# MACHINE LANGUAGE TRANSLATION SYSTEMS

# **DEMONSTRATION AND ASSESSMENT REPORT**



# **LIMITED USER EVALUATION** *EXERCISE ULCHI FOCUS LENS, KOREA AUGUST 2004*

# For Official Use Only

This document describes the Ulchi Focus Lens conducted by the Experimentation Center at U.S. Marine Corps Forces, Pacific (MEC) during August 2004. It reflects out observations, conclusions, and recommendations. It does not necessarily represent the formal position of the Marine Corps or the Department of Navy.

This report is for official use only and may not be released outside the Department of Defense without prior approval from the Marine Corps Forces Pacific Experimentation Center.

Shujie Chang, P.E.

Shujie Chang, P.E. Director, MEC September 3, 2004

# Contents

<u>1.0</u>	EXECUTIVE SUMMARY4
<u>2.0</u>	BACKGROUND5
<u>3.0</u>	PROBLEM STATEMENT/OPERATIONAL NEED5
<u>4.0</u>	GOALS AND OBJECTIVES FOR ULCHI FOCUS LENS 20046
<u>5.0</u>	LANGUAGE TRANSLATION TOOLS AND TECHNOLOGIES5
<u>6.0</u>	ASSESSMENT EXECUTION10
<u>7.0</u>	CONCLUSIONS
<u>8.0</u>	RECOMMENDATIONS



# **1.0 EXECUTIVE SUMMARY**

This report provides the results that Marine Forces Pacific Experimentation Center (MEC) derived from Ulchi Focus Lens 2004 (UFL '04). The UFL '04 assessment was conducted under the Language and Speech Exploitation Resources (LASER) Advanced Concept Technology Demonstration (ACTD) from 10 August to 1 September 2004 in the Republic of Korea (ROK). During the exercise, MEC completed a Limited User Evaluation (LUE) of four foreign language translation technologies: the Pharselator P2, Voice Response Translator (VRT), Speaking Multilingual Interactive Natural Dialogue System (S-MINDS) and the SpeechGear Expres (Expres). A limited military utility assessment of the Translingual Instant Messaging (TrIM) was conducted by Det 1 AFOTEC and will be covered in a separate report. The P2 and VRT are one way speech-to-speech devices that translate verbal input. The S-MINDS is a two-way speech-to-speech device that translates verbal input. All speech-to-speech devices were used to translate English to Korean and in the case of S-MINDS English to Korean and Korean to English. Expres was used to translate PowerPoint presentations from English to Korean and Korean to English.

UFL '04 is ROK-US Combined Forces Command (CFC), ROK government, simulation driven, OPLAN-oriented command post exercise (CPX) conducted annually. Ulchi Focus Lens is CFC's large scale warfighting command post exercise (CPX). It is an annual ROK-US combined forces government military exercise designed to exercise, evaluate, and improve crisis action measures and procedures for the combined war plans in the defense of the Republic of Korea in accordance with OPLAN and Supporting plans. It provides an opportunity for commanders and staffs to focus on strategic and operational issues associated with general military operations on the Korean peninsula. Ulchi Focus Lens is a CPX with the tactical situation portrayed through the use of computer simulation models and master scenario events list. Ulchi Focus Lens is the world's largest computerized command and control exercise. The exercise focuses on how U.S. and South Korean forces would defend against a North Korean attack. North Korea usually denounces the exercise, calling it a preparation for war. Held annually, Ulchi Focus Lens trains Combined Forces Command personnel and major component, subordinate and augmenting staffs using wargaming computer simulations and support infrastructures.

The P2 and VRT were used in a Force Protection scenario. U.S. Military Policeman trained and used the devices to determine whether Korean personnel were cleared for access to the Military Base. The S-MINDS was used by medical personnel to communicate with Korean speaking patients. The primary objectives during UFL '04 were to assess translation effectiveness, suitability, and mission impact of each device. During the exercise, MEC and Army Research Laboratory (ARL) data collectors observed Korean-speaking personnel being interviewed. They also used rating scales, questionnaires, and interviews to collect subjective data from foreign-speaking role players and warfighters during and after the exercise.

Results: The P2 results indicated that the technology was useful but required more individual training. The P2 accurately translated and users felt it accurately conveyed critical information. While the devices worked well users sometimes had difficulty communicating complex information. While the training received was adequate, users felt they need more practice to become more proficient. The P2 did "lock-up" on several occasions and had to be rebooted. Users would have had better success if they were able to create a favorites list on the P2.

The VRT can be used to convey critical information from English to a foreign language. Users did find the device sensitive to any speaker deviations. Some users felt that some phrases (Deadly Force) were not conveyed with enough vocal emphasis and that some phrases like "put your hand up in the air", could not be conveyed fast enough to preclude a life-threatening situation. They found the device very

easy to use and liked the device for its hands free capability. All users felt that they would use the VRT right now in day-to-day missions if it were available.

The S-MINDS was able to successfully recognize both English and Korean phrase with the appropriate verbal input. The device was easy to learn and little or no technical problems were noted. Best use of the device was for medical screening. Because of the requirement for the "patient" to use a hand set it would be difficult to communicate if the "patient" were in critical condition on a stretcher. Users felt they need more basic questions in order to diagnose the patient's medical problem.

Expres users reported mixed results. Primary concern continues to be associated with the translation of acronyms and abbreviations. Use of fragmented phrases and bullets confused the machine translation engine. Some users felt that Expres could save considerable time after the Translation memory is populated with acronyms and evaluations. Overall users felt that all technologies warrant further development and evaluation.

# 2.0 BACKGROUND

The Language and Speech Exploitation Resources Advanced Concept Technology Demonstration is a five-year (FY02-06) program to identify, integrate, test, evaluate, demonstrate and assess the military utility of language technologies for text-to-text translation, speech-to-speech translation, optical character recognition, training tools and cross-lingual information retrieval (data mining and management) in Military Utility Assessments (MUAs).

