Hitotsubashi Journal of Social Studies 35 (2003), pp.17-28.

GLOBALIZATION, CULTURE AND TECHNOLOGY: THE CASE OF KOREAN MACHINE TRANSLATION*

JONATHAN LEWIS**

About this Research

This is a preliminary report on the development of machine translation (MT) technologies in Korea. The report presents the findings of interviews with Korean MT researchers carried out in early 2003.

Machine translation refers to the automatic translation of human languages by computers. Machine translation is part of the wider field of natural language processing (NLP), which in turn is one area of artificial intelligence research. Other areas of natural language processing include summarization, information retrieval, speech recognition and synthesis, and handwriting recognition.

With the digitalization of large areas of human economic, cultural and political activity, and increasing Internet activity, many markets and communities are no longer delineated by geographical boundaries but rather by language. This means that language barriers are acquiring even greater economic and political importance than previously. The possibility of developing technologies to translate automatically between languages thus has enormous implications for all areas of society.

This research project aims to contribute to two areas of scholarship. The first area is science and technology policy. Many available studies on science and technology policy focus on macro-level political and economic factors. Mani (2002), for example, compares innovation and technology policies across eight countries, and develops models based on his findings for each country. His work builds on that of Lundvall (1988) and others on national innovation systems.

In contrast to those macro- and meso-level studies, this research investigates how policies impact on individual projects and researchers. The justification for this micro-level approach is the evidence that there is considerable variation between technological/scientific fields in the way policy is implemented. In particular, the interim findings regarding NLP presented here do not support Mani's optimistic conclusion that "the Korean system is rather unique in that

* Acknowledgements
This research was undertaken as part of a JSPS-funded project comparing policymaking in Japan and Korea. The project was initially led by the late Minoru Nakano of Meiji Gakuin University, and subsequently by Yutaka Tsujinaka of Tsukuba University. I would like to thank all members of the project, and in particular Jaeho Yeom of Korea University for his generous assistance. I am also greatly indebted to many Korean MT researchers for interviews and introductions. Due to the ongoing nature of this research I shall not name them in this report.

**Graduate School of Social Sciences, Hitotsubashi University, jonathan_lewis@mac.com. http://homepage.mac.com/jonathan_lewis/

it has managed increasingly to privatise the financing of R&D without experiencing the usual problem of underinvestment by the private sector" (2002, p.116). Since the mid-1990s, many private and public organizations in Korea have abandoned or reduced research on machine translation.

Of course there have already been many studies of technology policy that have delved beneath the generalizations based on national statistics. To give only a few examples with a Japanese focus, Nakayama's (1995) four-volume history of science and technology in Japan follows policy and research developments in many particular fields. Samuels (1994) investigates a number of different sectors of Japanese technology, only to identify an ideology of "technonationalism" that influences public and private policy decisions across the board. Noble (1998) compares the role of politicians in industrial policy networks in Japan, Korea and Taiwan. His cases include technology-oriented ones.

Policies for the research and development of machine translation have to deal with two key characteristics of the field. First, NLP technologies are to some extent language-dependent. In other words, Russian-to-English translation software will translate Russian into English but cannot be easily adapted to translate Japanese into English. (The question of how far MT technologies should be language-dependent is returned to below). As a result, the market for MT is largely limited to the population using a particular language. In this respect MT is clearly different from e.g semiconductors or space rockets. It is difficult to import ready-made technologies that can be quickly commercialized, improved upon, and exported — the classic model of catch-up industrial policy. This poses something of a policy dilemma for nations such as Korea where the number of overseas speakers of the national language is limited. It may be that supporting other technologies could foster competitiveness in global markets, whereas the market for MT would at most be a domestic one. On the other hand, if high-quality machine translation for Korean could be achieved, the benefits for Korean businesses could be very large. In addition, if MT for Korean falls behind that for other languages then there is a risk both of damaged competitiveness and of reduced prestige of the language, with its implications for national identity and state legitimacy.

