

PRESENTATION

Dr Sylviane CARDEY

Dr François-Claude REY

University of Bourgogne Franche-Comté, France

The contributions to this number of the BuLAG are dedicated to Peter Greenfield. The title "Languages analysis, comparison and generation, systems, models and applications" (Analyse, comparaison et génération des langues, systèmes, modèles et applications), reflects what Peter Greenfield thought about Natural Language Processing, that is "Before one can prepare formal models of linguistic phenomena, linguistic studies must be done with the view that the result will be subsequently modelled for some eventual use by mechanical mean". Also, "a theoretical approach which is mathematically based, whatever it is, ought to be able to accept linguistic formalisms. For this reason, such an approach has to be sufficiently flexible so as to enable the construction of models themselves founded on model theory using a constructive logic approach. These models must adapt themselves as the analysis proceeds and when new problems are uncovered. In order to describe a language, nothing ought to be fixed in advance, the risk being to find oneself tied into a situation where instead of analysing the language, one bends and deforms the said language to fit the supposed model; such a practice being not only against nature, but anti-productive, anti-descriptive, and contrary to and the inverse of all that is theoretical and experimental scientific practice. Linguistic is theory rich and it continue to be theory creative. Linguists working in the domain of producing grammars and meta-grammar for language engineering would like to be able to describe these using their own specialised languages without resorting to the overheads involved in programming". Peter Greenfield was a mathematician and computer scientist. Peter Greenfield received his BA from Cambridge and his MSc and PhD in computer sciences from the University of Birmingham where he presented "*A process and storage management scheme as a level within an operating system*". He was CEng and Member of the British Computer Society. He worked in big industries in England as scientific programmer in the aircraft industry (Structural Analysis Department, British Aircraft Corporation – now British Aerospace), and as a consultant and senior programmer in the computer manufacturing industry (Honeywell

BULAG N° 40, 2022
Languages Analysis, Comparison and Generation Systems,
Models and Applications

Information Systems, GEC Computers) and in the U.S.A. in Honeywell Computer Science Research Center in Minneapolis. Subsequently, he was tenured professor in Computer Science at the University of Birmingham. He became passionate by languages when being at GEC he was collaborating in the Alvey Programme, a British government sponsored *research programme* in information technology. Being so curious about languages, their grammar and the way they work, he decided to help the linguists working in NLP (Natural Language Processing) in England and when he came to the University of Besançon teaching and working in big national and international projects which he often directed. His researches in Industries and in Universities are linked to the following traditional domains:

- Scientific calculation (calcul scientifique).
- Operating systems (systèmes d'exploitation).
- Information systems and databases (systèmes d'information et bases de données).
- Computer engineering (génie informatique).
- Logic programming (programmation logique).
- Natural Language Processing (Traitement Automatique des Langues).

Among his publications we can find 84 papers and 1 book. He supervised PhD thesis in England and France.

We can now understand why the following contributions are about NLP and how Peter Greenfield saw this discipline. He himself contributed to the theoretical approaches, mathematical modelling, machine translation, controlled languages, information retrieval, dictionaries and the DecRIPT project.

The first contribution "Principles of Data Processing" is a presentation for a conference in Fribourg University in Switzerland by Peter Greenfield. It is followed by lecture notes taken by François-Claude Rey during Peter Greenfield's conference "computer where are we?" held at the University of Besançon. Both conferences being linked, François-Claude Rey's notes explain most of them both.

Sylviane CARDEY & François-Claude REY
Presentation

The concept of extended intelligence which corresponds to the reflection cited above about NLP is discussed by Pierre-André Buvet. It focuses on the complementarity between artificial linguistics and linguistic intelligence.

Systems and models, theoretical approaches, semantic modelling with the contributions of Sylviane Cardey, Hitoshi Isahara *et al.*, Rodolfo Delmonte, Walid el Abed, Oleg Kapanadze, Leonel Ruiz Miyares *et al.* and Bilal Shafei are presented with their defects and some better solutions are proposed.

All disciplines need to communicate in natural language, for example Kanjana Laosen proposes a tool for Thai software requirements specifications in controlled natural language because natural language causes the ambiguities of software requirements specifications where ambiguities are often encountered both syntactic and semantic. For the same reasons Paul Brillant Feuto Njonko, to avoid ambiguities, presents his controlled language for business rules specifications.

Many contributions deal with the problem of machine translation and comparisons between many different languages. Thus Farouk Bouhadiba explains why Arabic remains a rather complex language to manage in the field of Natural Language Processing (NLP) in the case, for example, when translating legislative texts from Arabic into other languages like French or English. Many examples are given showing the differences between languages and the variations in a same language. Bilal Shafei proposes a model to Resolve Deviation Cases Generated by Inquiries in Arabic Language, and Rosita Chan has created an algorithmic dictionary to help finding some variations in Spanish according to Spanish speaking countries. Jorge Baptista shows the specificity of Predicate Noun in Subject Position in European Portuguese, and Krzysztof Bogacki talks about how the concept of manipulation is used in different disciplines. Navid Atar Sharghi compares the adverbial group in Persian and French and Dominique Berlet Noel *et al.*, explain the differences of pronunciation between English and French which lead to big difficulties for learners.

Dictionaries play a big part in Natural Language processing as already shown with Rosita Chan and his algorithmic dictionary, Leonel Ruiz

BULAG N° 40, 2022
Languages Analysis, Comparison and Generation Systems,
Models and Applications

Miyares *et al.* with the unique Cuban Spanish corpus and Yves Bordet whose dictionary helps calculating the difficulties in texts.

Neology is represented by Igor Skuratov *et al.*, and constructed languages by Thierry Lecolinet whose paper starts with a good state of the art. Christine Bricard, a musician-linguist, explains what music is in comparison with languages. Peter Greenfield was also a musician.

We find also four papers concerning the European Interreg DecRpit project about personal data recognition, where are discussed the most recent state of the art by Arnaud Witschger, the theoretical model and its demonstration by Sylviane Cardey and Peter Greenfield, how to recognise Unnamed Entities or Indirect Named Entities by Axelle Seiler *et al.*, and an intermediate computing kernel for the applications by Abdoulaye Guisse *et al.*

But some remaining questions may be unanswered yet: "is it possible to replace the human brain by a machine, and can human users and robots communicate safely", "how to represent meaning and implicit linguistic items"? Kyoko Kanzaki *et al.*, Isuri Nanomi Arachchige *et al.*, Rodolfo Delmonte, Axelle Seiler *et al.* and Armelle Jacquet propose their views, and relatedly Laurence Dahan-Gaida examines the epistemological consequences of bigdata and the digital for literary theory and history.

Acknowledgments

We wish to thank all the contributors.

These works are published with the financial support of the Interreg France-Zwitzerland programme.