A primary LASER ACTD program objective is to provide and assess leading-edge technologies and concepts to reduce the language barriers experienced by operational and intelligence personnel. The LASER thrust is to improve interoperability, accuracy, and timeliness of translation for speech and documentation. Both Operations and Intelligence communities require speech and text processing capabilities in a wide range of foreign languages to support coalition/joint task force headquarters and field operations. Language related technology is a fundamental enabler in collection, processing, and exploitation of foreign language materials and sources.

The Marine Corps Forces, Pacific Experimentation Center (MEC), under the direction of Commander, Marine Forces Pacific, serves as a focal point for MFP transformation and experimentation throughout the Pacific and Central theaters. The MEC is co-Operational Manager (OM) of the LASER ACTD and as such seeks to employ and evaluate the LASER language technology tools in controlled and uncontrolled environments and joint/combined exercises within the United States Pacific Command (USPACOM). Use of these in exercises will provide realistic estimates of the usefulness of state-of-the-art tools under operational conditions. In-garrison use of the tools on a routine basis will permit users to become more familiar with the technology prior to LASER MUAs, assist them in bridging the language barriers encountered in the course of their daily duties, and will afford a mechanism for users to provide regular feedback to the LASER ACTD Operational Managers.

### **3.0 PROBLEM STATEMENT/OPERATIONAL NEED.**

The United States (US) Combatant Commands, Intelligence Community (IC), and Coalition partner nations conduct worldwide operations with widely diverse languages, often with insufficient numbers of language qualified analysts and translators to support existing mission requirements. Communications with our coalition partners and the local population is often a primary issue.

On the operational side, our ability to coordinate with partner countries and the local government is often hampered by language difficulties that are made more difficult by military jargon.

The US Military increasingly finds itself involved in coalition task forces and exercises. Areas of concern include the high volume of material to be analyzed and the diversity of languages encountered, the need for greater efficiency and speed in analysis and the limited numbers of language professionals.

Similarly in the intelligence gathering function where trained linguists do exist, we face operations routinely with insufficient numbers of language qualified analysts and translators to support mission requirements. In this area of intelligence support, major issues also include a high volume of material to be analyzed, diversity of languages, the need for greater efficiency and speed in analysis and a limited number of language professionals.

# 4.0 GOALS AND OBJECTIVES FOR ULCHI FOCUS LENS 2004 (UFL'04) LANGUAGE TECHNOLOGY (LT) LIMITED USER EVALUATION (LUE)

The overall intent of the MEC with respect to language translation technology is to:

- Introduce/socialize language technology (LT) / machine translation (MT) tools to operational units and use these tools in conjunction with unit level and task force operations both in-garrison and deployed.
- Assess, develop, and/or revise tactics, techniques and procedures (TTPs) to support the introduction of useful LT/MT tools into the Operational Architecture (OA). Engage the warfighter with developmental systems/tools/TTPs earlier and often.
- Develop and refine LT/MT requirements and capabilities (or required capabilities) as experience is gained from usage by operational forces. Develop and refine TTPs or concepts that enable LT/MT exploitation in both a garrison and field environment.
- Provide operational user assessment and feedback to MFP, OSD, and the development community.
- The expected outcome is to blend re-engineered processes with enabling technology in a spiral development process.

# 5.0 THE LANGUAGE TRANSLATION TOOLS AND TECHNOLOGIES

#### Phraselator P2

- One way phrase based speech to speech machine translator
- Input phrase is linked to a prerecorded output phrase and played through a speaker.
- User can provide information, give orders or directions, and ask simple questions in another language
- Speaker independent (any user without voice enrollment)
- Touch screen option to play phrases



Above photo displays a Phraselator P2

#### Voice Response Translator (VRT)

The VRT provides a one-way voice translation capability for crowd control, or directive type applications in an operational environment. A voice recognition algorithm recognizes a user's voice with near 100 percent accuracy even in high background noise environments; however, this algorithm does require individualized training for each user's speech pattern. Each device will retain up to 8 different user voice profiles. The VRT holds approximately 1000 15-word phrases and can support multiple languages in each unit. It is compact (4.75''l x 3.0''w x 0.75''d) and weighs approximately 1 pound. The design allows for hands-free, eyes-free operation.



Above: U.S Marine Military Policeman Demonstrating Voice Response Translator.

#### Speaking Multilingual Interactive Natural Dialogue System (S-MINDS)

Speaking MINDS is a portable speech-to-speech translation system that both translates and transcribes. S-MINDS provides 2-way (currently English to foreign language with limited foreign language to English response) speech translation in multiple languages, including Korean, Japanese, and Spanish. Additional languages can be added at short notice. It is interactive in that S-MINDS enables a conversation between 2 persons; the speakers alternate speaking in their respective languages and the system recognizes, translates and plays back their dialogue.

S-MINDS consists of a handheld notebook, together with a noise-canceling, hands-free microphone. With speaker-independent speech recognition no voice training is necessary. S-MINDS comes with a Rapid Interview Translation editor that allows the user to add new modules or topics in almost any new language in a matter of hours with the help of a linguist. It is implemented on an ultra-portable notebook PC, but can also work on any other Windows-based system. With an imbedded logging module, all utterances are recorded and a transcribed interaction report is created. An image processing module allows the user to take pictures of people, objects or places and asks questions about the pictures.



Above: Navy Corpsman conduct medical screening of ROK Marine.



Below: ROK Marine provides responses.

#### TrIM

TrIM provides trans-lingual instant messaging and presence information among coalition users by allowing speakers of different languages to chat with each other in their respective native languages. The capabilities of this instant messaging tool include: identify a set of buddies; determine whether they are online; verify their identity; and contact them to request a one-on-one chat session or invite to a multi-party chat room. The instant messaging client/server protocol used is the Simple Instant Messaging and Presence Protocol (SIMP; see simp.mitre.org for protocol details). The TrIM evaluation in UFL'04 was a Limited Military Utility Assessment (LMUA) conducted by Air Force Operaitonal Test and Evaluation Center (AFOTEC) ( ref. LASER ACTD TrIM LMUA Assessment Execution Document (AED)) and is the subject of a separate report.



