

**HOLD UNTIL RELEASED  
BY THE COMMITTEE**

**TESTIMONY OF**

**DR. ANDRÉ VAN TILBORG  
DEPUTY UNDER SECRETARY OF DEFENSE (SCIENCE AND  
TECHNOLOGY)**

**BEFORE THE SUBCOMMITTEE ON  
EMERGING THREATS AND CAPABILITIES  
OF THE  
SENATE ARMED SERVICES COMMITTEE**

**April 25, 2007**

## **Introduction**

Mr. Chairman, distinguished members of the Committee, thank you for this opportunity to appear before you to discuss the Department's science and technology (S&T) investments in the areas of language translation devices and cultural awareness training. I will use this opportunity to describe the Department's current capabilities in translation devices and cultural awareness training, the challenges in these areas, and some of the planned future work that supports our warfighters' needs for interaction with non-Western cultures. I am also pleased to have the chance to highlight in this hearing some of the new and expanded initiatives we are hoping to undertake in these areas to address the 2006 Quadrennial Defense Review (QDR).

I have organized this testimony into two sections, one dealing with language technology devices and the other with cultural awareness training. There is obviously some linkage between these two areas, but many of the research challenges in translation devices or socio-cultural awareness training are unique to each. For these areas, there is a clear need for increased skills and capabilities for all of the Combatant Commands (COCOM), and there are specific needs statements from at least two COCOMs, Central Command (CENTCOM), and Special Operations Command (SOCOM). The language and cultural terrain of each COCOM's Areas of Responsibility (AOR) define the technical challenges. However, we do recognize that the solutions are not solely technical in nature, and the Department must also address Manpower, Personnel, and Training as part of the solution sets.

I will address language translation devices research and engineering first. The Global War on Terror (GWOT), and, more generically, Irregular Warfare (IW)/counterinsurgency environments have forced the military to confront a situation in which all personnel need some abilities to operate effectively in settings where they must have skills to work in novel language and cultural dimensions. The requirement for a deeper understanding of the human environment in an AOR is now relevant, not just for Special Operations Units, but for all Soldiers and Marines who are deploying. At the strategic level, the GWOT has created the need not only to have the ability to communicate with indigenous peoples from diverse cultures, but also to be able to understand their written and media communications. The sheer volume of written text and other media (i.e., television broadcasts, internet postings) makes using individual human translators untenable. Given the realities on the ground in the CENTCOM AOR and the scale of the GWOT, it is evident that for our current and future forces we need to discover, develop and field technologies to augment our existing translation capabilities at the strategic, operational and tactical level, for multiple languages

and dialects, and for users that span a broad language skill level (novice to linguist). Some of the capability needs and technical challenges and the efforts to meet these challenges and field viable products to our forces are described below. The work listed below does have application to the general purpose force, to planners and operators, and to the Intelligence Community. However, the IC has their own specific, unique use scenarios that require linguistic capabilities to support intelligence analysis. In this testimony I will intentionally focus on DoD specific investments. However, please recognize that we are cognizant of the development being done in the IC, have numerous links to the IC linguists, and our investments complement their work.

### **Language Translation Devices**

From a technical perspective, there are some common underlying capabilities that serve both text and speech translation. However, each has unique features that make fielding operational devices difficult. The commercial sector has developed translation capabilities and technologies that meet some of the needs of the operational forces, and the pedigree of this technology (i.e., IBM's MASTOR) includes a history of Department of Defense funding. However, many of these available products are not robust enough to meet the scale, breadth and tempo required for the mission areas/needs of deployed forces. The threshold for effective use of speech to speech translation in the business world is lower than in the military, there is a need for 'street' level communication that accounts for unstructured/colloquial speech, varied sub-dialects, noisy environments, the need for hand's free communication, and the need for increased accuracy in real-time, tactical translation. Text and media translation has additional challenges that include bandwidth limitations in operational environments, and degraded signals such as smudged and handwritten text translation. Lastly, the military needs access to a large volume of spoken and/or written language to create the databases, scalable models, and training materials for some of the more obscure languages and dialects that occur across the globe. There are approximately 7000 distinct languages/dialects in the world. Many of the places we may operate have hundreds of subcultures and languages. The commercial technology investments are not focused on providing translation in these types of niche markets. DoD investments in this area are concentrated on improving existing technology to reach the translation accuracy necessary for our forces, and on expanding the reach of these tools to the socially and linguistically diverse regions in which the Department operates. What follows is a description of the most significant technology developments in the area of language. It should be evident from the descriptions below that the ongoing and planned programs within the services and other organizations, such as the Defense Advanced Research Projects Agency (DARPA), are truly complementary.

