The experience of using Digital Replay System for Social Science research

[1]Patrick Brundell, [2]Dawn Knight, [2]Svenja Adolphs, [1]Shaaron Ainsworth, [2]Ronald Carter, [1]David Clarke, [3]Andy Crabtree, [3]Chris Greenhalgh, [3]Asad Naeem, [1]Claire O'Malley, [3]Tony Pridmore, [3]Tom Rodden and [3]Paul Tennent

Email pat.brundell@nottingham.ac.uk

[1]School of Psychology, University of Nottingham, Nottingham, UK, NG7 2RD
[2]School of English Studies, University of Nottingham, Nottingham, UK, NG7 2RD
[3]School of Computer Science, University of Nottingham, Nottingham, UK, NG7 2RD

Abstract

This paper explores the ways in which the Digital Replay System (DRS, innovative social science software, constructed as part of the National Centre for e-Social Science's Digital Records for eSocial Science node) has facilitated research from two distinct social science based end-users; learning scientists and corpus linguists. We discuss how DRS has been designed with the flexibility to allow users from a different methodological backgrounds to collaborate and re-use data sets.

We set out to describe how DRS addresses the basic software requirements of social scientists, for the processes of organising, replaying, annotating, coding, re-representing and analysing data. The paper discusses the real-life experiences of using DRS in order to combine qualitative and quantitative research methodologies, allowing users to address key research questions and problems. We present the key features of the Digital Replay System (DRS), in relation to the support it offers for the research activities of learning scientists, linguists and ethnographers. We describe the underlying ontology and illustrate how DRS can be used to import both raw and structured data, replay synchronised multimodal data and allow further structuring and coding. We will further illustrate how DRS can be used to create databases that are flexible and configurable to suit the needs of individual research programs. Finally, we touch on some of the ethical issues that e-Social Scientists need to address when using multimedia data.

Introduction

The community of researchers investigating human-human and human-computer interaction collect rich data sets that are becoming ever larger and more diverse as digital recording technologies increase in availability and ease of use. A wide range of both quantitative and qualitative research methodologies are used to investigate these multimodal corpora including ethnography, social network analysis, protocol analysis, and interventional classroom and laboratory experiments amongst others. There is a wide variation in the temporal resolution of studies, with data of interest ranging from minutes, for example to analyses turn-taking behaviour in dialogue to ethnographic studies lasting for years. The size and range of data sets produced by varying methodological approaches mean that it can be time consuming and difficult to relate multiple analyses (for examples see Ainsworth and Burcham, 2007 and Forsyth et al., 2006). Software tools are increasingly essential to assist in the analytic processes of organisation, replaying, structuring and annotation) of these growing data sets. However, many of the analytic requirements of social scientists have not to date been met by existing tools.

The requirements of social scientists fall into the following key categories.

- *Organisation of projects and analyses*
- *Synchronizing and replaying imported data*
- *Coding and annotation*
- *Analysis and Re-representation*
- *Export of data, structuring and results*

DRS has been designed to address some of the shortcomings of existing tools and offer a flexible solution to the needs of social scientists. Some features have been described in detail elsewhere (e.g. Greenhalgh 2007a,b) and this paper will focus on more recent functional developments. The functionality of DRS is described and grounded with data from a number of social science studies described in the following sections

The VIRILE Study

The study focused on student's experience of using VIRILE: Virtual Reality Polymerisation Plant software. The VIRILE software is a highly realistic and interactive simulation of a large scale industrial chemical processing plant (Schofield, Lester and Wilson, 2004)

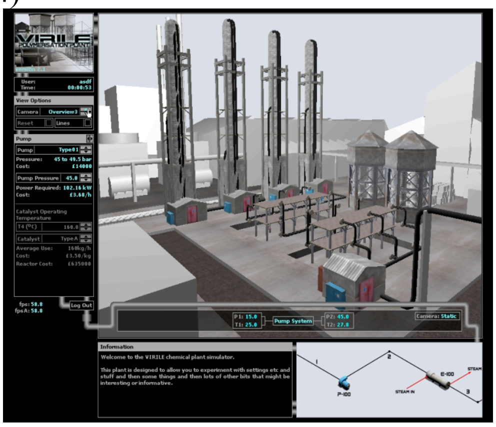

Figure 1. Screenshot of the VIRILE simulation

Though the specifics of the study are not relevant to this paper, to give an indication of the scope of data that DRS is intended to be able to assist in organising, synchronising, replaying, annotating and analysing, the study aims included; furthering students' understanding of how large scale industrial plants are designed and increasing their knowledge of what large scale processing components look like.

