code less readable and the control flow less transparent.

• Reusing code. As a result of SALT utilizing the three processes described above (multimodal support, programming model and programming logic) developers can easily reuse SALT code as well as port it to other devices That is, a SALT application designed for

the Pocket PC can be easily ported to a tablet PC. As stated earlier, SALT supports a new functionality such as adding multimodal features to an existing Web application. This also means that existing development environments and authoring tools can be used to generate SALT applications. Because of the differences between VoiceXML and accepted Web programming logic, custom VoiceXML tools need to be written.