Above: TrIM Screen Shots

#### SpeechGear 'Expres' PowerPoint Translator

Expres is a tool for a (partially) bi-lingual person, or an English-speaking and foreign counterpart pair to expedite the translation of briefings. An automated translation will require editing. The Korean version of Expres uses bi-directional COTS Machine Translation [MT] (provided by LNI Soft, a Korean software company) supplemented by Translation Memory [TM] phrase lookup tables. The TM files (Korean to English and English to Korean) must be built by the users, or contracted out. In this way, specialized phrases or sentences that are not contained in the COTS dictionaries, but which tend to reoccur from briefing to briefing can readily be translated accurately. Expres also currently supports limited Thai translation, but this was not used in UFL'04.

	👄 Workbench 📃 🗖 🔀
main menu —	-File Tools Help
document window -	-Open Documents: Clipboard
Source -> Target /	Source ->Target Suggestions
Suggestion tabs	Source Text:
Source Text	
	Target Text:
Target Text	
window	
command area —	- Write Add to TM Refresh



# 6.0 ASSESSMENT EXECUTION

**Training.** Prior to UFL'04, during Marine Expeditionary Force Exercise (MEFEX), representatives of the MarForPac Experimentation Center (MEC) and AFOTEC traveled to Palan, South Korea and Okinawa to provide server installation, user training, and to observe same on TrIM. During the period 10-31 August 2004, personnel supporting the MEC traveled to South Korea to provide installation and training on all the deployed language tools, and to evaluate the use of those tools at UFL'04. Training for TrIM and Expres was generally provided on an individual basis as the applications were installed on battle staff computers. Five formal training sessions were offered but of these only two (Expres to the C4 staff at Palan and one TrIM session at Pyongtaek) were well attended. Training on the Phraselator P2, VRT, and S-MINDS was provided, prior to use/role playing, by MEC personnel to small groups at Palan (CMFC MP and BAS), Pyongtaek (Combined Marine Expeditionary Force ((CMEF)), Military Police ((MPs)), Battalion Aid Station ((BAS)), Rear Area Operations Group ((LMCC)), and Rear Area Operations Group ((RAOG)) ), and Camp Humphreys (1<sup>st</sup> MAW MPs). Training efforts were constrained by several factors the foremost being the large number of reserve augmentees, the op-prep tempo prior to STARTEX, and delays caused by the 2-3 day shutdown of CMEF due to a Typhoon warning.

#### Deployment, System Set-up, Execution, and Data Collection.

From 23 August to 1 September 2004, MEC personnel provided additional installations and training, monitored tool employment, and performed data collection for subsequent assessment. The principle goal was to focus on employment venues and socialization of the systems, but also extract basic qualitative and quantitative measures of performance. The intent was for users to gain experience and familiarity with the tools/systems and to begin to discover, through operational use, the potential benefits and/or operationally significant constraints.

Survey questionnaires were used to gather background information about the users, such as rank, specialization, and experience using computers and collaborative tools, as well as to record data about the features and overall user reactions to the language translation tools. Information about usage was gathered through user-written comments in the questionnaires, interviews with the users, and notes taken by the observers. Open-ended interviews were conducted to gather contextual information about the situations in which the tools were used and how they fit into a larger picture of what users are trying to accomplish during the exercise. The emphasis of the observations was on how people interact with the technology and with each other, and on operational processes surrounding the use of these language technology tools.

Language tools were deployed as follows:

#### Combined Marine Forces Command (CMFC), Palan (also spelled like Baran)

- Phraselator P2 (medical and force protection)
- VRT (force protection)
- S-MINDS (medical)
- Trans-lingual Instant Messaging
- 'Expres' PowerPoint Translator

#### Combined Marine Expeditionary Force (CMEF), Pyongtaek (2<sup>nd</sup> ROK FLT)

- Phraselator P2 (medical and force protection)
- VRT (force protection)
- S-MINDS (medical)
- Translingual Instant Messaging (TrIM)

• 'Expres' PowerPoint Translator

#### <u>1<sup>st</sup> Marine Air Wing (1<sup>st</sup> MAW), Camp Humphreys)</u>

- Phraselator P2 (force protection)
- VRT (force protection)

# U.S. Army 599th Transportation Group (599th TG), Busan

• 'Expres' PowerPoint Translator

#### 6.1 ASSESSMENT RESULTS.

The personnel who were trained on and utilized the language translation tools were diverse. Most, but not all, completed tool evaluation surveys. The demographics of the personnel who participated are provided below. The tables that appear in the following sections reflect the composite numerical results of all the collected evaluation forms for that particular tool. Written comments provided by the users were reviewed by MEC evaluators, summarized, and recorded in the pertinent sections of these tables, as appropriate.

# 6.2 Phraselator P2 Observations.

The Phraselator P2 was used with Republic of Korea (ROK) Marine role players at Palan by seven MPs and two Corpsmen, all of whom completed evaluation questionnaires. In addition, approximately 6 MPs attached to the 1<sup>st</sup> MAW at Camp Humphreys were trained (3 of which completed evaluation forms), as were several MPs at the Logistics Movement Coordination Center (LMCC) from the 3<sup>rd</sup> FSSG at Pyongtaek. All the users were enlisted Marines. Although P2 phrase modules for both Force Protection (FP) and Medical categories supported multiple languages, only Korean was used.



Above photo displays a U.S. Marine using the P2 Phraselator

It was observed that, in the Force Protection scenario, several U.S. MP's tended to "freeze up" in the checkpoint role play with the ROK Marines. This might be expected on the first day but it also occurred on subsequent days, even though those Marines had practiced together prior to role playing. It appeared to take significant effort to stand face-to-face with the ROK Marine and operate the device. If simple operation of the P2 and memorizing command phrases is the only consideration, then these devices could be relatively easy to field, but actual use of the device in a scenario appears to be somewhat intimidating from an interpersonal perspective.

