NORTH ATLANTIC TREATY ORGANISATION

RESEARCH AND TECHNOLOGY ORGANISATION

AC/323(IST-011)TP/26

www.rta.nato.int

RTO TECHNICAL REPORT **TR-IST-011**

Implications of Multilingual Interoperability of Speech Technology for Military Use

(Les implications de l'interopérabilité multilingue des technologies vocales pour applications militaires)

This Technical Report has been prepared as a result of a project on "Implications of Multilingual Interoperability of Speech Technology for Military Use" for the RTO Information Systems Technology Panel (IST).

Published September 2004

NORTH ATLANTIC TREATY ORGANISATION

RESEARCH AND TECHNOLOGY ORGANISATION

AC/323(IST-011)TP/26

www.rta.nato.int

RTO TECHNICAL REPORT **TR-IST-011**

Implications of Multilingual Interoperability of Speech Technology for Military Use

(Les implications de l'interopérabilité multilingue des technologies vocales pour applications militaires)

This Technical Report has been prepared as a result of a project on "Implications of Multilingual Interoperability of Speech Technology for Military Use" for the RTO Information Systems Technology Panel (IST).

by

Dr. Timothy ANDERSON (Chairman), USA
Dr. Stéphane PIGEON, Belgium
Mr. Carl SWAIL (Secretary), Canada
Dr. Edouard GEOFFROIS, France
Ms. Christine BRUCKNER, Germany
Dr. David van LEEUWEN, The Netherlands
Prof. Carlos TEIXEIRA, Portugal
Mr. Ozgur ORMAN, Turkey
Mr. Paul COLLINS, United Kingdom
Mr. John GRIECO, USA
Dr. Marc ZISSMAN, USA

The Research and Technology Organisation (RTO) of NATO

RTO is the single focus in NATO for Defence Research and Technology activities. Its mission is to conduct and promote co-operative research and information exchange. The objective is to support the development and effective use of national defence research and technology and to meet the military needs of the Alliance, to maintain a technological lead, and to provide advice to NATO and national decision makers. The RTO performs its mission with the support of an extensive network of national experts. It also ensures effective co-ordination with other NATO bodies involved in R&T activities.

RTO reports both to the Military Committee of NATO and to the Conference of National Armament Directors. It comprises a Research and Technology Board (RTB) as the highest level of national representation and the Research and Technology Agency (RTA), a dedicated staff with its headquarters in Neuilly, near Paris, France. In order to facilitate contacts with the military users and other NATO activities, a small part of the RTA staff is located in NATO Headquarters in Brussels. The Brussels staff also co-ordinates RTO's co-operation with nations in Middle and Eastern Europe, to which RTO attaches particular importance especially as working together in the field of research is one of the more promising areas of co-operation.

The total spectrum of R&T activities is covered by the following 7 bodies:

- AVT Applied Vehicle Technology Panel
- HFM Human Factors and Medicine Panel
- IST Information Systems Technology Panel
- NMSG NATO Modelling and Simulation Group
- SAS Studies, Analysis and Simulation Panel
- SCI Systems Concepts and Integration Panel
- SET Sensors and Electronics Technology Panel

These bodies are made up of national representatives as well as generally recognised 'world class' scientists. They also provide a communication link to military users and other NATO bodies. RTO's scientific and technological work is carried out by Technical Teams, created for specific activities and with a specific duration. Such Technical Teams can organise workshops, symposia, field trials, lecture series and training courses. An important function of these Technical Teams is to ensure the continuity of the expert networks.

RTO builds upon earlier co-operation in defence research and technology as set-up under the Advisory Group for Aerospace Research and Development (AGARD) and the Defence Research Group (DRG). AGARD and the DRG share common roots in that they were both established at the initiative of Dr Theodore von Kármán, a leading aerospace scientist, who early on recognised the importance of scientific support for the Allied Armed Forces. RTO is capitalising on these common roots in order to provide the Alliance and the NATO nations with a strong scientific and technological basis that will guarantee a solid base for the future.

Published September 2004

ISBN 92-837-1118-1

Implications of Multilingual Interoperability of Speech Technology for Military Use

(RTO-TR-IST-011)

Executive Summary

Multilingual speech and language technology is becoming recognized as an important domain for international organizations, both civilian and military. For instance, one might want to use a speech coder optimized for French in Germany or Turkey. A native speaker of Spanish might want to use a speech recognizer trained for American English. Additionally with the explosion of multilingual text material on the Web, a British user might want to access Dutch documents using English search terms. For reasons such as these, a special task group of the NATO Research and Technology Organization (RTO) started a project on the development and assessment of multilingual speech and language applications.

To stimulate research and evaluation the NATO Research Study Group on Speech and Language Technology (IST-011/RTG-001) compiled a database of native and non-native speech. This database is called the NATO Native and Non-Native (N4) Speech Corpus. Studies conducted by participating NATO laboratories and discussed here suggest that many COTS speech systems, which were designed for native speakers cannot be effectively used for non-native speakers. The main findings and recommendations are:

- It is suggested that the effect of non-native speech on the quality of the speech produced is likely to be detrimental to the effectiveness of communication in general, in particular to the performance of communication equipment and weapon systems equipped with vocal interfaces (e.g., advanced cockpits, command, control, and communication systems, information warfare).
- Commercial off-the-shelf speech recognition systems are not yet able to handle the wide speaker variability associated with non-native speech.
- Databases obtained or compiled during this study have been distributed to all participating NATO countries, and most are available in CD-ROM format.
- Progress in the field of military based speech technology, including advances in speech based system design has been restricted due to the lack of availability of databases of non-native speech in military environments.
- It is foreseen that in the future it will be necessary to improve the coordination of multi-national military forces. The need therefore exists for planned simulation exercises involving military personnel using a wide range of speech technology.
- Military operations are often conducted under conditions of stress induced by high workload, sleep deprivation, fear and emotion, confusion due to conflicting information, psychological tension, pain, and other typical conditions encountered in the modern battlefield context. These conditions, combined with the effects of non-native speech will present challenges for speech technology for a long time to come.

Les implications de l'interopérabilité multilingue des technologies vocales pour applications militaires

(RTO-TR-IST-011)

Synthèse

L'importance des technologies du traitement multilingue de la parole et du langage est de plus en plus reconnue comme un domaine important par les organisations internationales civiles et militaires. Il se pourrait, par exemple, qu'un codeur vocal optimisé pour le français soit demandé en Allemagne ou en Turquie. De la même façon, il se pourrait qu'un hispanophone ait besoin d'un système de reconnaissance de la parole conçu pour l'anglais américain. En outre, avec le foisonnement de textes multilingues affichés sur le Web, il se pourrait qu'un utilisateur britannique veuille consulter des documents rédigés en néerlandais en se servant de termes de recherche anglais. Pour de telles raisons, un groupe de travail de l'Organisation pour la recherche et la technologie de l'OTAN (RTO) a lancé un projet sur le développement et l'évaluation d'applications à la parole et au langage multilingues.

Afin de stimuler la recherche et l'évaluation, le Groupe d'étude OTAN sur les technologies du traitement de la parole et du langage (IST-011/RTG-001) a créé une base de données de la parole autochtone et non autochtone. Cette base de données s'appelle le corpus OTAN de la parole autochtone et non autochtone (N4). Les études réalisées par les laboratoires OTAN participants et qui sont examinées ici indiquent que bon nombre de systèmes de parole COTS qui sont conçus pour des autochtones, sont inadaptés à des non autochtones. Les principales conclusions et recommandations sont les suivantes :

- Il est soutenu que l'effet de la parole non autochtone sur la qualité de la parole risque de nuire à l'efficacité de la communication en général, en particulier en ce qui concerne les performances du matériel de communication et des systèmes d'armes équipés d'interfaces vocales (par exemple les postes de pilotage avancés, les systèmes de commandement, contrôle et communications et de guerre de l'information).
- Les systèmes de reconnaissance de la parole disponibles sur étagère ne sont pas encore en mesure de gérer les grandes variations entre locuteurs associées à la parole non autochtone.
- Des bases de données obtenues ou recueillies au cours de cette étude ont été diffusées à l'ensemble des pays de l'OTAN participants, et la plupart d'entre elles sont disponibles sous forme de CD-ROM.
- Les avancées dans le domaine des technologies vocales militaires, y compris les avancées dans la conception des systèmes de parole, ont été freinées par la non-disponibilité de bases de données contenant des exemples de paroles non autochtones produites en environnement militaire.
- A l'avenir, il deviendra de plus en plus nécessaire d'améliorer la coordination des forces militaires internationales. Le besoin existe donc, de simulations du champ de bataille intégrant des personnels militaires internationaux et faisant appel à un éventail de technologies vocales.
- Les opérations militaires sont souvent conduites dans des conditions de stress induites par des charges de travail élevées, le manque de sommeil, la peur et l'émotion, la confusion due à des informations contradictoires, la tension psychologique, la douleur, et par d'autres conditions typiques du champ de bataille moderne. Ces conditions, associées aux effets de la parole non autochtone, continueront de poser des défis pour les technologies de la parole à l'avenir.

Table of Contents

List of Figures and Tables

Preface

Communications, command and control, intelligence, and training systems are increasingly making use of speech technology components: i.e. speech coders, voice controlled C^2 systems, speaker and language recognition, and automated training suites. Interoperability of these systems is not a simple standardization problem, as the speech of each individual user is an uncontrolled variable as in the case of non-native speakers using, in addition to their own language, an official NATO language. For multi-national operations, this may reduce performance or even cause malfunction of an action. Standardized assessment methods and specifications for both commercial-of-the-shelf (COTS) and for development of new technology are required. The work was separated into four tasks:

1) Collect native and non-native unclassified speech communications from training courses as used with inter and intra-ship communications,
2) Produce an annotated database that might be used beyond the confines of the Task Group,
3) Assess effects on performance of recognisers and communication equipment,
4) Relate derived results to military applications.