The study consisted of three parts. Part one included a pre-test of participant's ability to interpret and produce abstract, *process flow diagrams* and "realistic" drawings of plants. The second part was the learning experience of actively using VIRILE, consisting of two tasks to ensure the participants were able to use the software and had fully explored the visualisation, equipment and processes of the virtual chemical plant. The final task for the students was a test of their ability to interpret and produce process flow diagrams, scale drawings and their more general understanding of the VIRILE chemical plant.

In addition to tests of learning, recordings were made of students' use of the simulation through system logs and screen recordings. Video and audio recordings were also made of student's interactions with each other.

2-Party multimedia communication

Two strangers were paired together and asked each other a series of personal questions while communicating face to face, or using high quality audio and video or simply audio. The participants were allowed to lie or tell the truth in their answers. However, through a points scoring system, they were motivated to be believed at all times and were penalised if they were discovered lying. When asking questions, their task was to try and discriminate between the true and the false answers. It was recorded whether participants lied or told the truth for each question and also whether they were believed by their partner or not. High quality digital video and audio recordings were made of the participants in all communication media conditions. There were various experimental hypotheses, one of which was that success at lie detection would vary between media conditions. This hypothesis was supported, however, in order to try and explain this finding, it is desirable to analyse the behaviour and interactions of the participants.

2-Party academic supervisions

Multiple, digital and audio recordings were made of natural, unstructured conversation between a number of academic supervisors and their students. These recordings are part of a multimedia corpus created by linguists at the University of Nottingham, the NMC. This corpus comprises of 250,000 words of dyadic and single-party conversational data which has been recorded and transcribed.

Analysis using DRS

Project Management

Many of the problems found with collecting and organising social science data are a product of multiple methodologies and analyses, some of which are decided after data is collected and initially examined, but also due to the variation in temporal resolution of studies. Issues of concern common to all studies are;

- The huge variation in the kinds of data that the social sciences collect including (but not only) video and audio recordings of people's interactions with systems, environments and/or other individuals (including direct screen recordings), system log files, outputs from other recording devices such as eye-trackers and biophysical sensing equipment, still images and scans, dialogue transcripts and field notes
- The range of data used for analyses may extend from raw, unmodified data to tagged, annotated, coded or in other ways, structured data

The processes of acquiring, organising and accessing digital records need support for transparency and tractability. Therefore although some issues appear mundane, to extend the use and re-purposing of data and support tractability, digital data require extensive resources to be stored in it raw and modified states. For instance, the VIRILE data set consists of approximately 40 hours of digital video and audio of the participants, 20 hours of screen video recordings, 38 system log-files from and 76 sets of scanned test materials and dialogue transcriptions. Storage repositories need to be searchable, secure and well documented to allow for use, re-use and take account of ethical concerns.

DRS provides a project management mechanism whereby multiple files, annotation sets and coding schemes are organised and viewed as distinct "projects" and "analyses". A project might be all the records from a study or a trial. A DRS "analysis" is a set of related resources, usually co-temporal and may include audio and video (any format playable by Quicktime), transcripts, annotations, coding schemes and log files. DRS allows the linkage of data from individual and related studies whether recorded at the same time or related in other ways such as through experimental conditions, so that data and analyses may be accessed efficiently.

Synchronisation and replay of data

A time-line view is available for each analysis which gives a visual representation of the temporal extent of media files and associated annotation files, for example coding and transcriptions. The timeline is an abstract view of the files and annotations within an analysis. The temporal relationship between media and annotation files may be expressed automatically through explicit start dates/times or manually by specifying explicit start times in a synchronisation manager or by moving files in relation to each other along the timeline.

Transcriptions are stored internally as a set of non-overlapping free-text annotations and DRS uses time coding information to synchronise them with other media. Time coding may be performed within DRS through a dedicated transcription editor, within the track viewer as a coding track, or it can import files in Rich Text Format (RTF) created in other software such as Transana.

Playback of analyses is controlled via VCR-like controls allowing normal speed, slow-playback, looping and other such flexible control. Movement within the chronology is available through the time-line cursor and also through annotations files where the analyst may select text/codes and jump to the associated point in the records. Video, audio and transcription data from the VIRILE study are shown synchronised within a timeline viewer in figure 2 below.