Second, it has proved difficult to bring many of the key MT technologies to a level of performance where they can be successfully commercialized. This is a common feature across areas of natural language processing, and indeed artificial intelligence research as a whole. Despite fifty years of research, machine translation has not seen the spectacular advances and rapid commercialization that have taken place in hardware and other areas of software. (It has been argued that the whole area of software has seen only incremental improvements, as opposed to fundamental breakthroughs, since the 1960s. See Glass (2003, p.20)). So researchers and developers of NLP technologies have to develop a lot of costly, basic tools and knowledge domestically (due to language-dependency), and to do so within an institutional setting where they struggle to meet expectations of fast progress and early profitability. How flexible can nation- or sector-wide institutions be in supporting research with such atypical development trajectories?

The second area to which this study aims to contribute is the history of science and technology. Hutchins (1986) has documented the history of machine translation research, but readily admits that his coverage of Asia is insufficient. This report represents the first step towards a documentation of the research and development of machine translation in Asia, in order that the contributions of Asian researchers be recognized and also that comparisons can

be made between MT technologies developed in Asia and those produced elsewhere. The first generation of NLP researchers is long retired, and there is not much time left to interview them and bring together documentation on the early years of translation and other NLP research. The interviews carried out so far have been of mid-career researchers, but it is planned to extend coverage to include their predecessors. This project therefore aims not to stop at merely delineating technology policies and their impact on researchers, but to see which technologies were developed, and how those technologies reflected the social context in which they were produced.

The Start of MT Research in Korea

The IBM-SNU Project

Research started on a small scale in the early 1980s, in the computer science departments of Seoul National University (SNU) and government research institutes. The SNU research succeeded in developing a rudimentary English-to-Korean translation system by the mid-1980s. Researchers from the university presented their results at a conference in Japan, which by that time had become the most active place in the world for machine translation research and development. Some researchers from IBM Japan were attending that conference, and after hearing the SNU researchers' paper they proposed a joint project to develop a prototype English to Korean translation system.

IBM's joint research with SNU started in 1986 and lasted until 1999. The project was divided into three phases, and interviewees characterized the goals of each phase as realistic. Although English to Korean translation was the major focus of the project, work was also done on Korean to English translation, information retrieval and Korean grammar checking. Funding for the project, which was switched at an early phase from IBM Japan to IBM Korea, was described by all interviewees as the largest for any research project on Korean MT to date. In addition to producing tools and systems, the IBM-SNU project also provided a training-ground for many of Korea's NLP researchers.

There follows one example of the contribution of the IBM-SNU project to the development of NLP in Korea. One graduate student of SNU was assigned the task of developing a grammar checker together with two other researchers. The biggest challenge in developing a Korean checker is accurate lexical analysis. (This is different from English, where the trickiest problem is syntactical analysis.) He wrote a PhD thesis that established methods of lexical analysis for Korean, and produced an analyzer that embodied these methods. During his time as a graduate student he was employed for one year by IBM Korea, for whom he developed a commercial Korean speller. Under the terms of the agreement between IBM and SNU, however, he was not allowed to use the analyzer after graduating. Finding himself unemployed for a year after graduation, he wrote another analyzer from the ground up. This analyzer not only freed him from the restrictions placed on his previous work, but also contained improvements he had thought of during his graduate studies. He has since licensed his analyzer to a number of information retrieval companies, and it is used in several Korean web search engines.

Government research initiatives

At about the same time as this IBM-SNU project, more extensive research on NLP theory and applications began in three Korean government research institutes: The Electronics and Telecommunications Research Institute (ETRI), which coordinated the research; the smaller System Engineering Research Institute (SERI; absorbed into ETRI in the mid-1990s), and the Korea Advanced Institute of Science and Technology (KAIST). At ETRI there were five teams working on areas including information retrieval, machine translation, basic NLP technologies (including corpus and dictionary construction), and voice recognition. At SERI, research focused initially on translation between Japanese and Korean (a difference from the IBM project, which had no Japanese), and built to a considerable extent on work done in Japan. Later, SERI also started work on translation between English and Korean. KAIST was SERI's partner in this research.