#### Table6.1UFL'04 Phraselator P2 ENDEX Questionnaire

(12	Forms	completed.	Figures	in	the	form	body	indicate	number	of	respondents	giving	that
ansy	wer.)												

General Questions	Strongly	Disagree	Agree	Strongly	N/A	Notes
	Disagree			Agree		
The P2 selects the best translation for		1	11	1		1
English input.						
P2 translations accurately capture the		3	8	1		
intended meaning of the original						
message.						
It is easy to get my point across with		6	5	1		2
the Phraselator P2.						
P2 topic areas are sufficient to		4	7	1		
complete the mission.						
The P2 can be used to convey critical		2	10			3
information or terms from English to						
the foreign language.						
The P2 quickly translates from	1	1	8	2		4
English to the foreign language.						
The P2 reduces the time needed to	2	2	7	1		5
convey or acquire information.						
SUITABILITY Setup&Teardown						
The P2 is easy to set up and configure	1		6	4	1	
for use.						
The P2 easily recognizes individual	1	3	5		3	6
voices.						

Notes:

- 1. Occasionally mis-identifies a spoken input phrase.
- 2. Needs more phrases.
- 3. For critical information, users prefer human translators.
- 4. In an emergency situation, the operator would be in trouble. It's not fast enough for use in dangerous situations.
- 5. Users commented on how time consuming it was to search for phrases if one was not yet proficient with the P2.

6. Although P2 is supposed to be speaker independent, it does react differently to some voices. Phrase recognition for female voices is not as good as for male voices.

Occasionally mis- identifies a spoken input phrase for male and female speakers. Background noise sometimes interfered with phrase recognition.

Table 6.1 (Continued)

COMPATIBILITY/ INTEROPERABILITY	Strongly Disagree	Disagree	Agree	Strongly Agree	N/A	Notes
The P2 does not interfere with other warfighting equipment.	1	3	2	2	4	1
TRAINING						
Prepared me adequately for use of the P2 during the exercise.	1	2	8	1		
Given the training I received, I would be able to use the P2 after the exercise without the developers to assist me.	1	1	10			
If fielded, initial training on the P2 would best be conducted at the unit level.	2	1	8	1		2
Continued, expert use of the P2 will require periodic refresher training.		5	6	1		3
I would like to have more training in the following areas prior to using the P2.					12	
USABILITY/HUMAN FACTORS						
The P2 is easy to learn how to use.	1		8	3		
The P2 software interface (menus and commands) is user friendly.		2	9	1		4
It is easy to switch from one topic area to another in the P2.		1	10	1		
P2 hardware (e.g., buttons, stylus, mouse) is easy to use.		1	8	3		
The P2 manual is easy to use.		1	6	3	2	
P2 help menus are useful.		3	5	2	2	5

- 1. But background noise from other equipment sometimes interfered with phrase recognition. Requires use of both hands which is impractical under real world conditions. Hard to use while handling a weapon.
- 2. Some users recommended individual training.
- 3. Every six months.
- 4. One MP commented that the P2 was not user friendly.
- 5. Navigating menus is a little tricky at first.

#### Table 6.1 (Continued)

MAINTAINENCE AND FAILURES	Strongly Disagree	Disagree	Agree	Strongly Agree	N/A	Notes
The P2 operates with minimal failures.		2	6		4	1
The P2 is easy to troubleshoot when problems occur.		2	6		4	
I observed the following <i>hardware</i> malfunctions while using the P2.						2
I observed the following <i>software</i> malfunctions while using the P2.						3
DEPLOYABILITY						
P2 size/weight is suitable for our mission	1	1	7	3		4
P2 is rugged enough for the operational environment		1	7	3	1	5
P2 power requirements are suitable for an operational environment.	1	2	7	1	1	6

- 1. The P2 locked up on several occasions, requiring either a reset or complete power down by removal and reinsertion of the battery pack.
- 2. LCD display was difficult to read in outside daylight conditions.
- 3. Although P2 is supposed to be speaker independent, it does react differently to some voices. Phrase recognition for female voices is not as good as for male voices. Occasionally misidentifies a spoken input phrase for male and female speakers.
  - Background noise sometimes interfered with phrase recognition.
- 4. Too big/bulky for a combat load.
- 5. Not sufficiently field tested to answer this question definitively. Resistance to blowing sand questioned.
- 6. Battery discharges too quickly. Need min 8 hour battery life, since running from AC is not always feasible.

Table 6.1 (Continued)

MISSION IMPACT	Strongly Disagree	Disagree	Agree	Strongly Agree	N/A	Notes
Overall compared to the current system/procedures, The P2 improves my ability to complete a mission.		1	6	3	2	1
P2 enhances productivity (i.e., I can interview more personnel at a faster pace.	3	5	2	1	1	2
Current CONOPS can readily accommodate the P2.	1	1	7	1	2	
Individuals who have no foreign language skills can use the P2 effectively.		1	10	1		
A person with my skill level can easily use P2.			8	4		
P2 can be used to complete the mission without requiring the support of a linguist.	1	4	4	2		3
I would use P2 right now in my day-to-day mission(s) if it were available.		5	6	1		4

Notes:

- 1. Users generally preferred VRT over the P2. Easier to manage.
- 2. Additional training and use and further improvements, including additional modules will be needed before it will enhance productivity.
  - P2 does not effectively substitute for a human translator.
- Difficult to complete without a linguist since in most cases if the subject speaks no English it's very difficult for him/her to convey anything to the questioner. Useful for force protection questioning only if required questioning is not extensive.
- 4. Regular in-garrison training and use required. Users generally preferred VRT over the P2. Additional force protection phrases required: "I'm sorry, but you are not authorized to pass."
  - "It is necessary for me to take (confiscate) your..."
  - "I need to search your bag/person/vehicle."

Additional medical phrases required:

"Raise your right hand if you're here because you're sick; `left hand if injured."

