Speech Technology

Issue Date: November/December 2006, Posted On: 11/9/2006

Reaching Across the Language Barrier

The social and cultural landscape of our workplace is continually changing, with multilingual communication needs becoming commonplace. As our society becomes more multicultural, organizations need to develop and implement strategies that will accommodate it.

The residents of the ancient biblical city of Babel could readily attest that language barriers between people are nothing new. However, the impact of such barriers, and the need for developing effective solutions, is becoming more apparent each day. According to the 2000 U.S. Census Bureau numbers, 21.3 million people living in the U.S. have difficulty conversing in the English language. This represents a startling 52 percent increase over the 1990 Census Bureau data. Of those responding to the Census poll, 3.3 million spoke no English at all. Considering that the Census data is now six years old and that those who speak little or no English are very likely to be underreported in a poll, it would be a fair assumption that the language barrier in America is even greater today. As a result of these changing environments, rather than an "everyone should learn English" approach, businesses are learning to embrace foreign language skills to foster communication within the firm and with customers and clients.

Typically, most organizations manage to get by with a few dedicated bilingual employees or independent interpreters who do their best to bridge the language barrier. However, with an increasingly diverse population of immigrants entering into even our smaller to mid-sized communities, many corporations, government agencies, and other organizations are finding they cannot effectively communicate with their clientele. They simply do not have sufficient bilingual staff available, and the current demand to converse in many languages makes a reliance on office interpreters impractical.

Fortunately, there are emerging technologies that will greatly help. Although language translation systems have been in place for many years, up until now, they have only been marginally utilized. The early systems were awkward to use and not always effective, while the newest technology makes language translation for the correct applications amazingly effective, simple, and relatively affordable. It is now possible to converse with someone who speaks a different language in real time using bidirectional speech translation systems. The Probation Parole Division of New Mexico recently tested such a system in Albuquerque. The results of the experiment are discussed later in this article.

Early Translation Systems

Computer software for recognizing speech, translating text, and producing voice output has been under development for several decades. The key goal continues to be combining these items into a system that really works—that is, one that instantly translates anything you say or hear at anytime at any location for any language. As an initial step to achieving this goal, early translation systems focused just on translating text. Unfortunately, these initial systems often neglected to use context in generating the translation. For example, the word "right" can mean a statement of correctness as in *You are right*, the authority to do something as in *the right to vote*, or a statement of direction as in *Turn right at the corner*. It was common for these early systems to simply insert a word-for-word substitution using the most commonly used meaning (in this case translating *right* as a direction), thus often leading to completely misleading translations. It was a start but certainly did not provide the accuracy needed for practical use.

An alternative approach for early translation systems was using a database of preselected phrases that had been previously translated by a qualified interpreter. Here, the user is limited to only using phrases that have been previously entered into the system. When one of these phrases is selected, the preprogrammed translation is simply retrieved from the database. In general, this produces reliable translations, but this only works for the phrases that someone else has previously entered into the system. In this second approach, increased accuracy has been gained, but at the cost of significantly reducing the flexibility and applicability of the system. Examples of companies that use this approach are Franklin, Lingo, ECTACO, and SpeechGear.

Recent Technology Advancements

The U.S. military has long recognized the need for soldiers to communicate directly with the local populace. As a result of this need, they have been funding the development of new systems that are coming ever closer to meeting the stated goal–that of translating anything you say or hear at anytime at any location for any language. Examples include SpeechGear's Interact product (part of its Compadre suite of instant translation products) and IBM's MASATOR research effort. These systems are not perfect, but significant improvements have been made.

SpeechGear's Interact system, for example, supports an English vocabulary of over 200,000 words along with a foreign language vocabulary of around 50,000 words. These words can be combined into any sentence and

translated. The context of the words is used to both identify what was said as well as to generate the appropriate translation. For example, in the sentence, *Please **write** your name on the **right** side of the paper*, the second and seventh words are pronounced the same. Proper context is required to determine if the transcription should be *write* or *right*. The same applies to all homonyms such as *to, two, and too*; or *sea* and *see*. With an appropriate transcript, then the meaning of the sentence is identified and the appropriate translation generated.

How Accurate Is It?

When it comes to instant language translation, this is the most often asked and least understood question. The requirements for accuracy vary widely depending on the specific needs and expectations of the participants. Current instant translation systems are focused on accurately retaining the meaning of the original phrase. One approach that is commonly used to measure accuracy is to compare the translation generated by the system to one that was manually generated and verified. If this comparison is done using statistics, say a simple word-for-word comparison, this can lead to very misleading results. This is because pure statistics do not ensure that the meaning has been accurately communicated. For example, if a doctor says to a non-English speaking patient, *Make sure that you take two pills before you eat breakfast*, and it is translated into the foreign language as, *Make sure that you do not take any pills before you eat breakfast,* how accurate is it? Statistically, the percentage word match to the desired output gives an accuracy value of about 85 percent. However, the meaning is completely inverted due to the improper insertion of the word *not*. From the standpoint of successfully communicating the message, the translation can be viewed as being completely incorrect and thus it is a failure. On the other hand, the initial sentence could also be translated as, *You must swallow a couple of tablets prior to eating anything in the morning*. Here, only about 10 percent of the words match the desired output, however, the meaning has been accurately communicated and thus is a success.

