Deliverable D 1.1
Deliverable title: Benchmark

Project acronym:	Translate4Rail
Starting date:	01/12/2019
Duration (in months):	24
Call (part) identifier:	H2020-S2RJU-OC-2019
Grant agreement no:	881779
Due date of deliverable:	Month 2
Actual submission date:	17-08-2020
Responsible:	RNE
Dissemination level:	PU
Status:	Final

Reviewed: yes

Document history		
Revision	*Date*	*Description*
1	31-01-2020	First issue
2	21-08-2020	The updated version The Benchmark of the new industries and tools added

Report contributors		
Name	Beneficiary Short Name	Details of contribution
RailNetEurope	RNE	Consolidation of the benchmarking Assessment and analysis
Union Internationale des Chemins de fer	UIC	Initial benchmark research Consolidation of the report

Table of Contents

1. Executive Summary

Benchmark – Speech to Speech: State of the Art (Internet-based search)

Regulations in Directive 2007/59/EC makes it mandatory for train drivers operating across borders to be fluent in the language of the foreign country at level B1 (Common European Framework of Reference for Languages standards). This also leads to increased costs and is time-consuming for the Railway Undertakings. Currently, sector is coping with a general lack of train drivers in Europe. To be able to recruit more drivers and to be ready to easier cross borders, it is of great importance to find a solution where the railway industry can compete with other types of transports in Europe.

The concept of the Translate4Rail project is to offer train drivers a fully comprehensive set of predefined standardised messages which encompass all they have to exchange with an infrastructure manager traffic controller in normal or exceptional operational situations in a country where they do not understand nor speak the local language.

An IT tool will then be implemented to enable the driver and the traffic controller to understand each other even though each of them speaks in his/her native language. This will capitalise on the works already committed on this matter between Infrastructure Managers (IMs) and Railway Undertakings (RUs) at RNE and UIC level. These works have dealt with the analysis of the various types of operational situations needing exchanges between RUs drivers and IMs signallers. The project will enable to test these works and to further harmonise and standardise pre-defined messages in the light of the tests carried out. Such messages will be uttered by the driver or the traffic controller. They will then be identified, translated and spoken in the language of the other party by the given tool.

The project will define the functional characteristics of the tool which will create a frame for the exchanges between drivers and traffic controllers. This tool will use voice recognition and translation applications. The tool will be tested on pilot trains running on cross border sections of rail freight corridors where drivers have to use different languages. The project intends to at least maintain the level of safety, increase the traffic fluidity at borders and to increase the competitiveness of the rail sector.

Years ago, research has been started regarding the use of computers to recognize and render sign language. Currently, there are numbers of research projects focused also for text and speech automatic translation. Technology is rapidly changing and improving the way we know the world operates today and could in the future.

A typical speech-to-speech translation system consists of three components: speech recognition, machine translation and speech synthesis. Many techniques have been proposed for the integration of speech recognition and machine translation. However, corresponding techniques have not yet been considered for speech synthesis. [1]

Now, we are dealing with the question: *What are the areas in which you see artificial intelligence playing a role?* And there is a simple answer: *Whether we admit it or not, new technologies and artificial intelligence in the future, perhaps soon, will greatly affect our work and our way of life. It is up to us how we can take advantage of the opportunities that could positively influence our daily lives.*

For man, language is the most natural thing, and so it is the greatest challenge that machines face, just like humans.

Depends on the topic, this benchmark report is partly a compilation of internet-based search.

2. Abbreviations and acronyms

Abbreviation / Acronyms	Description
ABSR	Assistant Based Speech Recognition
AI	Artificial Intelligence
Agency	European Union Agency for Railways
AMAN	Arrival Manager
ANSP	Air Navigation Service Provider
API	Application
ASR	Automatic Speech Recognition
ATC	Air Traffic Control
ATCO	Air Traffic Controller
ATM	Air Traffic Management
CAT	Computer-Assisted Translation
CEFR	Common European Framework of Reference for Languages
CER	Community of European Railway and Infrastructure Companies
CMU	Carnegie Mellon University
CSD	Constant Speed Drive
DFS	Deutsche Flugsicherung
DST	Decision Support Tool
EIM	European Rail Infrastructure Manager
ERFA	European Rail Freight Association
EU	European Union
FLE	Foreign Language Speaker
GNMT	Google Neural Machine Translation
GSM-R	Global System for Mobile Communication – Railway
HLT	Human Language Technology
IATA	Air Transport Association
ICAO	International Civil Aviation Organisation
IM	Infrastructure manager
IMO	International Maritime Organisation
INTUIT	Interactive toolset for understanding trade-offs
IP	Innovation Programme
IT	Information technologies
IVR	Interactive Voice Response
IWSLT	International Workshop on Spoken Language Translation
MT	Machine Translation
NLU	Natural Language Understanding
NTCC	National Traffic Control Centre
PC	Personal Computer
PDA	Personal Digital Assistant
RNE	RailNetEurope
RU	Railway undertaking
SBC	Singe Board Computer
SESAR	Single European Sky – Research and Development
SIDE	Speech Integrated Development Environment
SMCP	Standard Marine Communication Phrases
SME	Small and Medium-sized Enterprises
SLU	Spoken Language Understanding
SNCF	French National Railways

SSE	Speech Signal Enhancement
SST	Speech-to-speech translation
STACL	Simultaneous Translation with Anticipation and Controllable Latency
STS	Speech-to-speech (S2S)
SW	Software
sWG	sub-Working group
TSI	Technical Specification for Interoperability
TTS	Text-to-text speech (T2T)
T4R	Translate4Rail project
UAS	Unnamed Aircraft System
UIC	International Union of Railways
VHF	Very High Frequency
WASN	Wireless Acoustic Sensor Network
WMT	Workshop on Statistical Machine Translation
WP	Work Package
WO	Written Order
XML	Extensible Markup Language

3. Background

The present document constitutes the Deliverable D1.1 "BENCHMARK" in the framework of the project titled "Translation for breaking language barriers in the railway field". This report is the first deliverable of WP1. The aim is to perform a complete benchmark of relevant activities related to automatic speech translation in other industries (e.g. air traffic management, military applications, etc.) as well as current initiatives in the rail sector to find what is the best performance that has been achieved in the field of speech to speech technology in order to analyse what would be the most suitable solution for the language programme.

Thanks to AI whose development is accelerating, the field of voice recognition is very active in many human activities. The aim of this subtask is to benchmark what is the best performance that has been achieved in the field of speech to speech technology – whether in different industries we could find a solution which could wholly or partly fit the need of the railway sector and could be tested in pilots in the framework of the revised Annex VI to Directive 2007/59/EC.

The aim of Translate4Rail is to develop a prototype language tool and carry out necessary tests to prove the concept with drivers who speak at least the B1 level (Common European Framework of Reference for Languages standards) and demonstrate, through the output of the project that we could do it with a lower level than B1.

The project outputs are also represented by technological digital solutions. **The technological digital solutions should demonstrate the ability to support safe, effective and efficient communication.**
The digital solution of the T4R project will be tested in a Pilot. To ensure that safety is not compromised, pilots will initially be run with the help of bilingual drivers with a language proficiency at the level of B1. Besides that, for all key stage of the project, a safety analysis will be carried out. These stages cover - operational situations, voice recognition, list of messages, translation of messages, transmission by GSM-R, installation of the equipment supporting the tool in the IMs traffic controller offices and the driver cabs and timing of the process in relation with the urgency of the reaction.
The laboratory and operational testing will improve and standardizes the messages exchanged and way of communications between train drivers and traffic controllers. The successful result of the testing and implementation lead to better competitiveness of the railway sector. The language tool will enable significantly enlarge the geographical scope of work of the train drivers. This will bring more flexibility and productivity to RUs.

All mentioned activities have its background in the European Commission Regulation 2019/554 which set up the condition for Pilot project. It states the one or several railway undertakings in cooperation with one or several infrastructure managers may **carry out pilot projects to test alternative means** of ensuring effective communication. In order to prepare the technical specifications for language tool development, analyses of language tools use in different sectors have been carried out. This analyse is the content of the next chapters.

4. LANGUAGE TRANSLATION TO BREAK LANGUAGE BARRIERS IN THE RAILWAY FIELD

The concentration of long-distance traffic on certain corridors has increased the need for removing the language barriers for train driver crossing national borders. The incident at Rasttat caused disruption of Corridor 1 for a significant time. This accident showed clearly how great were the consequences of capacity restriction where the only alternative was in a neighbouring country with different languages. The industry and the distribution centres mainly served by trains had seriously suffered during that period. The goods transporters, railway undertaking companies, had not been able to utilise the offered alternative train paths due to the lack of multilingual drivers. The Translate4Rail project aims to wipe off this language barrier to enable any driver to continue to drive safely its train in a country where he does not speak the national language.

4.1. Translate4Rail project objectives and benefits

The Translate4Rail project will capitalise on the works already committed on this matter between Infrastructure Managers (IMs) and Railway Undertakings (RUs) at RNE and UIC level. These works have dealt with the analysis of the various types of operational situations covering the interaction between RUs drivers and IMs signallers.
The project will enable to test these situations and further will help to harmonise and standardise pre-defined messages in the light of the pilot results. Such enhanced and test proved messages will be used for safe communication between the train driver and the traffic controller. These messages will be identified, translated and applied in the language of the other party by the given tool.

The main objectives of the project are:

- Enable any train driver to be safe when driving in a foreign country of which he does not know the language;
- Define specifications of a tool that enables the train driver to communicate with the infrastructure manager´s traffic controller;
- Define standard wordings for communication;
- Organise a tender for a tool;
- Test the tool within pilots;
- Check the safety;
- Define the training program for the drivers using the tool (list of predefined messages);
- Disseminate widely the results of the tests and the characteristics of the tool to the railway undertakings community and the Infrastructure managers community.

The uniqueness of the project is determined by the objectives themselves, which are designed to improve interoperability in railway transport operation.

5. OBJECTIVE/AIM

This document has been prepared to provide a general overview of existing language tools, which are used in the different sectors for automatic translation of speech to text/text to speech communication.

The aim of this benchmark report is to focus on existing possibilities and/or define the characteristics, which should be used for preparing the requirements of a digital solution under the objectives of Translate4Rail project. The benchmark presents the existing language tools overview that has been achieved by analysed sectors in the field of speech to speech technology to learn from and find what could be the suitable solution for the Translate4Rail project.

6. BENCHMARK IN RAILWAYS AND OTHER INDUSTRIES

This chapter consists of activities done in the sector and from the description of the specific project(s) which could be benchmarked to the Translate4Rail project.
All activities and projects mentioned in this chapter include a short description of their content, objectives and, in some cases, results.

6.1. Railways

6.1.1. Use of speech recognition in simulators

- ***A computer-based training system for the TGV radio-security procedures (1995)***

In parallel with the development of the high-speed railway network throughout Europe, the radio security system on board of the TGV (bullet train) had to be compatible with the different procedures.
Focusing on the language issues, to guarantee language independency purposes, the unit has contained mainly icons buttons. Besides this, since procedures dialogues are very often conducted in foreign languages, the dialogues have been modified and constrained like aeronautic radio procedures. This solution avoids misunderstandings and ensures optimal security of the communications.
The project proposed to integrate speech recognition in a multimedia computer-based training system. This system is used to teach French drivers from functionalities of the new radio control unit they are able to use on either the French, British or Belgium rail networks and to train them to the new radio procedures that they are executing in English. The system combines multimedia techniques like video images restitution or speech recognition, for simulation quality purpose. The important part of the unit is the user interface and the way how speech technologies have been used.

Table 1 TGV training system vs. “a T4R tool”

What is the main characteristic of the project	What can be used for the T4R translation tool
Modified and pre-set dialogues	The concept of the T4R project is to offer drivers a comprehensive set of PDMs
Icon buttons	An option for a T4R Tool in case of choosing on the “screen”

- ***LLoquendo Automates Railway Information Service (2002)***

The automated speech-interaction information service realized by LLoquendo for Italian national railway network & services company Trenitalia - has brought outstanding customer services results.
The service, which is known as FS Informa, provides customers throughout Italy with the very latest information on train travel over any phone (wireline or wireless) through flexible, user-friendly speech interaction.
Last minute information is automatically integrated with regular information in real-time to give customers information on timetables, services, train fares, industrial action, delays, etc. The system manages a steady flow of 30,000 calls every day - 90 percent of which are managed completely automatically.
The capacity to successfully manage extremely high call peaks was confirmed during the period leading up to the strikes that took place throughout Italy in May 2002. The service automatically managed 70,000 calls from customers requiring information on departing and cancelled trains.

Table 2 LLoquendo Automates vs. "a T4R tool"

What is the main characteristic of the project	What can be used for T4R translation tool
Mostly automatically managed calls	Not expected within T4R
One-way translation (for passengers in Italy) for a specific type of service	Divide PDMs into clusters (as a specific area/service)

- ***Railway company testing interactive speech translation system for foreign passengers (2018)***

Keikyu Corporation has **started testing an interactive speech translation system to give guidance to foreign train passengers in a bid to improve communication with station employees.**
The system can handle **four languages – Japanese, English, Mandarin Chinese and Korean.**
When foreign passengers ask questions to station employees, the staff will use a tablet installed with the multi-language speech translation application. The cloud-based system shows what the passengers say in their own language and in Japanese, and then the reply by the station staff member in both languages.
The railway company is using this system to give guidance to foreigners.
The system puts together the cloud-based system with a speech translation engine based on the app "VoiceTra".

