

Report on ESF Short Visit Grant – Bilateral preparatory action

Host institute: Dr. Marzi, Pisa, IT

European Network on Word Structure: Cross-disciplinary Approaches to Understanding Word Structure in the Languages of Europe

Emmanuel Keuleers
Department of Experimental Psychology
Ghent University, Belgium

This is a scientific overview of my visit to Dr. Claudia Marzi from June 10th to June 14th, 2013 at the CNR in Pisa. Dr. Marcello Ferro, Dr. Vito Pirelli, and Professor Walter Daelemans also participated in the meetings conducted during this visit.

Purpose of the visit

The visit's aim was to gain substantial mutual knowledge of the different architectures for computational modeling used by the group in Pisa (Temporal Self Organizing Maps) and the groups at the University of Antwerp and Ghent University (Memory-based learning), and to identify areas in which the models could be complementary (developing new architectures sharing ideas from both models) and competitive (identifying tasks in which the models could compete). In addition, I also addressed methodological topics and discussed experimental research being done at Ghent University, specifically in the area of Bilingualism.

Description of the main results obtained

A good part of the visit consisted of a thorough introduction to the Temporal Self-Organizing Map (TSOM) architecture. The colleagues from Pisa gave an overview of the classical Self-Organizing Maps (SOMs) developed by Kohonen and described how adding a temporal layer allowed transformed the SOM into a symbol-by-symbol word storage map, with a strong neurobiological analogue in brain cell assemblies. After the new architecture was sufficiently clear, we examined a series of computational experiments in which the architecture was subjected to different tasks (word recognition and word reproduction in Italian and German). A big advantage of the visit was that we were able to tweak the models directly and as such gain a better understanding of their functioning. I was impressed with the ability of the TSOM architecture to store words without position dependent information and with its ability to align words with each other, again without using slots or position information. Both of these features are very desirable in computational models of language processing, where word representation is crucial. Usually this problem is side-stepped by using “hacks” which invariably turn out to be very local solutions, i.e., applicable to only a few problems. In the TSOM architecture, both issues are solved in a principled manner.

The visit continued with an overview of the Memory-Based Learning (MBL) architecture. A good deal of the discussion of MBL centered on the less-known inductive components of the model, i.e. feature weighting by information-gain and feature-value similarities derived by feature-class association computations, using the Modified Value Difference Metric. We then discussed a series of simulations modeling different types of psycholinguistic experimental data, mostly in the field of the acquisition of inflectional morphology.

The following part of the visit focussed on identifying broad directions for the collaboration. As a result of this discussion, we identified that the two architectures are more complimentary than competitive, with the TSOM architecture having strong points in lexical storage and the influence of low-level processes related to storage (e.g., reading aloud, acquisition of foreign-language phonotactics), while the MBL architecture relates more to the analogical processes operating over

the stored representations, (e.g., generalization of inflectional patterns).

Future collaboration with host institution (if applicable)

The final part of the visit focused on identifying and selecting tasks for future collaboration. We decided on two tasks:

1. The implementation of a memory-based learning model of German plural inflection using similarities derived from a TSOM.
2. Using the TSOM architecture to predict human word-similarity judgments experimentally collected by Hahn & Bailey. (Bailey, T. M., & Hahn, U. (2001). Determinants of Wordlikeness: Phonotactics or Lexical Neighborhoods? *Journal of Memory and Language*, 44(4), 568–591.)

Projected publications/articles resulting or to result from your grant

Since this was a very short visit, there were no definitive publications targeted. Our research groups will continue to collaborate in the very near future, and we aim to publish the results of the simulations studies (the two tasks mentioned above).