

Voice Response Translators:

Comparing Three Units

By National Institute of Justice Staff

Consider the dynamics a police officer must weigh when responding to a call. What is the nature of the call? Who are the participants? What steps must be taken to reach resolution?

Now consider another dynamic officers are encountering with increased frequency: How does an officer handle the situation when the participants speak a different language?

The answer may lie in part in the emergence of voice translation devices. Although the idea of a voice translator is not new, technological progress has moved the device off the drawing board and into the hands of users. Demand for such a device has crept up due to an increasingly multilingual society and a world increasingly interdependent through business, communications and travel. A voice translator can save travelers the hassle of flipping through language dictionaries and fumbling through foreign syllables, but to a soldier or police officer, the device can save crucial seconds when lives hang in the balance. A voice translator lets an officer relay basic commands such as “open the door” and gather crime scene information from participants and witnesses.

Although supply and demand have increased, testimonials of marketed products had not been substantiated through objective laboratory testing. In an effort to evaluate the utility and applicability of the devices for criminal justice profes-

sionals, the National Institute of Justice teamed with NAVAIR ORL Training Systems Division to evaluate three devices that recognize simple voice commands and translates them into another language: the Phraselator, the Voice Response Translator (VRT) and the Universal Translator UT-103. The evaluation outlined the units’ performance capabilities, including operation within noisy environments, ease of use and other operational characteristics such as battery life.

VRT

Of the three units, the VRT was found to be the easiest, least intrusive device. Turn it on, set the phrase group and speak into the microphone. No additional adjustment is required, and a voice-activated headset microphone allows for hands-free operation. The push-to-talk microphone on the UT-103 and the Phraselator’s user interface screen required greater operator intervention. The VRT was also the fastest unit with response times of less than a second. The unit tested supported Spanish, Creole, Portuguese, Arabic and Darsi Farsi.

Once activated, the VRT supports 204 unique phrases and has the potential to support any spoken language. The unit does not require the user to speak the entire phrase; rather, keyword phrases provided on reference cards trigger the complete spoken phrase in the recorded language. Prior to operation, however, the VRT must be trained to recognize

Note: The views and opinions in this article are those of the authors and do not reflect an official position of the U.S. government. References to any specific commercial products by trade name, trademark, manufacturer or otherwise do not constitute or imply its endorsement, recommendation or favoring by the U.S. government.

a user. The VRT guides users through training, which involves the user saying each phrase on the reference cards to trigger the foreign language phrase. The unit tests the recorded wave quality during the training session, and when a phrase is not recognized, repeats the training until it is. From this evaluation, it was clear that training the device would take longer than 45 minutes.

The VRT also scored well during noise testing. This testing was conducted on each of the units with three forms of noise interference: “best case ambient” noise, which is noise under 50 decibels; high-pitched “white” noise, which makes a hissing sound; and “pink” background noise similar to natural sounds like wind and rushing water. Ten phrases were read 10 times each against pink and then white background noise at 60, 70, 80 and 90 decibels. The VRT was found to have the least amount of degradation as white and pink background noises were increased to 90 decibels.

One disadvantage was noted: the VRT does not have a mechanism to prevent the unit from broadcasting an incorrectly translated phrase. Twice during testing the phrase “stop vessel” was incorrectly translated to “vessel owner.” A mechanism to nullify incorrect translations would reduce confusion.

Maintenance consists of recharging the battery. The battery charger is a supplied AC-DC converter that plugs into standard wall outlets. The VRT required the least amount of recharging of the three units and

Although the idea of a voice translator is not new, technological progress has moved the device off the drawing board and into the hands of users.

used the least amount of power. With a recharging time of 30 to 45 minutes, the unit did not need recharging for an eight-hour evaluation period.

Phraselator

The Phraselator, which gained a measure of notoriety during Operation Enduring Freedom by enabling U.S. soldiers to communicate with captured Iraqi soldiers, was evaluated by two subjects through out-of-the-box testing. Neither tester had previous experience with speech recognition software. Both, however, were quickly able to operate it. While the users operated the device without much difficulty, the evaluators noted that the device did not recognize the phrase "Do you speak Arabic" for either tester.