- ***A railway locomotive combined control method and apparatus for speech analysis (2012)***

A kind of railway vehicle machine joint control speech analysis method and equipment, shows that joint control calls through result for carrying out analysis to the voice signal in station-locomotive joint control, comprise the steps:

1. the filtering process is carried out with received voice signal, carry out A/D transformation after retaining the voice signal of 300-3400 Hz and subsequent conversion to audio digital signals;
2. the sound stream information station-locomotive joint control is extracted from audio digital signals;
3. according to the setting of recognition strategy, this sound stream information is mated with the standard call model in station-locomotive joint control. Further steps are to adopt continuous speech recognition based on the language model and keyword spotting strategy.

The equipment based on this control method for speech analysis allows to record the voice samples, collects the signal and encode it, compresses and stores records, automatically identifies station-locomotive joint control voice content, judges whether that joint control communication is complete, replies and carries out data success statistics.

Use of speech recognition in simulators [2]

6.1.2. RNE and UIC ongoing projects

Language Programme and XBORDER

- ***ABATEMENT OF LANGUAGE BARRIERS***

The Language Programme was initiated by RNE in December 2017. During 2018 the programme defined distinct situations to focus on:

- **IM-IM communication at the national level**
- **IM-RU operational communication**
- **IM-IM communication at the regional level**
- **RU-RU operational communication**

The goals of the Language Programme are to:

- **Improve interoperability**
- **Lower costs**
- **Keep the system safe**
- **Contribute to increase cross border traffic**

The tasks of sub-Working groups are focused on supporting the language tool development and its pilot testing.

IM-RU COMMUNICATION of the Language Programme matches the same objectives as the T4R project.
Today, when Railway Undertakings (RUs) are operating across borders, they must relate to regulations in Directive 2007/59/EC. The Regulation makes it mandatory for train drivers operating across borders to be fluent in the IM language at level B1. This leads to high costs and is time-consuming for the Railway Undertakings. Besides that, there is also a general lack of train drivers across Europe.

In addition, **UIC´s project XBORDER** defined the first set of Predefined messages, in the cross-border traffic.
The UIC project started at the end of 2017 with the analysis of 3 border points: Venlo-Kaldenkirchen, Modane-Bardonecchia and Forbach-Saarbrücken regarding the whole handing process for freight trains and their actors involved.
The list of predefined messages was finalised which contained 99 messages. These messages were at this stage available in 4 languages, French, German, Italian and Dutch. To simplify the comparison and allow a common understanding, all sets of predefined messages were translated into English.

Table 3 RNE and UIC ongoing projects vs. "a T4R tool"

What is the main characteristic of the project	What can be used for T4R translation tool
The projects are dealing with Predefined messages and automatic translation for variables in the IM-RU communication	The T4R Translation/Language tool aims to work with the set of standardised PDMs (free speech is also possible)
The characteristics of integrating the Language tool are set up	Testing "a Tool" via T4R Pilot, to reach the objective of implementing an IT tool, which enables the driver and the traffic controller to understand each other even though each of them speaks in his/her native language

RNE and UIC sites [3]

6.2. Air traffic management

6.2.1. Activities in the field of speech recognition in air traffic

- ***Speech recognition technology for air traffic controllers (2018)***

The popularity of air transport continues to grow, placing an even greater workload on air traffic controllers (ATCOs) and their communication with pilots from different part of the world affecting their English pronunciation. This possible **predicament can be improved through an automatic speech recognition system closely integrated with an arrival manager developed by EU and SESAR funded researchers.**

One of the greatest hurdles to introducing higher levels of automation in air traffic management (ATM) is the intensive use of voice radio communication to convey air traffic control (ATC) instructions to pilots.

Automatic speech recognition, which converts human speech into texts, can provide a solution to significantly reduce ATCOs workloads and increase ATM efficiency.

The Horizon 2020 funded MALORCA project aimed to reduce the development and maintenance costs of assistant-based speech recognition (ABSR) by using machine learning instead of manual software programming. The project involves automatically learning local speech recognition and ATCO models from radar and speech data recordings. The project MALORCA is described in chapter 6.2.2.

Table 4 Air traffic speech recognition vs. "a T4R tool"

What is the main characteristic of the project	What can be used for T4R translation tool
Automatic speech recognition, which converts human speech into texts	The recognised speech will be transferred into the text for further confirmation and ex-post evaluation
Machine learning	A developer will be required to implement machine learning to reduce the software programming
Automatic speech recognition	Can be implemented also for B1 language level drivers to confirm the recognition, not in T4R but in later projects

- ***The Airbus Air Traffic control speech recognition 2018 Challenge: Towards ATC Automatic Transcription and Call Detection (2018)***

The recent advances in Automatic Speech Recognition (ASR) and Natural Language Understanding (NLU) technologies have opened the way to potential applications in the field of Air Traffic Control (ATC).
Aviation call signs (CS) are communication call signs assigned as unique identifiers to aircraft. **They are expected to adhere to the following pre-defined format:**

- an airline code followed by three to five numbers and zero to two letters.
  For instance, "ENAC School six-seven November" is a call sign in which ENAC school is a company name followed by two numbers (six and seven) and "November" stands for the "n" character in the aviation alphabet.

Speech material

The dataset used for running the challenge is a subset of the transcribed ATC speech corpus collected by Airbus. This corpus contains all the specific features of ATC mentioned above:

- non-native speech;
- bad audio quality;
- code-switching;
- high speech rate, etc.

Table 5 Airbus Air traffic speech recognition vs. “a T4R tool”

What is the main characteristic of the project	What can be used for T4R translation tool
Automatic speech recognition	For B1 language level drivers also to confirm the speech recognition (for now, not in T4R but possible in subsequent projects)

- ***Design and Development of Speech Corpora for Air Traffic Control Training (after 2016)***

The air traffic control (ATC) constitutes a crucial segment of the whole air traffic industry – the air traffic controllers communicate with the pilots almost continuously in order to ensure the fluent and safe flow of the aerial traffic.
The job of a controller is very demanding and requires – besides the specific personal prerequisites – intensive training. **This training is mainly focused on teaching and reinforcing the communication skills of the aspiring controller.**
The current state-of-the-art training procedure (at least in the Czech Republic) involves so-called pseudo pilots. These are usually retired pilots that prepare training scenarios and consequently act as pilots of a virtual plane (usually more than one at a time), communicate with the controller in training (trainee) and process the spoken prompts received from trainees to the form that can be entered into the software that simulates the plane movement on the radar screen.

Two major drawbacks were identified in such a training set.

- The length of the controller training (approx. 2 years on average) and the relatively high salaries of the pseudo pilots make the whole process very expensive. **This was actually the first incentive that sparked the idea of developing an automatic training simulator based on the intelligent spoken dialogue system.**
- The training environment lacks the noise that is massively present in the real-world VHF radio communication; this might lead to a drastic decrease of the unprepared controller’s ability to understand the communication once he is put into service. On the other hand, the human pseudo pilot usually handles several virtual airplanes. This might result in a confusion of the trainee as he hears the same voice from different simulated aircraft.

Table 6 Speech Corpora vs. “a T4R tool”

What is the main characteristic of the project	What can be used for T4R translation tool
Automatic training simulator based on the intelligent spoken dialogue system	Can be implemented to teach foreign languages or train pre-defined messages, not in T4R but in other projects

- ***Adacel – advanced speech recognition applications and operational air traffic management systems***

Adacel Technologies provides systems to commercial and military aviation, and the technology is being embraced around the world.
The company uses SRI's speech engine as its underlying technology but had to make many modifications for the transportation industry. Designing systems for this environment has its share of challenges, the least of which is noise. There are also language and accent considerations, coarticulation (the slurring of words together when users enunciate them consecutively and rapidly), and very specific industry terms that have to be recognized. The company developed a system that **produces 99.1 percent accuracy** in 120 decibels of noise while operating in a 6G (six times the weight of gravity) environment.

On the training front, in addition to its voice-driven systems for cockpit and tower simulators, **Adacel has also developed speech recognition technologies to help aviation personnel learn English. English is the official language of the air, as mandated by the International Civil Aviation Organization (ICAO),** and everyone must achieve proficiency in English before being allowed to fly or control an airplane anywhere in the world. The test that such personnel are required to pass also involves understanding industry-specific terminology.

Adacel developed a proprietary Speech Integrated Development Environment (SIDE), which combines a number of speech recognition technologies for the design, development, testing, optimization, deployment, and support of speech applications. SIDE includes tools for grammar development, dictionary development, automated batch grammar testing, grammar coverage queries, phonetic distance analysis, perplexity analysis, and co-articulation handling.

Table 7 ADACEL vs. "a T4R tool"

What is the main characteristic of the project	What can be used for T4R translation tool
The system that produces 99.1 percent accuracy in 120 decibels of noise while operating in a 6G (six times the weight of gravity) environment	Usually, the translation from/to English works well. It is not the same if one language is not English. The noise-cancelling can be the part which could be used in T4R but in other projects. The price of such a solution could be far over the allocated budget
SIDE combines speech recognition technologies for the design, development, testing, optimization, deployment, and support of speech applications. SIDE includes tools for grammar and dictionary development, automated batch grammar testing, grammar coverage queries, phonetic distance analysis, perplexity analysis, and co-articulation handling.	Needs further investigation, and can be used in subsequent railway language tool development

- ***Speech Technologies Take to the Skies in Newest Fighter Jet (2007)***

The DynaSpeak speech recognition engine was incorporated into onboard computer systems, allowing pilots to give voice commands to control both communication and navigation systems and enter data into the flight management system.
The DynaSpeak system will feature additional noise-cancelling technologies to filter out the many other sounds in the cockpit. Furthermore, speech recognition technology is speaker-independent, meaning that any pilot flying the plane could use the system immediately. The pilot speaks to the system through a microphone integrated into his flight helmet. A press-to-recognize switch, when depressed, tells the system to begin recognizing the pilot's voice inputs.
That system was flight testing. For this study, the benchmark is based on the noisy environment.

Table 8 Fighter Jet speech technologies vs. "a T4R tool"

What is the main characteristic of the project	What can be used for T4R translation tool
Allowing pilots to give voice commands to control both communication and navigation systems and enter data into the flight management system	Can be used for railways in the future, meanwhile, it is not the subject of T4R project

Activities in the field of speech recognition in air traffic [4]

6.2.2. MALORCA

Machine Learning of Speech Recognition Models for Controller Assistance (2016 – 2018)

MALORCA project proposed a general, cheap and effective solution to automate this re-learning, adaptation and customisation process to new environments, taking advantage of a large amount of speech data available in the ATM world. Machine learning algorithms using these data sources will automatically adapt the ABSR models to the respective environment.

MALORCA project covered a 24 months period and started in April 2016.

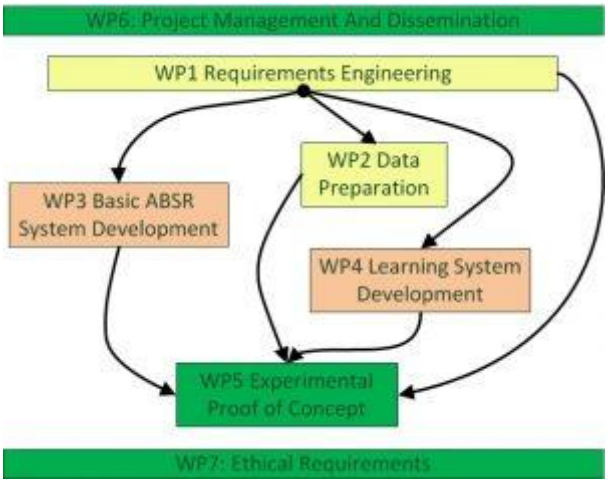


Figure 1 The schema of the MALORCA project

The project was divided into 5 main work packages (WP1-5) and the project management and dissemination in work package WP6 respectively the Ethical Requirements in WP7.

- WP1 analysed the requirements with respect to the users' needs and especially with respect to data availability;
- WP2 was responsible for delivering the first set of target-data (real speech recordings from air traffic controllers) for both selected approach areas (i.e. Prague and Vienna);
- WP3 the assistant based on automatic speech recognizer (ABSR-0) for a generic approach area together with an integrated basic recognition model was built;
- WP4 a significant amount of training data (approximately 100 hours of speech recordings) was provided by the two ANSPs each together with the associated context information from other sensors;
- WP5 was related to the experimental proof of concept. As controllers are end-users of speech recognition, their feedback is essential.

The project provides the aviation industry with a practical approach for developing and deploying a state-of-the-art speech recognition system and integrating it into today's voice communication systems for air navigation service providers.
The use of machine learning for speech recognition is only the first test case for its wider application in ATM. Its application in ATM can also help to reduce the adaptation and maintenance of other ATM tools. The adaptation for example of a generic Arrival Manager to specific airports could be the next step.

Table 9 MALORCA project vs. “a T4R tool”

What is the main characteristic of the project	What can be used for T4R translation tool
Automatic speech recognition, which converts human speech into texts, provides a solution to significantly reduce controllers´ workloads and increase ATM efficiency	The outputs how to set up automatic speech recognition (in case of converting human speech into text)
Possibility of automatically learning of local speech recognition and controllers ‘models from radar and speech data recordings	Possibilities of setting up “a tool”, in some case, as self-learning and/or know, how to feed the tool with new data
The structure of the project is similar to the T4R project	More details need to be gain from the project. We could learn from the testing approach

MALORCA project [5]

6.2.3. AcListant

Active Listening Assistant (until 2016)

The ability to listen to what another person says and adapt actions accordingly, that is one of the key points for every automatic translation machine.
This was achieved for communication between air-traffic controllers and pilots as part of the **AcListant project (Active Listening Assistant).**

The AcListant project shows, for a specific airport and its airspace, how to avoid this delay by analysing the controller-pilot-communication and using the gained information as an additional sensor.
The project aim is defined by achieving “recognize the controller concepts at a rate of at least 95%”.