**Speech to Speech Translation** The major speech to speech translation systems have taken advantage of previous commercial successes in translation. The largest program today is DARPA's Spoken Language Communication and Translation System for Tactical Use (TRANSTAC). TRANSTAC has developed PC-based translation systems that allow speech-to-speech translation between English and Baghdadi Arabic. Its current accuracy, in controlled noise environments, is between 70-80 percent. The TRANSTAC system is being field tested in Iraq for specified use domains (i.e., medical care, vehicle checkpoint, and joint Iraqi-Coalition Force missions). The program has had early successes and has contributed to the development and fielding of a number of products such as the PHRASELATOR system and IBM's MASTOR system. The MASTOR system recently made headlines when IBM offered to donate over 10,000 software licenses and 1000 devices to the Department. The TRANSTAC program is attempting to develop and field hands-free two-way speech to speech translation systems that can provide accurate translation in urban military environments. A secondary goal is to expand the domains and accuracy of the existing system. Other programs, such as the Instant Language Translation project under the Office of Naval Research, are expanding the capability of portable translation systems by including multi-mode inputs (spoken, written, images) and additional languages and dialects from other regions of the world such as Korea and Somalia.

**Text to Text and Media to Text** The largest DoD S&T investment in text to text translation is DARPA's Global Autonomous Language Exploitation (GALE) program. The program's goal is to translate and distill foreign language material (e.g. television shows and websites) in near real-time, highlight salient information, and produce targeted query responses. The program will deliver the capability to translate both structured and unstructured text and speech, with a goal of delivering an accuracy of 95% for text and 90% for speech. Other investments include improvements in tactical document translation system accuracy and capabilities. It should be noted that the media-speech to text is working in a domain where the speech is controlled with rather predictable vocabulary. The GALE program has already accomplished much in the way of improved accuracy in translating text (55 to 75% accuracy) and media (35 to 65% accuracy), but they are continuing to improve the technology to a maturity level to deliver the capability to translate both text and speech at 90-95% accuracy.

**Language Databases** The current method to support developing new language modules for existing translation devices is costly and time consuming, requiring the collection, transcription and translation of large amounts of training data (written and spoken language). The Air Force Research Laboratory has an ongoing project that will provide a rapid turnaround on developing linguistic data sources for new languages and domains of interest. The goal is to use innovative techniques to take languages and dialects in which we have limited data, with less

than 10 hours of speech data or 20,000 words or less of text, and produce useful spoken or written translations.

**Fielded Technology** Three of the technologies described above are in use by military units. The PHRASELATOR, handheld one way speech to speech device, provides tactical level communication for soldiers within specific domains, such as checkpoint and medical operations. The device was originally developed as part of a DARPA SBIR effort. There are currently over 2000 PHRASELATORS in the field in Iraq. The DOCEX system provides the capability to process and exploit captured documents for actionable intelligence within tactical time scales. The DOCEX was developed as part of a DDR&E Advanced Concept Technology Demonstration project, transitioned to the National Ground Intelligence Center and fielded. There are systems currently deployed in Iraq. Finally, the GALE program transitioned structured text/media translation technologies to 12 USG organizations, and 2 systems are currently fielded in Iraq.