The results of the study are presented in this report. Preliminary results were also presented and discussed at a satellite workshop of the International Speech Communications Association (ISCA) held in September 2001 at the Aalborg Conference Centre, Denmark under the responsibility of ISCA and IST-011/RTG-001.

Préface

Les communications, le commandement et contrôle, le renseignement et les systèmes d'entraînement font de plus en plus appel à des composants issus de la technologie vocale : il s'agit de codeurs vocaux, de systèmes C^2 à commande vocale, de la reconnaissance du locuteur et du langage, ainsi que de programmes automatisés d'entraînement. L'interopérabilité de ces systèmes ne se présente pas comme un simple problème de normalisation, car la voix de chaque utilisateur en particulier est une variable non maîtrisée, comme dans le cas de locuteurs non autochtones qui s'expriment dans une langue officielle de l'OTAN, en plus de leur propre langue. Dans le cas des opérations multinationales, ce problème peut entraîner des performances réduites, voire même l'échec d'une action. Il y a lieu de définir des méthodes et des spécifications d'évaluation normalisées, tant pour les produits du commerce (COTS), que pour le développement de nouvelles technologies. L'atelier a été organisé en 4 sessions, à savoir :

1) La collecte de communications vocales autochtones et non autochtones non classifiées à partir de stages de formation, comme dans le cas des communications internavires et intranavires,
2) La création d'une base de données annotée pouvant être exploitée par des personnes à l'extérieur du groupe de travail,
3) L'évaluation des effets des systèmes de reconnaissance de la parole et du matériel de communication sur les performances,
4) L'établissement d'un lien entre les résultats obtenus et des applications militaires.

Les résultats de l'étude sont inclus dans le rapport. Des résultats préliminaires ont également été présentés et discutés lors d'un atelier satellite de l'Association internationale de la communication verbale (ISCA), organisé en septembre 2001 au Centre international de conférences à Aalborg, au Danemark, sous la responsabilité de l'ISCA et de l'IST-011/RTG-001.

Foreword

Efficient speech communication is recognized as a critical and instrumental capability in many military applications such as command and control, aircraft and vehicle operations, military communication, translation, intelligence, and training. The former NATO research study group on speech processing (AC243(Panel 3)RSG10) conducted since its establishment in 1978 experiments and surveys focused on military applications of language processing. Guided by its mandate, the former RSG.10 initiated in the past the publication of overviews on potential applications of speech technology for military use and also organized several workshops and lecture series on military-relevant speech technology topics. Recently the group continued under the IST panel as AC232/IST/RTG-001.

In recent years, the speech R&D community has developed or enhanced many technologies which can now be integrated into a wide-range of military applications and systems:

- Speech coding algorithms are used in very low bit-rate military voice communication systems. These state-of-the-art coding systems increase the resistance against jamming;
- Speech input and output systems can be used in control and command environments to substantially reduce the workload of operators. In many situations operators have busy eyes and hands, and must use other media such as speech to control functions and receive feedback messages;
- Large vocabulary speech recognition and speech understanding systems are useful as training aid and to prepare for missions;
- Speech processing techniques are available to identify talkers, languages, and keywords and can be integrated into military intelligence systems;
- Automatic training systems combining automatic speech recognition and synthesis technologies can be utilized to train personnel with minimum or no instructor participation (e.g. Air traffic controllers).

This report is the result of a project on “Implications of Multilingual Interoperability of Speech Technology for Military Use” with contributions of all Task Group members, which represent nine NATO countries (Belgium, Canada, France, Germany, the Netherlands, Portugal, Turkey, United Kingdom, and the United States). Because speech technologies are constantly improving and adapting to new requirements, it is the intention of the Task Group to initiate projects on military applications of speech technology. Therefore the group appreciates any comment and feedback on this report.

Information Systems Technology Task Group 001
"Speech and Language Technology"

CHAIRMAN

Dr. Timothy Anderson
Air Force Research Laboratory
AFRL/HECA
2255 H Street
Wright Patterson AFB, OH 45433-7022
USA

SECRETARY

Mr. Carl Swail
Flight Research Laboratory
Building U-61
Montreal Road
Ottawa, Ontario K1A-0R6
CANADA

MEMBERS

Dr. Stéphane Pigeon
Koninklijke Militaire School
Leerstoel voor Telecommunicaties
Royal Military Academy
Renaissancelaan 30
B-1000 Brussels
BELGIUM

Mr. Ozgur Devrim Orman
TUBITAK-UEKAE, National Research Institute
of Electronics & Cryptology
P.K. 74, 41470 Gebze. Kocaeli
TURKEY

Dr. Edouard Geoffrois
CTA/GIP
16 bis avenue Prieur de la Côte d'Or
94114 Arcueil Cedex
FRANCE

Mr. Paul Collins
Room G007, A2 Building
DSTL Farnborough
Hampshire GU14 0LX
UNITED KINGDOM

Ms. Christine Bruckner
Bundessprachenamt – SM1 –
Horbeller Strasse 52
50354 Huerth
GERMANY

Mr. John J. Grieco
AFRL/IFEC
32 Brooks Rd.
Rome, NY 13441
USA

Dr. David A. van Leeuwen
TNO Human Factors
P.O. Box 23
3769 ZG Soesterberg
Kampweg 5
3769 DE Soesterberg
THE NETHERLANDS

Dr. Marc Zissman
Information Systems Technology Group
MIT Lincoln Laboratory
244 Wood Street
Lexington, MA 02420-9108
USA

Prof. Carlos Teixeira
INESC-ID, Lisboa
Spoken Language Systems Lab
L2F – Laboratório de Sistemas de Língua Falada
R. Alves Redol, 9
1000-029 Lisboa
PORTUGAL

PANEL EXECUTIVE

From Europe:
Lt.Col. A. GOUAY, FAF
RTA-OTAN
IST Executive
BP 25
F-92201 Neuilly sur Seine, Cedex
FRANCE

From the USA or CANADA:
RTA-NATO
Attention: IST Executive
PSC 116
APO AE 09777

Telephone: +33 (1) 5561 2280 / 82 - Telefax: +33 (1) 5561 2298 / 99

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. 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-RTG-001 recognized the need to perform research and studies on this topic to better understand, detect, and mitigate the effects of non-native speech production. 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. 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 environment.

1.3 WORK PROGRAM

In the past, TG-001 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.

The project discussed in this report is focused on the interoperability issues of speech (communication) technology as applied to a wide range of military coalition operations. As in previous RTG-001 projects,

different scientific aspects of this topic began with the organization of a scientific workshop in cooperation with the European Speech Communication Association (ESCA)[1]. This international workshop on "Multi-lingual Interoperability in Speech Technology" (MIST) was held in Leusden, The Netherlands in September 1999. This very successful workshop demonstrated the necessity for a coordinated international effort to support NATO interests in this area. Speech data was then collected in four countries and a database developed to further study and foster research on multilingual, non-native speech. This data set is very representative of military type communication in a ship-to-ship scenario, and was used for evaluation and modification of Automatic Speaker Recognition and Word Spotting. This database also focused research on non-native speech issues, which lead to a special session at the international speech and language conference, Eurospeech 2001.

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 multilingual and non-native databases, which were considered for the workshop in Leusden. Also included in this chapter is a detailed description of the database developed by this TG in the ship-to-ship communication scenario. An overall description of each database and its content, such as task, amount of data, language, non-native type, and characteristics is included.

Chapter 3:

This chapter focuses on the problems of detection, classification, and assessment of non-native, accented speech. Here, previous work and analysis of various speech production issues are considered for the speech data discussed in Chapter 2. This chapter briefly reviews work in the field, and presents representative findings obtained on accent and non-native speech research.

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 multilingual and non-native speech has on military speech technology and its application is presented.

[1] Now ISCA, previously ESCA.

Chapter 2 – NON-NATIVE SPEECH DATABASES

2.1 INTRODUCTION

Over the years, pronunciation variation due to non-native speech has been the interest of many phoneticians. A most remarkable number of researchers have studied the production and perception of the famous 'l-r confusion' for Chinese and Japanese natives. It is not an exaggeration to claim[1] that more than 50% of the research papers on non-native speech deal with this interesting subject of /l/ and /r/.

Long after phoneticians were drawn by the subject, non-native speech slowly started to become an issue in speech technology. Since speech databases are an invaluable resource for researchers in the field of speech technology, soon the first non-native speech databases were recorded. One of the problems with non-native speech is that there are potentially so many different kinds: if N is the number of languages in the world[2], the number of non-native accents is close to N^2. This number only considers speech production. If the perception of speech is taken into account too, the number of possible language combinations scales with N^3.

It is clear, that in a field of research that has only recently started and with so many possible language combinations, the coverage by speech databases is rather limited. Yet, there are a number of interesting recordings available.

2.2 TERMINOLOGY

2.2.1 Language Specification

In non-native speech communication there are at least three languages of importance:

1) The native language of the speaker (S),
2) The language that is spoken, or the target language (T),
3) The native language of the listener (L).

In literature one often finds the symbols L1 for native language and L2 for the target language, but this is not always used consistently and, especially in listening experiments, this terminology might lead to confusion. It is always a good exercise to really understand the configuration of languages in a description of non-native language experiments in a research paper. Van Wijngaarden [van Wijngaarden 2001] proposes the notation for communication between two persons:

$$S > (T) > L$$

Meaning that a speaker who's native language is S speaks in language T to a listener who's native language is L. For the purpose of speech databases for speech technology, it usually suffices to specify S and T only, because L is either not considered or the technology takes the role of the listener, as is the case for speech, speaker, language and accent recognition.

[1] Based on JASA abstracts.

[2] N is estimated to be about 6000; the bible has been translated into 2000 languages alone.

2.2.2 Language Proficiency

One of the most important parameters in non-native speech communication is the language proficiency of the speaker and listener. There is a NATO standard, STANAG 6001 that classifies the language proficiency of people into five levels:

1) Elementary
2) Fair (Limited working)
3) Good (Minimum professional)
4) Very good (Full professional)
5) Excellent (Native/Bilingual)

These levels define both speaking and listening proficiency.