Annotation and coding

Many social scientists develop their own coding scheme as a result of different analytical and theoretical commitments. In the majority of existing software, coding schemes usually need to be determined before coding starts. This is a critical issue with data used in the social sciences. It is generally necessary to make an initial "quick and dirty walk through" the data before constructing the coding scheme. The main purpose of this walk through is to make qualitative annotations (such as, timings, interesting features, analytic properties, et cetera.). The ability to annotate the data directly and use these annotations to find points of interest to later apply a finalised coding scheme is a particularly desirable feature of DRS allowing an iterative process of data structuring from qualitative through to quantitative descriptions to be used in multivariate statistical analyses.

The model of annotation defined within the DRS ontology is an extension of that suggested by Bird and Liberman (2001) in which an annotation consists of content associated with a subject. The content of annotations within DRS is text or user defined codes. The main subject of annotations is a region on the analysis timeline, the length of which is defined by the analyst.

Coding

DRS facilitates the structured coding of data. The user typically defines a hierarchical scheme of codes, which are descriptions of behaviour, dialogue etc. (see top left of figure 2 for an example coding scheme used to annotate the VIRILE data. The user then creates a coding track within the track viewer and applies the codes (see the track viewer at the bottom of figure 2 with media, transcription and coding tracks). Codes are defined with a range of temporal attributes and are applied to data either during playback (through keyboard shortcuts) or whilst static (by selecting timespans with the mouse). Code schemes may be modified at any point during the coding process. This flexibility supports the typical work patterns of interaction analysts where rapid development of an initial coding scheme that is applied to some or all of the data, is subsequently modified and re-applied depending on the descriptive results or visualisations of the data.

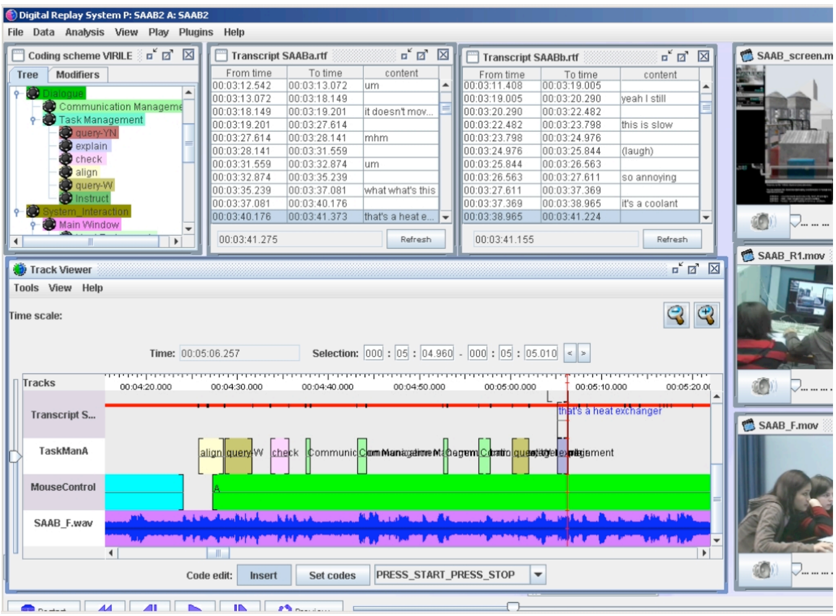

Figure 2. Coded VIRILE data

Searching text & annotations

The DRS interface has recently been integrated with a concordance tool facility, which provides the analyst with the capacity for interrogating data constructed from textual transcriptions anchored to video or audio, and from coded annotations. The concordancer can utilise specific words, phrases, or other lexical tags in addition to any other codes, as a 'search term'. Once presented with a list of occurrences, and their surrounding context (within the concordancer window), the analyst may jump directly to the temporal location of each occurrence within the video or audio clip. Further to this, the concordancer provides the commonplace frequency utility, giving raw counts of the frequency-of-use of each search term, again providing an invaluable impetus for the quantitative exploration of data.

However, for many interaction researchers, it is the complex *interactions* between dialogue, non-verbal behaviour and system activity that are of most interest. There has been an increasing interest in the relationship between certain nonverbal behaviours and the linguistic context in which they occur as researchers recognise that they can be tightly synchronised in natural language (e.g. to convey shared meanings, Bavelas and Chovil, 2006) The study of these relationships will lead to a greater understanding of the characteristics of verbal and non-verbal behaviour in natural conversation and the specific context of learning, and will allow social scientists to explore in more detail the relationships between linguistic form and function in discourse, and how different, complex facets of meaning in discourse are constructed through the interplay of text, gesture and prosody (building on the work of McNeill, 1992 and Kendon,1990, 1994). This interest coupled with the development of tools such as DRS is leading to a need for linguists, psychologists and learning scientists to share methodological approaches, results and data.