Although figures are not available, interviewees suggested that the funding for this government research was probably not as plentiful as that provided by IBM. Interviewees who were not employed in these government institutes noted that researchers in government labs receive funding for relatively short periods (maximum three years). In order to obtain funding for the next three years they have to propose a new project with a different title, and this means that they have to shift the focus of their work. Nevertheless the government laboratories can boast some achievements: one of them was the first Korean CD-ROM encyclopedia, which was published in 1993. Many academic researchers have also participated in ETRI-led projects. KAIST is also an academic institution in its own right, and PhD students from its natural language processing laboratory are now researching and teaching in universities around Korea.

Government funding for NLP research increased somewhat in 1997 when the Ministry of Information and Communications started to support the area. The MIC had completed the nationwide information superhighway project the previous year, and was shifting its focus to new areas. This development meant that the primary responsibility for NLP research shifted from the Ministry of Science and Technology (MOST) to the deeper-pocketed MIC. At the MIC, however, NLP research had to compete with many other areas of information technology. Officials from ETRI and elsewhere made weekly visits to the new ministry to explain the importance of NLP and to offer predictions of the potential market for NLP applications. Despite their efforts, annual government funding did not rise beyond about $1.2 million.

Computer Scientists and Linguists

One recurrent problem in the history of machine translation research has been the tendency for institutional and intellectual gaps to develop between linguists and computer scientists. Institutionally, the two disciplines are located in different university faculties, each with its own evaluation, admission and promotion structures. It is often not straightforward for the holder of an undergraduate degree in linguistics to enter a graduate program in computer linguistics, or vice versa. The barriers are even higher when it comes to academic appointments. Hutchins (1986) has noted that this institutional divide can be accompanied by an intellectual gap: engineers can tend to seek pragmatic solutions that provide adequate

translations in many cases, whereas linguists can tend to focus on abstract or esoteric features of language that are difficult to capture in working MT systems.

This institutional gap has also been a problem in Korea. It is important to note, however, that many individuals and some organizations in the MT research community have made great efforts to combine the ideas and efforts of the two disciplines. The director of the NLP program in ETRI, for example, had studied for his PhD in France (rather unusually for a Korean computer scientist), and had been very impressed by the interdisciplinary nature of the laboratory in which he studied there. He subsequently played a major advisory role in the planning of the Sejong Project, but his attempts to get linguists and computer scientists working together were to meet with some frustration.

Although the IBM-SNU and ETRI-centered projects were both computer science projects, a number of Korean linguists became interested in computers during the period before 1995 and have since made significant contributions in the research and development of machine translation. One of them was writing his PhD on old Korean syntax in the early 1990s when he read a book about the development of artificial intelligence in Japan under the Fifth Generation Computer project. This sparked his interest in natural language processing, and over the past decade he has developed a number of technologies. His ultimate objective is to achieve a system where humans and computers can communicate seamlessly by voice in different languages. While he does have an academic post, he has developed these technologies in the private sector using venture capital.

Another leading private-sector developer of machine translation also started his academic career in linguistics. No professors in his department were interested in the use of computers for linguistics research, but a group of students organized themselves to build a rudimentary English to Korean MT system. The student switched to a department of cognitive science for his graduate studies, and after graduation started a company to commercialize the system. According to its developer, the system that started out as an undergraduate club project now has a performance comparable with MT engines in other countries.

These two cases show that those determined to participate in the research and development of machine translation will find a way of doing so whatever their disciplinary backgrounds. However, they also suggest that large academic and government research institutions find it difficult to accommodate trans-disciplinary researchers. Both men have formed their own companies to pursue their work, despite the severe funding difficulties this entails.