For speech databases, it is very important that the language proficiency of the individual speakers be known, because the quality and character of the speech is very dependent on this. None of the databases discussed in this chapter has classified the speakers according to STANAG 6001 levels, but there is generally information about the speaker's non-native language acquisition. Important information includes:

Native language: The mother tongue of the speaker,

Age of acquisition: The speaker's age when the non-native language was learned,

Experience: The number of years that the speaker has been regularly using the language.

Generally, an age of acquisition of over 6 years is considered to always lead to a noticeable non-native accent. The higher the age of first learning, the stronger the non-native accent will be. Of course, these parameters are not the only important factors in language proficiency; there are also matters such as willingness to learn another language, level of exposure, talent, etc. There are numerous cases where 'expatriates' live in a foreign country for decades without being exposed to the local native language at all.

Databases differ in the information that is specified about the speakers, even though there has been an effort to guide the database meta-data collection, such as through the EAGLES handbook [Howell 1997].

2.3 A SELECTION OF AVAILABLE NON-NATIVE SPEECH DATABASES

In this section an overview is given of some non-native databases that are available. There are many more recordings made of non-native utterances under a wide range of conditions, for all the different studies made in literature. Here we only list the databases that are publicly available and have some relevance to speech technology research.

2.3.1 Translanguage English Database

In a daring plan Joseph Mariani, then at LIMSI-CNRS, proposed to record the speeches of speech researchers made at Eurospeech '93 in Berlin and distribute the data amongst the speech laboratories, with the aim of being able to automatically transcribe speeches in 1995. This plan led to the recording of 224 speeches by mostly non-native speakers of English. The data was published in 1995, through the efforts of the Institute of Phonetics at the University of Munich and LIMSI-CNRS. The project was partially financed by Eurococosda, Relator and ELSNET.

The European Language Resources Distribution Agency (ELDA) now distributes 188 speeches of the data collection, along with speaker information and text material related to the conference (proceedings papers and oral transcriptions).[3] In 2002 the Linguistic Data Consortium (LDC) published detailed transcripts of 39 of the speeches, mostly of non-native speakers. The database is known as the 'Translanguage English Database' but is often referred to as the 'terrible English database.'

About 28% of the speeches were made by native speakers of English (American and British), the other speakers native language include: German, French, Italian, Spanish, Japanese, Dutch, Swedish, Polish, Greek, Danish, Chinese, Atungsiri, Serbo-Croatian, Norwegian, Korean, Catalan, Bulgarian and Arabic. The database contains 15 minutes of speech for each presentation. The speech can be considered spontaneous (as opposed to read) and under slight stress. The speaker information was obtained by a questionnaire containing 14 questions related to the speaker's language skills.

2.3.2 Australian National Database of Spoken Language

Since 1992 the Australian Research Council has supported data collection of spoken language in Australia. This has led to the recording of the Australian National Database of Spoken Language (ANDOSL).[4] The database consists of recordings of both native and non-native (Australian) English. The database includes read speech, isolated words and spontaneous speech.

There are three different groups of speakers, 108 native Australian English speakers from three accent groups, 48 Lebanese Arabic speakers and 48 South Vietnamese speakers. All speakers were older than 17 years and had spent at least 4 years in Australia [Kumpf 1997]. The native speakers read 200 phonetically rich sentences and the non-native speakers read 50 different phonetically rich sentences. Spontaneous speech was obtained by giving pairs of speakers a MAP task.

2.3.3 Strange Corpus

The Bavarian Archive for Speech Signals[5], an institution hosted by the University of Munich, has recorded two non-native databases called Strange Corpus (SC). The first, SC1, contains read speech by 16 native and 72 non-native German speakers. The second corpus, SC10, also contains dialogue and monologue spontaneous speech recordings by 3 native and 67 non-native German speakers.

ELDA distributes these databases under references S0029 and S0114. SC1 contains read utterances of the story "The north wind and the sun" in German, a short story that has been translated into many languages and has a tradition for being used as material for linguistic studies. SC10 adds numbers, phonetically balanced sentences, and a dialog with a native German and re-telling of a story as speech material. The speaker's native tongue include: Arabic, English, Finish, French, Greek, Italian, Japanese, Dutch, Polish, Portuguese, Russian, Spanish, Swedish, Turkish, Hungarian, Hebrew, Persian, Rumanian, Bulgarian, Hindi, Nepalese, Vietnam, Korean and African languages.

[3] http://www.elda.fr/cata/speech/S0031.html

[4] http://andosl.anu.edu.au/andosl/ANDOSLhome.html

[5] http://www.phonetik.uni-muenchen.de/Bas/

2.3.4 Interactive Spoken Language Identification

In the European ISLE project a database was collected to aid in developing pronunciation-training tools for second language learning [Menzel 2000]. This database contains 23 German and 23 Italian intermediate learners of the target language: English. The database is annotated on various detailed levels: word, phone and stress. The phone and stress levels can be used to study pronunciation errors.

The database is distributed through ELDA (reference S0083) and consists of read speech of varying complexity. The read material was designed to concentrate on specific linguistic issues, such as vocabulary coverage, problem phones, weak forms and stress. The speech material was then automatically annotated at the word, phone and stress level. After this process the phone and stress annotation was manually checked and pronunciation differences were corrected in the annotation. Two teachers of English as a foreign language estimated proficiency levels of the speakers.

2.3.5 Multilingual Interoperability in Speech Technology

In 1999, an ESCA/NATO-RTO Tutorial and Research Workshop was organized on the subject of Multilingual Interoperability in Speech Technology (MIST) [van Leeuwen 1999]. In order to stimulate the discussion at the workshop, a non-native speech database was distributed amongst speech researchers. This was carried out before the workshop took place, allowing attendants to carry out different experiments on the same data. The data was collected as part of a read speech Dutch database in 1996 at TNO Human Factors in the Netherlands.

All speakers of the MIST non-native database are native Dutch and they read sentences in three languages: English, French and German [van Leeuwen 1999]. Most of the speakers were recruited from the TNO institute, in which 60% of the employees have an academic background and 20% a higher technical education. In total 74 speakers spoke in at least one of the three languages: 71 in English, 66 in German and 60 in French. In each language recorded, speakers uttered 10 sentences: 5 that were the same across all speakers and 5 that were unique to the speaker. The sentences were selected from newspapers, *Wall Street Journal*, *Frankfurter Rundschau* and *Le Monde*, the same sources of text for the development and test utterances in the SQALE project [Young 1997]. As a reference, for each speaker there were also 10 Dutch (native) utterances, again 5 the same for all speakers and 5 unique to the speaker. Transcriptions are available in all languages, but only the Dutch have been manually checked due to personnel limitations.

2.3.6 NATO Native and Non-Native Speech Corpus

The NATO RTO/IST-011/RTG-001 research task group, the originator of this report, sponsored research into non-native speech for military applications. The group recorded a non-native speech corpus in four countries, targeting NATO naval procedures in English [Benarousse 2001]. The native language of the speakers is Canadian English, German, Dutch and British English. In the Canadian part of the database there are both native English and native French speakers represented.

The naval communication recordings amount to 1.6–3.0 hours of speech data, depending on the country of origin. Apart from these recordings each speaker read the text of “The north wind and the sun” in English and his native language. These short stories amount to about 1 hour for non-native and 1 hour for native English speech. Transcription of all material has been performed in the counties where the recording took place, by local (possibly non-native) transcribers. After that, the transcription was normalized across all four contributing countries in order to have uniform call signs and other idioms.

2.4 REFERENCES

Benarousse, L., Geoffrois, E., Grieco, J., Series, R., Steeneken, H., Stumpf, H., Swail, C. and Thiel, D. (2001). "The NATO Native and Non-Native (N4) Speech Corpus", In Proc of the Workshop on Multilingual Speech and Language Processing, Aalborg, Denmark. (Available as NATO Publication RTO-MP-066, April 2003).

Howell, P. (1997). Handbook of Standards and Resources for Spoken Language Systems, Chapter 9: "Assessment Methodologies and Experimental Design", pp. 344-380, Mouton de Gruyter.

Kumpf, K. and King, R. (1997). "Foreign Speaker Accent Classification Using Phone Phoneme-Dependent Accent Discrimination Models and Comparisons with Human Perception Benchmark", In Eurospeech, pp. 2323-2326, 1997.

Menzel, W., Atwell, E. Bonaventura, P., Herron, D., Howarth, P., Morton, R. and Souter, C. (2000). "The ISLE Corpus of Non-Native Spoken English", *LREC*, pp. 957-963, Vol. 2, Athens, Greece.

van Leeuwen, D. and Orr, R. (1999). "Speech Recognition of Non-Native Speech Using Native Acoustic Models", In Proc of the MIST Workshop, pp. 23-28. (Available as NATO Publication RTO-MP-28, August 2000).

van Leeuwen D., Wijngaarden, S. and Steeneken, H., editors. (1999). Multi-lingual Interoperability in Speech Technology, TNO Human Factors, ISBN 90-76702-01-02.

Wijngaarden, S., Steeneken, H. and Houtgast, T. (2001). "Methods and Models for Quantitative Assessment of Speech Intelligibility in Cross-Language Communication", In Proc. of the Workshop on Multilingual Speech and Language Processing, Aalborg, Denmark. (Available as NATO Publication RTO-MP-066, April 2003).

Young, S., Adda-Decker, M., Aubert, X., Dugast, C., Gauvin, J., Kershaw, D., Lamel, L., van Leeuwen, D., Pye, D., Robinson, A., Steeneken, H. and Woodland, P. (1997). "Multi-lingual Large Vocabulary Speech Recognition: The European SQALE Project", Computer, Speech and Language, Vol. 11, pp. 73-89.

