The translation project CASMACAT (Cognitive Analysis and Statistical Methods for Advanced Computer Aided Translation) is just one of the activities in which Dr Michael Carl and his colleagues in Copenhagen are involved.

Their general mission is to establish how foreign language skills are best acquired and how language is used profitably in the globalised world. They also aim to establish how technology can be used to make language processes and the sharing of knowledge about language more efficient.

Background
Integration in Europe and, indeed, globalisation, increases interaction across borders. But while political borders become less significant, there is a danger that the world will remain fractured by linguistic boundaries. Addressing each individual in a language that s/he speaks - ideally his or her native language - requires a huge amount of translation work. There have been significant improvements in machine translation technology in the past decade, but most of this work is targeted towards bulk translation that falls short of the precision often needed.

The productivity of translators can undoubtedly be increased with computer-aided tools. Translation memories are standard in the industry, but post-editing machine translation output is only slowly growing in use. The current integration of machine translation technology into translators' work processes is often done in too simplistic a way, Dr Carl believes. It breaks their work practices, and hence it is widely resisted by translators.

Therefore the CASMACAT project is carrying out an in-depth study of translator behaviour to tailor the tools to the requirements of translators, and not the other way around.

The future
In the not-too-distant future computer-assisted translation (CAT) systems will be developed, which put the translator in front of a computer screen on which a large number of translation options and translation resources will appear. The translator can then choose from the proposals, and compose, edit or maybe even speak a translation.

Translation tools have to better support humans when they produce translations by taking machine translation processes into account and investigate their interaction.
Increasingly, questions will be produced and answered based on the technological possibilities and the suitability of CAT tools for the human translator.

The IBC
The Department of International Business Communication (IBC), which was recently founded by merging the Department of International Language Studies and Computational Linguistics (ISV) and the Department of International Culture and Communication (IKK) at Copenhagen Business School, is responsible for educating experts in language, translation, language teaching and language technology for companies, organisations and educational institutions. IBC actively contributes towards the debate on the importance of language and language processes for the Danish business community and Danish society in general.

The department is fostering dialogue with the business world, and learning about their challenges. Intercultural and interlingual communication is the department’s main priority: it is looking at challenges such as how to enable communication to include people with different cultural backgrounds – for instance, what may happen when a Danish company communicates with a Chinese company in English – a foreign language to both departments.

It gives special attention to issues faced by Danish small and medium enterprises, which need to reach out for business but do not have the same resources as large Danish companies to operate multiculturally, so their conditions for growth are challenged by cultural and linguistic barriers.

Dr Carl
Dr Carl has been involved in many natural language and machine translation projects. Since 2008 he is associate professor at the Copenhagen Business School where his work is very much focused on human translation processes. He says there is a huge potential and need for further research into how automatic systems can best complement human language capacities.

As well as CASMACAT, Dr Carl is member of a five-year WCRE (World-Class Research Environment) – funded by the Copenhagen Business School – and the Centre for Research and Innovation in Translation and "This technology has substantially changed translation work, compared with 30 years ago when translators used typewriters, or, as is still common in India and China, where translators work with pencil and paper while typists type in the translations into the computer"
Translation Technology (CRITT) both of which have the mission to help businesses to address language problems efficiently through a deeper understanding of languages and its processes, language technology, and how they are used in commerce. The project and the centre focus on translation processes and translation systems.

Projects asked Dr Carl how technology has changed translation work over the years. “Tremendously. Translation in Europe is a 4 billion-euro industry, growing 7 per cent a year. Much of this can only be achieved with language technology, grammar and spell checker, translation memories, termbanks, collocation tools, web searches, etc.”

“This technology has substantially changed translation work, as compared with 30 years ago when translators used typewriters, or, as is still common in India and China, where translators work with pencil and paper while typists type in the translations into the computer. In the West, a translator today is also a typist, a reviser, a layouter, and maybe even an editor. In the future (or even now) s/he may also need to have some computer knowledge, and know how to set up a translation memory, load a termbank, etc.”

“There are business and domain-specific terminologies, corporate languages and so on, which are different in different companies,” he says. “And technological solutions exist, such as termbanks, controlled languages, grammar and style checkers.

“The content of these tools has to be specified and populated by their users, but the same tools can be used in different environments, because there are terms and other language conventions in every domain. However, language technology solutions can only be deployed to the extent that language is predictable.”

How is Dr Carl helping the translation industry? “There is a huge potential and need for further research into how automatic systems can best complement human language capacities,” he says. “Within CASMACAT - and also as one of the goals of CRITT- we intend to build a database of translation process data and make it publicly available for other researchers very soon.

“This database will be initially filled with almost 500 translation sessions that we have collected over the past five years or so. We expect that more translation process data will be added, and that analysing the data will give further insights into how humans translate, and how they make use of the translation technology.

“We are interested in the behavioural differences of translators: classification of translation styles and translation preferences. This will put us into a position to develop differentiated and targeted translation aide to optimally support translators during the translation processes,” he says.

CASMACAT
CASMACAT, the language translation project that is developing the translators’ workbench, is a three-year EU Framework 7 project which began in November 2011 and finishes in October 2014.

Researchers on the project carry out cognitive studies of translator activity based on keyboard logging and eye tracking. They examine the data acquired to determine translator types and styles, and to build a cognitive model of the translation process. Working with these insights, they are developing new types of assistance for human translators and integrating them into a new workbench – which consists of an editor, an MT server and analysis and visualisation tools. The workbench is being designed in a modular fashion and can be combined with existing computer-aided translation tools.

CASMACAT is the first European project to bring together the leading groups spanning the whole range from cognitive modelling to statistical machine translation and computer-aided translation technology research and development, and finally deployment. The participants expect significant synergy from this effort.

Owing to the integration and co-ordination of computer-aided translation technologies and the ambitious goals, the project budget is estimated to be around 3.3 million EUR with a requested funding of 2.5 million EUR.

The project will demonstrate the workbench’s effectiveness in extensive field tests of real-life practice of a translation agency. In addition, it will reach out to the wider language service industry and online volunteer translation platforms. The outcome of CASMACAT will be made available as open-source software to industry, academia, and to individual users.