The Korean government has taken one major initiative to bridge the institutional gap between linguistics and computer science. In 1997 the Korean Ministry of Culture and Tourism formed a six-member committee to consider the future direction of Korean NLP research. Three members were experts on the Korean language, and three were computer scientists. The report submitted by the committee formed the basis of the ten-year Korean Language Information-Oriented Project (KLIOP), which has three phases:

constructing Korean language databases (corpora)
improving the language processing ability of computers
applying the results in e.g. machine translation and voice sensor systems

These aims are being implemented by the National Academy of the Korean Language in the 21st Century Sejong Project, which started in 1998 and is scheduled to run until 2007. Research itself is undertaken by academics, and funding has been approximately $1 million per year. The Sejong Project website (http://sejong.or.kr/english/) shows the progress that has

been made to date, particularly in the area of corpora. The project has significantly increased the available knowledge on Korean linguistic phenomena, a field that previously was receiving perhaps $200,000 of research funding per year.

The initial plan for the Sejong project envisaged linguists and computer scientists working together. The corpora assembled by linguists would serve as resources for MT developers. However, the project has experienced difficulties in its efforts to bridge the disciplinary gap. The primary difficulty is in funding. Research in natural sciences and engineering generally requires much larger budgets that that in the humanities, and natural language processing is no exception. The annual budget for the Sejong Project has been approximately $1 million per year, which is a considerable amount for linguists but after being divided up is insufficient for major computer science research. As a result, linguists have tended to play a dominant role.

The dominance of academic linguists has further undermined the cross-disciplinary aims of the project by shifting the emphasis of research. Linguists have been concerned, with the language itself rather than the use of computers to process it. For example, linguists on the project have built corpora that enable them to compare language use in South and North Korea, and to investigate languages spoken and written 100 years ago. Such databases are of little interest to developers of MT systems.

The Switch to Information Retrieval

By 1995, ten years of research on Korean machine translation had reached the stage where systems could be sold. IBM released its first English-to-Korean MT system at that time, and many other companies followed suit. The release of the products coincided with the rapid spread of the Internet through Korea, and the late 1990s saw something of a craze for machine translation software. At that time a very large proportion of resources available on the Web were in English, which made automatic translation an attractive technology for many Koreans. Unfortunately, certain software vendors made exaggerated promises regarding the quality of translation that their applications would provide. The word soon spread among Korean computer users that the quality of automatic translations was much lower than the hype suggested, and sales contracted. The resulting mistrust in machine translation remains a problem for software vendors and MT researchers alike; government agencies as well as computer users are now skeptical of new initiatives in MT development, despite incremental advances in translation quality. The rapid bursting of the Korean MT bubble caused many companies to abandon research and development in the area, and divert their resources to information retrieval.

Information retrieval is the other area of NLP that experienced rapid growth as a result of computerization in general and the Internet boom in particular. In order to retrieve textual information from a database, that information has to be broken down into its constituent sentences and words, and furthermore words reduced to their basic forms (so that, for example, a search for “person” also returns hits for “people”). In smarter systems, semantic analysis is also done on the text in order to increase the relevance of results. Search engines such as Google also use algorithms to rank search results in order of their popularity. The point here is that many of the basic technologies for information retrieval are the same as those used for analyzing source texts in machine translation. This remains the case, even though the

focus of research on information retrieval has recently shifted to some extent from search engine technologies to document classification, clustering, and Q&A systems.

Starting in 1999, the market for information retrieval software experienced tremendous growth. According to a report compiled with the support of the MIC, the market grew from $4 million in 1999 to $10 million in 2000 and $20 million in 2001. The prospect of immediate returns on investment in information retrieval software made development in this area more attractive than the much thornier path of machine translation. Computer science students were also drawn to the study of information retrieval; one academic NLP researcher interviewed said that all of his graduate students were working on information retrieval rather than machine translation. The NLP departments of most Korean universities are said to be concentrating their efforts on information retrieval.