Chapter 3 – LANGUAGE, DIALECT AND ACCENT RECOGNITION

3.1 INTRODUCTION

For an automatic speech system that needs to be able to work worldwide, a first task would be to identify the language that is being spoken (*T*).

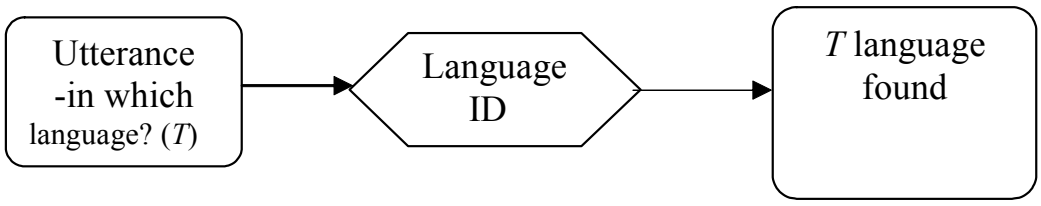

Figure 3.1: Block Diagram of Language Identification System.

Following the first identification phase, it should be determined whether the speaker is a native speaker of that language or not. If the speaker is native, it is important to use specific speech models suitable for his native accent (here referred as dialect). A dialect recognizing system can then be used for selecting the closest models. If the speaker is a non-native, an accent identification system can be triggered in order to identify the speaker's mother/first language (*S*). For an automatic speech recognition (*ASR*) system this can either be used to select the appropriate models for language *S* or the closest models for the specific accent of the language *T*.

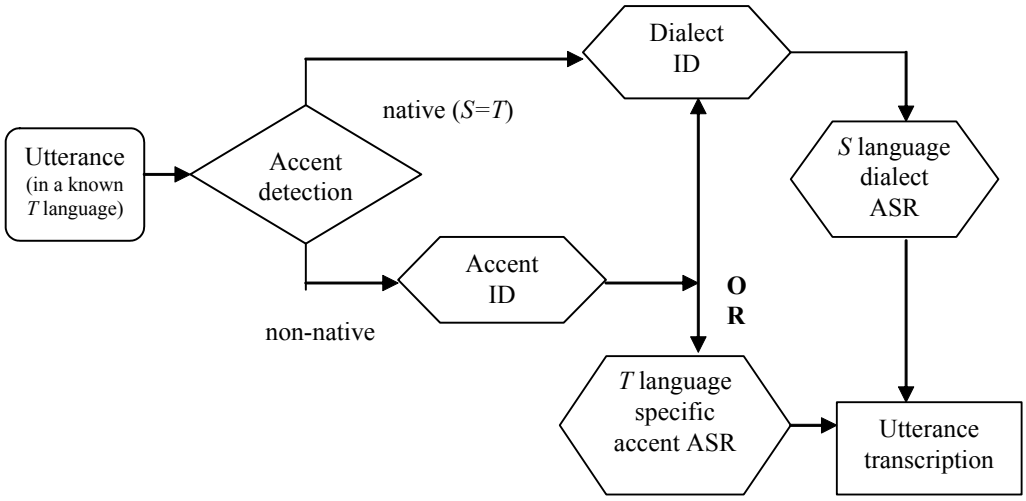

Figure 3.2: Block Diagram of Dialect/Accent Identification System.

Obviously, some of these identification stages are often merged or even omitted depending on the particular implementation. A language, dialect or accent recognition system can select an appropriate speech recognizer or a human interpreter. Suitable applications are in surveillance for medical assistance, law enforcement and military communications. Also, some criminal departments in the world have accent experts that could be assisted by these systems.

3.2 LANGUAGE RECOGNITION

A language identification system is used to identify the language of the speech utterance. Automatic language identification systems generally work by exploiting the fact that languages have different phoneme inventories, phoneme frequencies, and phoneme sequences. These features can be obtained, although imperfectly, using the same spectral analysis techniques developed for speech recognition and speaker recognition. The use of higher-level features such as the prosodic and the use of expert knowledge about the target languages should also contribute to the language identification task, but to date, the best results have been obtained with systems that rely mainly on statistical analysis of spectral features.

There are practical issues that must be considered in putting together a system for a real application. Performance is of course a primary concern. This must be weighed against issues such as system complexity, the difficulty in training the system, and the ease with which new languages can be added. For example, the type and amount of data required for training could be very important. Some systems can be trained given only examples of conversations in a given language. Others require detailed phonetically marked transcriptions in each language. The relative importance of these issues will differ depending on the constraints of the particular application. A survey article by Muthusamy (1994) describes many of the techniques.

Some of the factors that make the language identification problem easier or harder are the following:

- The quality of the speech and the channel over which it is received.
- The number of possible languages from which the system must choose.
- The length of the utterance on which the decision must be made.
- The amount of training data that is available for making the models for each language. Both total duration of the training speech and the number of different training speakers are important factors.
- The availability of transcripts of the training speech, text samples from the language, and phonetic dictionaries for the language to assist in the creation of the models for the language.

Language identification continues to be an area of research. A series of tests was coordinated by NIST in the mid 1990s comparing the performance of language identification systems on a standard set of test data. In 1996 the evaluation focused on long-distance, conversational speech. Forty, 30-minute conversations per language were available for training in each of 12 languages. 20 conversations per language were used for testing. Test utterance size varied from 3 seconds to 30 seconds. The best systems exhibited 25% closed-set error rates on the 12-alternative, forced-choice problem on 30-second utterances. Average error rates of 5% were measured for the language-pair (language A vs. language B) [Zissman 1997] experiments.

Current areas of research include reducing the training requirements, reducing the size of test utterances, and reducing the computational complexity. Recent publications reported the use of high-order GMMs with shifted-delta-cepstra features. These systems show a performance comparable to the more complicated parallel phone systems, while having only a fraction of the computational complexity [Torres-Carrasquillo 2002].

3.3 ACCENT AND DIALECT RECOGNITION

Preliminary results have indicated a 15% drop in recognition performance for non-native speakers when compared to native speakers using a native speech recognition system [Teixeira 1992]. Selecting acoustic and phonotactic models, adapted to a specific accent, can improve the recognition performance significantly.

However, this means the development of as many accent specific recognizers as accents identified for a particular language. Use of an accent specific system gave improvements of up to 60% when compared to the use of a conventional recognizer for native speakers. Although similar improvements can be found by adding the representative accents in the acoustical recognizer training material, the modularity of the proposed approach provides important advantages. Specific accent recognizers, which were previously developed and tuned according to the best suitable methodologies, can simply be integrated together in order to cover a wider range of speaker's accents [Teixeira 1993, 1996, 1997, 1998]. Moving further away from the speaker independent recognition paradigm, better results may be found with more general approaches such as speaker adaptation [Tomokiyo 2001].

As has been proposed for language identification, an accent identification system may also consist of a set of parallel speech recognizers. The one with the most plausible recognition output identifies, at the same time, the most plausible speaker accent. The following table represents the confusion matrix obtained from one of this type of system trained and tested with European male speakers for 200 English words. The overall accent identification rate was 85.3% providing an overall word recognition rate of 86.8%. However the recognition rate for the native speakers (UK) was 98.9%. Some of the figures in the confusion matrix can raise interesting explanations. Consider, for example, the first row of figures. Danish speakers provided the bigger number of non-native utterances classified as native (7.6%). They were the most difficult to identify as Danish, mainly because 19.3% of their words got a higher score from the Spanish recognizer. It is interesting to note that Danish people are generally fluent with English. On the other hand, these speakers were from Jutland having a strong tendency to transform the alveolar fricatives /s/ into palatal fricatives (/sh/) which is also common among Spanish speakers.

Table 3.1: Confusion Matrix (%) for an Accent Identification System (Teixeira, 1998)

Accent	Danish	German	English	Spanish	Italian	Portuguese
Danish	63.6	1.0	7.6	19.3	2.0	6.6
German	0.0	99.6	0.0	0.0	0.0	0.4
English	2.2	0.6	88.1	5.0	0.3	3.9
Spanish	3.7	1.7	4.9	83.7	0.0	6.0
Italian	0.0	3.7	1.2	0.5	91.9	2.6
Portuguese	2.0	2.7	7.4	3.7	2.5	81.8

The accents to be identified should be as well defined as possible. This will guide the collection of a representative speech corpus for training and testing and for choosing suitable features and classification methods [Benarousse 2001]. A first basic distinction should be made between non-native speech and dialects [Chengalvarayan 2001], or language varieties. Dialect differences are often significant and speakers do not generally attempt to conform to a standard variant. Non-native speakers, on the other hand, show different degrees of reading and pronunciation competence [Mengel 1993]. Their knowledge of the

grapheme-to-phoneme conventions of the foreign language may vary a lot, as well as their ability to pronounce sounds, which are not part of their native sound inventory [Trancoso 1999]. These differences justify the distinction of dialect or variant identification from accent identification that is here strictly addressed for non-native speakers [Brousseau 1992, Zissman 1996].

Non-native populations of speakers can be categorized in two different scenarios. One scenario considers immigrants or refugees. These speakers are under a long-term influence to acquire vocabulary and fluency according to the language variety used by the local population. A second scenario considers occasional travelers, such as businessmen, tourists and military personnel. These speakers usually have had a limited exposure to a standard variety of a foreign language (usually English) that will be needed in relatively few, but sometimes-crucial circumstances.

Speaker classification according to their first/mother language can also be used for selecting a recognizer. This was also one application area for language identification, when the speaker actually uses his first language. However, for language identification, there is a vast amount of knowledge available about each specific language (phoneme inventories, grammar, etc.), which cannot be used in a straightforward manner for accent detection. This can be considered among the reasons why accent identification is generally considered a more difficult task than language identification.

