Joint Report on Tokyo University of Foreign Studies, Centre of Excellence

October 2006

1) Project planning

Since the specific formulation of the objectives of the Centre of Excellence research progamme set out in 2002, has not been made available to the panel, we are not in a position to make fully informed evaluations. We can, however, project back to what was planned, at least to some extent, on the basis of the outcomes presented to us. The quality of these outcomes is such that it is difficult not to assume that the initial proposal must have displayed some highly positive features: a creditably broad vision firstly, of the linguistic areas to be covered: phonetics, morphology, syntax and lexicology, but also pragmatics and sociolinguistics; secondly, of the methods to be used, particularly the integration of CALL materials into the overall teaching programme; thirdly, of the essentially practical goals related to the more effective delivery of language teaching.

With regard to research on linguistics, four international conferences/workshops with four volumes on Linguistic Informatics already published and a fifth in the pipeline are indicative of high-quality and sustained effort. The panel was impressed by the excellent standard of organisation and co-ordination displayed in bringing the September 2006 workshop together.

2) Originality

Although the term Linguistic Informatics goes back at least to the 1980s, the TUFS project uses it with a new meaning, that of a new interdisciplinary area, combing language learning, linguistic theory and computing. While clearly originality cannot be claimed for certain components of the project, the idea of combining research in theoretical, applied (computational) and corpus linguistics in order to produce a multilingual CALL platform for languages as different as the 17 languages covered in the COE programme is certainly original. This combination of theory-driven linguistic analysis and specific applications to teaching practice is indeed unusual. For instance, explanations on the use of lexical items founded on a corpus-based analysis of the most frequent context in which it occurs. In the opinion of the panel, this approach both increases the effectiveness of the teaching materials and sharpens the formulation of theoretical questions.

The multilingual approach adopted provides a useful overarching framework for the languages already covered and any that might be added. This provides a valuable comparative tool that is of considerable use to both researchers and students. The cross-linguistic grammar modules assist students to acquire their chosen target language(s) by making two-way contrasts between first and second languages. For researchers, the shared framework over a wide range of languages, including a number of Asian languages facilitates wide-ranging comparative and contrastive projects. Of note too is the fact that the framework maximises potential parallels while avoiding the cultural interference associated with parallel texts, construed in the sense of multiple translations of single or small number of source texts. The resultant combination of comparability and adaptability to individual languages is to be particularly applauded.

3) Scientific quality

Clearly the principal application, the web-based language learning programme is of considerable interest, as already observed. The panel is, however, limited in what it can say about use interface design, since the interface language is Japanese, but looks forward to the addition of interfaces in English and other languages.

The UBLI volumes bring together fascinating work by scholars from many countries, some of whose contributions are cutting edge, while others give valuable overviews of various fields and sub-fields. It appears that the format laid down has restricted to some degree the development of the theoretical aspects of specific research projects. The panel notes a relative paucity of references to some current major linguistic theories, for instance generative grammar, cognitive linguistics, construction grammar, functional grammar and pattern grammar. It is highly likely that contributors, respecting prescribed time constraints, and anxious to focus on the generally rich findings of their analyses of corpora, condensed theoretical arguments for lack of time and space. The contents of the volumes are also dictated to a degree by the expertise of the scholars who were able to accept invitations, resulting in some degree of unevenness of coverage of various sub-fields. It may be worth considering writing or commissionning theoretical articles introducing volumes or sections of volumes in future UBLI publications.

As regards pedagogical applications, it has to be recognised too that in some areas, the state of the art in computational linguistics is not sufficiently developed to be integrated into modules for language learning. This should be seen as an opportunity rather than a criticism. (see "results").

The organisation of international conferences and workshops was, from the invitees' perspective as in the case of the September 2006 workshop, an extremely worthwhile exercise to meet and exchange views with both Japanse colleagues and graduate students as well as colleagues from various countries. For the graduate students, it is an important and valuable opportunity to be able to approach established scholars at close hand.

4) International Contribution :

The international contribution of the TUFS COE programme may be evaluated on a number of levels.

Firstly, the international character of the range of languages coverd by the didactic applications. In this respect, the programme is highly international, since a large number of typologically different languages are treated using a unified methodological and systematical approach.

Secondly, the use of international research results. For instance, in French linguistics, the project has integrated the most relevant institute working with corpora of spoken French (Aix-en-Provence). According to the programmes of preceding workshops, other contacts concern the domain of syntactic lexicons (LADL) and syntactic parsing (Intex). Although the precise level of integration of these resources is not entirely clear, the programme has undeniably created an invaluable international network.

