

## Chapter 1 – INTRODUCTION

### 1.1 MILITARY IMPORTANCE

As speech-processing technology becomes mature, the potential to utilize the technology for speech-enabled military systems strongly increases. The technology can be embedded in military communication, command and control, intelligence, and training systems. Interoperability of these systems is paramount to the success of NATO multi-national operations. This, however, creates interesting and unique problems in the successful implementation of speech technology, where multi-national forces working in a coalition environment exist. In this environment, speech-processing equipment designed by one country must be used by soldiers from another. Unlike other military systems, where interoperability could be created by simply rewriting a user's manual in the native language for a particular soldier, speech systems must be created and measured for effectiveness before deployment. Interoperability of military systems such as speech coders, voice controlled C2 systems, speaker and language recognition, and automatic training suites are not a simple standardization problem. The speech of each individual user is an uncontrolled variable. The use of speech systems by non-native speakers speaking the official NATO languages, French and English, may cause reduced performance or even complete malfunction of a system, especially in a battlefield environment. Standardized assessment methods, specifications, and training techniques are required for both commercial-off-the-shelf (COTS) and for the development of new technology-based military systems.

### 1.2 TECHNICAL CHALLENGE

The IST-RTG013 recognized the need to perform research and studies on this topic to better understand, detect, and mitigate the effects of native and non-native speech production in military battlefield environments. Minimal research had been conducted in this area prior to the initiation of this project. Commercial systems were built with little regard for non-native speech production and battlefield noise and channel conditions. As a result, interoperability of systems developed for specific languages becomes an issue, especially when military forces are pressed into action often with short notice. Examine the case where in a particular operation a native speaker of Dutch speaking Dutch must use a speech coder in a secure communication device, which was optimized for British English. Imagine the case where a native speaker of German might need to use a speech translator trained for Spanish. Interoperability of speech systems is an important issue for many applications of modern speech technology in the coalition environment. For this reason, the NATO Research and Technology Organization (RTO) under the Information Systems Technology (IST) Panel authorized a task group to identify the application of and assess the use of multilingual speech technology in the military battlefield environment.

### 1.3 WORK PROGRAM

In the past, TG01 constructed projects which studied the various effects of military environments in relation to the performance of speech technology. Examples are the effect of noise on speech recognition, the effect of stress induced by workload, sleep deprivation, and battlefield stress. The biggest impact of these projects was the creation of datasets representative of the military environment, which fostered interest in the academic and industrial scientific communities. This has shaped the development and evaluation of speech technology for the harsh military environment.

Speech data was collected in three conditions representative of military battlefield conditions to foster research on multi-lingual, non-native speech in battlefield conditions. This data set is very representative of military type

communication in a military battlefield scenario, and was used for evaluation and modification of Automatic Speaker Recognition and Word Spotting. These databases also focused research on non-native speech and robustness issues, which led to a special session at the international speech and language conference, Interspeech 2007 in Antwerp, Belgium.

### **1.4 REPORT ORGANIZATION**

This report is organized into five chapters. Below is a description of the content in each chapter.

#### **Chapter 1:**

This chapter contains an introduction to the project and describes how the report is organized.

#### **Chapter 2:**

This chapter presents the various military databases which were considered and/or collected for this project. Also included in this chapter is a detailed description of these databases. An overall description of each database and its content, amount of data, language, non-native type, and characteristics is included.

#### **Chapter 3:**

This chapter presents an experimental plan on using the database (nnMATC) to measure the performance of speech processing systems on the problems of detection, classification, and assessment of non-native, accented speech in a realistic battlefield environment.

#### **Chapter 4:**

The issues and findings of various speech systems are presented.

#### **Chapter 5:**

In this chapter conclusions are drawn. A discussion of the impact that multi-lingual and non-native speech in a realistic battlefield environment has on military speech technology and its application is presented.