# 6.2 VRT Observations

The VRT was used with ROK Marine role players at Palan by six MPs, all of whom completed evaluation questionnaires. In addition, approximately 6 MPs attached to the 1<sup>st</sup> MAW at Camp Humphreys were trained (4 of which completed evaluation forms), as were several MPs and some LMCC Marines with the 3<sup>rd</sup> FSSG at Pyongtaek. All the users were enlisted Marines. Although the VRT phrase modules for Force Protection supported multiple languages, only Korean was formally evaluated. An informal test was performed using Thai with a Thai speaking reservist attached to the CMEF Rear Area Operations Group (RAOG). The same comments pertaining to Phraselator P2 users freezing during the Force Protection (FP) role play also apply to the VRT.

# Table 6.2.1 UFL'04 VRT ENDEX Questionnaire

(1)	) Forms completed	Figures in the f	orm body indicate	the number of respon	dents giving that answer)
(1)	i onno compretea.	i iguies in the r	onn oouy maleute	the number of respon	donto giving that anower.

USABILITY/HUMAN FACTORS	Strongly Disagree	Disagree	Agree	Strongly Agree	N/A	Notes
VRT is easy to learn how to	_		8	2		
use.						
Easy to locate/remember the		2	5	2	1	1
command I need in order to						
convey my message.						
It is easy to add new					10	2
commands.						
The VRT headset/	2	3	4	1		3
microphone is comfortable.						
VRT manual is easy to use.			8	1	1	
SUITABILITY						
SETUP-TEARDOWN						
VRT is easy to set up and			7	2	1	
configure for use.						
Easy to train VRT to recognize	1	2	4	1	2	4
my voice.						
VRT does not interfere with	2	1	2		5	5
other warfighting equipment.						

- 1. The laminated phrase cards are useful in this regard. Many of the key phrases/commands did not relate well to the complete phrase, and were
  - difficult to remember.
- 2. Users did not have access to this feature.
- Needs to have adjustable headset to accommodate female hairstyles, and head shape/size. Also needs replaceable ear and mouth pieces.
   Uncomfortable with Keyler helmet
  - Uncomfortable with Kevlar helmet.
- 4. Training for female voices is much more difficult than for male voices. Even male voice commands often needed repeating.
- 5. But background noise from other equipment sometimes interfered with phrase recognition. Interferes with Kevlar helmet and gas mask.

#### Table6.2.1 (Continued)

TRANSLATION	Strongly Disagree	Disagree	Agree	Strongly Agree	N/A	Notes
	Disagi ee	1		ngite		1
VRI easily recognizes and	2	1	5	2		I
translates my English commands.				1		
I estimate that the VRT		2		1		
recognizes approx. $40-95$ % of						
my English commands on the first						
attempt.				1		
Easy to get my point across with		2		1		
the VRT.						
The commands currently		2	8			
available in VRT are sufficient to						
complete the mission.						
The VRT can be used to convey		2	8			2
critical information from English						
to a foreign language.						
VRT quickly translates from		1	8	1		3
English to a Foreign language.						
VRT reduces the time needed to		1	8	1		4
convey information.						
TRAINING						
The training prepared me		1	8	1		5
adequately for use of VRT during						
the exercise.						
Given the training I received, I		1	7	1	1	6
would be able to use VRT after						
the exercise without the						
developers help.						
If fielded, initial training on VRT		4	5	1		7
would best be conducted at the						
unit level.						
Continued, expert use of VRT		3	6	1		
will require periodic refresher						
training.						

- 1. For male voices some phrases were difficult to recall-VRT very sensitive to the trained phrases and speaker deviation from that, due to stress or other factors causes recall problems. Background noise sometimes interfered with phrase recognition.
- 2. Critical information, such as deadly force warning, does not convey with the vocal emphasis needed. For critical information, human translators are generally needed.
- 3. Emergency phrases such as "Put your hands in the air and drop your weapon") wouldn't be fast enough to preclude a life-threatening situation.
- 4. Can be time consuming.
- 5. Laminated phrase cards help as a memory aid.
- 6. Without lots of practice and the laminated phrase cards, first time use can be difficult.
- 7. Initial training needs to be with senior leadership prior to any unit level training.

Some users thought individual training was best.

#### Table6.2.1 (Continued)

MAINTAINENCE AND FAILURES	Strongly Disagree	Disagree	Agree	Strongly Agree	N/A	Notes
VRT operates with minimal failures		3	5	1	1	
The VRT is easy to troubleshoot when		2	5	1	3	
problems occur		2	5		5	
I observed the following malfunctions						1
while using VRT						
DEPLOYABILITY						
VRT size/weight is suitable for our		2	5	3		2
mission.						
VRT is rugged enough for the operational		1	4	1	4	3
environment						
VRT power requirements are suitable for		1	6		3	4
an operational environment.						
MISSION IMPACT						
Overall compared to the current system/			5	2	3	5
procedures, VRT improves my ability to						
complete a mission.						
VRT enhances productivity (i.e., it helps		1	6	2	1	6
me do my job more efficiently.)						
Current CONOPS can readily			5		4	
accommodate VRT.						
Individuals who have no foreign language			5	3	2	
skills can use the VRT effectively.						
A person with my skills level can easily			6	4		
use VRT.						
VRT can be used to complete the mission		2	6	2		
without requiring the support of linguist.						
I would use VRT right now in my day-to-			7	1	2	7
day mission(s) if it were available.						

Notes:

- 1. Sometimes voice commands are not recognized.
- 2. Size good and could be hands free, but need a method to attach externally to uniform or gear. Would be good if VRT pouch could attach to war belt.
- 3. Not sufficiently field tested to answer this question definitively.
- 4. Need min 8-12 hour battery life, since running from AC is not always feasible on shift.
- 5. Not efficient for female voices, but agree for male voices.
- 6. Difficult to complete without a linguist since in most cases if the subject speaks no English it is very difficult for him/her to convey anything to the questioner.
- 7. Some mistranslations noted by linguists

Additional force protection phrases required:

"I'm sorry, but you are not authorized to pass."