Information retrieval enjoys the advantage over machine translation that it is a highly visible technology. On the web, everybody uses Google and other search engines, and users frequently discuss the relative merits of different search sites. By comparison, translation of web pages tends to be a last resort when the user does not find the desired resources in her own language. Web translation services are offered as a supporting service of the main search engine rather than a sophisticated application in their own right. The very simple interface of web translation engines compared to stand-alone MT applications can also convey the impression that they are not performing a technically very difficult task. This supplementary, hidden nature of automatic translation makes it less attractive to computer science students who dream of creating the next Google, and also to government agencies with a remit to fund projects that offer a clear route to commercial applications.

Despite the high visibility of information retrieval, the diversion of resources from MT to IR research and development does not seem likely to bring all the returns expected of it. The first reason is the rush of new entrants to the information retrieval business, which means that even in expanding markets profits are being squeezed. One company reduced the price of its search engine from $50,000 to $15,000 in 2002 but still struggled to find buyers. Second, the burst of the dot-com bubble means that the market is not expanding as much as had been hoped; the estimate of another year-on-year doubling of the market to $40 million in 2002 is unlikely to have been realized. The third reason has to do with technological development. Although IR researchers apply basic NLP technologies in their products, they themselves tend to specialize in developing probabilistic methods for calculating the most relevant hits. While the researcher who finds some new probabilistic method may strike the jackpot and give the world its next Google, in the long term the accuracy of information retrieval is likely to depend on incremental improvements in NLP technologies. Because the benefits of such improvements are felt only indirectly, however, research on basic NLP technologies is not attracting much funding. Researchers and students, in turn, are moving away from basic NLP research. The ironic result of the rush to develop information retrieval applications is that the future development of information retrieval itself is being compromised.

Faced with stagnation in the MT market and a crowded IR market, companies are trying to find other outlets for their NLP technologies. One company at least is trying to capitalize on the demand for English-language teaching by offering computer-driven language lessons for Korean children. The system analyses sentences up to about 20 words long that do not contain difficult words, and looks for errors that are commonly made by Korean learners of English, e.g. “I am go to school”. The company claims a success rate in error detection of over 90%.

While still at an early stage in its marketing of the product, the company hopes that young Koreans' openness to new technologies will lead many of them to start using it. The company points out that the product's NLP technology is only one part of the product; the educational conception has to be right, and a lot of work also needs to be done on interface design. If the product becomes profitable then the company will be able to cross-subsidize its research on machine translation.

Key Companies in Korean MT Today

Following the shake-out in the late 1990s, private-sector development of MT in Korea is concentrated in six companies.

The first is IBM Korea, which in about 1999 parted ways with SNU and started its own project. This project reportedly involves close collaboration with IBM Japan and IBM China and uses IBM's common English parsing system. IBM Korea has developed its own Korean generation system, and has employed a lot of part-time workers to make transfer dictionaries.

Microsoft Korea has a small Korean language resource team in Korea. The team is led by a former engineer at KAIST, described by one interviewee as "a brilliant woman." Interviewees outside Microsoft thought that the team's activities are limited to localization and to developing of Korean spell checkers. Microsoft's work on machine translation is carried out by its large Natural Language Processing Group in Seattle.

LNTech is the company set up to commercialize the technologies developed by Seoul National University's NLP lab. It offers English-Korean and Japanese-Korean translation software and online services.

Language and Computer Corp. (LNC) sells an English to Korean translation application, and is developing a Korean-to-English translation application.

ClickQ markets English-to-Korean and Korean-to-English machine translation software, both of which products are adopted from MT engines developed by Japanese companies. (The linguistic similarities between Korean and Japanese make such adaption feasible).

LNI Soft markets an English-to-Korean package, the engine of which it claims to have developed itself (although one interviewee suspected a Japanese influence).

The reader will note the absence of larger Korean electronics companies. Samsung was said to have begun research on machine translation but to have given it up when it found the prospects for commercialization insufficiently promising. This situation is in start contrast to Japan, where major electronics companies such as Toshiba, Hitachi, Fujitsu and NEC have all commercialized MT products.