Overall, the evaluation found the Phraselator to be best suited for quieter situations such as in medical settings or during inmate interrogation in a designated room. Unlike the VRT, the Phraselator must be held in the user's hand for operation and kept about four to six inches from the mouth while speaking. Although the push-to-talk microphone means the unit cannot be used hands free, it can be operated with one hand after its recognition program has been started and the phrase set selected. One advantage of the Phraselator is that the user need not learn phrase sets, as required for the VRT. Instead, the user can look up the phrase on the unit's screen. In addition, users can bypass the voice recognition feature by selecting the desired phrase from the screen. Another advantage is that the Phraselator is speaker-independent; unlike the VRT, it does not need to be trained to recognize a specific person.

The Phraselator as tested supports Arabic, Dari, Pashto and Urdu. Because the unit plays back fixed, stored responses, the potential exists to support any spoken language, limited only by the amount of disk space required to hold the playback recordings.

The Phraselator was the slowest of the three units. Response time

took four to five seconds, and response time slowed significantly as the batteries wore down. Also, although the Phraselator had the highest number of correct translations of the three units against ambient noise, correct translations fell off markedly as phrases were tested against increased levels of pink and white noise.

Maintenance consists of recharging the battery pack. Batteries are charged with the AC-DC converter supplied with the unit. The charger uses a standard 120 volt wall outlet. The unit seemed to consume the most power of the three, lasting six hours after being recharged. Because of this, the evaluation recommends using the device near a readily available power source such as a wall outlet.

Universal Translator UT-103

The Universal Translator UT-103 was designed for translations centered on transportation and tourist-related interactions, and the evaluation it best suited for that purpose. While the VRT and Phraselator had phrases and languages meant for law enforcement or military use, the UT-103's phrase/language set would be most helpful to users trying to communicate transportation and lodging needs in Europe. The UT-103 translates English to Spanish, German and French. The UT was included in testing primarily to evaluate how the technology worked in a noisy environment.

The manufacturer claims that the unit contains about 3,000 unique phrases. The domains they are subdivided into, however, are fairly restrictive. Domains include phrases for planes, railways, taxis, ordering at restaurants, etc. With only about 10 phrases per domain, the user must scroll through many on the device's display before finding the one that contains the desired phrase. This technique boosts translation accuracy but diminishes ease of use.

The UT-103's best noise testing results occurred against ambient

noise. The unit translated correctly at least 80 percent of the time on nine of 10 phrases; however, it was not able to translate "I'm a tourist" on any of the 10 attempts for that phrase. The unit's accuracy had a steep drop off against 90 decibels of white background noise, as it was unable to translate eight of the phrases on any attempt.

The UT-103 is battery operated only and retained power throughout the eight-hour evaluation. Response time was roughly three to four seconds, slowing significantly as the batteries wore down.

Summary

The Universal Translator is clearly best suited for tourist travel, for which it was designed. It was included to see if it could work in a noisy environment. Had it done well in those tests, the technology behind it may have been a candidate for further development for criminal justice use.

Considering all of the test results, the VRT seems to be the easiest, least intrusive unit to use once it is trained to a user's voice. But the Phraselator has a large phrase set, the ability to bypass voice recognition if needed, and had the highest correct translation rate against ambient noise.

Each unit tested has pluses and minuses and each may work best in a given situation. Agencies interested in purchasing such units should consider what the units will be used for and what the setting is likely to be like.

The information in this article was taken from the NIJ Grantee Report, "Voice Recognition Evaluation Report," which is available for download from the National Criminal Justice Reference Center Web site at www.ncjrs.org. The report was prepared by the Naval Air Systems Command, Training Systems Division, under Interagency Agreement #2002-LB-R-045 awarded by the U.S. Department of Justice, Office of Justice Programs, National Institute of Justice.

Reprinted from the April 2004 issue of *Corrections Today*, Vol. 66, No. 2 with permission from:



The American Correctional Association, 4380 Forbes Blvd., Lanham, Maryland 20706