"It is necessary for me to take (confiscate) your...."

"I need to search your bag/person/vehicle."

Additional medical phrases required:

"Raise your right hand if you're here because you're sick; left hand if you're injured.

1<sup>st</sup> MAW stated desire to use VRT Japanese in-garrison in Okinawa.

# 6.3 S-MINDS Observations

S-MINDS was used by two (2) Corpsmen interacting with six (6) ROK Marine role players from the ROK Marine Headquarters communications battalion at Palan, and five (5) Corpsmen using ad hoc Korean role players at Pyongtaek. In all cases the Korean/English language pair was used. Evaluation forms were collected from all participants at Palan. At Pyongtaek, evaluation forms were received from all five U.S. participants and one of the Korean participants. Since the evaluation forms that were printed in Korean varied somewhat from those used by the U.S. participants, the results are presented in two separate tables.

# Table 6.3-1 UFL'04 S-MINDS ENDEX Questionnaire - U. S. Personnel

# (14 Forms completed. Figures in the form body indicate number of respondents giving that answer)

EFFECTIVENESS	Strongly Disagree	Disagree	Neither Agree	Agree	Strongly Agree
	Disagice		or		Agree
			Disagree		
S-MINDS successfully recognized		1	4	8	1
English utterances during the evaluation.					
S-MINDS successfully recognized		1	4	7	1
Korean responses during the evaluation.					
The range of anticipated English		6	4	3	1
utterances was appropriate for the task.					
The range of anticipated Korean		4	5	4	1
responses was appropriate for the task.					
S-MINDS is effective for screening		1	8	3	2
Korean speakers needing medical					
treatment.					
S-MINDS is effective for rudimentary		2	5	6	1
spoken communication with Korean					
speakers.					
The Korean translations of English		1	8	4	1
utterances were understood by Korean					
speakers during the evaluation.					
The English translations of Korean			3	8	3
responses were understood by English					
speakers during the evaluation.					
S-MINDS is suitable for initial medical			7	4	2
screening of Korean-speaking persons.					
S-MINDS software is suitable for use on		2	4	5	2
computers that accompany liaison					
personnel for Korean units.					
S-MINDS software is suitable for use on			7	4	2
computers that accompany medical					

personnel operating near Korean units.			

Table6.3-1 (Continued)

EFFECTIVENESS (Continued)	Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree
S-MINDS s/w is suitable for use on com- puters that accompany medical personnel operating near Korean local nationals.			6	5	2
Computers with S-MINDS are suitable for use in an operational environment.		4	5	1	
S-MINDS is ready for immediate deployment.		4	7	1	1
USABILITY					
The training I received on S/MINDS was adequate for this demonstration.			3	9	1
It is easy to learn to operate S/MINDS.			3	9	1
It is easy to conduct medical screening of Korean-speaking personnel with S/MINDS.		5	3	5	
It is easy to hear audio information played by S/MINDS.		1	4	7	1
It is easy to read text information displayed by S/MINDS.				9	4
Graphical information displayed by S/MINDS is useful.			5	3	5
The layout of information in S/MINDS is easily customized for special situations		4	2	2	5

User Comments:

-Need to have questions form Mosby and Bates physical exam book in S-MINDS.

-Questions that should be included in S/MINDS. "Are you Sick?" or "Are you injured?"

-Need more basic questions. There's currently not enough in S-MINDS to diagnose the problem.

-S-MINDS is susceptible to background noise, such as that from tent/room air-conditioners,

causing translation problems. (Several users commented on this.)

-S-MINDS would be well suited for a NEO.

-S-MINDS is an effective tool for communication.

# Table 6.3-2 UFL'04 S-MINDS ENDEX Questionnaire Republic of Korea Personnel

(7 Forms completed. Figures in the form body indicate number of respondents giving that answer. Note that on the eval forms, 1 role player stopped answering after question 15, another after question 30, and another checked N/A for everything after question 9.)

EFFECTIVENESS	Strongly Disagree	Disagree	Agree	Strongly Agree	N/A
The topics available in S-MINDS effectively support operations.			7		
The questions within each topic are sufficient.		3	2		2
It is easy to conduct queries via S- MINDS.		1	3	1	2
S-MINDS translations accurately capture the intended meaning of the original message.			6		1
The word choice in S-MINDS is accurate.			6	1	
The word order in S-MINDS translations is accurate.			4	2	1
S-MINDS provides timely translations.		2	4	2	
S-MINDS reduces the time needed to convey or acquire information.			5	1	1
SUITABILITY Set-up & tear down					
S-MINDS is easy to set up.			3		4
S-MINDS is easy to tear down.	1		1	1	4
S-MINDS is easy to pack.		1	1	1	4
S-MINDS software is easy to configure for use.			3		4
S-MINDS easily recognizes individual voices.	1		3	1	2
DEPLOYABILITY					
S-MINDS size/weight is suitable for mission operations.	1		3	1	2
S-MINDS is rugged enough for the operational environment.			2	1	3
COMPATIBILITY/ INTEROPERABILITY					
S-MINDS translation software does not conflict with other software			2		4
S-MINDS does not interfere with any other warfighting equipment.			3		3

#### Table6.3-2 (Continued)

TRAINING	Strongly Disagree	Disagree	Agree	Strongly Agree	N/A
The training prepared me adequately for use of S-MINDS			3	1	2
Given the training I received, I would be able to use S-MINDS after the exercise without the developers to assist me.	1		3		2
If fielded, initial training on S-MINDS would best be conducted at the unit level.		1	1	1	3
Continued, expert use of S-MINDS will require periodic refresher training.			1	3	2
I would like to have more training in the following areas prior to using S-MINDS.					6
USABILITY/HUMAN FACTORS					
S-MINDS is easy to learn how to use.		1	2	1	2
S-MINDS software interfaces (menus and commands) are user friendly.			3		3