Since the concordance tool treats textual and coded annotations in the same way internally, the search functionality extends to all media within an analysis. This allows an investigation of all types of annotation in the same manner, applying the same skills and techniques used in the analysis of traditional corpora (for an example see Scott, 1999). This multimodal concordancer has led to the need for developing new approaches for coding and tagging language data in order to align textual, video and audio data streams (see Adolphs and Carter, 2007 and Knight, 2006).

5 hours of video data (from 2-party academic supervisions) was been coded for 5 categories of head nod types. The search facility allows these nonverbal codes to be searched in conjunction with a word or phrase and the instances with which these codes co-occur are highlighted. This function can be seen in the top right corner of the concordancer window seen in figure 3 below. This provides an easy to use frame of reference between the existence of verbal and non-verbal data.

This approach is also being used to encode and search sequences of hand movements in extended samples from the NMC. This will allow the linguists to explore the possible relationships between, firstly, the use of discourse markers in conversation, their attributed function and the co-occurrence of specific patterns of gesticulation. So for example we are investigating whether the discourse marker *anyway* which is commonly used as a conversational management token (so perhaps to signal a topic shift) is more frequently used with beat type gestures (see McNeill, 1992 for details) whereas the token *Oh* (which is commonly used as an interpersonal discourse marker, to signal that the listener is paying attention but perhaps requires more information from the speaker) is more frequently used with more cohesive or iconic forms of gesticulation. Through the input and searching of hierarchical codes of gesture form in relationship to specific discourse markers, such hypotheses can be approached more systematically than with manual methods of analysis alone.

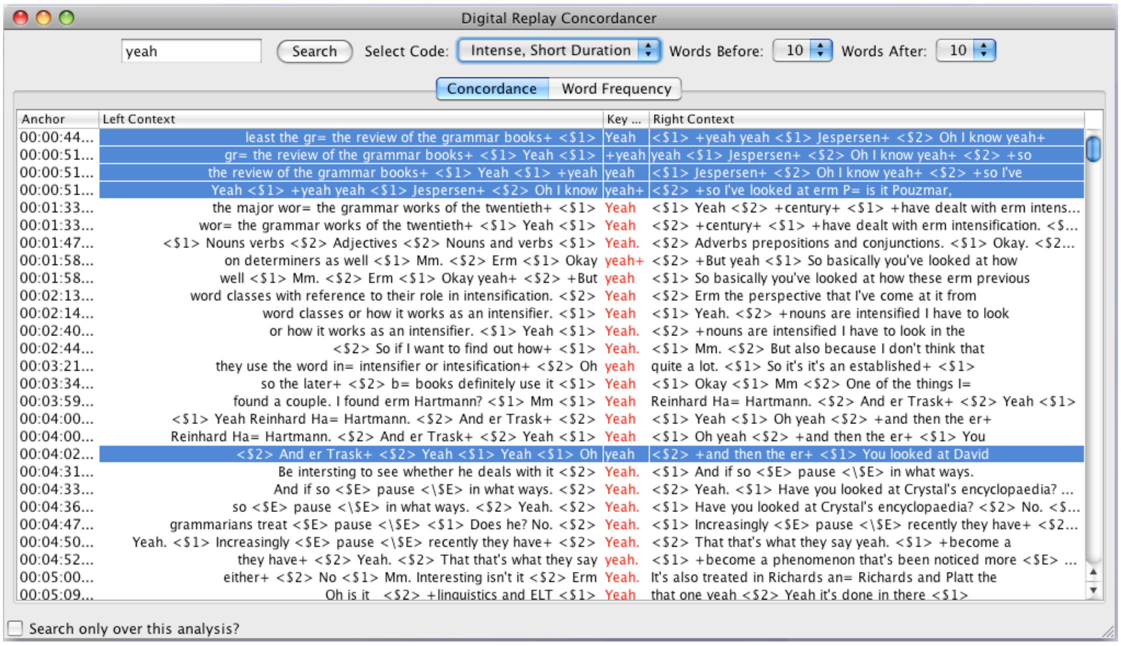

Figure 3. The DRS concordance tool

In Figure 3, occurrences of the word *yeah* have been searched for within the transcribed videos of a recorded academic supervision. Each occurrence is highlighted and provided in context within the concordance tool, and the turns which feature intense, short head-nods have been highlighted.

