



Voice-Enabling Forms-Based Applications

A Vanguard Voice Whitepaper

For More Information Contact:

Phil Walker
Vanguard Voice Systems, Inc.
22382 Ave Empresa
Rancho Santa Margarita CA 92688
949.888.5262
pwalker@vanguardvoice.com

Introduction

When portable computing devices (eg: PDAs, laptops, tablets) are used in a hands-free / eyes-free environment, the ability to communicate using speech will supply a compelling advantage over use of supplied styli or add-on keyboards.

To apply speech as a truly useful business tool, the following capabilities are required:

- The ability to navigate thru all “forms” of the application, including taking into account possible hierarchies of forms and sub-forms.
- The ability to provide users with “prompts” to guide them thru forms absent visual reference.
- The ability to “fill in” fields in a form by speaking an acceptable word, phrase or number.
- The ability to update lists of acceptable words, phrases, or numbers on a dynamic basis.
- The ability to define lists of acceptable words, phrases or numbers and relate them to a specific field in a form.
- The ability to capture long strings of speech (such as subjective observations, eg: “The patient presented with purple and green spots on her lower left ankle...”) for later transcription or processing.
- The ability to manage the above capabilities in a multi-user environment.
- The ability to add the above capabilities to existing or newly created applications.

Up until now, adding these speech capabilities to existing applications (even building them into newly designed applications) has been a daunting task, typically requiring weeks-to-months of effort. Vanguard Voice Systems, Inc. has addressed this difficulty, producing an integrated suite of development tools that allow the job to be done in hours-to-days.

This white paper describes **Vanguard’s AccuSPEECH® FormSuite**, which allows enterprises to easily add speech to their forms based applications.

Structure

AccuSPEECH FormSuite is a development and runtime environment. It consists of two required components and an optional component. The required components, discussed more fully below, are:

- AccuSPEECH Client – which resides in each device to running the voice-enabled application
- AccuSPEECH Studio – a graphic development environment which automates and simplifies most of the tasks involved in voice-enabling applications.
- An optional component, AccuSPEECH Configuration Server, is used in large enterprise applications where dynamic updating of lists is required. As an example, in a large medical center, patient lists have to be updated frequently. The Configuration Server is discussed in a separate white paper.