### **Cultural Awareness Training**

The need for improved cultural awareness training was identified in the early phases of Operation Iraqi Freedom. It was realized that the general Joint Force needed some of the same cultural awareness competency that our Special Operations Forces have traditionally maintained. Military operations in complex, multicultural environments require more than just being culturally sensitive to the do's and don'ts of a society. Such operations also require an awareness and knowledge that can be applied to improve operator interactions and shape the outcome of the interactions. Each of the Services have established cultural awareness training centers that are developing content, sharing this content, and have begun training their personnel on the specific knowledge necessary to support their military missions. Fortunately, these centers have access to the extensive on the ground experiences of the Soldiers and Marines returning from Iraq, combined with the relatively well known academic knowledge of Iraq's religious and sectarian history. Providing the same level of 'understanding' and training for data-poor, less studied socio-cultural environments, such as the mountainous tribal regions of Afghanistan or the multicultural regions of Indonesia is much more difficult.

The ultimate goal is to achieve an acceptable baseline for cultural competency across our forces. As mentioned above, the first generation capabilities in this area are being derived from the best academic and professional subject matter experts (SMEs) providing schoolhouse content. The next generation will likely be computer-mediated training and mission rehearsal in relevant venues. The third generation will be embedded within more immersive, dynamic

environments. Methodologies to collect, package, and understand knowledge of cultural landscapes will be needed to support the generation of content that will fill these curricula and training systems. What follows is a description of the current technology developments in the area of culture awareness training for second and third generation capabilities and planned future investment in the area of socio-cultural understanding.

**Ongoing Efforts** The Combating Terrorism Technology Support Office's Technical Support Working Group (TWSG) is developing training support packages that focus on the operational and tactical applications of cultural awareness, with a specific focus on Indonesia. The training material is being developed in coordination with SOCOM and the U.S. Army Training and Doctrine Command.

A number of Small Business Innovation Research (SBIR) and Small Business Technology Transfer (STTR) Projects have begun in the area of cultural awareness. They include projects to train at the tactical and strategic level and all are focused on developing computer-based awareness training. Again, at the tactical level, an SBIR project will develop a system that will train how to read Middle Eastern non-verbal cues and develop an understanding of what those cues say about a person's intentions. Another will create more accurate and realistic non-U.S. entities that can be used in convoy commander and ground troop training simulations. At the strategic level, there are two new projects. One will support the training of planners and senior leaders in developing and assessing metrics for effects based operations in complex conflict environments. The second project will develop a computer-based tool to support leader training on interagency goals and progress in non-western conflict environments. All of these have been initiated under the new Human Social, Culture and Behavior (HSCB) Modeling Initiative led by DDR&E.

The HSCB initiative has sprung from the lesson learned in the on-going GWOT. That lesson learned is that the DoD has capability gaps in software tools and decision aids that will allow U.S. commanders to better understand different cultures. The QDR highlighted these lessons in stating that current and future military operations will require enhanced capability to understand social and cultural "terrains" as well as various dimensions of human behavior. The HSCB initiative will develop the required scientific base and will field matured technologies that support human terrain understanding and forecasting across a span of missions and geographic regions. The DDR&E staff worked with the military components and intelligence community in 2006 to identify capability needs in 75 areas; there were gaps in roughly 70 of these areas. The HSCB initiative will address these gaps and integrate complex human factors into the pre-planning, planning and execution cycle of military operations. HSCB modeling is

focused on filling capability gaps within data collection/infrastructure and knowledge management, and then developing the models to forecast societal and cultural behaviors. In addition to delivering software modules that are fully integrated into DoD command and control systems, the HSCB effort will help to create the infrastructure (simulations and content (data, models and theories)) to support tactical through strategic training, mission rehearsal and experimentation using valid cultural entities and models.

**Fielded Technology** DARPA's Tactical Language and Culture Training system was designed to provide our warfighters with some basic spoken language and cultural proficiency with only limited (2 weeks) laptop computer training. The system is currently available in Iraqi Arabic and Pashto and provides basic language and cultural awareness skills training for troops. There are currently 800 copies of the software installed in bases here in the U.S. as well as in theater. Over 6000 troops have used the system for initial skills training.

In conclusion, the need for a robust DoD S&T program in language and cultural awareness and associated capabilities is a central element to fighting the global war on terror. The ongoing and future efforts of Defense S&T will support the training and equipping of today's force, tomorrow's force, and the future's force. We believe these efforts are meeting this challenge, and we truly appreciate the continued support of this committee in providing us the tools and resources to carry out this vital mission.