Export of structured coding

In common with some other coding tools, individual DRS code tracks may be exported as CSV format files. Along with features of the concordancer, this allows traditional code and count analyses. In the 2 party communication studies this can produce quantitative measures of, for example, the frequency of smiles and hand gestures, the number of words used in answers, the latency between question and answer or anything else that may have been coded in the textual, prosodic and non-verbal features of the interaction. However, as mentioned earlier, it is the relationship between dialogue and non-verbal behaviour that may be of interest to many. In the example DRS track viewer screen shots shown below (figure 4), video, audio, transcriptions and coding tracks are shown from the 2-party study investigating truth and lying under varying communication conditions.

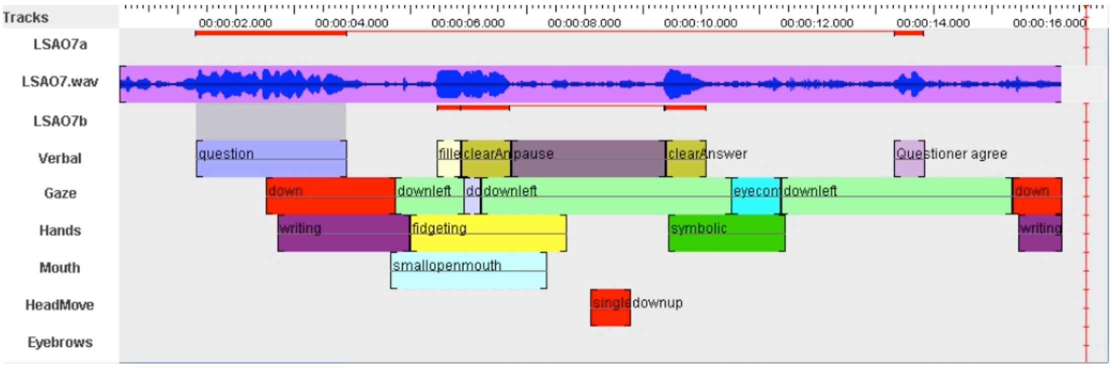

Figure 4. Multi-track coding of interactional data

Previous research into the detection of deception has suggested that there may be variations in nonverbal behaviour when people lie or tell the truth. There is evidence of an increased frequency of smiles (Vrij, Edwards and Bull, 2001) and fidgeting behaviour (DePaulo et al 2003). There have also been shown to be verbal behaviour changes such as more speech errors (Vrij and Mann, 2001) and fewer self-references (*I*, *me*, *my*, *we*) in deceptive communications (Newman et al, 2003). To date there appears to be little research that has investigated how the nonverbal and verbal behaviours of lying and truthful communications may vary together.

The track viewer provides a useful visualisation of the relationships between elements of non-verbal and verbal behaviour, however the ability to export multiple related annotation sets (coding tracks and/or transcripts) is a function of DRS that allows quantitative analysis of the data within excel or SPSS. The export processor takes an individual analysis as input and outputs (under user control) the subject and content of all annotations. Practically, this means that text, code and timing information is outputted into a tabular view window and in a form that allows the use in external statistical analysis packages. The main phases of export are;

- Segmentation – which determines what each row of the output table will contain and will be dependent on the analysis to be performed with the data. Segmentation is either *fixed period*, which divides the analysis into user defined time periods and outputs the contents of each coding track and transcription at each time period. This essentially shows what is occurring at regular intervals throughout an analysis. *Annotation-based* segmentation outputs data determined by the start or end (or start and end) of annotations and is effective for showing the times at which events start and end.
- Calculation – which determines the values each row of the output table contains which are usually either *textual* or *numerical*. *Text* output, where values are the names of the codes applied at the relevant times. Transcription tracks are always output in this manner. In *Numerical* output, the columns are named after the codes tracks and the numerous output values include, “duration”, the time within the row’s start and end time that the code is applied and “countstart”, the number of annotations labelled with this code that start during the time covered by the row.

The exploration of our data is in the early stages, however, we have transcribed and coded some 40 question and answer interactions for verbal and nonverbal behaviours. We expect to find some revealing relationships between the visible and non-visible experimental conditions, the behaviours of the participants and the judgements of truthfulness.