In between these two measurements of obvious success and failure are a multitude of translations that may be close enough for some users, but not for others. For example, if the translation was, *Make sure you take your medicine before you eat breakfast*; the translation is marginally correct but certainly not precise. Here the generic word *medicine* has been substituted for the more precise description of *two tablets*. Depending on the particular users and context, this may or may not be a successful translation.

So, what exactly does this mean for the person looking for instant language translation solutions? First, understand exactly what it is you need. If you are translating legal documents or negotiating an important business deal, your requirements for retaining accurate grammar, as well as meaning, are going to be much more stringent and may require the services of a trained interpreter. If, on the other hand, you're interested in communicating freely, without being restricted to specific phrases, and are more concerned with ensuring that meaning is communicated than having perfect grammar, then these new technologies could very well meet your needs.

The Albuquerque Test

As a border state with a large population of Spanish-speaking residents, New Mexico is challenged to provide appropriate services to its growing population of non-English speaking individuals. Having relied solely on bilingual probation/ parole officers to handle Spanishspeaking cases, the Probation Parole Division of New Mexico found itself in a difficult position. Many bilingual officers point out that offenders who do not speak English are much more likely to have critical needs. They often have difficulty understanding the court system and the requirements of their sentences. Unless significant time is invested and great care is taken, these offenders are more likely to recidivate.

The bilingual officers that supervise Spanish-speaking caseloads are given a small pay differential to compensate them for their much needed language skills. Unfortunately, with tight budgets prevailing, the extra pay is often inadequate to attract and keep these valuable employees. When speech translation systems began to gain credibility, the Division took steps to begin an evaluation process of the technology. They contacted SpeechGear, Inc. who agreed to assist the Division with a trial of its Interact speech translation system. The trial, which is discussed in the remainder of this article, was conducted in Albuquerque in June 2006 and illustrates the performance of instant language translation systems in real-world applications.

The parameters of the study were rather straightforward. A bilingual officer in Albuquerque was asked to use Interact to communicate with his Spanish-speaking caseload. The trial was limited to the Mexican Spanish offenders as the software used was engineered for this dialect. In the test, the metric, "Was the meaning correctly translated?" was used. The officer evaluated the effectiveness of the equipment in translating each portion of his interviews by rating the communication into one of four areas:

- ***Successful*** No problems. The translation was encountered and the concept was translated without any difficulty.

- ***Successful after rephrasing*** The concept was translated appropriately after rephrasing or repeating the initial sentence.

- ***Unsuccessful communication*** The officer could not successfully communicate the concept even after rephrasing.

- ***Equipment failure*** The officer was unable to communicate the concept due to problems with the equipment.

Trial Results

The trial provided interesting results that are generally applicable to all conversational instant translation systems. The English-to-Spanish translation was extremely effective with a 100 percent success rate, while the Spanish-to-English communication was effective approximately 70 percent of the time. There were three primary reasons that account for this disparity.

First, the English-speaking probation officers assigned to the trial had the advantage of becoming familiar with the software. As they continued to use the system, they identified words that it did not translate correctly (for example, when the acronym PO was spoken, it was translated as Post Office instead of the desired Parole Officer) and avoided using these words in their speech. Because the Spanish-speaking offenders had no previous exposure to the system, there was no such learning opportunity.

Second, the officers took the time to train the software to recognize the nuances of their speech. Anyone who has used commercially available computer dictation software is familiar with this process. Users typically read a script, which helps the computer learn their speech patterns. The Spanishspeaking offenders who reported to the officer again did not have the opportunity to perform such training, and thus used Interact's reduced vocabulary (50,000 words) speaker-independent engine.

Finally, the probation officers learned the fine points about using Interact, such as the most effective rate of speech and the optimal distance to hold the microphone from their mouth, while the Spanish-speaking individuals did not have this advantage. The probation officers became comfortable with the equipment and learned how to speak clearly, precisely, and with minimal stammering, while many of the offenders seemed intimidated by the equipment, which resulted in an increased level of stammering.

Will the Systems Work for You?

In conclusion, today's speech-translation systems have made significant strides in overcoming language barriers. They are effective tools for communication and are no longer limited to specific phrases or usage domains; however, they are not perfect. It is important to use these speech translation systems in appropriate settings. Where precise language is required, these systems are not yet up to the task. Court proceedings and legal discussions are examples of situations in which professional translators are still required. However, there are hundreds of applications where the use of language translation technologies is totally appropriate. Understanding the systems' capabilities and limitations is the key to deploying a successful solution for your needs.

Lisa Cochran oversees the corporate and commercial marketing and business development activities for SpeechGear Inc. Contact her at lcochran@speechgear.com.