#### 6.4 SpeechGear Expres PowerPoint Translator Observations

The 'Expres' PowerPoint translation tool was widely dispursed at both CMFC (Palan) and CMEF (Pyongtaek). Ten installations were performed at CMFC with training provided to eight individuals (two Captains, four Majors, and two civilians) and a few additonal staff in the initial group training. In some cases, operational tempo precluded training with the applicable staff member, and some declined training. Operators required 'power user' privileges on the CMFC computers in order to use Expres due to system administrators imposed security constraints. Eleven copies of Expres were installed at CMEF with individual training provided to all (Sergant through LtCol). One copy was used by the Army's 599<sup>th</sup> Transportation Group at Busan. In all, fifteen (15) evaluation forms were collected – two (2) from the Rear Area Operations Group (RAOG) (one from a U.S. member, one from a ROK member) and three (3) from Combined Marine Expeditionary Force (CMEF) C3 Future Operations (FUTOPS), one of which dealt with a varient of Expres which also translated MS Word documents. The remaining ten consisted of one each from Combined Marine Forces Command (CMFC) C3(SITREP), C4, and C6, /SITREP, and CMEF C3/IMO, C3 Watch Clerk, C2 OPSO, C2 Watch Officer, C4 OPSO, C1 Ops Clerk, and the 599<sup>th</sup> Transportation Group.

#### Table 6.4 UFL'04 Speechgear Expres ENDEX Questionnaire

GENERAL QUESTIONS	Strongly Disagree	Disagree	Agree	Strongly Agree	N/A	Notes
Exprés language dictionaries effectively support operations.	1	5	8		1	1
Exprés translations accurately capture the intended meaning of the original PowerPoint slides.	3	6	5			2
The word choice in Exprés translations is accurate.	1	5	8			3
The word order in Exprés translations is accurate.	2	1	8		1	4
Exprés can easily handle acronyms.	6	5	3			5
Exprés translations can be used to identify critical information or terms in the PowerPoint slides.	4	2	7		2	
Exprés can be used to quickly translate PowerPoint slides.		2	9	3	1	6
Exprés reduces the time needed to convey or acquire information in briefings.	1	7	6	1		7

#### (15 forms completed. Figures in the form body indicate number of respondents giving that answer.)

- 1. Errors encountered mostly in military and technical terms.
  - The extent to which abbreviations and terse sentence structure is used in the operational environment causes severe problems with the translations.
- 2. If the author articulated his/her intent carefully, but this is too time consuming for an operational environment.
- 3. Translation accuracy is about 60%
- 4. The word order in Exprés translations is accurate, given the proper input sentence structure. Fragments/bullets seem to confuse the engine.
- 5. Evaluators comment-Policy calls for U.S. acronyms to be passed through untranslated. Many times what appears as a term acronym is instead an abbreviation of the sentence structure (i.e. IOT-in order to..) which will be lost in the translation.
- 6. Most useful for longer presentations-saves about ten minutes per slide vs. human translation from scratch. (CMEF/RAOG).
- 7. Very time consuming to figure out if the translation is appropriate and to correct if not.

# Table 6.4 (Continued)

SUITABILITY	Strongly Disagree	Disagree	Agree	Strongly Agree	N/A	Notes
Exprés software is easy to install.		1	7	1	6	1
It is easy to configure the system to support multiple languages.		1	7	1	6	2
COMPATIBILITY/ INTEROPERABILITY						
Exprés does not conflict with other software programs (e.g., operating system, audio system) on the computer			9	2	3	
TRAINING		· ·				
The training prepared me adequately for use of Exprés during the exercise.	2	4	8	1		3
Given the training I received, I would be able to use Exprés after the exercise without the developers to assist me.		4	10	1		
If fielded, initial training on Exprés would best be conducted at the unit level.		3	8	3	1	4
Continued, expert use of Exprés will require periodic refresher training.		3	8	3	1	
I would like to have more training in the following areas prior to using Exprés.						5
MAINTENANCE AND FAILURES						
Exprés operates with minimal failures.		2	9	2	2	
Exprés is easy to troubleshoot when problems occur.		3	5	7		
I observed the following <i>hardware</i> malfunctions while using Exprés.						
I observed the following <i>software</i> malfunctions while using Exprés.						

- Only 599<sup>th</sup> TG personnel performed their own installation.
   Only Korean and English used, but switching translation directions was easy.
- 3. Training should have been provided prior to deployment.
- 4. On-line, interactive tour/tutorial recommended.
- 5. Populating the Translation Memory files and verifying with the Koreans that the entries are correct.

#### Table 6.4 (Continued)

USABILITY/HUMAN FACTORS	Strongly Disagree	Disagree	Agree	Strongly Agree	N/A	Notes
Exprés is easy to learn how to use.			15			1
Exprés software interfaces (menus and commands) are user friendly.		1	13	1		
Exprés manual is easy to use.		3	8	3	1	2
Exprés help menus are useful.			11	1	3	
Exprés does not interfere with original formatting of PowerPoint slides.	1	2	9	3		
MISSION IMPACT						
Overall, compared to the current system/procedures, Exprés Improves my ability to complete a mission.		6	7	2		3
Exprés enhances productivity (e.g., I can do my job more efficiently).		7	6	2		4
Current CONOPS can readily accommodate Exprés.	1	5	8	1		
Individuals who have no foreign language skills can use Exprés effectively.						5
A person with my skill level can easily use Exprés						
Exprés can be used to complete the mission, without requiring the support of a linguist.						6
I would use Exprés now in my day-to- day mission(s) if it were available.						

- 1. Only trick to learning is to remember to lowercase text so that words are not mistaken for acronyms and passed through untranslated, and to keep the 'Workbench' open.
- 2. Introduces wild card characters (square boxes) at the end of each translation block, that needs to be moved.
  - Translations can spill over and obscure other text boxes or graphics, unless you pre-plan the layout to avoid this.
- 3. Allowed greater integration of Korean LNO's into the working cell.
- 4. Not yet too time intensive to edit.
- 5. But requires careful editing by a native language speaker.
- 6. Translated presentations still need to be edited by a linguist.