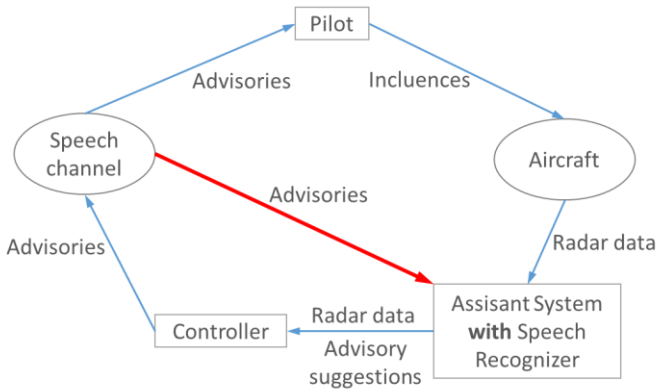


Figure 2 The workflow of the AcListant

Follow-up project AcListant-Strips

In the AcListant project, the performance of approach management systems (AMAN) **was improved through the use of assistant-based speech recognition (ABSR).** ABSR is however also considered to have great potential as an input device to replace mouse and keyboard commands given by air traffic controllers. This was the reason for starting follow-up project AcListant Strips in May 2015. Air traffic control centres have increasingly replaced or are replacing paper flight progress strips with electronic solutions.
It quantifies the advantages which result when speech recognition is used as the primary input device for tracing electronic flight progress strips. Mouse and keyboard inputs are then only necessary in cases where speech recognition does not work.

Table 10 AcListant project vs. "a T4R tool"

What is the main characteristic of the project	What can be used for T4R translation tool
Speech recognition is used for improving the management system (the acceptable rate of speech recognition was set up)	Set up the acceptable rate of speech recognition also in railway transport/communication
The content of the project deals with collecting and preparing speech data used in the operational environment for speech recognition and testing in a demonstration	Similar principles defined in work packages of T4R

AcListant projects [6]

6.3. Maritime traffic

6.3.1. Activities in the field of speech recognition in maritime traffic

- ***Speech recognition for safe maritime traffic (2016)***

Robust even under difficult acoustic conditions
Intelligent speech input systems are increasingly being used as a user interface in human-technology interaction. The scientists use current scientific findings from psychoacoustic and psychophysical fundamental research to develop algorithms with a minimum false recognition rate. The phonologically based speech recognisers can even be adapted to the reduced articulation of stroke patients.

Range of Services

- Robust speech recognition in far-field and acoustically difficult environments;
- Application-specific adaptation of vocabulary;
- Microphoning and signal pre-processing;

- Implementation of recogniser algorithms in embedded systems;
- Evaluation studies.

The project of the speech recognition of Maritime traffic is described in Chapter 6.3.2.

- ***VTS-Bot: Analysis and Implementation of a Student-centred Learning Approach by using a ChatBot Computer Programme (2016)***

In order to have effective communication with each other, the International Maritime Organisation (IMO) **chose English as the 'lingua-franca'** (a word that means trade or "bridge" language) of the sea.
For many seafarers, this leads to the need to learn a second language (English) as well as very specific nautical terms **(Standard Marine Communication Phrases or SMCP).** Many seafarers struggle with this part of their education, and this can lead to communication problems.
The research team have **developed a "chat-bot"** – a computer dialogue system to improve language education in the maritime industry. The chat-bot is programmed with several communication situations relevant to the safety of sea-going ships. Internet-based, computer-controlled dialogues will be available independently from teaching modules. These will enable students to train standard dialogues including the SMCP.

In the future, the chatbot will also be able to examine other research ideas that are associated with verbal communication. These include how seafarers manage risks, or even why some people are more quickly frustrated by ambiguous or distracting verbal exchanges than others.

Table 11 VTS-Bot vs. "a T4R tool"

What is the main characteristic of the project	What can be used for T4R translation tool
Computer dialogue system to improve language education in the maritime industry	It could be used in the railway sector, but for T4R is not relevant as it is educating the English language

Activities in the field of speech recognition in maritime traffic [7]

6.3.2. Fraunhofer IDMT: Improved Maritime Communication

Improved Maritime Communication
Miscommunication is a major trigger for maritime incidents. To counter language barriers in multilingual crews, the IMO Standard Marine Communication Phrases (SMCP) have been introduced. **Fraunhofer IDMT presents a dialogue-oriented training system for SMCP based on automatic speech recognition.** The research objective is to further improve ship-to-ship and ship-to-shore communication by integrating speech recognition in communication systems.

Speech recognition for safe maritime traffic

Communication problems are the cause of more than forty percent of accidents at sea. In over ninety percent of all reported incidents, it is an aggravating factor. Verbal communication on the ship's bridge, in the engine room or by radio is a particular challenge: on the one hand due to the noise level on board, and on the other, because the crew on ships, tugs and in the vessel traffic service centres belong to many different nationalities.

By using acoustic signal processing and computer-based speech recognition, researchers at the Fraunhofer IDMT aim to contribute to improved safety in maritime traffic.

One of the projects is oriented to the training platform for so-called Maritime English. Through speech input and output, the system permits realistic, dialogue-based training of the Standard Marine Communication Phrases (SMCP).

Table 12 Maritime communication project vs. "a T4R tool"

What is the main characteristic of the project	What can be used for T4R translation tool
Dialogue-oriented training system based on automatic speech recognition	The system uses the English language what is not the aim of the T4R project
Training platform so-called Maritime English	Predefined messages as a "training platform"

Fraunhofer IDMT: Improved Maritime Communication [8]

6.4. Military applications

6.4.1. Activities in the field of speech recognition in military applications

- ***Use of Speech and Language Technology in Military Environments (2003)***

The key military applications areas for speech and language technology are:

- command and control;
- communications;
- computers and information access;
- intelligence;
- training;
- and multinational forces.

In principle, all the speech and language technology areas are of some relevance to all fields of application.

Speech Synthesis

Speech synthesis is, in a very simplified definition, the capability of a machine to produce speech. How the task is performed, and what the system is capable of speaking is quite variable.

Techniques can be:

- the use of pre-stored digital messages;

- concatenation of pre-recorded sentences or words generating new sentences that had not been entirely uttered before;
- generation of speech from unlimited text using text-to-speech (TTS) synthesis.

There are many existing applications of this technology in the telephone network. Drawbacks are the need for recording, labelling, and careful editing. This technique is not fitted for applications where there is a need for a large number of pre-stored segments, such as names. Naturalness can also be a problem if the segments are not carefully matched to the intended sentence in terms of prosody.

Text-to-Speech = Linguistic Processing + Speech Generation

Text-to-speech (TTS) synthesis allows the generation of any message from the text. This generally involves the first stage of linguistic processing, in which the text-input string is converted into an internal representation of phoneme strings together with prosodic markers, and the second stage of sound generation on the basis of this internal representation. The normalisation of the text is necessary, for example, to take care of numbers, converting them to dates, years and amounts. The system has also to deal with abbreviations, acronyms, etc.

Speech Recognition

Automatic speech recognition (ASR) is the capability of a machine to convert spoken language to recognized speech units, for example, a word or phrase. The action, which is a function of the application, could be for example the tuning of a radio receiver, a request for information or the conversion of a spoken input to text. Whatever the action, ASR can be valuable where the user's hand and/or eyes are busy performing other tasks, such as a pilot flying an aircraft.

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 is available for making the models for each language. Both the 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.

- ***Speech-to-speech translation technology developed to support military operations in Europe, the Middle East and Southeast Asia (2019)***

Eyes Free, Compact Translation System

The "SQU.ID®" is Voxtec's newest language translation system. The system offers eyes-free language translation capability, allowing users to focus attention on their tasks, not their equipment. The system also offers access to the entire online library of language modules, giving military, law enforcement, first responders and correctional officers unparalleled communication control.

Activities in the field of speech recognition in the military sector [9]

6.4.2. MFLTS Project

Portable speech-to-speech translation on an Android smartphone: The MFLTS system (2018)

The MFLTS project's origins can be traced back to DARPA's "Translation System for Tactical Use" (TransTac) program, which aimed to spur research in the feasibility of **running a full speech-to-speech system on a portable device.** Partially in response to earlier systems that worked on the basis of choosing from a fixed set of phrases (and the limitations arising from that), the goal of the research project was to allow for free-form responses from both the soldier and the foreign speaker.
The **application works in an "interview-style"** fashion where the soldier is the driver of the conversation. Initially, he/she will start the conversation by speaking into the application, and then waiting for it to recognize, translate and output the translation via text-to-speech (TTS). All this happens close to real-time, with latencies from end-of-speech to begin-of-TTS on the order of 1000 milliseconds on the Android system.
On the much more powerful Windows desktop system, there is virtually no latency. Depending on the conversation, the foreign language speaker "FLE" might then be the one responding in turn, just as before speaking into the application and waiting for the translation. It is important to note that it is the soldier who signals through body language to the FLE that it is his turn to speak; not using the application to queue the FLE (e.g., with a TTS prompt) was a key realisation that later became a design principle of the application.

In its current form, there are two ways to interact with MFLTS S2S:

- Smartphone only: When pressed, start recording speech from the phone's internal microphone. Upon lifting the finger, the translation starts.
- Peripheral: A peripheral microphone device was created, alongside a battery-driven speaker, greatly enhances interaction.

The MFLTS S2S application has shown its value as a translation application under real-life constraints, and development is ongoing.

Table 13 Military communication project vs. "a T4R tool"

What is the main characteristic of the project	What can be used for T4R translation tool
Full speech to speech system on a portable device able to work autonomously. The device is able to react to user gestures	The aim of the T4R project is to use a similar concept of speech recognition but without gestures recognition features. The allocated budget hardly can cover the price of the military developed system. T4R will develop its own simplified system using all civil available technologies on the market within the T4R project budget

MFLTS System [10]

6.5. Automotive

6.5.1. Activities in the field of speech recognition in Automotive

- ***SoundHound Inc. Unveils Hybrid Voice AI and Natural Language Understanding System for Cars (2019)***

Houndify Voice AI platform is the solution that enables real-time responses to voice queries in vehicles, even without internet connectivity. This is achieved with **high speed and accuracy through a hybrid speech recognition system that processes voice requests both in the cloud and locally on the embedded system** to return fast responses.

The platform combines the system, tools, and algorithms to enhance the driver's situational awareness, assist in driving functions and provide intelligent interactions between the vehicle and its occupants. With Houndify, drivers can interact with hundreds of domains—programs that provide users with relevant information or actions related to their queries.
These include:

- navigation and weather;
- stock prices
- sports scores;
- flight status;
- local business searches, and hotel searches with complex criteria, etc.
- ***Nuance Automotive Powers Geely's GKUI Smart Ecosystem (2019)***

Nuance Communications, Inc. announced that its automotive platform provides **AI-powered voice recognition for select car models** from Geely, China's fastest-growing automotive manufacturer, in its GKUI smart ecosystem powered by ECARX, a technology company under Geely Group. GKUI is Geely's innovative digital cockpit system, integrating infotainment, connectivity, and vehicle management into one smart in-car system.

Nuance Automotive powers a multitude of features that are core to the GKUI experience on some of Geely's models:

- **High-speed, voice-powered interaction** – Nuance's platform enables an extremely responsive human-machine interface (HMI) that satisfies the demanding standards of the Chinese market, allowing drivers and passengers to interact with GKUI naturally, just as they would with another person, to ask questions and complete requests around all key in-car functions: point of interest search, navigation, temperature control, etc.
- **Speech signal enhancement (SSE) –** A suite of signal processing technologies removes noise from the microphone input and sends out a cleaner signal, improving speech recognition and hands-free communication with GKUI.

Table 14 Automotive project vs. "a T4R tool"

What is the main characteristic of the project	What can be used for T4R translation tool
Allowing drivers and passengers to interact with GKUI naturally, just as they would with another person, to ask questions and complete requests around all key in-car functions: point of interest search, navigation, temperature control, etc.	The use of this system proofs that on the market exists software modules in good quality for speech recognition. In T4R project we ask the developer to choose the right module for railway needs
Signal processing technologies removes noise from the microphone input and sends out a cleaner signal	The similar system needs to be used in the railway sector to reduce the noise, especially in the train driver cabin

Activities in the field of speech recognition in Automotive [11]

6.6. Health sector

A health sector, or medical industry, belongs to progressive industries for the speech-to-speech solutions (translation tools/applications). Days we are living, offer us a lot of opportunities to travel, spend our lives in foreign countries or just work there. This, of course, is related to health care and in case you are a new somewhere, some supporting tool can be really helpful.

Generally, there are three basic types of translation services/applications in the health industry:

- **Predefined medical phrase translator (apps)** – require the user to search for or find a question or statement in order to facilitate a conversation. A health care provider can choose fully predefined sentences, which then can be played or read back to the patient in the chosen language. The Canopy Speak and Universal Doctor Speaker could be an example, here (both available from the Apple iTunes and Google Play for free).
- **Medical Dictionary apps** – require the user to search for a medical term in one language to receive a translation in another language. These apps can help providers find and define specific terms in a given language, not chose a whole / predefined sentence.
- **General language translator (apps)** – require the user to enter a term, statement, or question in one language and then provide a translation in another language. Google Translate, Microsoft Translator and others, are examples here.

6.6.1. Activities in the field of speech recognition in the Health sector

- ***Global Speak Translation***

The Global Speak Translation company offers for the medical sector a professional medical translator. This certified medical translator can translate consent forms, prescription information and other important medical documents. For effective patient care, live communication between patients and healthcare professionals is facilitated.