Instead of finding some qualities in the non-native speech such as the effects of the mother language, one might be interested in classifying it according to some kind of measure of performance in relation to a standard pronunciation in the second language. Giving feedback on the degree of nativeness of a student's speech is an important aspect of language learning. Skilled native speakers can easily discriminate at least five different ordered scores for classifying a student's utterance – average correlation between raters was once measured as 0.8 [Franco 2000a]. In computer-aided language learning, this task has been addressed by many studies focusing on the segmental assessment of the speech signal [Neumeyer 1996, Franco 1998, 2000ab]. Recently, several studies have used suprasegmental speech information for computer-assisted foreign language learning [Delmonte 2000]. Some of these systems were able to obtain an average correlation between human and machine scores similar to the one obtained between different human scorers [Teixeira 2000, 2001].

Other related research issues can bring new approaches to this area [Angkititrakul 2002]. Knowing the effects of noise on accented speech [Weil 2002] is also important, namely in situations in which critical information is being communicated by individuals with different language backgrounds, such as air traffic control and military applications [Anderson 1998]. In a world where globalization is an inevitable reality, accent identification will become a more important aspect of research in speech processing.

3.4 REFERENCES

Anderson, T. (1998). "Applications of Speech Based Control." Paper presented at the RTO Lecture Series on Alternative Control Technologies: Human Factor Issues. Ohio, USA. (Published in RTO-EN-3, October 1998).

Angkititrakul, P. and Hansen, J. (2002). "Stochastic Trajectory Model Analysis for Accent Classification", In Proc. of 7th Int'l Conf. on Spoken Language Processing (ICSLP), Denver, USA.

Arslan, L. (1996a). "Automatic Foreign Accent Classification in American English", PhD thesis, Duke University, Durham, North Caroline, USA.

Arslan, L. and Hansen, J. (1996b). "Language Accent Classification in American English", Speech Communication, 18(4): 353-367.

Benarousse, L., Geoffrois, E., Grieco, J., Series, R., Steeneken, H., Stumpf, H., Swail, C. and Thiel, D. (2001). "The NATO Native and Non-Native (N4) Speech Corpus", In Proc of the Workshop on Multilingual Speech and Language Processing, Aalborg, Denmark. (Available as NATO Publication RTO-MP-066, April 2003).

Berkling, K., Zissman, M., Vonwiller, J. and Cleirigh, C. (1998). "Improving Accent Identification through Knowledge of English Syllable Structure", In Proc. of 5th Int'l Conf. on Spoken Language Processing (ICSLP), Sydney, Australia.

Blackburn, C., Vonwiller, J. and King, R. (1993). "Automatic Accent Classification Using Artificial Neural Networks", In Proc. of the European Conf. on Speech Comm. and Tech., Volume 2, pp. 1241-1244, Berlin, Germany.

Brousseau, J. and Fox, S. (1992). "Dialect-Dependent Speech Recognizers for Canadian and European French", In Proc. of the 2nd Int'l Conf. on Spoken Language Processing (ICSLP), Volume 2, pp. 1003-1006, Banff, Canada.

Chengalvarayan, R. (2001). "HMM-Based English Speech Recognizer for American, Australian and British Accents", In Proc. of the Workshop on Multilingual Speech and Language Processing, Aalborg, Denmark. (Available as NATO Publication RTO-MP-066, April 2003).

Delmonte, R. (2000). "SLIM Prosodic Automatic Tools for Self-Learning Instruction", Speech Communication, Vol. 30, pp. 145-166.

Franco, H., Abrash, V., Precoda, K., Bratt, H., Rao, R., Butzberger, J., Rossier, R. and Cesari, F. (2000a). "The SRI EduSpeak System: Recognition and Pronunciation Scoring for Language Learning", Proc. of Integrating Speech Technology in Language Learning.

Franco, H., Neumeyer, L. Digalakis, V. and Ronen, O. (2000b). "Combination of Machine Scores for Automatic Grading of Pronunciation Quality", Speech Communication, Vol. 30, pp. 121-130.

Franco, H. and Neumeyer, L. (1998). "Calibration of Machine Scores for Pronunciation Grading", In Proc. of 5th Int'l Conf. on Spoken Language Processing (ICSLP), Sydney, Australia.

Fung, P. and Liu, K. (1999). "Fast Accent Identification and Accented Speech Recognition", In Proc. Int'l Conf. on Acoustic Speech and Signal Processing (ICASSP), Phoenix, Arizona, USA.

Hansen, J. and Arslan, L. (1995). "Foreign Accent Classification Using Source Generator Based Prosodic Features", In Proc. of the Int'l Conf. on Acoustic Speech and Signal Processing (ICASSP), Volume 1, pp. 836-839, Detroit, USA.

Humphries, J. and Woodland, P. (1998). "The Use of Accent-Specific Pronunciation Dictionaries in Acoustic Model Training", In Proc. Int'l Conf. on Acoustic Speech and Signal Processing (ICASSP), Seattle, USA.

Kumpf, K. and King, R. (1997). "Foreign Speaker Accent Classification Using Phoneme-Dependent Accent Discrimination Models and Comparisons with Human Perception Benchmarks", In Proc. of the European Conf. on Speech Comm. and Tech., pp. 2323-2326, Rhodes, Greece.

Mengel, A. (1993). "Transcribing Names – A Multiple-Choice Task: Mistakes, Pitfalls and Escape Routes", In Proc. 1st ONOMASTICA Research Colloquium, pp. 5-9, London, UK.

Muthusamy, Y., Barnard, E. and Cole, R. (1994). "Reviewing Automatic Language Identification", IEEE Signal Magazine, October, pp. 33-41.

Neumeyer, L., Franco, H., Digalakis, V. and Weintraub, M. (1996). "Automatic Text-Independent Pronunciation Scoring of Foreign Language Student Speech", In Proc. of the 4th Int'l Conf. on Spoken Language Processing (ICSLP), pp. 1457-1460.

Teixeira, C., Franco, H., Shriberg, E., Precoda, K. and Sömnez, K. (2001). "Evaluation of Speaker's Degree of Nativeness Using Text-Independent Prosodic Features", In Proc. of the Workshop on Multilingual Speech and Language Processing, Aalborg, Denmark. (Available as NATO Publication RTO-MP-066, April 2003).

Teixeira, C., Franco, H., Shriberg, E., Precoda, K. and Sömnez, K. (2000). "Prosodic Features for Automatic Text-Independent Evaluation of Degree of Nativeness for Language Learners", In Proc. 6th Int'l Conf. on Spoken Language Processing (ICSLP), Beijing, China.

Teixeira, C. (1998). "Reconhecimento de Fala de Oradores Estrangeiros", PhD thesis, DEEC (IST-UTL), Lisboa, Portugal.

Teixeira, C., Trancoso, I. and Serralheiro, A. (1997). "Recognition of Non-Native Accents", In Proc. of the European Conf. on Speech Comm. and Tech., Rhode, Greece.

Teixeira, C., Trancoso, I. and Serralheiro, A. (1996). "Accent Identification", In 4th Int'l Conf. on Spoken Language Processing (ICSLP), Philadelphia, USA.

Teixeira, C. and Trancoso, I. (1993). "Continuous and Semi-Continuous HMM for Recognizing Non-Native Pronunciations", In Proc. of IEEE Workshop on Automatic Speech Recognition, Snowbird, Utah, USA.

Teixeira, C. and Trancoso, I. (1992). "Word Rejection Using Multiple Sink Models", In Proc. of 2nd Int'l Conf. on Spoken Language Processing (ICSLP), 2:1443-1446, Banff, Canada.

Tomokiyo, L. and Waibel, A. (2001). "Adaptation Methods for Non-Native Speech", In Proc. of the Workshop on Multilingual Speech and Language Processing, Aalborg, Denmark. (Available as NATO Publication RTO-MP-066, April 2003).

Torres-Carrasquillo, P., Singer, E., Kohler M., Greene, R., Reynolds D. and Deller, J. (2002). "Approaches to Language Identification Using Gaussian Mixture Models and Shifted Delta Cepstral Features", In Proc. of 7th Int'l Conf. on Spoken Language Processing (ICSLP), Denver, USA.

Trancoso, I., Viana, C., Mascarenhas, I. and Teixeira, C. (1999). "On Deriving Rules for Nativised Pronunciation in Navigation Queries", In Proc. of the European Conf. on Speech Comm. and Tech., Budapest, Hungary.

Weil, S. (2002). "Comparing Intelligibility of Several Non-Native Accent Classes in Noise", In Proc. of 7th Int'l Conf. on Spoken Language Processing (ICSLP), Denver, USA.

Wong, E. and Sridharan, S. (2001). "Fusion of Output Scores on Language Identification System", In Proc. of the Workshop on Multilingual Speech and Language Processing, Aalborg, Denmark. (Available as NATO Publication RTO-MP-066, April 2003).

Zissman, M., Gleason, T., Rekart, D. and Losiewicz, B. (1996). "Automatic Dialect Identification of Extemporaneous Conversational Latin American Spanish Speech", In Proc. Int'l Conf. on Acoustic Speech and Signal Processing (ICASSP), Volume 2, pp. 777-780, Atlanta, USA.

Zissman, M. (1997). "Predicting, Diagnosing and Improving Language Identification Performance", In Proc. of the European Conf. on Speech Communication and Technology, Rhodes, Greece.

Chapter 4 – EXPERIMENTAL RESULTS

4.1 INTRODUCTION

The presence of multilingual and non-native speech complicates the task faced by those who wish to apply automatic speech processing technology to military applications. Most automatic speech processing algorithms (e.g. speech recognition, speaker recognition, language recognition) operate in two phases. During the *training phase*, statistical models are created from labeled training speech utterances. During the *recognition phase*, the statistical models built during training are used to hypothesize the words (or speaker, or language, etc.) of a new test utterance. Mismatched situations in which the training speech and test speech are spoken in different languages, or in which the training speech is spoken by native speakers but the test speech is spoken by non-native speakers, typically cause a degradation in performance of automatic speech processing systems vs. the performance obtained when only single-language, native speech is processed. Degradation can also be caused by the increased rate of disfluencies and other similar errors made by non-native speakers.

