

Development of Voice Control in Multiple Languages

SUPPORT FOR MULTIPLE LANGUAGES is key for European companies to reap the benefits of voice control, already exploited in the US. As an example, Charles Schwab, a U.S. stock broker, handles half of its 100,000 daily calls from customers seeking stock quotes using automatic speech recognition (ASR). It has been estimated that savings of \$5-\$15 per call are possible. This technology has been slow to spread to Europe, partly because of the many European languages.

Voxi's patent pending technology, and its Voximizer™ development tool can save development time and cost for a user friendly multi-lingual service.

This paper will analyze the cost of developing a multilingual natural language speech telephone interface using a stock trading example.

Natural Language

For voice control to be user friendly, the natural dialog between humans must be the goal. This requires three main elements:

1. Complete transaction in one sentence. (“Book a meeting with Bob Friday at three”)

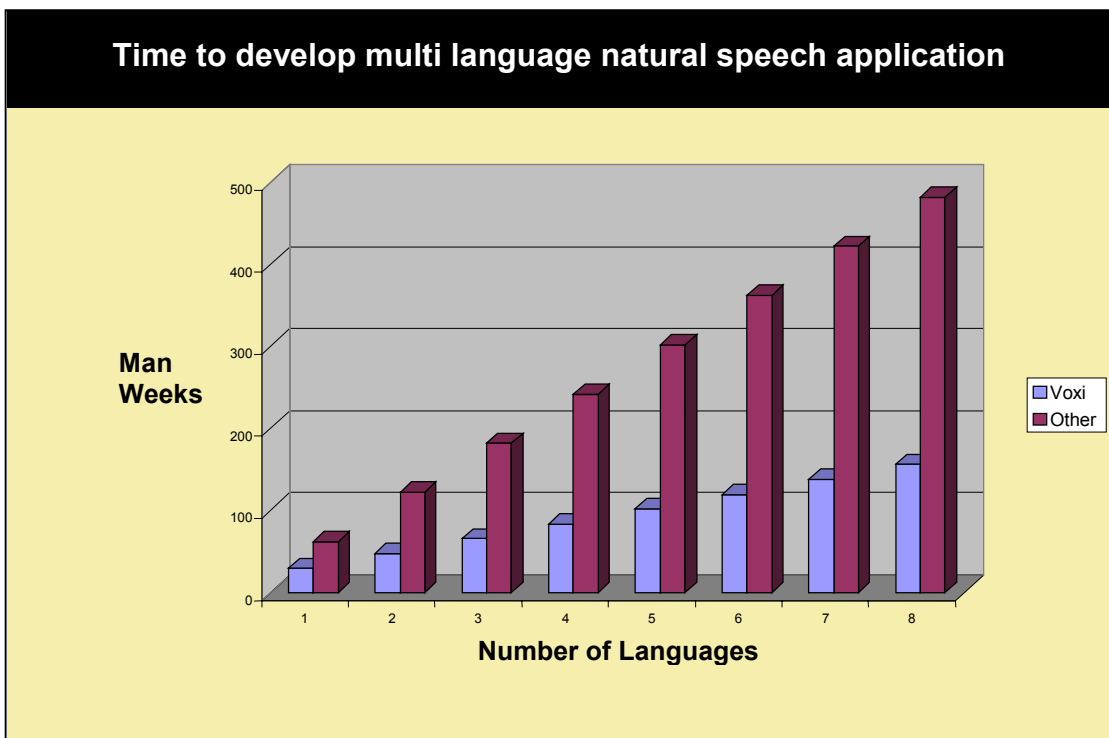
2. Several ways to create one transaction. (“Increase the temperature somewhat” or “Make it warmer”)
3. Ability to jump between parts of the application and between applications without accessing a “main menu”.

One major speech technology company recommends its customers to hire their in-house linguists when implementing a natural language speech interface. In Voxi's approach the difficult language processing is solved once per language and included in the platform. Any professional systems engineer can develop natural speech interface with the Voximizer™, and does not need to be trained in linguistics.

As in any systems development, a good awareness of user interface issues is important to ensure a successful launch and widespread user acceptance.

Development Process

We will compare the cost of developing a stock trading application using Voxi's technology with our knowledge of tools from other companies.



Estimates are made of the effort required for a consulting company to implement a relatively complex stock trading system with a natural language interface, given a specification for the system. It is assumed that basic functionality (functions for getting stock quotes or placing trades) has been implemented, perhaps in a touch-tone system or web interface.

An important part of speech interface development is user trials. They are used to discover which words users actually use, and to subsequently add the new words to the system.

The system development process will contain the following steps:

- Define the words and phrases that the user can use.
- Program the interface to the stock trading computer system.
- Program the user interface responses.
- Produce simple user documentation.
- Deploy the system for a limited number of test users.
- Follow up the results of the test deployment, adding new words and pronunciations to the system.
- Deploy the system on a large scale.

Below is the estimated effort for a multi-language system:

| Step | 1:st Language | | New language | |
|------------------------|---------------|-----------|--------------|-----------|
| | Voxi | Other | Voxi | Other |
| Define words & phrases | 4 | 12 | 2 | 12 |
| Program Interface | 4 | 4 | 0 | 2 |
| User Interface code | 8 | 16 | 2 | 16 |
| User documentation | 4 | 4 | 4 | 4 |
| Test Deployment | 2 | 2 | 2 | 2 |
| Modify System | 8 | 24 | 8 | 24 |
| Total | 30 | 62 | 18 | 60 |

The main advantages of the Voximizer™ are the built in grammar for each language and the program-by-example user interface. Extending the vocabulary with corresponding pronunciations to the application as a result of user trials is very easy.

Since much of the language support is built into the Voxi platform, adding a new language with the Voxi tools only consists of selecting the new grammar and substituting the words.


Note the difference in the time estimates for adding each additional language:

Note also how the program interface does not need to be modified for each new language. The basic structure of the service, and its functionality doesn't vary between languages, and thus doesn't have to be re-implemented in each language when using Voxi's Intelligent Speech Interfaces™ platform.

Summary

To summarize, an eight-language stock trading system would, given a good specification, take 482 man-weeks to implement using other tools on the market, and only 156 man weeks using the Voximizer™.

The savings in development time for an eight-language system is about 68%.

These considerable savings are particularly important for global companies with a need to have services in multilingual environments. 

For more information, please contact Voxi:

WWW: www.voxi.com
 E-mail: information@voxix.com
 Telephone: +46 8 453 9050