Healthcare documents that can be translated include:

- Consent authorization forms;
- Medical marketing brochures and Website translations;
- Operation manuals;
- Patient information forms;
- Compliance documentation;
- Medical questionnaires and Pharmacological studies;
- Contracts;
- Training materials;
- Medical data sheets translations;
- Simultaneous interpretation;
- Scientific journal articles;
- Drug registration forms.

- ***Care To Translate***

Care to Translate, earlier called Språk i vården, was founded in the fall of 2015 by medical students from Karolinska Institute. Today it has become one of the most used translation tools in the Swedish healthcare sector and **is used in over 100 countries worldwide.**
Care to Translate offers medically correct communication with translations verified by native speakers. It can be used in acute situations or in daily work at the ward. Care to Translate serves as a complement to human interpreters in healthcare and is available 24/7 in the smart device.

- ***Pair A Phrase***

Pairaphrase for Medical Document Translation is a productivity enhancer for medical document translation, offering to the healthcare companies who hire human translators to provide high-accuracy medical translations. Can be used to translate medical literature, patient correspondence or any content that demands **high-quality translations while protecting confidential data.**

Benefits of Pairaphrase Translation Software For Healthcare Providers:

- Speed up the process of translating medical documents accurately;
- Translate medical documents securely & confidentially;
- Reduce the costs & risks associated with healthcare translation services.

Table 15 Health sector projects vs. “a T4R tool”

What is the main characteristic of the project	What can be used for T4R translation tool
Works with pre-defined messages	The same approach was chosen for T4R
Medical Dictionary apps	Railway jargon will be implemented into the T4R language tool
General language translator	Except for pre-defined messages, the free speech translation will be implemented into T4R
Offers medically correct communication with translations verified by native speakers	Pre-defined messages are prepared by railway experts, text to text translation will be recorded in text format for ex-post check and translation improvement by experts
Medical documents translation	Document translation is out of the scope of the T4R project (can be checked in case of WO translation)

Activities in the field of speech recognition in the Health sector [12]

6.7. Construction industry

The construction sector is one of the next fields, where the international teams' cooperation is more than often. Large construction companies realising their project almost all over the world, very often together with smaller – local constructors and working crews. The variety of nations cause that professional translation services are needed, to ensure that all aspects are well understood by every member of the team.
The translation in the construction industry can cover several aspects, including building plans, surveyor specifications and legal contracts, specifications and manuals and daily communication in offices or on-site. In general, all documentation and processes of the specific construction phase and communication, which are needed to be well translated to reach the specific part of the construction project.

Generally, several of the most used translation applications are well designed also to be used in the construction industry (e.g. Google Translator, iTranslate, Microsoft Translator). Therefore, activities of the translation in the construction industry are more-less oriented to professional translation services.

6.7.1. Activities in the field of "translation" in the construction industry

- ***Optima Translation – Optilingua International***

Provides low-cost translation services, carried out by combining artificial intelligence with translation tools or CAT, databases, and post-editing. The post-editing is a professional translator, specialising in the sector concerned, working in their mother tongue, and often in the same country where the target language is used.

Optima Translation´s benefits:

- Low prices (a combination of technology and human translation);
- A team of post-editors;
- A team of professionals who specialise in specific fields of service es (e.g. legal, medical, financial, commercial etc.);
- Translation services in many languages (French, English, German, Spanish, Portuguese, Italian, Dutch, Polish, and Russian);
- A high level of professionalism and confidentiality is guaranteed for each service.

- ***GlobalSpeak Translations – Your Business in Every Language***

The GlobalSpeak Translations are not just professionals with a native language proficiency (both source/target language), but also construction industry experts. These services are provided in line with four key points – Efficiency, High level of Safe, Compliance with Local Regulations and take advantage of Technological Development.

Construction (and Engineering) Translation includes:

- Engineering specifications
- Valuations and Patents;
- Regulatory documents and Operation manuals;
- Legal contract translations;
- Emergency response plans;
- Safety training materials;
- Architect planning and AutoCAD drawings;
- Website and documentation translation;
- Annual reports translation, etc.

The translation process works as the following:

- Request and Upload (upload documents to GlobalSpeak Translation servers);
- Evaluation and Comprehension;
- Translation process (paying attention to context and industry-specific language);
- Review (by experts);
- Submit and Revise (revisions are made until the translation satisfies customer´s standards);
- Finalize.

Table 16 Construction industry projects vs. “a T4R tool”

What is the main characteristic of the project	What can be used for T4R translation tool
Combination of technology and human translation	Possibility in the testing phase, but not exactly the scope of the T4R project
Translation of documents	Can be checked in case of WO translation, but not the main aim of the T4R project (full document´s translation)
Translation process considers the industry-specific language	Translation process should consider railway-specific language

Activities in the field of “translation” in the Construction sector [13]

6.8. General Language projects

This chapter contains a selection of projects that can be included in the area called “General Language projects”.

6.8.1. EMIME

Effective Multilingual Interaction in Mobile Environments (2008 – 2011)

The goal of the EMIME project was to enhance such speech-to-speech translation devices by producing personalised speech synthesis. In order to demonstrate this, a speech-to-speech translation system has been created in which the output translated speech sounds like the input speaker.

The EMIME project will help to overcome the language barrier by developing a mobile device that performs personalised speech-to-speech translation, such that a user's spoken input in one language is used to produce spoken output in another language while continuing to sound like the user's voice.
The project builds on recent developments in speech synthesis using hidden Markov models, which is the same technology used for automatic speech recognition. Most current speech recognition systems use hidden Markov models (HMMs) to deal with the temporal variability of speech and Gaussian mixture models (GMMs) to determine how well each state of each HMM fits a frame or a short window of frames of coefficients that represent the acoustic input. [14]
Using a common statistical modelling framework for automatic speech recognition and speech synthesis will enable the use of common techniques for adaptation and multilingualism. Significant progress will be made towards a unified approach for speech recognition and speech synthesis: this is a very powerful concept and will open up many new areas of research. In this project, the use of speaker adaptation across languages is explored so that, by performing automatic speech

recognition, we can learn the characteristics of an individual speaker, and then use those characteristics when producing output speech in another language.

The general objectives of the project:

- Personalise speech processing systems by learning individual characteristics of a user's speech and reproducing them in synthesised speech;
- Introduce a cross-lingual capability such that personal characteristics can be reproduced in a second language not spoken by the user;
- Develop and better understand the mathematical and theoretical relationship between speech recognition and synthesis;
- Eliminate the need for human intervention in the process of cross-lingual personalisation;
- Evaluate of the research against state-of-the-art techniques and in a practical mobile application.

Table 17 EMIME project vs. "a T4R tool"

What is the main characteristic of the project	What can be used for T4R translation tool
To enhance speech-to-speech translation devices by producing personalised speech synthesis	Personalised voice is not going to be used in T4R. The neutral voice will be used in the project to ensure better understanding

EMIME project [15]

6.8.2. TC STAR

Technology and Corpora for speech to speech translation (2004 – 2007)

TC-STAR was envisioned as a long- term effort focused on advanced research in all core technologies for speech-to-speech translation (SST): **speech recognition, speech translation and speech synthesis.**
The objectives of the project: **making a breakthrough in SST research to significantly reduce the gap between human and machine performance.**
The project focused on the development of new, possibly revolutionary, algorithms and methods, integrating relevant human knowledge which is available at translation time into a data-driven framework.
TC-STAR was planned for a duration of six years. The first three years were targeted a selection of unconstrained conversational speech domains -i.e. broadcast news and speeches - and a few languages: native and non-native European English, European Spanish and Chinese.
The second three years, were target more complex unconstrained conversational speech domains and European languages.
Key actions:

- the implementation of an evaluation infrastructure;
- the creation of a technological infrastructure;
- the support of knowledge dissemination.

Table 18 TC STAR project vs. “a T4R tool”

What is the main characteristic of the project	What can be used for T4R translation tool
The project developed new algorithms and methods integrated relevant human knowledge which is available at translation time	The new technologies will be required in the tender
The technological infrastructure was created (for language purposes)	Similar principles defined in work packages of T4R

TC STAR project [16]

6.9. The Benchmark of the industries

The benchmarking process with industries provides a comparison of the products or services. The benchmark of the railway sector with other sectors (projects, tools) can show us, how the language translation could be solved.

The major objectives of this benchmarking are:

- Identifying the types of project where the translation process was key for future development;
- Comparing the results.

The focus has been set on language projects which are comparable to the railway sector. It is important to mention, that every tool, developed by the specific project, is able to remove the noise from microphone input. For the operation conditions, it is one of the highest priorities.

Table 19 Sector projects

Sector (different projects in sector)	Common language for a whole sector is used	Speech recognition	Multi-language speech translation (e.g. application)	Training system for users	Existing metrics for evaluation	Information for customers via "any portable device"
Railway	No	Yes	Yes	Yes	-	Yes
Aviation	Yes	Yes	Yes	Yes	Yes	Yes
Maritime	Yes[1]	Yes	Yes	-	Yes	-
Military	-	Yes	Yes	Yes	-	Yes
Automotive	-	Yes	Yes	-	-	Yes
Health	No	Yes	Yes	Yes	-	Yes
Construction	No	Yes	No[2]	Yes	-	Yes

Note:

- *Could not be compared (e.g. because of some specification in Maritime or Military sector; competition at the market)*

[1] *The specific project where IMO chose English as a "bridge language" of the sea*

[2] *General translation applications are designed to be used in the sector*

There are two possible ways, how to solve the language task, according to the result of the benchmark of the language projects in different sectors.

- One of the possible ways is to set up English as a “lingua-franca” also for the railway sector. But this could increase the cost of educating the railway staff, which would not be with intentions of the Translate4Rail project and the language aim in the railway sector. The language tools also could be used for language education to improve English knowledge.
- On the other side, there are some successful projects, which could be as a model for railways. The specification of speech recognition, multi-language speech translation, a training system for users in the areas, which could be in the first line of improving railway transport.

The results of projects from different sectors have led to the development of various translation tools. Even if we have been still comparing “the transport”, there are some significant technical and operational differences between the transport modes.
To define the requirements for the translation tool for the railway sector, the specific project should be realised. Then the results can be compared with the possibility of using the existing translation tool or develop a new one, according to sector specifications.

7. KEY SOLUTION PROVIDERS

This topic is related to benchmark the companies which we can describe not only as providers of the automatic translation tools but also as providers of text to text translation. The chapter “Key solution Providers” does not suggest that we have focused our benchmark to specific providers. We tried to analyse relevant models, which could fit for the T4R solution.

7.1. Key aspects for a language solution

The Translate4Rail key aspects for language solution, to fulfil the main objective of the research project, are based on the opportunity for IMs and RUs to abatement the language barriers in the railway sector, according to set up the translation tool based on the principle of pairing messages and turn speech to text and back, text to speech. In this case, the translation tool could be the solution as a tool to increase the transport volume, reduce the costs, keep the system safe and improve the interoperability in the railway sector.
Improving the interoperability means to provide and use the tool across Europe and enable any train driver to be safe when driving in a foreign country of which he/she does not know the language. This entails the capacity of the train driver to exchange necessary standard information in written or voice messages with the traffic controller for a safe run of his train and good information on the traffic controller.

The solution should be based on the IT tool recognizing speech, pairing the predefined messages and retelling in paired language. The language tool will be gradually tuned during the laboratory and field tests, to reach the level for operational use in the railway sector and potential development for the GSM-R system.

Requirements which are in the first step defined, then analysed and/or researched are the following:

- define specifications of a tool that enables the train driver to communicate with the infrastructure manager traffic controller by GSM-R and to be understood (definition set the functional characteristics of the information tool and of the software of the voice recognition and the transfer of the voice messages to be sent to the traffic controller);
- the language tool will recognize the input speech and will be able to translate the input speech to the required language as output speech;
- the translation tool will be worked with the database of predefined messages organized in specific clusters according to the railway terminology and operational situations and with the variables, which should be translated via automatic translation module;
- the technical specification for the tool should be set up and the necessary technical and safety documentation should be prepared;
- fine-tuning of the tool, after the test the tool via Pilot;
- cooperate in the process of setting up the training program for the train drivers using the tool, is also one of the tasks.

Developing the IT tool is in many ways related to the developing of Artificial Intelligence (AI). AI is a field within computer science which has a great focus on making machines behave and act like humans.
Following are examples were the AI has been developed for intelligence machines:

- Speech recognition;
- Learning;
- Planning;
- Problem-solving.

Translation technology

Developing of the translation tool should be based on specific translation technology. In this case, machine translation is when words and sentences are automatically being translated into another language. When the tool is translating a text, it needs to reinstate the context of the original language to the target language. Up until now, the biggest challenge for the machine translation tools has been the quality. On the other hand, the machine remembers the logic of translations and can use them again, while in most cases human translators will translate similar text differently.

- **Rule-Based Machine Translation technology**
  The Rule-based machine translation technology has its foundation on many linguistic rules and dictionaries for the languages available in the tool. With the rules and dictionaries, the tool translates the text word by word. The disadvantage of the rule-based machine translator is the hard work to find the correct translation and trying to describe the language. Otherwise, it would be a good translation machine to use when reaching for high translation of words.
- **Statistical Machine translation technology**
  Statistical machine translation technology is built on analysing monolingual and bilingual data (dictionary words) when translating. The tool is translating word by word and has a high quality when translating words. Nevertheless, statistical machine translation tools have difficulties to use the correct grammar. The advantages of using the tool are the memorization of already existing translations. With this solution, the tool can connect several old translations to receive an as high quality of the translation as possible.
- **Neural Machine Translation technology**
  The neural machine translation technology is different from the other two mentioned. First of all, they are much easier to implement. Secondly, they create their own logic references to optimize the translations. By doing this, the tool can be compared to a human brain, where translations are similar to human translations. Due to the ability to translate a bigger amount of data, neural machine translation is a slower alternative than both rule-based and statistical machine translations. Additionally, the tool is translating whole sentences instead of the only word by word.