As part of the NATO Multilingual Speech and Language Processing Project, two speech corpora were collected, labeled and distributed that allowed researchers to measure the effectiveness of speech processing systems on multilingual and non-native speech. The MIST corpus, collected at TNO Human Factors in the Netherlands, contains speech spoken in English, French and German – all spoken by native speakers of Dutch (See Section 2.3.5). The N4 corpus, collected in Canada, Germany, the Netherlands and the UK, contains primarily English speech spoken by both native and non-native speakers (See Section 2.3.6).

The rest of this chapter of the report describes experiments performed on both the MIST and N4 corpora. These experiments showed the impact of multilingual and non-native speech on speech recognition, speaker recognition and language recognition performance. In some cases, the effect on recognition accuracy was modest. In other cases, it was moderate or severe.

4.2 THE IMPACT OF MULTILINGUAL AND NON-NATIVE SPEECH

Multilingual and non-native speech impacts the performance of speech processing systems in a variety of ways.

4.2.1 Impact on Speech Recognition

Many hundreds of hours of transcribed training speech can be required to train acoustic and language models for speech recognition. When it is likely that non-native speech will be encountered during recognition, and when little non-native speech is available for testing, a dilemma is faced: is it better to train on a small amount of non-native speech to avoid a training/testing mismatch but incur the penalty of poorly trained models, or is it better to use large amounts native speech yielding well-trained, but mismatched, models. Depending on the circumstances, it may also be possible to adapt the well-trained native models to the non-native speech, to use acoustic models from one language to perform speech recognition in another, or to use multilingual acoustic models.

4.2.2 Impact on Speaker Recognition

Most conventional speaker recognition systems hypothesize the speaker of an utterance through extraction of features from the speech signal that is related to the speaker's vocal tract shape. To the extent that many languages share a common set of sounds and to the extent that speakers of one language have vocal tracts that are generally similar to speakers of another language, one might predict *a priori* that the complexities of multilingual and non-native speech would have a less severe impact on speaker recognition vs. speech recognition. But other factors such as speaking rate, phone frequency of occurrence, hesitations, etc. could cause multilingual and non-native speaker recognition performance to degrade relative to performance on single-language, native speech.

4.2.3 Impact on Language Recognition

Language identification systems use both acoustic and phonetic measurements to hypothesize the language of a speech utterance. Just as in the case of speech recognition, one would expect non-native speech to degrade performance vs. native speech because of acoustic and language-model mismatches.

4.3 EXPERIMENTS ON THE MIST CORPUS

The MIST corpus contains read sentences spoken in English, French and German by native speakers of English [van Leeuwen 1999]. The prompts for the foreign language sentences were taken from various newspaper texts. Of the 74 subjects recorded, 71 spoke English, 66 spoke German and 60 spoke French. For each subject for each language, five sentences were spoken that were the same for all speakers of that language, and five sentences were spoken that were unique for each speaker of that language. As a reference, ten Dutch sentences (five common, five unique) per subject were also made available.

4.3.1 Speech Recognition Experiments on the MIST Corpus

Some initial experiments performed at TNO in the Netherlands addressed some of these issues [van Leeuwen 1999]. The purpose of these experiments was to measure speech recognition performance of non-native speakers (Dutch speakers speaking English) using a variety of different modeling approaches. The baseline systems were trained using native English acoustic models and pronunciation dictionaries. Both US English and UK English models and dictionaries were tested. Contrast systems used acoustic models produced from Dutch speech with US or UK English pronunciation dictionaries. The study concluded that best performance was obtained when the acoustic models were trained using speech from native UK speakers rather than either Dutch speakers or US. Additionally, the error rate for non-natives speaking English was twice as high as for native English speakers.

4.3.2 Speaker Identification Experiments on the MIST Corpus

Experiments performed at Faculté Polytechnique de Mons in the Netherlands on the MIST corpus investigated two facets of speaker recognition: cross-language speaker ID and same-language non-native speaker ID [Durou 1999]. In the cross-language experiments, speaker models were trained on speech spoken in Dutch and were tested on speech spoken in Dutch and the other three languages. Generally, speaker ID error rate in the cross-language condition (train in Dutch, test in non-Dutch) was two to six times higher vs. the within-language condition (train in Dutch, test in Dutch). In the same-language non-native experiments, speaker recognition accuracy was computed for non-native speakers of English (train in English, test in English), non native speakers of French (train in French, test in French) and non-native speakers of German

(train in German, test in German). These results were compared to a native experiment (train in Dutch, test in Dutch). Generally, non-native speaker recognition performance in English and German was approximately the same as in the Dutch native condition, while the non-native French error rate was approximately three times higher.

4.3.3 Language Identification Experiments on the MIST Corpus

Experiments measuring the performance of language identification systems on non-native speech were carried out in France at DGA and LIMSI [Wanneroy 1999]. Models were trained on a large corpus of data containing phone calls spoken in English, French and German. Three-way, closed-set, forced-choice language ID experiments were conducted used test data from both the MIST corpus (non-native speakers) and the SQALE corpus (native speakers) [Young 1997]. Aside from the nativeness of the speakers, the MIST and SQALE corpora are very similar. Language ID error rates on the MIST corpus were on average 2.8 times higher vs. error rates on the SQALE corpus. Adaptation of the non-native acoustic models resulted in a modest error rate reduction.

4.4 EXPERIMENTS ON THE N4 CORPUS

The N4 corpus contains speech collected at naval training schools within several NATO countries. The speech utterances comprising the corpus are primarily short, tactical transmissions spoken in English and typical of NATO naval communications. Speech collected in the UK and Canada was primarily native, while speech collected in Germany and the Netherlands was non-native. A typical transmission is:

```
papa alfa zulu juliett this is papa alfa sierra zulu
Reporting into net over
```

Speech from 115 total speakers was collected (Canada: 22, Germany: 51, Netherlands: 31, UK: 11). Read speech from each subject was also collected in English and in his native language.

4.4.1 Speech Recognition Experiments on the N4 Corpus

Experiments measuring speech recognition accuracy and call sign ID accuracy on the N4 corpus were run in the US at the Human Effectiveness Directorate, Air Force Research Laboratory [Williamson 2002]. A commercial speech recognition system trained on US English speech was run over the Canada, Netherlands and UK segments of the corpus. Word error rates of 29.7%, 24.6% and 22.1% were obtained for Canada, Netherlands and UK, respectively. Call-sign error rates of 65.9%, 79.1% and 47.8% were measured, respectively. The relatively low word error rate obtained on the Netherlands segment and the relatively high error rate on the Canada segment were surprises. Aside from non-nativeness, other factors impacting error rates included frequency of disfluency and call-sign complexity/variability.

4.4.2 AFRL Speech and Audio Processing Group

The speech and Audio Processing Laboratory of the Air Force Research Laboratory evaluated speech recognition on the N4 corpus [Lawson 2003]. Accuracy within each accent group was first benchmarked using 38 ‘round-robin’ style experiments, where a certain set of speakers is removed for testing and the rest of the speakers are used for training. In this way the same speakers never appear in both test and train data. This procedure is used to cycle though the speaker sets until all the data has been tested.

Table 4.1: Average AFRL-Speech Group Intra-Accent Word Accuracy Results

	COR	SUB	DEL	INS
German	73.74	13.48	12.75	2.27
British	69.45	16.13	14.43	3.05
Dutch	81.19	11.77	7.05	2.81
Canada	74.34	16.85	8.8	5.08
Average	**74.68**	**14.92**	**10.09**	**3.65**

Benchmark results show that accuracy within an accent group is high, with word accuracy ranging from 74% to 81%. Keywords (the alpha-numeric components of callsigns) and whole callsigns had very high accuracy within accent groups, averaging 89% and 78% respectively.

Table 4.2: Average Keyword and Callsign Accuracy in Intra-Accent Experiments

	Keyword Accuracy	Callsign Accuracy
Canada	90.69	77.85
British	80.66	63.73
Dutch	94.73	87.43
German	89.73	81.27
Average	**88.95**	**77.57**

Evaluating across accent groups required a great deal more experimentation, in total 114 experiments were run and evaluated for word accuracy, callsign accuracy and keyword accuracy across the British, Canadian, German and Dutch data. The average results of each phonetic model (PM) tested on other accents clearly demonstrate that speech recognition across accents reduces word accuracy significantly.

Table 4.3: Average Cross-Accent Word Accuracy Results

	COR	SUB	DEL	INS
Canada PM on British	43.68	38.38	17.98	8.08
Canada PM on Dutch	48.37	44.14	7.44	15.57
Canada PM on German	47.37	41.53	11.06	9.61
Dutch PM on Canada	43.9	39.51	16.62	5.01
Dutch PM on British	34.49	40.95	24.53	3.51
Dutch PM on German	50.75	33.43	15.83	4.78
British PM on Dutch	36.18	55.36	8.48	15.05
British PM on German	30.8	51.29	17.93	6.48
British PM on Canada	41.06	43.77	15.18	5.78
German PM on Dutch	49.99	39.23	10.78	7.91
German PM on British	28.54	42.4	29.06	3.46
German PM on Canada	40.11	42.47	17.45	5.17
Average	**41.27**	**40.9**	**15.52**	**8.36**

Cross-accent callsign and keyword accuracy was reduced in an even more dramatic fashion, with same-accent models being almost twice as accurate (89.0% to 49.3%), on keywords, and many times more accurate for whole callsigns (77.6% to 16.3%).

These results argue for small, specialized phonetic models that target specific accent groups as being the most accurate approach to robust callsign identification. A small vocabulary system is ideally suited to tasks where a compact and highly predictive models needs to be built quickly with a very limited training set, as was the case with the N4 corpus.