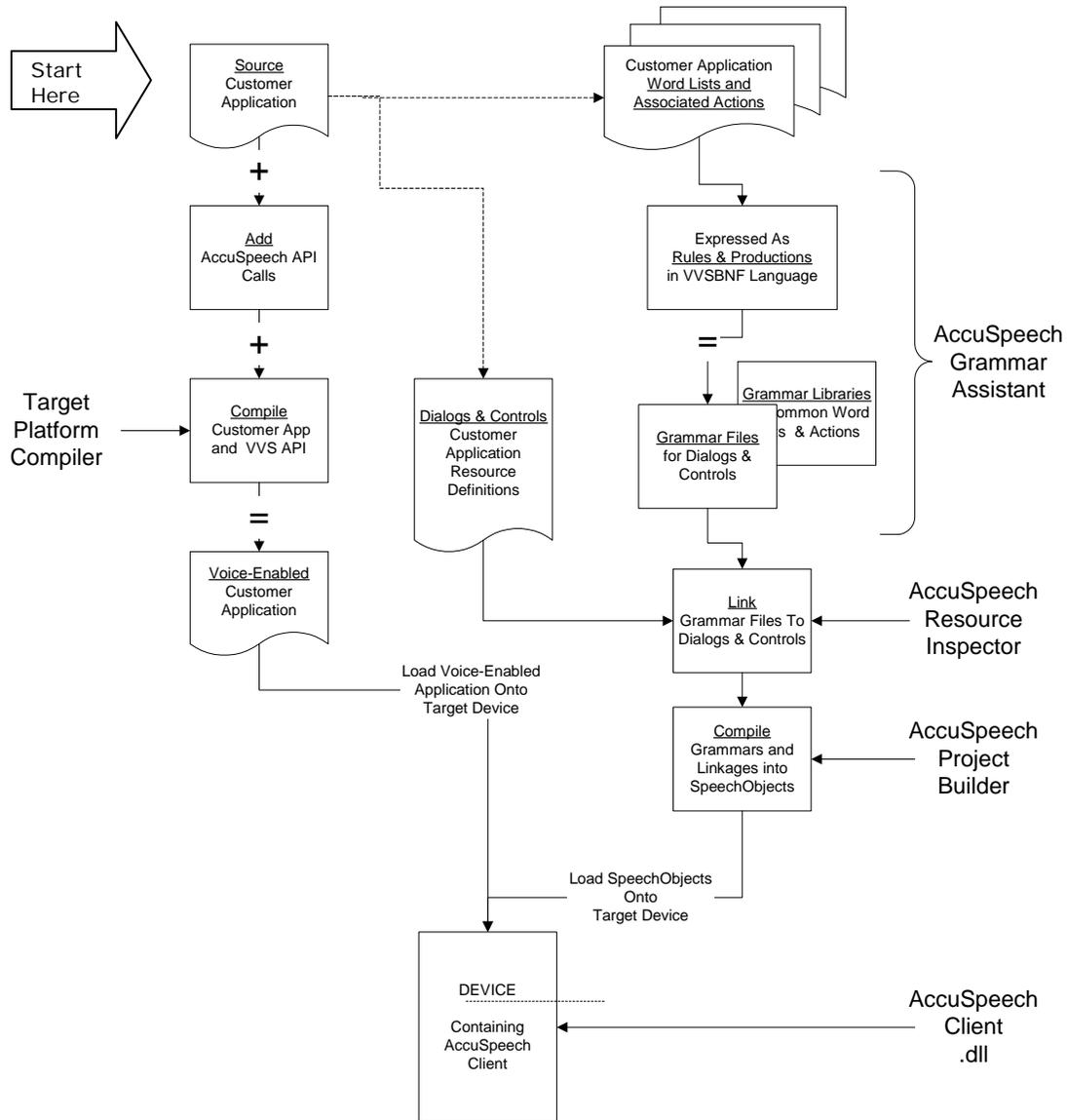
Supported Platforms

AccuSPEECH FormSuite is supported on three widely used platforms:

- Win32
- Web and .NET
- Java

Voice-enabling the application

At its essence, the process of voice enabling a forms based application consists of identifying all objects to be activated or controlled by speech. . Once identified, they must have their “speech object” attached to them. Speech objects are usually prompts, lists of words to be recognized and actions to be performed upon recognition. This flowchart shows the process and identifies the AccuSPEECH tools used.



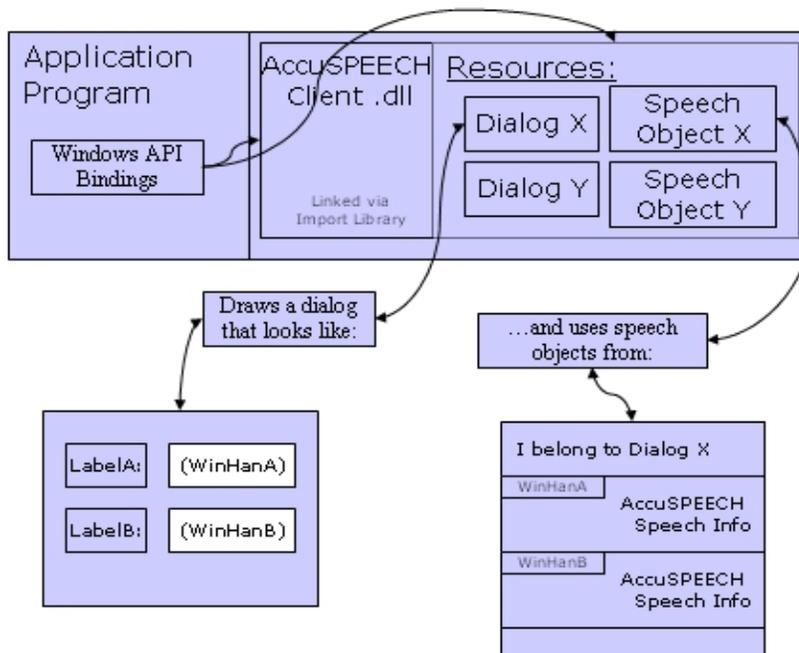
In detail, the following “procedure” describes what is accomplished by the flowchart above:

1. Identify all the forms in the application that need to be speech-enabled. For each form to be speech-enabled, perform the following:
 - a. Identify all the objects within each form that need to be speech-enabled. For each object to be speech-enabled, perform the following:
 - i. Determine the type of speech service required:
 1. Recognition?
 2. Synthesis?
 3. Capture?
 - ii. For recognition or synthesis, determine the vocabulary applicable to the object.
 - iii. Link the object ID to the vocabulary
2. Compile the vocabularies and object IDs into an AccuSpeech™ Resource Block. Assign a unique ID to that block.
3. Insert two pieces of code into code that controls each form:
 - a. An initialization call to the AccuSpeech™ Client API containing:
 - i. Type
 - ii. Owner’s ID number
 - iii. AccuSpeech™ Resource Block ID number.
 - b. An ‘On Focus Change’ call to the AccuSpeech™ Client API

The output of this procedure produces a linked structure consisting of:

- Speech Objects
- AccuSPEECH Client .dll
- Application code with API calls inserted

...as shown below in this figure.