Ethical concerns

As corpora move from text to multimedia there are considerations as to how researchers may best able to explore and share the research opportunities of these data sets, whilst respecting the rights of the participants. It is relatively easy to anonymise textual records, it is considerably harder when records are in the form of audio or video. Indeed, anonymising multimedia through audio filtering or pixilation may render the content unusable. We propose that through informed consent, transparency in our relations with participants and taking seriously the role of "gatekeepers" digital records may used and re-used both successfully and ethically.

Conclusions and future plans

In this paper we have presented some features of the Digital Replay System, a tool designed to assist social science research. We describe some of the main features including; data management; synchronisation and replay of multimodal data; structuring, annotation and coding of data, and some of the tools to interrogate raw and structured data. We have touched upon some ethical concerns and suggest some ways forward.

Digital Replay System is in a rapid development cycle and within the next release due in September 2008 shall be an improved user interface along with refined and extended concordance and multi-track export tools.

Acknowledgments

This work was supported by the ESRC, UK through the grant "Understanding New Forms of Digital Records for e-Social Science" (DReSS node of the NCeSS) and by the EPSRC, UK, through grant EP/C010078/1, "Semantic Media" – Pervasive Annotation for e-Research and EQUATOR IRC, grant GR/N15986/01. Thanks to the other members of the DReSS node for their collaboration and also to staff from the School of Chemical and Environmental Engineering for their help in conducting the VIRILE study.

References

Adolphs, S. and Carter, R. (2007) "Beyond the word: new challenges in analysing corpora of spoken English", European Journal of English Studies, vol. 11 (2), pp. 114-128.

Ainsworth, S., and Burcham, S. (2007). The impact of text coherence on learning by self-explanation. Learning and Instruction, 17(3), 286-303

Bavelas and Chovil (2006) Nonverbal and Verbal Communication. Hand Gestures and Facial Displays as Part of Language Use in Face-to-Face Dialogue. In V. Manusov & M. Patterson (Eds.), Handbook of nonverbal communication (pp.97-115) Thousand Oaks, CA: Sage.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K. & Cooper, H. (2003) Cues to Deception. Psychological Bulletin. 129, 74-118

Forsyth, R., Ainsworth, S., Clarke, D., Brundell, P. and O'Malley, C. (2006) "Linguistic-computing methods for analysing digital records of learning", Online

Proceedings of the 2nd International Conference on e-Social Science, June 28-30, University of Manchester: ESRC NCeSS.
Greenhalgh, C., French, A., Tennant, P., Humble, J. and Crabtree, A. (2007a) From ReplayTool to Digital Replay System , Online Proceedings of the 3 rd International Conference on e-Social Science , October 7-9, Ann Arbor, Michigan: ESRC/NSF.
Greenhalgh, C., French, A., Humble, J. and Tennant, P. (2007b) Engineering a replay application based on RDF and OWL , Online Proceedings of the UK e-Science All Hands Meeting 2007 , September 10-13, Nottingham: NeSC/JISC.
Kendon, A. (1990) Conducting Interaction. Cambridge: Cambridge University Press.
Kendon, A. (1994). Do gestures communicate? A review. Research on Language and Social Interaction 27, 3: 175-200.
Knight, D. (2006) “Corpora: the next generation”, Corpus Building and Investigation for the Humanities (Online Introduction to Corpus Investigative Techniques), AHRC. www.humcorp.bham.ac.uk
Knight, D. and Adolphs, S. (2007) “Multi-modal corpus pragmatics: the case of active listenership”, Corpus and Pragmatics (ed. Romeo, J.), Berlin and New York: Mouton de Gruyter.
McNeill, D. (1992). Hand and Mind. Chicago: The University of Chicago Press.
Newman, M. L., Pennebaker, J. W., Berry, D. S. and Richards, J. M. (2003) Lying Words: Predicting deception from linguistic styles. Personality and Social Psychology Bulletin. 29, 665-675
Schofield, D., Lester, E. and Wilson, J.A., (2004) VIRILE: Virtual Reality Interactive Laboratory Experiments, Proceedings of Innovation, Good Practice and Research in Engineering Education Conference, Wolverhampton, 7th – 9th June.
Scott, M. & Johns, T. (1993) MicroConcord. Oxford: Oxford University Press.
Scott, M. (1999). Wordsmith Tools. Oxford: Oxford University Press
Vrij, A., Edwards, K. and Bull, R. (2001) People’s insight into their own behaviour and speech content while lying. British Journal of Psychology, 92, 373-389
Vrij, A. and Mann, S. (2001) Telling and detecting lies in a high-stake situation: The case of a convicted murderer. Applied Cognitive Psychology. 15, 187-203