Table 4.4: Average Keyword and Callsign Accuracy in Inter-Accent Experiments

	Keyword Acc.	Callsign Acc.
Dutch PM on British	38.9	10.74
Dutch PM on Canada	51.87	21.74
Dutch PM on German	61.2	30.45
Canada PM on Dutch	55.02	17.01
Canada PM on British	52.19	15.1
British PM on Canada	58.62	23.6
British PM on Dutch	40.73	5.44
Canada PM on German	55.94	23.75
British PM on German	36.09	9.62
German PM on Canada	51.29	18.07
German PM on Dutch	54.32	14.75
German PM on British	35.03	5.22
Average	**49.27**	**16.29**

4.4.3 Speaker Recognition Experiments on the N4 Corpus

Three sites ran speaker recognition experiments on the Netherlands segment of the N4 corpus [Zissman 2001]. These sites were MIT Lincoln Laboratory (US), TNO Human Factors (Netherlands), and Information Directorate, Air Force Research Laboratory (US). Given sufficient training data, high-performance speaker recognition was obtained on the short tactical utterances from the NL data set, with single-transmission equal error rates often measured below 5%. There were some statistically significant differences in performance among the speaker recognition algorithms that were evaluated. System performance was largely determined by the complexity of the model (e.g. number of parameters) employed, with simpler systems having somewhat higher error rates that increased the speaker recognition error rate. Cross-language training/testing had a modest system-dependent impact on error rate.

4.4.4 Language Recognition Experiments on the N4 Corpus

Language identification experiments on the N4 corpus were performed at DGA in France [Benarousse 2001]. Training was performed on large quantities of broadcast news speech spoken in English and French. The native, read English and read French portions of the N4 corpus were used for testing. Two-alternative, forced-choice error rates generally decreased as the test duration increased, ranging from about 20% for two seconds of speech to less than 2.3% with 20 seconds of speech. Because the French speech in the training set was European, whereas the French in the N4 corpus is Canadian, it seems that language recognition is rather robust to accent variations, at least for French.

4.5 CONCLUSIONS

A variety of experiments measuring the impact of multilingual and non-native speech on automatic speech processing accuracy have been performed. The results vary depending on the type of technology employed, the way in which in the data are used, and the experimental methodology. Generally, however, we see that speech-processing performance degrades somewhat as we move from single-language, native applications to multilingual, non-native applications. Research efforts seeking to close this gap are underway at many sites worldwide.

4.6 REFERENCES

Benarousse, L., Geoffrois, E., Grieco, J., Series, R., Steeneken, H., Stumpf, H., Swail, C. and Thiel, D. (2001). "The NATO Native and Non-Native (N4) Speech Corpus", In Proc. of the Workshop on Multilingual Speech and Language Processing, Aalborg, Denmark. (Available as NATO Publication RTO-MP-066, April 2003).

Durou, G. (1999). "Multi-lingual Text-Independent Speaker Identification", In Multilingual Interoperability in Speech Technology. pp. 115-118, September 1999. (Available as NATO Publication RTO-MP-28, August 2000).

Lawson, A., Harris, D. and Grieco, J. (2003). "Effect of Foreign Accent on Speech Recognition in the NATO N4 Corpus", submitted to Eurospeech 2003.

van Leeuwen D. and Orr, R. (1999). "Speech Recognition of Non-Native Speech Using Native and Non-Native Acoustic Models", In Proceedings of the MIST Workshop, pp. 23–28, September 1999. (Available as NATO Publication RTO-MP-28, August 2000).

Wanneroy, R., Bilenski, E., Barras, C., Adda-Decker, M. and Geoffrois, E. (2001). "Acoustic-Phonetic Modeling of Non-Native Speech for Language Identification", In Multilingual Interoperability in Speech Technology. (Available as NATO Publication RTO-MP-066, April 2003).

Williamson, D. (2002). "Performance Assessment of a COTS Speech Recognition System on the N4 Database", Air Force Research Laboratory, AFRL-HE-TR-2002-49.

Young, S., Adda-Decker, M., Aubert, X., Dugast, C., Gauvin, J., Kershaw, D., Lamel, L., van Leeuwen, D., Pye, D., Robinson, A., Steeneken, H. and Woodland, P. (1997). "Multi-lingual Large Vocabulary Speech Recognition; the European SQALE Project", Computer Speech and Language, Vol. 11, pp. 73-89.

Zissman, M., van Buuren, R., Grieco, J., Reynolds, D., Steeneken, H. and Huggins, M. (2001). "Preliminary Speaker Recognition Experiments on the NATO N4 Corpus", In Proc of the Workshop on Multilingual Speech and Language Processing, Aalborg, Denmark. (Available as NATO Publication RTO-MP-066, April 2003).

Chapter 5 – RECOMMENDATIONS AND CONCLUSIONS

The field of military communications requires the integrated use of speech technology for command, control, and communications. In addition for multinational environments, it is necessary for a wide range of protocols from participating countries to be integrated together for safe and effective operations. Speech technology offers the promise of more direct and effective communications, verification of personnel, and allowing operators to have seamless access to information. The problem of non-native speech, however, raises a serious obstacle for the transition of commercial off-the-shelf (COTS) speech technology for speaker recognition, speaker verification, synthesis, and coding. Studies conducted by participating NATO laboratories and reported here suggest that performance of COTS speech technology is degraded when used by a non-native speaker of that language. The performance will only be degraded more when that non-native speaker is in a stressful, noisy environment characteristic of most military environments. Advances in basic research to address this problem have not kept up with the demand for more wide spread application of speech technology. It is hoped that this report will serve to focus the speech community on the important issue of speech and language variability due to non-native speech. A database collected during this study has been distributed to all participating NATO countries and is available on CD-ROM format for those interested [Benarousse 2001]. Below we summarize the main finding and recommendations.

1) Military operations are often conducted in which multi-national coalition partners must communicate in a non-native language. These conditions are known to cause problems especially in stressful, noisy military environments.

2) These factors are detrimental to the effectiveness of communications in general, as well as to the performance of communications equipment and weapons systems equipped with vocal interfaces (e.g., advanced cockpits, command, control, and communications systems) trained for the native language.

3) Commercial off-the-shelf speech recognition systems are not yet able to address the wide variability associated with a non-native speaker.

4) Progress in the field of military based speech technology has been restricted due to the lack of availability of database of non-native speech in a military communications scenario.

5) It is certain that in the future it will be necessary to improve the coordination and effectiveness of multi-national military forces. The need therefore exists for planned simulations and exercises requiring coordinated emergency and/or emergency personnel using a wide range of speech technology. Such settings will have to address effective communications between multi-national forces using the same speech systems.

6) The success of the four-year effort by IST-011/RTG-001 has underlined the necessity to further invest coordinated international effort to support NATO interests in understanding speech production and perception and our ability to implement speech systems that are robust to the realities of everyday military speech.

7) In order to share the most recent advances in this field NATO IST/RTG-001 has a web page located at http://extranet.if.afrl.af.mil/ist_slt/ Information found here includes an overview of activities, collected and available speech databases.

REPORT DOCUMENTATION PAGE

1. Recipient's Reference	2. Originator's References	3. Further Reference	4. Security Classification of Document
	RTO-TR-IST-011 AC/323(IST-011)TP/26	ISBN 92-837-1118-1	UNCLASSIFIED/ UNLIMITED

5. Originator

Research and Technology Organisation
North Atlantic Treaty Organisation
BP 25, F-92201 Neuilly-sur-Seine Cedex, France

6. Title

Implications of Multilingual Interoperability of Speech Technology for Military Use

7. Presented at/Sponsored by

The RTO Information Systems Technology Panel (IST/RTG-001) as a result of a project on "Implications of Multilingual Interoperability of Speech Technology for Military Use".

8. Author(s)/Editor(s)	9. Date
Multiple	September 2004
10. Author's/Editor's Address	**11. Pages**
Multiple	40

12. Distribution Statement

There are no restrictions on the distribution of this document.
Information about the availability of this and other RTO unclassified publications is given on the back cover.

13. Keywords/Descriptors

Automated language processing
Automated Speech Recognition (ASR)
COTS (Commercial Off The Shelf)
Data bases
Dictionaries
Digital techniques
Human factors engineering
Information systems
Intelligibility
Interoperability
Languages
Machine translation
Military applications
Multilingualism
NATO forces
Pattern recognition
Precision
Speech analysis
Speech recognition
Speech technology
Terminology
Voice communication
Voice recognition
Words (language)

14. Abstract

Military operations are often conducted under conditions of stress induced by high workload, sleep deprivation, fear and emotion, confusion due to conflicting information, psychological tension, pain, and other typical conditions encountered in the modern battlefield context. These conditions are known to affect the physical and cognitive abilities of human speech characteristics, and this study was intended to determine the actual effects of stress on voice production quality. It is suggested that the effect of operator based stress factors on voice is likely to be detrimental to the effectiveness of communication in general, in particular to the performance of communication equipment and weapon systems equipped with vocal interfaces (e.g., advanced cockpits, command, control, and communication systems, information warfare). Progress in the field of military based speech technology, including advances in speech based system design has been restricted due to the lack of availability of databases of speech under stress. In particular, the type of stress which an operator may experience in the modern battlefield context is not easily simulated, and therefore it is difficult to systematically collect speech data for use in research and speech system training. It is foreseen that in the future it will be necessary to improve the coordination of multi-national military forces. The need therefore exists for planned simulations with military personnel using a wide range of speech technology and addressing factors such as high workload, sleep deprivation, fear and emotion, confusion, psychological tension, pain, etc.

NORTH ATLANTIC TREATY ORGANISATION

RESEARCH AND TECHNOLOGY ORGANISATION

BP 25
F-92201 NEUILLY-SUR-SEINE CEDEX • FRANCE
Télécopie 0(1)55.61.22.99 • E-mail mailbox@rta.nato.int

DIFFUSION DES PUBLICATIONS

RTO NON CLASSIFIEES

Les publications de l'AGARD et de la RTO peuvent parfois être obtenues auprès des centres nationaux de distribution indiqués ci-dessous. Si vous souhaitez recevoir toutes les publications de la RTO, ou simplement celles qui concernent certains Panels, vous pouvez demander d'être inclus soit à titre personnel, soit au nom de votre organisation, sur la liste d'envoi.