What we see is thus an extreme polarization of development of MT in Korea between the local corporations of multinational companies and small companies specializing in MT. The following section describes these small companies' attempts to coordinate their efforts and to obtain government support.

Coordination and Government Support

The start of an MT research association

In 2002 a large number of small Korean NLP companies started an association for MT research. The association aims to develop core dictionaries and corpora.

The association faces two problems. The first is lack of resources in member companies to carry out research on MT (the construction of dictionaries, for example, is labor-intensive). According to one interviewee, the association succeeded in obtaining $5 million in funding from the MIC in 2002, but the number of member companies was so large that the money was spread very thinly. The second problem is that most companies' research levels are not high but they cannot admit this. To start research projects the association needs to select some core project members with the best technologies, but the selection process is politically very difficult.

Government support for NLP through KOSA

In February 2002 a study on the planning of R&D in the language information industry was submitted to the Ministry of Information and Communications by ETRI and the Korean Software Industry Association. The goal of the report was to explain the NLP market and technologies to the government in order to obtain support for R&D projects. This report was thus in some ways a continuation of ongoing lobbying work to persuade the MIC of the strategic importance of NLP. With regard to machine translation, the report had the tough job of explaining why the government should fund an area that large Korean companies had judged to be commercially unviable. The approach taken was to argue that a large demand for MT does exist, but that the technologies at present are unable to make the market a reality. Government funds are therefore needed to raise the levels of NLP technology.

The report started by situating the language information industry within the government's Cyber Korea 21 initiative, which had started in 1999 with the modest aim of "building a creative, knowledge-based nation." It presented data on the increasingly multilingual nature of the Internet, and an August 1999 survey by the Korean IT portal site Electronic Times showing strong domestic demand for MT.[1]

The market for NLP technologies was presented as consisting of three parts: information retrieval (the largest), machine translation translation, and text mining. The recent difficulties in the information retrieval market were noted, as well as the tiny size of the markets for machine translation and text mining.

After presenting the state of the art of NLP technologies abroad, the report outlined the prospects for future development in the various sub-fields of NLP, and finally got down to the nitty-gritty matter of government support. $8 million dollars were needed for 2002, split four ways: $2 million each for language resources and basic software, machine translation, information retrieval, and text mining. Funding should be increased to $13 million in 2003 and $19 million in 2004.

---

[1] Cyber Korea 21 documentation available from www.ipc.go.kr/ (search Google for "Cyber Korea 21"). Electronic Times: http://www.etienglish.com/

Needless to say, the MIC did not grant the funding request in full. It gave KOSA $1.5 million in research support funding for NLP in 2002. Prospects for funding in 2003 were made increasingly uncertain when the section head in charge at MIC changed, necessitating a further round of explanatory visits.

In order to avoid the coordination problems experienced by the MT research association as described above, KOSA proposed that a research center should be established through which funds could be channelled into the development of common resources such as corpora. This would avoid the danger that companies receiving direct subsidies would divert the funds to other other purposes.

Globalization and Language-dependent MT

A common issue that emerged from the interviews was a tension between language-independent and language-dependent approaches to machine translation. In simple terms, a language-independent approach aims to construct tools or systems that function the same way whatever the source or target language. Of course dictionaries for the source and target languages have to be specific to those languages, but many components of the translation system can either be designed to work with any language (language-independent) or to work with one particular language (language-dependent). The attraction of language-independent approaches is that they can be applied in systems translating between any given pair of languages. The potential strength of language-dependent approaches, by contrast, is that they can be tailored to reflect and render the unique characteristics of a particular language, thus giving potentially better results.