The Benchmark of each translation tool

After the description of each translation tool mentioned in Chapters 7.2 – 7.11, the Benchmark tables are elaborated. The evaluation method is based on the rule that, as far as the tool meets the requirements, it will be awarded as many points as possible. In the table *"How the tool fits to T4R requirements"*, the indicator of safety is evaluated according to knowledge from this research. The "safety level" should be analysed for the railway conditions, when the database of the tool are feed with the vocabularies and messages used in railway communication (railway jargon).

The subject of further analysis should also be a possibility to connect the tool to the GSM-R network. For these purposes, the future deployment of "a translation tool" will be the main condition of the analysis. For the purpose of this benchmark, the SW support of the tool is an evaluation point.

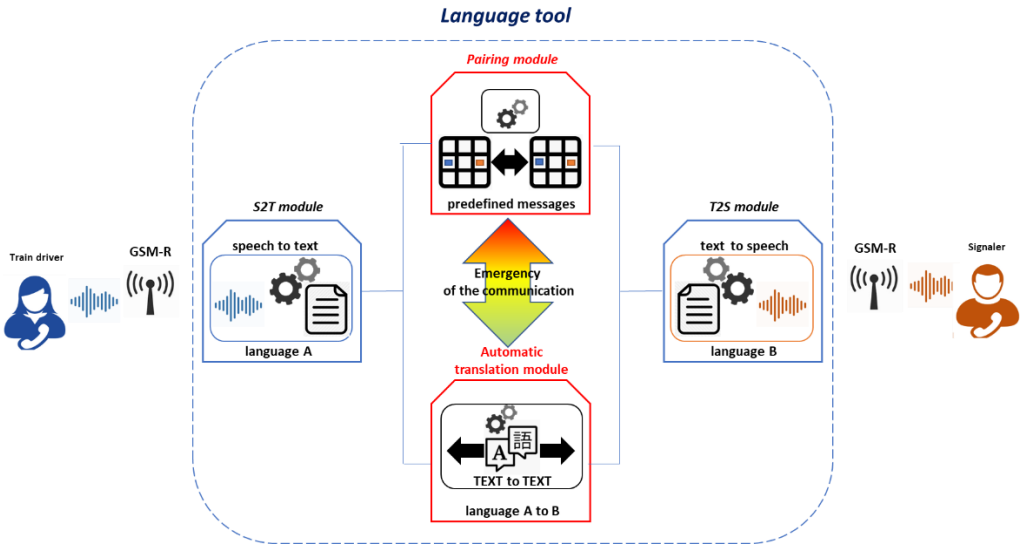


Figure 3 The Language tool schema

7.2. Google Translate

Google Translate is a free multilingual machine translation service developed by Google, to translate text. It offers a website interface, mobile apps for Android and iOS, and an application programming interface that helps developers build browser extensions and software applications. Google Translate supports over 100 languages at various levels, claimed over 500 million total users, with more than 100 billion words translated daily. During a translation, it looks for patterns in millions of documents to help decide on which words to choose and how to arrange them in the target language.
In November 2016, Google announced that Google Translate would switch to a neural machine translation engine - Google Neural Machine Translation (GNMT) - **which translates "whole sentences at a time, rather than just piece by piece.** It uses this broader context to help it figure out the most relevant translation, which it then rearranges and adjusts to be more like a human speaking with proper grammar". GNMT is used in all 103 languages in Google Translate.

Functions
Google Translate can translate multiple forms of text and media, which includes text, speech, and text within still or moving images. Specifically, its functions include:

- **Written Words Translation** – A function that translates written words or text to a foreign language;
- **Website Translation** – A function that translates a whole webpage to selected languages;

- **Document translation** – A function that translates a document uploaded by the users to selected languages. The documents should be in the form of: .doc, .docx, .odf, .pdf, .ppt, .pptx, .ps, .rtf, .txt, .xls, .xlsx;
- **Speech Translation –** A function that instantly translates spoken language into the selected foreign language;
- **Mobile App Translation –** In 2018, Google Translate has introduced its new feature called "Tap to Translate," which made instant translation accessible inside any apps without exiting or switching it;
- **Image Translation –** A function that identifies text in a picture taken by the users and translates text on the screen instantly by images;
- **Handwritten Translation –** A function that translates language that is handwritten on the phone screen or drawn on a virtual keyboard without the support of keyboard.

For most of its features, Google Translate provides the pronunciation, dictionary, and listen to the translation. Additionally, Google Translate has introduced its own Translate app, so the translation is available with mobile phone in offline mode.

Table 20 Google Translate vs. "a T4R tool"

What is the main characteristic of the translation tool	What can be used for T4R translation tool
Translate whole sentences at a time, rather than just piece by piece	Translation of full sentence
Support over 100 languages at various level	EU languages (IMs / RNE´s members languages) are covered

How the tool fits T4R requirement	
No additional equipment for GSM-R users	●●●●●
Set up for predefined messages	●●●●○
Support IM languages	●●●●●
Possibility to connect into GSM-R	●●○○○
Automatic translation	●●●●●
Safety level	●●●○○

Google Translate [17]

7.3. Microsoft

In 2007, the Microsoft translator was created by Microsoft for internal use. Four years later, in 2011, the tool became available on the market for text translation. The tool keeps developing and the neural machine translation technology was the news. In recent years, also speech translation has become more developed.

When speech translation is being translated, the machine will first recognise the language in written form. As a second step, it will correct possible grammar/wording mistakes done during recognition. Thirdly, the written text will be translated in written form to the wanted language. As the last step, the translation will be read out orally.

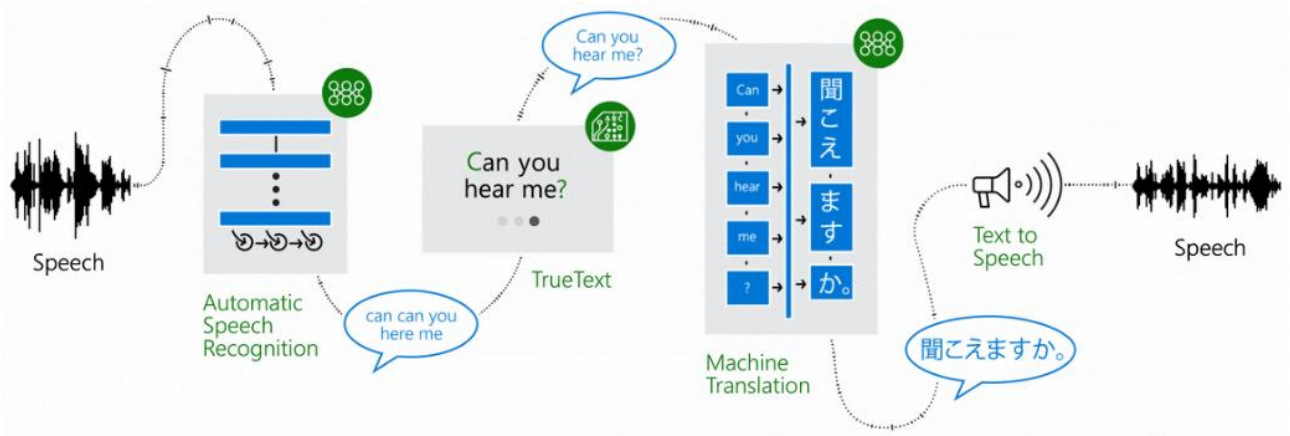


Figure 4 Microsoft Translator – speech translation

When using speech translation, it is important to teach the tool to exclude noises, which is common within the railway industry. To receive the best translations, later on, it is recommended to collect much data (oral communication) for each language where it is noisy.

- **Skype Translator**

Skype Translator is a speech to speech translation application developed by Skype, which has operated as a division of Microsoft since 2018. Skype Translator Preview has been publicly available since December 15, 2015. Skype Translator is available as a standalone app and, as of October 2015, is integrated into the Skype for Windows desktop application.

Skype Translator was built on developments in deep neural networks for speech recognition and Microsoft Translator's statistical machine translation technology. Users converse in their native languages, and the speech is translated from one language to the other in “near real-time”, with the output translation presented by computer-generated speech synthesis. The current version supports speech translation to and from English, French, German, Chinese (Mandarin), Italian, Spanish, Portuguese, Arabic, Japanese and Russian.

The application also features on-screen text transcripts of the spoken phrases in their original language along with the translation. In addition, Skype Translator supports Skype Instant Message, which can translate users' instant messages into more than 60 languages supported by Microsoft Translator.

Table 21 Microsoft Translate vs. "a T4R tool"

What is the main characteristic of the translation tool	What can be used for T4R translation tool
Multiple datacentres around the world. Excellent covering	Cloud app.
Good customisation (provide API, customer build own tool)	Creation and usage of a database with railway jargon

How the tool fits to T4R requirement	
No additional equipment for GSM-R users	●●●●●
Set up for predefined messages	●●●●●
Support IM languages	●●●●●
Possibility to connect into GSM-R	●●○○○
Automatic translation	○○○○○
Safety level	●●●○○

Microsoft Translator [18]

7.4. IBM

Speech to speech translation
A Real-time and Mobile Solution to Mitigate Language Barriers

IBM TRANSTAC speech-to-speech translation systems that address various issues arising from dealing with resource-constrained tasks, which include both limited amounts of linguistic resources and training data, as well as limited computational power on mobile platforms such as smartphones. [19]
Audibly and visually translates between two languages:

- 2-way translation of "free form" conversational speech;
- Target for instantaneous & highly accurate Speech-to-Speech (S2S) translation on mobile devices (handhelds, smartphone etc.);
- No need for server connectivity;
- Robust & speaker-independent large vocabulary speech recognition, accommodating differences in tone and accent with online speaker adaptation;
- Data-driven statistical machine translation Machine learning techniques ubiquitously applied, which enables rapid development for new languages & domains.

Figure 5 IBM – speech translation

Table 22 IBM – Designed Domains (example of the areas)

English – Chinese	Englisch – Arabic
Travel and Tourism	Business (Rental service, Advertisement, Tax, Import and Export)
Airport (Customs, Duty-free goods, Visa, Money Transfer)	Job interview – Introduction and greeting
Hotel reservation – Dining, Food and Restaurant	Health and Medical Care (Diagnose, Pharmacy)
Culture and Entertainment	Sports (Balls, Aquatic, Track, Olympic Games and Venues,…)
Traffic and Directions	Auto and Ground Transportation
Sightseeing and Shopping	…more…

Table 23 IBM translation vs. "a T4R tool"

What is the main characteristic of the translation tool	What can be used for the T4R translation tool
2-way translation of "free form" conversational speech	"free form" conversational speech (group call, e.g. in an emergency situation should be solved)
No need for server connectivity	No (additional) data needed

How the tool fits the T4R requirement	
No additional equipment for GSM-R users	●●●●●
Set up for predefined messages	●●●●○
Support IM languages	●●●●○
Possibility to connect into GSM-R	●○○○○
Automatic translation	●●●●●
Safety level	●●●○○

IBM translation [20]

7.5. Baidu

In Beijing, the company unveiled an A.I.-enabled tool it claims **can translate English into Chinese and German almost instantly.** Baidu calls its interpretation technology STACL, short for "Simultaneous Translation with Anticipation and Controllable Latency." As the name (sort of) implies, Baidu believes that what distinguishes its system from existing online translation services is a combination of speed and flexibility.

While products like Google Translate allow users to say or write a sentence and receive a translation after a lag, Baidu claims its new translation tool allows for sentences to be deciphered **in real-time, matching the best human interpreters.** Users dealing with closely related languages, like French and Spanish, can choose to start translation after a single word, while those dealing with unrelated languages, like Chinese and English, can instruct the system to wait longer to improve accuracy.

Table 24 Baidu translator vs. "a T4R tool"

What is the main characteristic of the translation tool	What can be used for the T4R translation tool
Sentences deciphered in real-time	Possibility to use the tool for translation of variables
Baidu prefers sites that are hosted in China	Need to specify the maintaining processes

How the tool fits the T4R requirement	
No additional equipment for GSM-R users	●●●●●
Set up for predefined messages	●●●○○
Support IM languages	●●●○○
Possibility to connect into GSM-R	●●○○○
Automatic translation	●●●●●
Safety level	●●●○○

Baidu translator [21]

7.6. Carnegie Mellon University

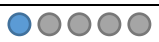
CMUSphinx collects over 20 years of the CMU research. The main advantages are:

- State of art speech recognition algorithms for efficient speech recognition. CMUSphinx tools are designed specifically for low-resource platforms;
- Flexible design;
- Focus on practical application development and not on research;
- Support for several languages like US English, UK English, French, Mandarin, German, Dutch, Russian and ability to build a model for others;
- BSD-like license which allows commercial distribution;

- Commercial support;
- Active development and release schedule;
- Active community (more than 400 users on Linkedin CMUSphinx group);
- Wide range of tools for many speech-recognition related purposes (keyword spotting, alignment, pronunciation evaluation).

Table 25 CMU translator vs. "a T4R tool"

What is the main characteristic of the translation tool	What can be used for the T4R translation tool
Algorithms for efficient speech recognition	Speech recognition is considered for communication between different IMs
Ability to build a model for other languages, which are not directly supported, currently	The database for some languages have to be created

How the tool fits the T4R requirement	
No additional equipment for GSM-R users	
Set up for predefined messages	
Support IM languages	
Possibility to connect into GSM-R	
Automatic translation	
Safety level	

Carnegie Mellon University translator [22]

7.7. Systran

From the early beginning after founded the Systran is focused on automatic translations. During the cold war, Systran customised a tool which was able to translate text from Russian to English. The customisation was done together with the US Air Force.
The Systran working team consist of i.e. linguistics and statistics which are continuously improving the technology. Their solution is available on Linux, Solaris and Microsoft Windows.