Les publications de la RTO et de l'AGARD sont également en vente auprès des agences de vente indiquées ci-dessous.

Les demandes de documents RTO ou AGARD doivent comporter la dénomination « RTO » ou « AGARD » selon le cas, suivi du numéro de série. Des informations analogues, telles que le titre est la date de publication sont souhaitables.

Si vous souhaitez recevoir une notification électronique de la disponibilité des rapports de la RTO au fur et à mesure de leur publication, vous pouvez consulter notre site Web (www.rta.nato.int) et vous abonner à ce service.

CENTRES DE DIFFUSION NATIONAUX

ALLEMAGNE
Streitkräfteamt / Abteilung III
Fachinformationszentrum der
Bundeswehr (FIZBw)
Friedrich-Ebert-Allee 34, D-53113 Bonn

BELGIQUE
Etat-Major de la Défense
Département d'Etat-Major Stratégie
ACOS-STRAT – Coord. RTO
Quartier Reine Elisabeth
Rue d'Evère, B-1140 Bruxelles

CANADA
DSIGRD2
Bibliothécaire des ressources du savoir
R et D pour la défense Canada
Ministère de la Défense nationale
305, rue Rideau, 9e étage
Ottawa, Ontario K1A 0K2

DANEMARK
Danish Defence Research Establishment
Ryvangs Allé 1, P.O. Box 2715
DK-2100 Copenhagen Ø

ESPAGNE
SDG TECEN / DGAM
C/ Arturo Soria 289
Madrid 28033

ETATS-UNIS
NASA Center for AeroSpace
Information (CASI)
Parkway Center, 7121 Standard Drive
Hanover, MD 21076-1320

FRANCE
O.N.E.R.A. (ISP)
29, Avenue de la Division Leclerc
BP 72, 92322 Châtillon Cedex

GRECE (Correspondant)
Defence Industry & Research
General Directorate, Research Directorate
Fakinos Base Camp, S.T.G. 1020
Holargos, Athens

HONGRIE
Department for Scientific Analysis
Institute of Military Technology
Ministry of Defence
H-1525 Budapest P O Box 26

ISLANDE
Director of Aviation
c/o Flugrad
Reykjavik

ITALIE
Centro di Documentazione
Tecnico-Scientifica della Difesa
Via XX Settembre 123
00187 Roma

LUXEMBOURG
Voir Belgique

NORVEGE
Norwegian Defence Research Establishment
Attn: Biblioteket
P.O. Box 25, NO-2007 Kjeller

PAYS-BAS
Royal Netherlands Military
Academy Library
P.O. Box 90.002
4800 PA Breda

POLOGNE
Armament Policy Department
218 Niepodleglosci Av.
00-911 Warsaw

PORTUGAL
Estado Maior da Força Aérea
SDFA – Centro de Documentação
Alfragide
P-2720 Amadora

REPUBLIQUE TCHEQUE
DIC Czech Republic – NATO RTO
LOM PRAHA s. p.
o.z. VTÚL a PVO
Mladoboleslavská 944, PO BOX 16
197 21 Praha 97

ROYAUME-UNI
Dstl Knowledge Services
Information Centre, Building 247
Dstl Porton Down
Salisbury
Wiltshire SP4 0JQ

TURQUIE
Milli Savunma Bakanlığı (MSB)
ARGE ve Teknoloji Dairesi Başkanlığı
06650 Bakanliklar – Ankara

AGENCES DE VENTE

NASA Center for AeroSpace Information (CASI)
Parkway Center, 7121 Standard Drive
Hanover, MD 21076-1320
ETATS-UNIS

The British Library Document Supply Centre
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
ROYAUME-UNI

Canada Institute for Scientific and Technical Information (CISTI)
National Research Council
Acquisitions, Montreal Road, Building M-55
Ottawa K1A 0S2, CANADA

Les demandes de documents RTO ou AGARD doivent comporter la dénomination « RTO » ou « AGARD » selon le cas, suivie du numéro de série (par exemple AGARD-AG-315). Des informations analogues, telles que le titre et la date de publication sont souhaitables. Des références bibliographiques complètes ainsi que des résumés des publications RTO et AGARD figurent dans les journaux suivants :

Scientific and Technical Aerospace Reports (STAR)
STAR peut être consulté en ligne au localisateur de ressources uniformes (URL) suivant:
http://www.sti.nasa.gov/Pubs/star/Star.html
STAR est édité par CASI dans le cadre du programme NASA d'information scientifique et technique (STI)
STI Program Office, MS 157A
NASA Langley Research Center
Hampton, Virginia 23681-0001
ETATS-UNIS

Government Reports Announcements & Index (GRA&I)
publié par le National Technical Information Service
Springfield
Virginia 2216
ETATS-UNIS
(accessible également en mode interactif dans la base de données bibliographiques en ligne du NTIS, et sur CD-ROM)

NORTH ATLANTIC TREATY ORGANISATION

RESEARCH AND TECHNOLOGY ORGANISATION

BP 25
F-92201 NEUILLY-SUR-SEINE CEDEX • FRANCE
Télécopie 0(1)55.61.22.99 • E-mail mailbox@rta.nato.int

DISTRIBUTION OF UNCLASSIFIED RTO PUBLICATIONS

AGARD & RTO publications are sometimes available from the National Distribution Centres listed below. If you wish to receive all RTO reports, or just those relating to one or more specific RTO Panels, they may be willing to include you (or your Organisation) in their distribution.

RTO and AGARD reports may also be purchased from the Sales Agencies listed below.

Requests for RTO or AGARD documents should include the word 'RTO' or 'AGARD', as appropriate, followed by the serial number. Collateral information such as title and publication date is desirable.

If you wish to receive electronic notification of RTO reports as they are published, please visit our website (www.rta.nato.int) from where you can register for this service.

NATIONAL DISTRIBUTION CENTRES

BELGIUM
Etat-Major de la Défense
Département d'Etat-Major Stratégie
ACOS-STRAT – Coord. RTO
Quartier Reine Elisabeth
Rue d'Evère
B-1140 Bruxelles

CANADA
DRDKIM2
Knowledge Resources Librarian
Defence R&D Canada
Department of National Defence
305 Rideau Street
9th Floor
Ottawa, Ontario K1A 0K2

CZECH REPUBLIC
DIC Czech Republic – NATO RTO
LOM PRAHA s. p.
o.z. VTÚL a PVO
Mladoboleslavská 944, PO BOX 16
197 21 Praha 97

DENMARK
Danish Defence Research Establishment
Ryvangs Allé 1
P.O. Box 2715
DK-2100 Copenhagen Ø

FRANCE
O.N.E.R.A. (ISP)
29, Avenue de la Division Leclerc
BP 72
92322 Châtillon Cedex

GERMANY
Streitkräfteamt / Abteilung III
Fachinformationszentrum der Bundeswehr (FIZBw)
Friedrich-Ebert-Allee 34
D-53113 Bonn

GREECE (Point of Contact)
Defence Industry & Research General Directorate, Research Directorate
Fakinos Base Camp, S.T.G. 1020
Holargos, Athens

HUNGARY
Department for Scientific Analysis
Institute of Military Technology
Ministry of Defence
H-1525 Budapest P O Box 26

ICELAND
Director of Aviation
c/o Flugrad, Reykjavik

ITALY
Centro di Documentazione Tecnico-Scientifica della Difesa
Via XX Settembre 123
00187 Roma

LUXEMBOURG
See Belgium

NETHERLANDS
Royal Netherlands Military Academy Library
P.O. Box 90.002
4800 PA Breda

NORWAY
Norwegian Defence Research Establishment
Attn: Biblioteket
P.O. Box 25, NO-2007 Kjeller

POLAND
Armament Policy Department
218 Niepodleglosci Av.
00-911 Warsaw

PORTUGAL
Estado Maior da Força Aérea
SDFA – Centro de Documentação
Alfragide, P-2720 Amadora

SPAIN
SDG TECEN / DGAM
C/ Arturo Soria 289
Madrid 28033

TURKEY
Milli Savunma Bakanlığı (MSB)
ARGE ve Teknoloji Dairesi Başkanlığı
06650 Bakanliklar – Ankara

UNITED KINGDOM
Dstl Knowledge Services
Information Centre, Building 247
Dstl Porton Down
Salisbury, Wiltshire SP4 0JQ

UNITED STATES
NASA Center for AeroSpace Information (CASI)
Parkway Center, 7121 Standard Drive
Hanover, MD 21076-1320

SALES AGENCIES

NASA Center for AeroSpace Information (CASI)
Parkway Center
7121 Standard Drive
Hanover, MD 21076-1320
UNITED STATES

The British Library Document Supply Centre
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
UNITED KINGDOM

Canada Institute for Scientific and Technical Information (CISTI)
National Research Council
Acquisitions
Montreal Road, Building M-55
Ottawa K1A 0S2, CANADA

Requests for RTO or AGARD documents should include the word 'RTO' or 'AGARD', as appropriate, followed by the serial number (for example AGARD-AG-315). Collateral information such as title and publication date is desirable. Full bibliographical references and abstracts of RTO and AGARD publications are given in the following journals:

Scientific and Technical Aerospace Reports (STAR)
STAR is available on-line at the following uniform resource locator:
http://www.sti.nasa.gov/Pubs/star/Star.html
STAR is published by CASI for the NASA Scientific and Technical Information (STI) Program
STI Program Office, MS 157A
NASA Langley Research Center
Hampton, Virginia 23681-0001
UNITED STATES

Government Reports Announcements & Index (GRA&I)
published by the National Technical Information Service
Springfield
Virginia 2216
UNITED STATES
(also available online in the NTIS Bibliographic Database or on CD-ROM)

ISBN 92-837-1118-1