The language-independent approach has particular attractions for companies operating in a large number of national markets where there are demands for translation software and language tools in different languages. IBM and Microsoft are just such companies, and the information gathered to date suggests that they do indeed seek to develop "reusable" components in their translation and other NLP software. The fact that three of IBM's top MT researchers left to join Microsoft's NLP Group in the early 1990s would further lead us to except considerable similarity in the two companies' approaches. IBM Korea was said to be using the same English parsing system as its sister companies in Japan and China; it would be irrational to do otherwise. One interviewee, however, said that "Microsoft's approach is not exactly language independent: rather it is to find the best method for English, then apply the methodology to other languages." Leaving aside the question of how accurate this statement about Microsoft's approach is, the concern it voices is that the linguistic differences between Korean and English are so large that even a so-called language-independent approach may produce sub-optimal results for Korean.

These observations about IBM and Microsoft should not be interpreted as technonationalist attacks on US or global capital. For the interviewees, what mattered was not language-dependence or independence but the quality of results obtained from tools or systems. In fact, one interviewee felt that much greater pressure to adopt language-independent approaches was exerted in Korean universities. His explanation went as follows: as a result of recent reforms, Korean academics have to get as much of their research as possible published in international journals and conference proceedings. However, it is difficult to get papers using language-

dependent methodologies published in international journals. And it is easier to get an article accepted for an international journal when the test data is in English. At most international conferences, too, only a handful of participants are interested in Korean. The result of these pressures is that many Korean researchers pursue language-independent techniques. This suggests that globalization, in the shape of a drive by Korean universities to achieve "global standards" in research, is having an impact on the direction taken by Korean MT research. However, this impact is exerted less through the activities of global IT companies than through the exposure of Korean academics to the globalized marketplace of ideas.

Summary of Main Findings and Future Research

While the global war of terror has contributed to a modest revival in machine translation research in the United States, times are hard for MT researchers in Korea. The software-buying public is leery of the industry's products. Government research funding is low and with some notable exceptions seems to be channelled towards short-term research projects to develop commercializable tools rather than basic technologies. The role of universities is unclear, and the rush towards information retrieval has reduced the number of researchers in basic NLP and machine translation. Academic reforms, however necessary, risk skewing research methods in favor of language-independent techniques. While a number of small companies continue to develop MT technologies with remarkable energy, their efforts are poorly coordinated and even more poorly funded. Larger Korean companies have given up on MT entirely. All in all, this is a far cry from Suni's depiction of the Korean innovation system as one that "has managed increasingly to privatise the financing of R&D without experiencing the usual problem of underinvestment by the private sector." Of course it would be difficult to find researchers in any field who would complain of over-investment in their field. More international comparisons of research funding are needed before we can reach firm conclusions about the extent of underfunding in Korean MT. But given the huge economic and social impact that the use of computers to translate text and speech is widely expected to have during the next decades, the low levels of funding suggest that the Korean private and public sectors are failing to invest adequately in an important area of R&D.

More research is also needed to flesh out the sketch of MT development policies presented above. The progress of the MT research association and the KOSA initiative should be monitored. From the point of view of the sociology of science and technology, it is necessary to investigate the tools and systems that have been developed, to see how how they reflect the social framework within which their authors found themselves. How do MT technologies resolve the institutional divisions between computer science and linguistics, or the tension between language-dependent approaches and methods developed for Korean? The research and development of machine translation is a site of complex interaction between globalized research institutions and national cultures, and the resulting technologies embody these complexities.

REFERENCES

Glass, Robert L. (2003). *Facts and Fallacies of Software Engineering*. Addison-Wesley.
Hutchins, W. John (1986). *Machine Translation: Past*, *Present*, *Future*. Ellis Horwood.
Lundvall, B.A. (1988). *National Innovation Systems*. Pinter.
Mani, Sunil (2002). *Government*, *Innovation and Technology Policy: An International Comparative Analysis*. Edward Elgar.
Nakayama, Shigeru (1995). *Tsûshi Nihon no Kagakugijutsu*. Gakuyô Shobô.
Noble, Greg (1998). *Collective action in East Asia: How Ruling Parties Shape Industrial Policy*. Cornell University Press.
Samuels, Richard (1994). "*Rich Nation*, *Strong Army*". Cornell University Press.