Systran refers to language pairs, which are two languages that can be translated into each other. **The automatic translation tool being provided by Systran can be customised with railway words and sentences.** All available data can be fed into the tool before it is being distributed to the customer. To still be able to feed the tool with relevant data, Systran can give access to some users, which would be able to add, change and delete part of the data.

Currently, Systran has no solution for oral communication but works with various partners which are reading audio communication. The oral translation is, comparing to the written translation, quite new technology. Therefore, Systran would today recommend starting with written

translations. Another option would be to speak in a native language and the tool will translate into written translation.

Table 26 Systran translator vs. "a T4R tool"

What is the main characteristic of the translation tool	What can be used for the T4R translation tool
High level of safety/security	Safety level as a key point in railway operation
Available for more than 140 languages, the tool can be customised	Covering IMs languages, the data can be fed into the tool before the tool is used for the first time (predefined messages)

How the tool fits the T4R requirement	
No additional equipment for GSM-R users	●●●●●
Set up for predefined messages	●●●●●
Support IM languages	●●●●●
Possibility to connect into GSM-R	●●○○○
Automatic translation	●●●●●
Safety level	●●●○○

As a next software solution, Systran is offering also the medical translation software, to cover the need to translate medical documents for patient information daily, to speed up the translation process and offer patients documentation in their native language.

Systran offers translation in over 52 language pair combinations including Spanish, French, German, Portuguese, Italian, Swedish, Dutch, Russian, Japanese, Korean, Traditional Chinese and Simplified Chinese.

The most popular software packages, offered by Systran, used by Hospitals are:

- Systran Professional Premium
- Systran Volume Licensing (Premium Version)
- Systran Enterprise Solutions (Premium Enterprise Version)
- Systran Links – Translate the website for patients and staff

Table 27 Systran Medical translation SW vs. "a T4R tool"

What is the main characteristic of the translation tool	What can be used for the T4R translation tool
High level of safety/security (as an SW is being used in the health sector)	Safety level as a key point in railway operation
Available for more than 52 language pairs	Covering a majority of IMs languages, the data can be fed into the tool before the tool is used for the first time (predefined messages)

How the tool fits the T4R requirement	
No additional equipment for GSM-R users	●●●●●
Set up for predefined messages	●●●●●
Support IM languages	●●●●○
Possibility to connect into GSM-R	●●○○○
Automatic translation	●●●●●
Safety level	●●●○○

Systran: TranslationSoftware4u [23]

7.8. DeepL

DeepL is a tool which was developed from the machine dictionary Linguee in August 2017. Linguee was launched in 2009 and already from start, the goal was to only focus on a few languages to make sure the translations were accurate. In 2016, the company decided to develop a neural network system which would be able to translate longer texts. The name of the company "DeepL" corresponds to what the machine is focusing on; **deep learning research.**

At the moment, DeepL only provides the customer with an API. The customer has to have/create their own solution from where the API will do its translations. DeepL is working on an additional tool where the customers will be able to feed the tool with specific data.

Up until now, DeepL has focused on written translation for seven languages. Nevertheless, the possibility to do speech translation will be available in the future, since there are companies working on it at the moment. As a result of this, for the moment DeepL's translation tool would be used together with a speech tool and cannot translate oral communication by itself.

DeepL have two different solutions/tools to offer to the customers; on free version or DeepL Pro. For the free version, all text and information entered by the users are being stored to train the machine. When using DeepL Pro, no data is being stored. By using the REST API, the data will be securely translated.

Table 28 DeepL translator vs. "a T4R tool"

What is the main characteristic of the translation tool	**What can be used for T4R translation tool**
Provides the customers with an app., need internet access – cannot be used in offline mode	The app. can be downloaded and used in offline mode (when users require dictionaries where there is no Internet connection)
Two different solution – a free version and Pro version (pay for), but with the database from DeepL, current customers are not able to feed the tool with own messages (this is in progress)	Free version could be used in Pilot testing (if the DeepL could be able to feed the database with predefined messages from customer´s side)

How the tool fits the T4R requirement	
No additional equipment for GSM-R users	
Set up for predefined messages	
Support IM languages	
Possibility to connect into GSM-R	
Automatic translation	
Safety level	

7.9. iTranslate

iTranslate is an application, which was launched in 2009. The application is available for different kind of stakeholders, where travellers, business professionals and students are some of them.

iTranslate **can translate text, websites and oral communication into more than 100 languages.** Available application is a general application which has the same data/vocabularies for all users. iTranslate is working on a business solution where the customer will be able to create its own vocabulary (available approx. from 2020).

With iTranslate both oral and written translation is possible. Additional to these two translation methods, there is a possibility to take photos with iTranslate Lens, which will translate the text in a photograph.

When using the speech translation, the user needs to press somewhere on the mobile screen and the say the messages. The speech will first be recognized in written form, translated into text and then pronounced in the translated language.

Table 29 iTranslate vs. “a T4R tool”

What is the main characteristic of the translation tool	What can be used for the T4R translation tool
Translation into more than 100 languages, possible for oral and written form, too. The speech will first be recognised in written form, translated into text and then pronounced in the translated language	Speech recognition, S2T, T2T, T2S modules. Open point it the translation of variables
Customers can create its own vocabulary; another customisation is not possible	Could be used for the creation of a database of predefined messages

How the tool fits the T4R requirement	
No additional equipment for GSM-R users	●●●●●
Set up for predefined messages	●●●●●
Support IM languages	●●●●●
Possibility to connect into GSM-R	●●●○○
Automatic translation	○○○○○
Safety level	●●●○○

7.10. SDL Government

SDL Government is a company which was founded in 1992. Their focus is to provide language translation solutions in several areas where security and sensitive information needs to be kept within the organisation of the customer. The company is specified in translating informal and social media languages but has customers within other areas as well. The mission of the company is to remove language as a barrier to communication.

The company has been focusing on translations to and from English. Recently, other language pairs have been created. There are different levels of customisation. The first step is to add customer's dictionary in all possible languages, even for those language pairs not currently supported by the SDL Government. If this step is not enough for the customers, in order to have proper translations, the company can customise the engine itself.

SDL Government has its focus on written translation but collaborates with partners, which are specialised in text-to-speech and speech-to-text technology.

Figure 6 SDL Government translation schema

Table 30 SDL Government translation vs. “a T4R tool”

What is the main characteristic of the translation tool	What can be used for the T4R translation tool
The tool works with various partners – can only do text-to-text (only this module), cooperation with a text-to-speech, speech-to-text technology companies are needed	Using only one module in a combination with other providers
Focusing on translations to and from English	Other IMs languages could be communicated to be added in the future. This will need additional time for developing “a T4R tool”

How the tool fits the T4R requirement	
No additional equipment for GSM-R users	●●●●●
Set up for predefined messages	●●○○○
Support IM languages	●○○○○
Possibility to connect into GSM-R	○○○○○
Automatic translation	○○○○○
Safety level	●●●○○

7.11. KantanMT

KantanMT is an automation translation tool provider which has its focus on written translation. They believe in customisation and have developed a feature, where the customer can create its own translation platform on KantanMT’s website. Therefore, the company can only provide text-to-text translations.
Currently, the company is providing translations into 80 languages in their standard tool. These 80 languages are both from countries within Europe but also outside.
The company is encouraging its customers to customise their own tools. The customers will feed the tool with their own phrases and vocabularies for better and more precise translations.

The provider can offer two translation solutions; either cloud-based or on-premises. Both these options are well secured. Nevertheless, the big customers have chosen to have an on-premises solution to be absolutely sure no data is being leaked outside of the company.

Table 31 KantanMT translation vs. “a T4R tool”

What is the main characteristic of the translation tool	What can be used for the T4R translation tool
Automation translation tool focused on written translation	Provide the company with a text file. The company use the training algorithm to feed the tool
Cloud and on-premises solution, can be used in offline mode	When used in offline mode, the users will have to read what has been written – translated

How the tool fits the T4R requirement	
No additional equipment for GSM-R users	🔵🔵🔵🔵🔵
Set up for predefined messages	🔵🔵🔵🔵⚪
Support IM languages	🔵🔵🔵🔵⚪
Possibility to connect into GSM-R	⚪⚪⚪⚪⚪
Automatic translation	🔵🔵⚪⚪⚪
Safety level	🔵🔵🔵⚪⚪

7.12. List of speech to speech applications

Just simple applications, that the smartphone can become a personal translator. This topic is mostly focused to describe the application of the specific translation tool.

1. Google Translate

Google Translate is **one of the best-known translators** due to its versatility and simplicity. And you can access even more features in Google Translate's mobile app. **The service supports over 100 languages for translations of typed or handwritten text.**

When you have to communicate with someone who speaks another language, forget awkward hand gestures and pointing—the Google Translate app does the work for you. As you talk into the app, it repeats what you say in the language of your choice. When your conversation partner responds, the app translates it back to you. In the offline mode is also possible, with the translation of 59 languages.

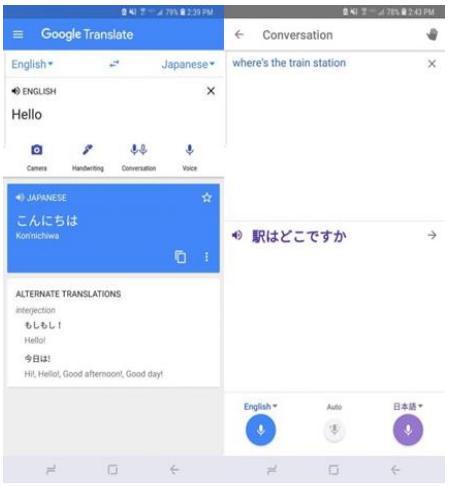


Figure 7 Google Translate

2. Microsoft Translator

Microsoft Translator is another big name in translation apps and offers an even simpler interface than Google Translate. Four big bubbles float on the home screen of the app.
Their functions are self-explanatory:

- voice translation;
- conversation translation;
- photo translation;
- text translation.

Microsoft Translator supports fewer languages than Google Translate. What makes Microsoft Translator so innovative is its ability to translate your message for a large number of listeners, making it perfect for a presentation.

Figure 8 Microsoft Translator

How does it work? After you start up conversation mode, you'll receive a code. Give this code to your listeners, and they can join your conversation. As you talk or type into the app, your listeners will see a translation of your speech in their own language.

3. Reverso Translation Dictionary

Reverso is definitely limited when it comes to supported languages (it only works with 11), but it provides more detailed translation information than other apps.

How does it work? Choose a language and type a word into the search bar. On the results page, the word is presented in the language you choose. It provides several definitions that show how to use the word in different contexts as well. Reverso is also a language learning app, as it helps you memorise and translate words.

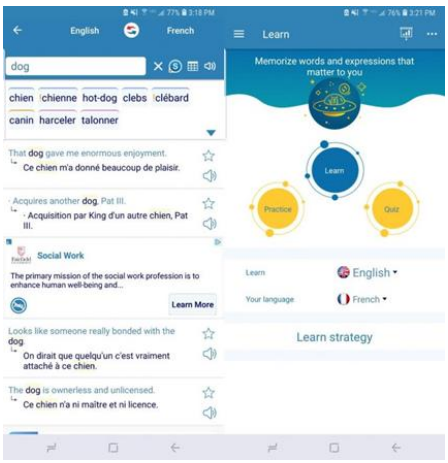


Figure 9 Reverso Translation Dictionary

4. Naver Papago Translator

When you open Papago, you'll notice an area to type the text you want to translate, along with three colourful bars that represent your translation options. **It supports 13 languages that you can translate via text, voice, and photo.**
With Papago Mini, the translation parallel with surfing on the web is possible. A tiny box along with a parrot icon will appear in the corner of your browser.

In addition to all these features, Papago also stores your frequently used phrases, gives you access to a dictionary, and translates entire websites.

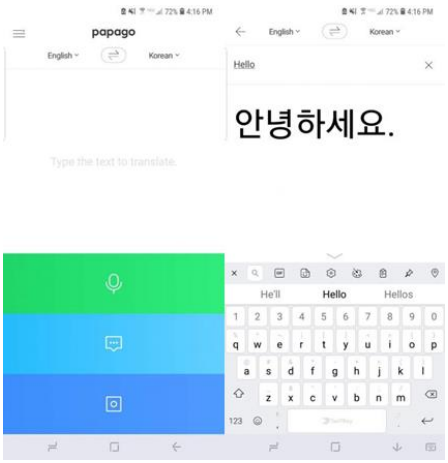


Figure 10 Naver Papago Translator

5. SayHi Translate

SayHi is made specifically for translating speech and text, which means it's a bit limited in other features. However, it does support a variety of languages and a large number of Spanish dialects.
This app makes it easy to have a conversation in another language. Tap the microphone icon to speak into the phone or hold the icon to open the keyboard. The speech or text gets translated instantly and appears on the screen as two bubbles – the speaker's words are on top, while the translation is on the bottom. It is possible to share the application to Twitter, Facebook, or SMS.

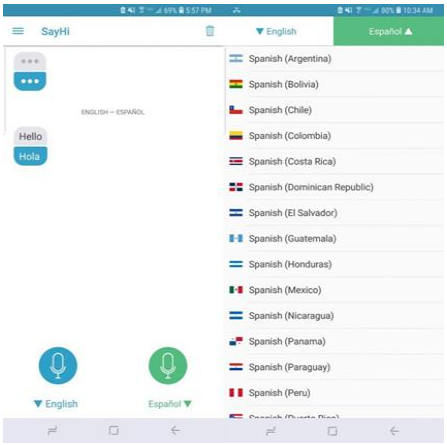


Figure 11 SayHi Translate

6. Dictionary Linguee

It's a reliable translation tool that puts foreign words into context. **The app supports several languages and allows you to conduct bidirectional searches.** This means that you can search for translations in both English to Chinese or Chinese to English (for example). As soon as you type a word into the search bar, Dictionary Linguee presents you with a definition. This comes from its editorial dictionary assembled by a team of lexicographers, so you know you can trust it. The results from the translation search engine can be found below the editorial definition. These examples show you how the word is used in online translations.