Thirdly, the international availability of research results. The series of volumes on *Usage-Based Linguistic Informatics* (John Benjamins) has made the work of both home-grown researchers and members of UBLI's international network widely available to scholars throughout the world.

According to the information delivered at the panel's request, further significant steps forward will be/have been taken. Results will be made publicly available for international research, i.e. corpora will be distributed on demand, software will be published on an open-source basis, and language modules will be made accessible for research institutions.

Finally, the overall international visibility of the project. Here, the COE programme could achieve a greater international contribution, simply by presenting the research results on the web pages more systematically and in English (at the time of evaluation, only an "outer shell" of web pages was available in English, the lower levels presenting more detailed information were in Japanese).

8

5) Feasibility (Results)

Although during our discussion on September 15, the panel expressed misgivings about giving an *a posteriori* assessment of the feasibility of the programme, it is now our general feeling that the outcomes achieved, for instance, in contrastive analysis, typological decription and applications to language teaching, however these may line up with the plans formulated in 2002, are more than sufficient to justify adopting a positive view of the TUFS COE programme.

It would of course be instructive to carry out analyses of the quality of learning achieved by students using the CALL programme.

6) Scientific results

In Section 2 we have pointed out the originality and in Section 4 the international contribution achieved by the TUFS COE programme.

In the presentation on September 15th it was stated that a considerable amount of corpora, written and oral, have been compiled, annotated and made available to the COE project. The interaction between the COE language modules and these "real-data modules" was demonstrated for several languages during the UBLI-project reports. Considering the fact that the primary interest of TUFS lies in CALL aplications, these results are quite impressive. However, it was not clear to what extent the present results and the future objectives depend on the progress in NLP technologies. For example, at the time of the evaluation, the grammar modules were still being developed, and the integration of technologies like parts-of-speech tagging or syntactic parsing, still less semantic disambiguation could not be demonstrated. On the one hand, interactive CALL modules require such technologies, on the other, the development of such tools for a considerable number of languages goes far beyond the stated objectives of the COE, despite the clear potential to expand what has already been achieved for 17 languages. However, if the project is to achieve greater things than simply more of the same, co-operation with competent partners in appropriate areas should be considered for the next phase of the project. In this as well as in other aspects of the project, it would seem sensible to take account of comparable work carried out in other countries. For instance, the illustration of IPA sounds could be related more to specific languages in which they occur as is the case for Peter Ladefoged's *Vowels and Consonants* or John Wells on-line. Moreover, some commercially available packages are capable of giving sophisticated feedback on supra-segmental aspects of pronunciation.

The UBLI project has brought together researchers from a number of countries and the edited volumes published contain some reports of cutting-edge research.

All in all it was a bold decision to develop an ambitious programme in linguistics more or less entirely with home-grown talent, save of course for the invited contributions of foreign scholars. Many western universities in comparable circumstances would have sought to buy in established scholars in attempt to give an immediate boost to the profile of the project outside the home institution. The policy of developing from within, as far as can be discerned from the evidence presented, probably meant that 5 years was too short a period of time to achieve apparently far-reaching goals, although once again our view is limited by our ignorance of the situation, when research grants were awarded to TUFS. Given what we have been able to see, it is our sincere hope that the project will continue to be resourced in order that it may achieve its full potential.

7) Fostering young researchers

The panel gained a very positive impression from the workshop as regards the development of young researchers. Co-operation between postgraduates and supervisors seems very good. We noted with approval that a number of young researchers were (or had on other occasions been) given the opportunity to present papers in an autonomous way (that was certainly the case at the September workshop) in international fora.

Considering the detailed list of conference participations, we note that no fewer than 21 students, whether singly or in groups, have had the opportunity to present their work to an audience of international scholars, in some cases on a number of occasions. Such occasions are extremely valuable in the training of young researchers.

Peter Blumenthal, Universität zu Köln, Germany
Jean-Philippe Dalbéra, Université de Nice, France
Jean-Michel Eloy, Université de Picardie, Jules Verne France
Pierre Kunstmann, University of Ottawa, Canada
Chantal Lyche, University of Oslo, Norway
Yves-Charles Morin, Université de Montréal, Canada
Tim Pooley, London Metropolitan University, United Kingdom
Achim Stein, Universität Stuttgart, Germany