#### 6.5 TrIM Observations

The TrIM evaluation in UFL'04 was a Limited Military Utility Assessment (LMUA) conducted by the Air Force Operational Test and Evaluation Command (AFOTEC) (ref. LASER ACTD TrIM LMUA Assessment Execution Document (AED)) and is the subject of a separate report.

# 7.0 CONCLUSIONS

All of the language translation tools were capable of basic translation functions. Full utilization was hampered because they were not part of the exercise training objectives or scripted into the UFL'04 scenario. Part of the utilization problem, with the MT-based tools, clearly rests with the state-of-the-art of Korean machine translation software and the completeness of the dictionaries (with respect to containing sufficient and proper military terms. For all of the technologies, use might have been more widespread and consistent during the exercise if the tools had been used regularly, in-garrison, prior to deployment or if users had been utilizing them during the workup phases of UFL'04 (TrIM was used during MEFEX '04).

The Phraselator P2 and VRT were readily accepted by the users and basic training was completed in one to two hours. Of the two devices, the smaller size and potential for hands free operation of the VRT made it the preferred tool over the P2 for the MPs. The Phraselators still frequently lock up and require the user to reset or sometimes to completely power down by removing and re-inserting the battery. This is a long-standing problem that needs to be fixed. Since the P2 and VRT only have the capability for uni-directional translation, their use is limited to routine/structured situations. S-MINDS was well suited to office style medical interviews and with the exception of a desire for more phrases (echoed for all three speech devices), especially ones from standard medical diagnosis books, the users found S-MINDS an effective means for communicating with non-English speakers. All three tools suffered interference from background noise in loud environments that caused missed inputs, or improper identification of the spoken phrase. While the users were ultimately successful in using the technology to converse with non-English speaking Korean personnel, success was limited in that it was achieved under supervised conditions with some coaching. The "freeze up" reaction observed in some of the FP scenarios suggests that there needs to be a training element for these devices that includes significant actual role playing with linguists and real time feedback to build user confidence.

Various elements of CMEF/C2 and C3, as well as CMFC/C3 and C4, realized the potential of the Expres PowerPoint translator and are supportive of the technology, but found the resultant product still too time consuming to edit versus having a linguist do the translation. The 599<sup>th</sup> TG and the CMEF/RAOG found the tool most useful and were positioned well to use it in the manner intended/prescribed (i.e. as a tool for generating an approximate translation for U.S./ROK counterpart editing for accuracy to final form). The Expres PowerPoint translator might have been utilized more during the actual exercise if the users had been exposed to it prior to deployment and used the tool during the workup phases to the exercise. Although available, the military dictionaries for the underlying LNI Soft translation engine used in Expres were absent and this likely affected the quality of the translation. The lack of any prepared custom Translation Memory file(s) for Expres, combined with the fragmented sentence structure found in briefings, also impacted the accuracy of the translations. Current staff estimate of the accuracy of the underlying translation engine used in this product is about 60%. The practice of abbreviating sentence structure in text during operations will require this tool, and other text translation tools, to distiguish between term acronyms (like Combined Marine Forces Command, CMFC) and abbreviations (such as BTW for by the way, IOT for in order to, or proper abbreviations like abbrev., etc.) and to automatically expand and translate the abbreviations. By policy, English acronyms are usually not translated.

All five tools brought to UFL'04 were demonstrated for General LaPorte (CG CFC & USFK), Lt Gen. Gregson (CG MFP), and his Chief of Staff, Colonel O'Neal. Both Lt Gen. Gregson and Col. O'Neal had prior exposure to many of the products. The demonstration to Gen. LaPorte took place on 30 August and included TrIM, Expres PPT Translator, S-MINDS, Phraselator P2, and the VRT. He seemed especially interested in the inter-personal communications tools (TrIM and the speech devices).

# 8.0 RECOMMENDATIONS

The following recommendations are made:

Speech to Speech devices

- Simplify the procedure for phrase addition in the field.
- Fix the Phraselator lock up problem.
- Increase the degree of speaker independence of these devices, specifically to eliminate VRT voice training and the resultant issues that arise when a speaker is sick or under stress, and in general to respond better to female voices.
- Re-examine VRT microphone/headset design for better comfort and compatibility with helmets, gas masks and women wearers.
- Prepare additional modules (Marine in a foreign land) per request of 1<sup>st</sup> MAW and provide additional P2s/VRTs for their FP use in garrison (Japanese).
- Provide training for these devices that includes significant actual role playing with linguists and real time feedback using graded tasks to build user confidence.

PowerPoint translator

- Identify units that have a requirement and the willingness to use this tool. Distribute licenses for routine use.
- MEC coordinate with MFP and III MEF staffs to compile phrases and translations likely to be encountered in briefings for use in populating Translation Memory [TM] files. Complete prior to exercise COBRA GOLD '05 for Thai and UFL '05 for Korean and provide to the MEC for submission to the vendors NLT 30 January 05 and 30 March 05 respectively.
- Prepare a list of operational requirements for machine translation software (i.e. use of TM, handling of abbreviations, handling of term acronyms, adjustment for format changes, accuracy, etc.).

#### General

- Train with language tools routinely in-garrison, or at a minimum train one month prior to STARTEX with additional training provided on-site at the exercise. Training sessions for the language technology tools at the exercise should be incorporated with / appended to approve training classes for the operational systems to maximize attendance.
- Training for foreign nation partner staff needs to be accommodated, and arranged early in the exercise planning cycle with buy-in from their leadership.
- Institute a formal plan for advertising the availability of computer based language tools to the exercise participants to include exposure at the planning conferences, postings on the exercise WAN web home page, and all-hands email announcements a couple of days prior to STARTEX.

# 9.0 SUMMARY

In summary, when given the right environment, coordination with the appropriate users, (Intelligence Community operatives, military law enforcement agencies, medical and civil operations

personnel, perspective coalition partner nations), and the active support of the recipient commands at all levels, these tools/technologies can be valuable assets for easing the language barrier.