

## **Scientific Report Short Term Scientific Mission COST ISO703 - Veerle Baaijen**

*March 15<sup>th</sup> until the 26<sup>th</sup> of March 2010*

### **Purpose of the visit - Writing to learn, part 2**

The first study in my PhD research set out to investigate whether development of ideas varied as a function of planning and self-monitoring. In addition, process data was collected in order to investigate the processes responsible for the development of understanding through writing. Data collection for this study has been completed and the effects of planning and self-monitoring on developments in writers' subjective understanding and on idea change have been analysed.

The purpose of the STSM was to examine how these changes in ideas and subjective understanding relate to writing processes. Preliminary analyses of the process data indicated that changes in subjective understanding depend on the process by which writing is carried out. Specifically, the extent to which writers increased their understanding in the synthetic planning condition was strongly related to the extent to which they modified their texts as they were produced. There was no such relationship within the outline planning condition. During the STSM we aimed to develop a coding scheme for more detailed analyses of text production processes. In addition, we aimed to develop a programming script with Visual Basic in order to automate the analysis of production and revision phases in writing.

### **Description of the work carried out during the visit**

During the STSM the following tasks were carried out:

- (i) A coding scheme for the initial pauses was developed. The initial pause is defined as the time that writers pause before the first keystroke is logged. Within this analysis we have made a distinction between pauses that occur before the title is produced and pauses that occur before continuous prose is produced.
- (ii) I investigated whether ideas that were classified as new ideas on the list that was produced after writing were produced during outlining or during text production. This analysis was carried out in order to distinguish the process by which new ideas were produced during writing.
- (iii) A literature review of research on problem solving and memory related issues was carried out. In addition, David Galbraith gave an extremely helpful tutorial about the differences between the classical models of writing (e.g. Flower & Hayes and Bereiter & Scardamalia) and his alternative dual process model of writing. Also the predictions about text production processes made by the two models were discussed.
- (iv) Papers by Hayes (2009) and Chenoweth & Hayes (2001) were discussed in order to explore the possibilities for analysing keystroke logging data in terms of production and revision phases.
- (v) Some time was invested in developing a graphical representation of production and revision phases during writing. This graph should show the number of characters produced, the position of the cursor and the length of the document as a function of time. However, we were unsuccessful in the production of such a graph. We are currently discussing

how to resolve these issues with the authors of Inputlog which is the keystroke logging software that we used for our data collection.

(vi) We spent the majority of the STSM hand-analysing a small sample of key-stroke logs and the associated final texts in order to see how the recursive nature of the writing process can be captured, and in order to test the validity of the global text modification index that we have been using in previous analyses. The following markers and units were developed as a means of identifying different kinds of processing within the texts :

P: pause longer than 2 seconds  
PT: pause followed by text production  
TP: text production terminated by a pause longer than 2 seconds  
PR: pause followed by revision  
RP: revision terminated by a pause longer than 2 seconds  
M: movement  
Rins: revision of the kind insertion  
Rdel: revision of the kind deletion

These units of analysis were then used to identify features of the production and revision phases taking place during writing. Production and revision phases were distinguished by the point in the text at which they occurred, and the size (number of consecutive P and R bursts per phase) and frequency of the phases were calculated.

(vii) As a first step in automating the analysis, a literature review and some pilot studies were carried out in order to develop a programming script for the analyses of the keystroke logging files. I had some success in automatically identifying units within the key-stroke log, however further work is needed to distinguish whether these units occur within production or revision phases. This is currently in progress.

### ***Description of the main results obtained***

During this STSM the following results have been obtained:

- (i) Detailed coding schemes were developed: (a) to classify initial pauses in writing; (b) to identify production and revision phases in key stroke logs; (c) to identify new ideas in outlines; and (d) to parse texts into idea units. These coding schemes will be used to analyse the full set of texts and key-stroke logs. Further work is needed to automate the analysis of the key-stroke data.
- (ii) A range of questions about the key-stroke logging output were identified. These will be discussed with the authors of the software (Inputlog).
- (iii) I have improved my knowledge about writing models and the predictions of these models as they relate to my data.

### **Future collaboration with host institution**

This short term scientific mission has proven to be very productive. The opportunity to work together with David Galbraith from the Staffordshire University has helped me with the development of my understanding of possible ways to analyse process data as well as my

understanding of how ideas can be coded in outlines and texts. Furthermore, this Short Term Scientific Mission also gave me the opportunity to develop my understanding of writing research and it has helped me to develop my PhD project. Future collaborations will involve the development of further analytical methods, the preparation of jointly authored papers, and the design of further research.

**Projected publications/ articles resulting or to result from STSM**

A paper has been accepted for publication in the Proceedings of the 32th Annual Conference of the Cognitive Science Society. (Baaijen, Galbraith & de Glopper, (2010), Writing: The Process of Discovery). Several further papers about this research will be submitted to refereed international journals over the summer months.

**Confirmation by the host institute of the successful execution of the mission**

I confirm that the mission was productive and has led to preparation of several joint publications as well as the development of detailed plans for future collaborations between the two research teams. I can be contacted at [d.galbraith@staffs.ac.uk](mailto:d.galbraith@staffs.ac.uk) if further details are required.

**Other comments (if any)**