Figure 12 Dictionary Linguee

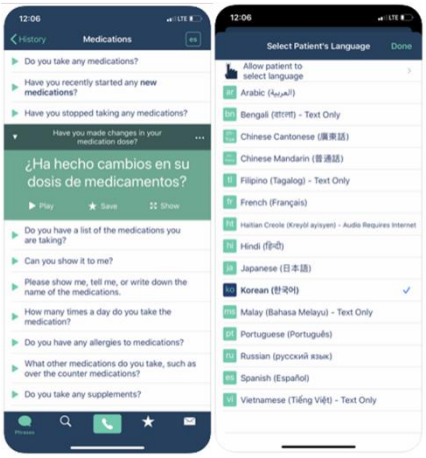
7. Care To Translate

Care to Translate offers medically correct communication with translations verified by native speakers. It can be used in acute situations or in daily work at the ward. Care to Translate serves as a complement to human interpreters in healthcare and is **available 24/7 in the smart device.** This tool can be used in any situation when something needs to be communicated fast, regardless whether it is in an acute situation or in daily work at the ward and offer a medically correct communication with translations in the following languages: Albanian; Arabic; Bosnian/Croatian/Serbian; Bulgarian; Chinese (Mandarin); Dari; English; Finnish; French; German; Greek; Hungarian; Italian; Kurmanji; Lule Sami; Norwegian; Pashto; Persian/Farsi; Polish; Portuguese; Romanian; Russian; Somali; Sorani; Spanish; Swedish; Tigrinya and Turkish.

Figure 13 Care To Translate

Care to Translate reduces the risk of misdiagnosis, malpractice, complications, readmissions and hospitalization time by providing safer and more accurate communication between patients and healthcare providers.

8. Canopy Speak



Canopy Speak is a Free Medical Translation Mobile App available in over 15 languages, including Spanish, Mandarin, Cantonese, French, Arabic, Japanese, Korean, Russian, Portuguese, etc. It was designed to overcome logistical difficulties inherent in accessing current language assistance tools. Canopy Speak enables clinicians to communicate freely with Limited-English Proficiency patients for common, straightforward, routine communication. **The one-touch interpreter call button connects the provider directly to the facility's over-the-phone interpreter line.** The phrase library consists of pre-translated medical phrases, organized by frequently encountered procedures and medical specialities.

Figure 14 Canopy Speak

9. Universal Doctor Speaker

It is the multilingual medical translation tools to breakdown language barriers in healthcare. Universal Doctor Speaker Translator can be customised for doctors, nurses, pharmacists and midwives. Universal Doctor Speaker is a multilingual application providing key medical phrases translated across 17 languages with audios to facilitate communication between patients and healthcare professionals. The app can be used for the following languages: Arabic; Somali; English; Spanish; French; German; Chinese Mandarin; Simplified Japanese; Moroccan Arabic; Portuguese; Catalan; Russian; Romanian; Polish; Italian; Norwegian; Portuguese (Brazil). All the translations and audios are on a device for offline use. Universal Doctor Speaker comes with over 500 commonly used medical phrases with accompanying audios by native speakers.

Figure 15 Universal Doctor Speaker

10. Pairaphrase Translation Software

Pairaphrase has a secure, confidential File Translator built-in that allows you to upload a medical document (including a scanned document). Based on Dynamic Machine Learning technology, the software can save the improvements done by using the translator and automatically reuse them, when the system identifies repetition of that same text segment across other documents, what can reduce translation time significantly.

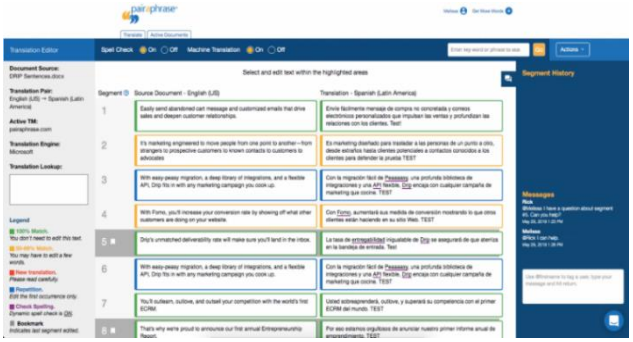


Figure 16 Pairaphrase SW

Pairaphrase can memorize the medical terms and phrases that are post-edited. In case, the translator needs to edit a specific medical phrase, so that it makes more sense in the target language, the system will remember this correction. In effect, in the future translations, the phrase will be remembered, and there is no need to make that correction again for any file. Pairaphrase includes a text-to-speech translation feature that allows to securely upload text and receive a spoken translation available for download in .wav or .mp3 format, in a female or male voice. These synthetic voiceovers are available in 18 different languages. The document translation process is the following:

- Select the source language and target language;
- Upload a file or a batch of files;
- Receive a “first draft” translation from the software;
- Invite a certified medical translator into the software to post-edit the text (improve the accuracy). After all, machine translation is far from perfect.

Speech to speech applications (described in the report) [24]

7.13. List of translation devices

A separate chapter in this benchmarking report is translation devices, which could be called **"Pocket foreign language translation devices".**

These devices can be used in many situations when you want to travel, educate yourself, do the business or work.

In this chapter, there are some translation devices listed. Still, for the purpose of the Translation4Rail project, it is not possible to be used, according to the idea of no additional equipment in the communication of railways´ staffs.

Example of "Pocket devices"

- ***Travis Touch***
  105 languages are available for online translating and 15 languages are supported without an Internet connection (limited translations; basic words, and phrases). Travis, with the help of AI technology, uses 14 different translation engines, including Google Translate, to provide the most accurate translations for each language. Its dual noise-cancelling microphones and high-quality speakers allow translating outdoors, on the move or in noisy environments.

Figure 17 Travis Touch

- ***Pocketalk***
  74 languages at hand. Pocketalk is able to translate long sentences, therefore it is suitable for travelling, face-to-face conversations. Due to the dual noise-cancelling microphones and speakers, outdoor translations are possible even in a loud environment. The translating history stores up to 10,000 past translations and they are accessible in the timeline. There are two options to purchase the device: with the built-in mobile data or without. Pocketalk doesn't support offline translations at all, meaning that the device has to be connected to the Internet at all times in order to work.

Figure 18 Pocketalk

- ***ili***

  ili is an offline translation device. It supports three languages: Japanese, Mandarin, and Spanish. A beta version of the Korean language has been recently added. The most distinctive feature of ili is that it offers only one-way translations.

Figure 19 ili handy translator

- ***WT2 Plus***

  Two wireless headphones come in the charging case to be shared between two speakers. The earbuds are connected to the Internet via a mobile app and require a phone to be present at all times. WT2 Plus claims to have three translation modes available – Simul Mode, Touch Mode and Speak Mode. Each mode is designed to optimize the ways of translating on different occasions or environments. WT2 Plus supports 36 languages together with 13 English accents, 15 Arabic accents, and 18 Spanish accents.

Figure 20 WT2 plus

- ***Muama Enence***

  Requires an additional app and phone in order to translate (connected via Bluetooth). The device itself functions as a glorified microphone, all the translations and processing are done by your phone and the app. This allows the device to have a longer battery life, but you will find yourself fumbling with a phone, an app and the translator device, all while trying to have a conversation with someone.

Figure 21 Muama Enence

List of translation devices [25]

7.14. The Benchmark of the Translation tool and providers

The benchmark of the most comparable translation tool related to the objectives of Translate4Rail.

The proposal for the T4R tool concentrates on the requirements MVP (minimum viable product) for such a Pilot with the following priorities:

- **Integrates the most important functions for a translation tool** (functional experience);
- **Integrates easily into existing IM infrastructure and operational rules** (easiness of pilot participation);
- **Omits costly development without critical added value for the Pilot** (probability of cost management).

Table 32 The Benchmark of the Translation tool providers

Company	Translation machine	Customisation	Languages	Feeding possibility	Extra	Usability for "a T4R tool"
Google Translate	Neural MT	Yes	Support over 100 languages at various level	-	Translate whole sentences at a time, rather than just piece by piece	●●●●○
Microsoft	Neural MT	Yes	10 languages	-	The first step is to recognise the language in a written form	●●●○○
Skype Translator (operated as a division of Microsoft)	Neural MT	Yes	The current version supports speech translation to and from English, French, German, Chinese (Mandarin), Italian, Spanish, Portuguese, Arabic, Japanese and Russian	-	Is integrated into the Skype for Windows desktop app	●●●○○
IBM	Statistical MT	-	Multilingual (database of words depends on chosen language)	-	Audibly and visually translates between two languages, need mobile devices	●●●●○
Baidu	Neural MT	No	Multilingual (English, Chinese, German, other languages could be supported)	-	Baidu prefers sites that are hosted in China	●●●○○
CMU	-	-	Support for several languages like US English, UK English, French, Mandarin, German, Dutch, Russian and ability to build a model for others	Possibility to create a model for new language	Focus on practical application development and not on research	●●○○○
Systran	Neural MT	-	More than 140 languages	Feeding also possible with oral words	Some kind of computer is needed when using this tool	●●●●○
Systran (healt sector)	New Neural MT (PNMT™)	Yes	More than 52 language pairs	Learning from repetitive information - many documents is often a template that is reused	Translate full documents	●●●●○
DeepL	Neural MT, Artificial Inteligence	-	7 languages (English, German, French, Spanish, Dutch, Italian, Polish). Main languages in Linguee. Four additional languages from 2018: Portugese, Japanese, Chinese, Russian	-	Need to have additional eqiupment	●●○○○
iTranslate	-	Yes	More than 100 languages	-	The speech will first be recognized in written form, translated into text and then pronounced in the translated language	●●●○○
SDL Government	Statistical MT and custom language pairs	Yes	Mostly native English-speaking people which can translate into other languages. Custom language pairs (hundreds of language pairs). Over 100 lanuage combinations	SDL can implement centralized translation memories and terminology databases	Can do different kind of translation (text, website, oral - working with partners..)	●○○○○
KantanMT	MT	Yes	Use language pairs and most pairs are available with English or German	Need to use translation memory interchange files (TMX) as training data. Terminology interchange files (TBX) for training terminolgy	Customers can build their own tool on the website	●●○○○

Hereabove benchmarked translation tool/providers represent the usability of the existing tools, according to the Translate4Rail objectives for translation tool development and testing. The Translate4Rail research project will try to apply the most relevant characteristics of the benchmark in the definition of requirements for the digital solution.

Table 33 The Benchmark of the Translation tool providers vs. T4R “a translation tool”

Translation tool	Speech recognition	S2S (e.g. application)	S2T / T2T / T2S	SW solution - possible to be implemented into the railway network "directly"[1]	Information for customers via "any individual device"
Translate4Rail "a translation tool"	**Yes**	**Yes**	**Yes**	**Yes**	**No**
Google Translate	Yes	Yes	Yes	Yes	No
Microsoft	Yes	No	Yes	Yes	No
IBM	Yes	Yes	Yes	Yes	No
Baidu	Yes	Yes	-	Yes	No
CMU	Yes	Yes	-	Yes	No
Systran	Yes	Yes	Yes	Yes	No
DeepL	-	Yes	Yes	-	No
iTranslate	-	Yes	Yes	Yes	No
SDL Government	No	No	Yes	-	No
KantanMT	Yes	No	Yes	Yes	No
Care To Translate	Yes	Yes	Yes	Yes	No
Canopy Speak	Yes	Yes	Yes	Yes	No
Travis, Pocketalk, ili,…	Yes	Yes	Yes	No	Yes

Note:

S2S *"Speech to speech" automatic translation*

S2T *"Speech to text" translation*

T2T *"Text to text" translation*

T2S *"Text to speech" translation*

1 *No bilateral/multilateral agreement or licenses are taken into consideration, in this case of the benchmark*

8. CONCLUSIONS

The conclusions reached and highlighted in this benchmarking report are the input to the next follow-up of the Translation4Rail project.
The Benchmark shows that in the same vein as the railway sector, several other sectors are confronted and dealing with similar issues. The reducing of language barriers could bring new opportunities for a lot of sectors, not only transportation.

The results from the benchmark are influenced by these factors:

- Tools were evaluated according to the facts published on the Internet;
- The physical test has not been the subject of the benchmark;
- Any of the tools has not published a quality benchmark according to the standardized methodology.

The aspect of language pairs

- One of the languages is English
  - Some sectors agreed on using the English language as a common one. In this case, the translation is done between the English language and the other one. Several tools are also used to train users in the English language and learn pre-defined phrases typical for the sector and professional communication.

- Two non-English languages
  - This is the most usual cause in the railway sector. The set of offered language pairs is reduced, focused mainly of world languages not covering all European language variations necessary for seamless interoperable train driving.

The aspect of translation quality

The translation quality is one of the most important factors especially for the sector with high safety demands. Several translation-quality methodologies are defined but not used and published with offered products. Without a commonly agreed approach how to rate the translation engines it is challenging to compare the products and choose the most accurate one fitting the project needs. To partly tackle this fact, we can stick to the pre-defined translations using already translated and approved sentence pairs. The situation starting to be more complex in case that such paired translations using variable words represented by values, names, etc. The free speech module can help on it, translating the variables in the sentence pairs.