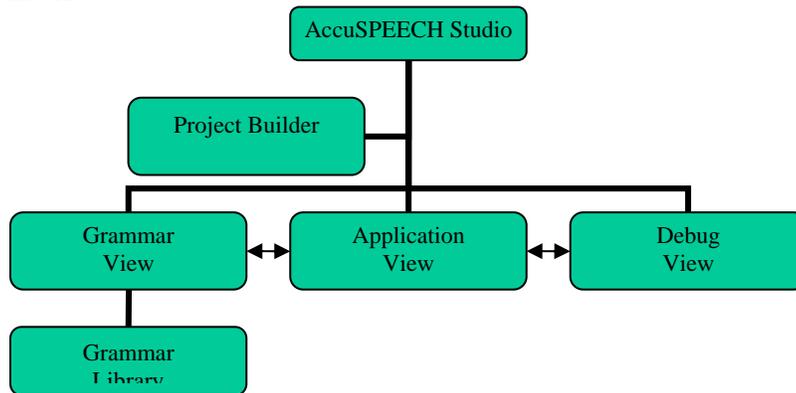


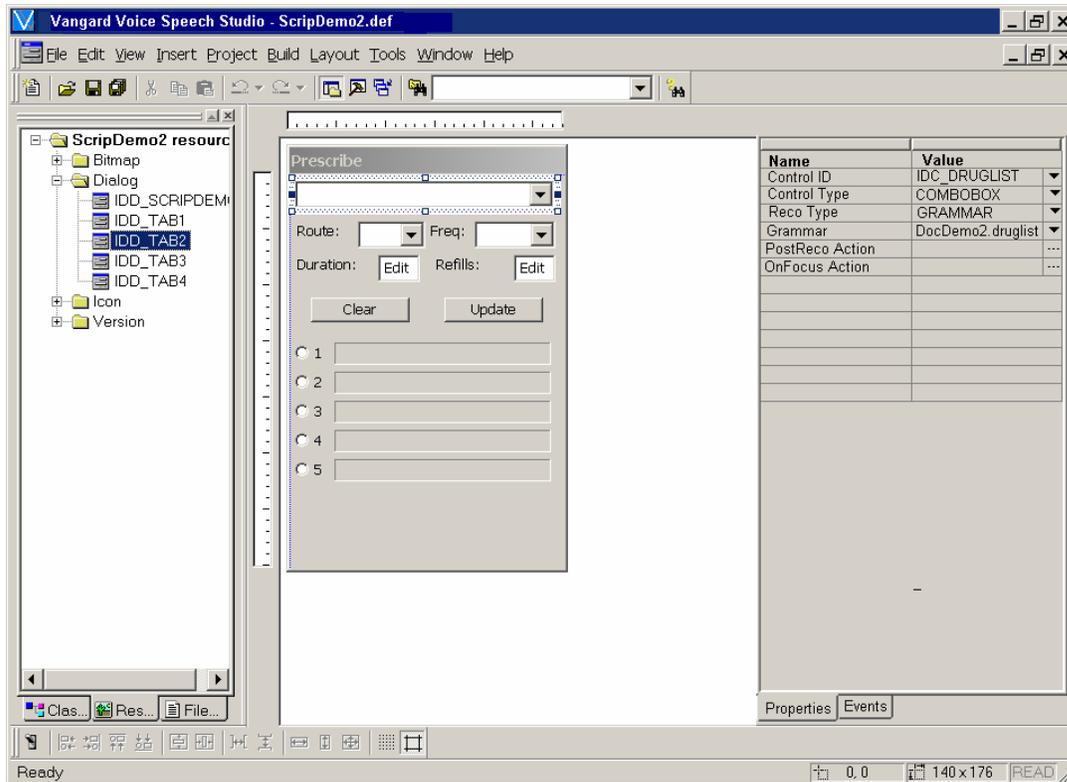
As a Practical Matter...

Doing all of this manually takes a long time and the process is prone to mistakes. Since Vanguard Voice Systems has done quite a bit of voice-enabling, we developed a tool set to facilitate the process described above. The toolset, called the AccuSPEECH Studio, encapsulates development and test tools in a graphics environment similar to Microsoft's Visual Studio. The Studio provides the developer with the ability to:

- Create a project – gather the application code and the grammar list together
- Manage grammars – create new grammars or use existing grammars
- Inspect objects within the application – attach grammars to objects in applications
- Test and debug – ensure that the grammars added to objects do the “right thing” within the application logic.

The structure of the Studio is shown here, followed by a representative screenshot of the Studio interface.





The Bottom Line

Using the tools supplied by the commercial recognizer / speech synthesizer vendors to add speech to applications by is an arduous task. Voice-enabling an application of moderate complexity can take several weeks to a few months. The amount of code that must be inserted into the application is large; on the order of 1k lines of source

FormSuite encapsulates and automates almost all the tasks involved in interfacing these speech components to an application program, saving time and greatly errors.

Summarizing the advantages of the AccuSPEECH FormSuite:

- No long learning curve for a complex recognizer/synthesizer API
- Grammar creation done by wizards. Enter a word and rule, the Studio converts it into grammars for the recognizer/synthesizer.
- Automated attachment of grammars and rules to application program objects.

A specific example: Vanguard Voice Systems speech-enabled a PDA-based prescribing program consisting of 5 screens, 60 objects and some lengthy word lists (eg: drug lists).

Our elapsed time? THREE DAYS.

Contact us for a demonstration.

AccuSPEECH® is a registered trademark of Vanguard Voice Systems, Inc.
All other trademarks are the property of their owners.