The aspect of the translation module and dictionaries

Generally, two main possibilities of benchmarked tools have been recognized. The first one is fully autonomous, and all necessary modules and dictionaries are part of the closed language tool without the need for permanent connection to the Internet. Advantages of such tools is that it can be personalized and work without any need of internet connection. The quality of translation strictly depends on the tool itself. The second possibility is the tool connection to the network using its all features where the translation can be done on high-performance servers with powerful artificial intelligence for self-improving. The requests to the tool performance are minimalised

what is balanced with the need for network connectivity. Other possibilities can be implemented as a combination of mentioned one putting the priority based on required functionality.

Best achieved performance in different sectors

Based on the published facts about existing translation tools we can estimate the best performance of the benchmarked tools/applications in the sectors as the following:

Railway

- The railway sector has not offered a language solution which could be used within the project. Communication in operational situations is covered by international and national rules. To simplify and unify the rules, the idea has been partly implemented via a common set of pre-defined operational messages and dialogues. This approach should be followed by a gradual development of language tools fitting the needs of the sector keeping high safety level.
- To enhance and adapt of an existing set of PDMs in parallel of language tool development should be the key point of the T4R project. This approach establishes a firm place in the sector among industries developing language tools.

Air traffic

- Tools are strictly focused on the English language which was agreed as a common one. The tools development is progressing thanks to the global coverage of the tool and attractivity of the sector. The project could benefit from the adaptation of existing tools using also non-English pairs of languages. As the existing tools are not presenting this case it is difficult to assess the possible quality.
- The projects show us that automatic speech recognition, which converts human speech into texts is crucial for every translation tool. Only the correct recognised speech, transferred into the text can be subsequently evaluated and confirmed for the translation. This voice recognition module seems to be the part which the T4R project could benefit from.
- Air transport is considered as one of the safest. We assume that also for speech recognition, the machine learning continuously improves the quality of this part of the translation/language tool, as well as for translation and text to speech part in other projects. For that reason, the developer would be asked to implement machine learning (any other way of improving the quality of the tool) to reduce the SW programming.

Maritime traffic

- In the maritime sector, there is presented a project aimed to implement English as a “bridge language of the sea”.
- Besides that, presented maritime tools are more or less oriented to the computer dialogue to improve language education, what currently does not fit to the aim of the T4R project.

Military

- As a result, in the military sector, the fully operational speech to speech system on a portable device able to work autonomously is the solution here. The device is also able to react to user gestures.
- The text generated from the input message is being normalised. It means it takes care of numbers, converting them to dates, years and amounts. The system has also to deal with abbreviations, acronyms, etc.
- The ability of the T4R translation/language tool to deal with numbers and abbreviations would be very useful in the railway operation. The T4R project could benefit from the military solution which

seems to be more advance, in case that it is allowed to implement publicly and the bidder with that tool will be successful.

Automotive

- These types of tools are determined by their ability to work also offline what should be one of the T4R project requirements. One of the main advantages is removing the noise from the microphone input and send out a cleaner signal.
- For the T4R project, the function of the noise removing/reduction can be taken into account, especially in the train driver cabin as well as an offline solution of the tool.

Health sector

- Tools are close to the railway sector with language pairs and used predefined phrases. The safety aspect can be an issue as the presented tools in the benchmark are not presenting the results from the real emergency cases.
- The health sector is one of the sectors, where specific tools, as well as commonly available translators, are being used (can be adapted with the specific phrases/words). The advanced capabilities of the tools in this sector are the ability to translate even complete documents (here, the tool can learn and remember often used medical phrases).
- The approach of the PDMs was chosen for the T4R translation/language tool with implementing the specific railway jargon into the tool database.

Construction sector

- Generally, several of the most used translation applications are well designed also to be used in the construction industry. Therefore, activities of the translation in the construction industry are more-less oriented to professional translation services, also for the reason of industry-specific language.
- Combination of the technology and human translation is not foreseen as a part of the T4R translation/language tool but can be used in the testing phase as a possible safety level.

General language projects

- Projects, identified in this Benchmark as the general, were realised to overcome the language barrier by developing e.g. mobile device (personalised STS translation) and created the technological infrastructure.
- The T4R tool aims to use a neutral voice to ensure better understanding.

All tools being benchmarked are more or less similar in the modularity when speech to text-translation text to speech modules are used for translation. Taking into the account the presented information is not convincing enough to choose the right tool which could be adapted easily and reliably to railway sector needs.
According to the benchmark results related to the best performance and possible connection with a T4R translation/language tool, the most suitable solution is to develop own tool based on the requirements defined for the tender, taking also into the accounts the recommendation of this benchmark. It seems to be for a benefit that the tools are using different modules from between speech to speech chain and can be combined choosing the most appropriate one for a given language. It is recommended to address existing key providers as well as other possible developers, including SME to enable them to combine the best modules to get the sector language tool with required performance and quality.

Based on the benchmark we can state that:

- The market offers many translation tools tailor-made for the purpose of their use, unfortunately not exactly developed for railway sector;
- The market offers several solutions and approaches which should be used in the tool definition;
- The translation technologies are progressing with huge steps and it is very probable that the implementation of new technologies can significantly improve the translation quality and reached the setup standards;
- The sector should permanently take part in the translation tool development and play a stable role at this field based on its own developments, solutions and results of testing;
- Removing the language barriers gives a significant competitive advantage to the leader sector;
- The benchmarked tools/applications could not be directly used to cover all railway specifications. All technical and operational requirements related to the T4R translation/language tool need to be defined;
- The tools offer many features which are a good base to learn from them and prepare a description of sector tailor-made requirements and specific features for sector language tool;
- Presented information about the tool are general and are not presenting the testing result based on a common methodology. For the implementation within the project, the possibility for the use of the existing/benchmarked tools needs to be analysed and tested.
- The detailed analysis of the existing tools would require a separate "new" project, where the physical tests of the tools are a part of the project and the results would by detailed compared with the railway sector demands and needs based on multicriterial analyses; Generally, the translation from/to English done by the benchmarked tools/application seems to be sufficient for daily use. To make the tool fit the specification of the relevant industry, the tool should be fed with the specific jargon (e.g. tool for medical purposes) and tested the accuracy of translation with safety assessment;
- An evaluation methodology should be jointly agreed as a standard to enable any tool assessment or its modules to get a clear view of the reached quality;
- Results of the development and testing can contribute to other sector activities related to communication (education, communication with customers, partners);
- That provider can be found on the market to develop a suitable solution to address the issues tackled by T4R.

The main recommendations for the T4R translation tool based on the benchmark are the following:

- Take part in the cooperation with other industries to learn and exchange best practices;
- Use the modules offered by the market to get the required translation quality;
- Learn from the solution which is successful in noise cancellation to reduce the impact of surrounding noise in locomotive cabin and signaller office;
- For the needs of the railway sector, a T4R tool prototype has to be defined from the early beginning. There is no “box full of Lego bricks”, from which we could build “the tool” only “putting bricks together” even off-shelf solution;
- At the beginning be focused on a personalized tool which can reliably recognize its owner voice to minimalize the voice recognition mistakes;
- To start with simple predefined messages and continue with more complex one using variables in the text;
- Enhance the list of predefined messages as it is possible creating clusters for easier application;
- Implement an algorithm which matches the speech with agreed pre-defined messages to get the required safety level;

- The ability of tool training and adapting to specific railway jargon should be necessary property for the railway translation/language tool;
- Try to adapt existing dictionaries to accelerate the tool translation quality and implement an own railway dictionary and reduce the synonyms not relevant to railway sector;
- Train the tool with predefined messages and gradually start using free speech translation;
- Gradually try to take the approach of the pre-defined messages centralisation and distinguish between country-related and Europeans ones;
- Progressively migrate from individual solution to centralised one with self-learning algorithms to improve the translation.

The research carried out in the context of this work package and this specific Benchmark study lead us to the following recommendations:

- The safety aspect is the prerequisite for the entire project;
- The consortium partners which will define the requirements for digital solution should take into account the tool benchmarking assessment detailed in this report;
- Carry out thorough pilot testing for the T4R tool development and improvement;
- Parts of existing translation tools/translators may be inspiration to compose a customised sector language tool;
- To continuously follow the trends in general translation tools development.

9. References

The area of artificial intelligence, speech recognition and speech translation changing very fast. This represents one of the key points to use the internet as one of the main areas for references (web sites of specific companies, magazines, workshops, etc.).

[1] Kei Hashimoto,Junichi Yamagishi,William Byrne,Simon King,Keiichi Tokuda; Speech Communication; Impacts of machine translation and speech synthesis on speech-to-speech translation; ELSEVIER 09/2012

[2] Use of speech recognition in simulators
http://www.fandry.net/pub/GAVIGNET_ANDRY_WHCS94.pdf
http://www.speechtechmag.com/Articles/News/Speech-Technology-Digest/LLoquendo-Automates-Railway-Information-Service-31353.aspx
https://mainichi.jp/english/articles/20180520/p2a/00m/0na/003000c
https://patents.google.com/patent/CN102862587B/en

[3] RNE and UIC sites
http://rne.eu/tm-tpm/language-programme/
https://uic.org/projects/xborder-quick-wins

[4] Activities in the field of speech recognition in the air traffic
https://phys.org/news/2018-09-speech-recognition-technology-air-traffic.html
https://arxiv.org/pdf/1810.12614.pdf
https://www.aclweb.org/anthology/L18-1450
http://www.adacel.com/
http://www.speechtechmag.com/Articles/News/Speech-Technology-News-Features/Speech-Technologies-Take-to-the-Skies-in-Newest-Fighter-Jet-39947.aspx

[5] MALORCA project
http://www.malorca-project.de/

[6] AcListant projects
http://www.aclistant.de/wp/
https://www.dlr.de/fl/en/desktopdefault.aspx/tabid-1150/1741_read-36258/

[7] Activities in the field of speech recognition in the maritime traffic
https://www.idmt.fraunhofer.de/en/Press_and_Media/press_releases/2016/speech_recognition_maritime_traffic.html
https://www.jade-hs.de/en/the-university/departments/maritime-and-logistics-studies/research/vts-bot/

[8] Fraunhofer IDMT: Improved Maritime Communication
https://www.idmt.fraunhofer.de/en/events_and_exhibitions/smm_2018/jcr:content/contentPar/tradefair/teaserParsys/teaser_587047972/linklistParsys/downloadcomponent/file.res/20180625%20SMM%20Folder%202018%20EN.pdf
https://www.idmt.fraunhofer.de/en/Press_and_Media/press_releases/2016/speech_recognition_maritime_traffic.html

[9] Activities in the field of speech recognition in the military sector
https://stephanepigeon.com/Docs/TR-IST-037-ALL.pdf
http://voxtec.com/?page_id=18

[10] MFLTS System
https://www.aclweb.org/anthology/W18-1923
https://slator.com/technology/us-military-equips-soldiers-field-machine-translation/

[11] Activities in the field of speech recognition in Automotive
http://www.speechtechmag.com/Articles/News/Speech-Technology-Digest/SoundHound-Inc.-Unveils-Hybrid-Voice-AI-and-Natural-Language-Understanding-System-for-Cars-130916.aspx
http://www.speechtechmag.com/Articles/Archives/Industry-View/Nuance-Automotive-Powers-Geelys-GKUI-Smart-Ecosystem-131959.aspx

[12] Activities in the field of speech recognition in the Health sector
https://mhealth.jmir.org/2019/4/e11316/

https://www.globalspeaktranslations.com/services/certified-translation-services/healthcare-translations/
https://caretotranslate.com/#whyus
https://www.pairaphrase.com/about-pairaphrase/who-uses-pairaphrase/translation-software-for-healthcare/

[13] Activities in the field of "translation" in the Construction sector
https://optima-translation.com/en/news/translations-construction-industry
https://www.globalspeaktranslations.com/services/certified-translation-services/construction-translations/

[14] Geoffrey Hinton, Li Deng, Dong Yu, George E Dahl, Abdel-rahman Mohamed, Navdeep Jaitly, Andrew Senior, Vincent Vanhoucke, Patrick Nguyen, Tara N Sainath, Brian Kingsbury; IEEE; Signal processing magazine 2012/11; pp 82 – 97

[15] EMIME project
https://cordis.europa.eu/project/rcn/85525/factsheet/en

[16] TC STAR project
https://cordis.europa.eu/project/rcn/71389/factsheet/en

[17] Google Translate
https://en.wikipedia.org/wiki/Google_Translate

[18] Microsoft Translate
https://en.wikipedia.org/wiki/Skype_Translator

[19] Bowen Zhou,Xiaodong Cui,Songfang Huang,Martin Cmejrek,Wei Zhang,Jian Xue,Jia Cui,Bing Xiang,Gregg Daggett,Upendra Chaudhari,Sameer Maskey,Etienne Marcheret; Computer Speech & Language; The IBM speech-to-speech translation system for smartphone: Improvements for resource-constrained tasks; ELSEVIER 02/2013

[20] IBM translation
https://researcher.watson.ibm.com/researcher/view_page.php?id=2323

[21] Baidu translator
http://fortune.com/2018/10/24/baidu-looks-to-amaze-with-a-i-translation-in-real-time/

[22] Carnegie Mellon University translator
https://cmusphinx.github.io/

[23] Systran: TranslationSoftware4u
https://www.translationsoftware4u.com/medical-translation.php

[24] Speech to speech applications (described in the report)
https://www.makeuseof.com/tag/best-mobile-translation-apps/
https://caretotranslate.com/#whyus
https://www.universaldoctor.com/
https://www.pairaphrase.com/about-pairaphrase/who-uses-pairaphrase/translation-software-for-healthcare/

[25] List of translation devices
https://www.travistranslator.com/choosing-the-best-translator-device/