Introducing DRS (The Digital Replay System): A tool for the future of Corpus Linguistic research and analysis

Knight, D. and Tennent, P.
The University of Nottingham
The School of English Studies, The University of Nottingham, University Park, Nottingham, NG7 2RD, UK
E-mail: aexdk3@nottingham.ac.uk, pxt@Cs.Nott.AC.UK

Abstract

This paper outlines the new resource technologies, products and applications that have been constructed during the development of a multi-modal (MM hereafter) corpus tool on the DReSS project (Understanding New Forms of the Digital Record for e-Social Science), based at the University of Nottingham, England. The paper provides a brief outline of the DRS (Digital Replay System, the software tool at the heart of the corpus), highlighting its facility to display synchronised video, audio and textual data and, most relevantly, a concordance tool capable of interrogating data constructed from textual transcriptions anchored to video or audio, and from coded annotations of specific features of gesture-in-talk. This is complemented by a real-time demonstration of the DRS interface in-use as part of the LREC 2008 conference. This will serve to show the manner in which a system such as the DRS can be used to facilitate the assembly, storage and analysis of multi modal corpora, supporting both qualitative and quantitative approaches to the analysis of collected data.

1. Introduction

This paper provides an overview of the developments made during the 3-year NCeSS (National Centre for eSocial Science, funded by the ESRC; Economic and Social Research Council) funded DReSS project that was based at the University of Nottingham. It highlights the outcomes of a specific ‘driver project’ hosted by DReSS, which sought to combine the knowledge of linguists and the expertise of computer scientists in the construction of the MM corpus software, known as DRS. DRS presents ‘data’ in three different modes, as spoken (audio), video and textual records of real-life interactions, accurately aligning them within a functional, searchable corpus setting (known as the Nottingham Multi-Modal Corpus: NMMC herein). The DRS environment therefore allows for the exploration of the lexical, prosodic and gestural features of conversation and how they interact in everyday speech.

2. Mono-Modal and MM Corpora

The impetus behind the development of the NMMC lies is the notion that current spoken corpora have a fundamental shortcoming: the fact that they represent all features of communication in the same format, that of transcribed textual records (see Knight, 2006). Since ‘the reflexivity of gesture, movement and setting is difficult to express in a transcript’ (Saferstein, 2004: 213), current spoken corpora therefore provide the linguist with little utility for the exploration of the non-verbal, gestural aspects of interaction (for examples of spoken corpora see the 5 million word CANCODE; Cambridge and Nottingham Corpus of Spoken Discourse, and spoken elements of the BNC; British National Corpus). Due to the fact that ‘we speak with our vocal organs, but we converse with our whole body’ (Abercrombie, 1963: 55), it is appropriate to call for a new generation of corpora to be developed, to allow for a more comprehensive view of the characteristics of language ‘beyond the text’ to be rendered.

So whilst existing corpus software tools have reached a high degree of sophistication insofar as they are capable of quickly searching through large databases of text and manipulating that text to produce word lists and concordance lines and a variety of statistical calculations related to the text, they are limited in the fact that the data that they can read has to be in text files (although there is a move towards xml in professional corpora). This means that the formatting in a corpus text is often a simplified version of the original text it represents.

Other software tools with the ability to store and represent MM data, that are currently available, tend to focus upon managing data or the processes of annotating previously collected data. They do not provide the utility for integrating all of the processes involved in corpus development, supporting the research process through each stage of data collection, mark-up, storage, management and analysis.

In addition to this, they lack the integration of concordancing software, which is traditionally used to allow the user to calculate frequencies, analyse collocates and often calculate statistical measures of the strength of word associations’ in order to provide the impetus for exploring the characteristics of specific tokens, phrases and patterns of language usage in their co-text and context (Conrad, 2002: 77-83).

Examples of such tools appear both commercially and academically. *Transana* (http://www.transana.org) is perhaps the most widely used, and focuses primarily on transcription of both audio and video. Developed by Noldus, *The Observer* (http://www.noldus.com) was designed originally for studying animal behaviour patterns, but has been adopted as a more general coding solution within the social sciences. Mangold International’s *INTERACT* (http://www.mangold-international.com) is another observational analysis solution that supports the process of coding videos, then provides some simple visualisation tools to support analysis of the coded data. One final commercial tool supporting real time coding of video is *Studiocode* (http://www.studiocodegroup.com), this time developed for Apple's OSX platform. Academic offerings include *I-Observe* (Badre et al., 1995: 101-113), an early project which used a video tape based system and made use of captured event streams to synchronise time-stamped events with the time-code on a video. The ever poular *ANVIL*, Developed in 2001 by Michael Kipp at the University of the Saarland was designed as a video annotation tool specifically for the purpose of analysisng multimodal corpora (Kipp, 2001: 1367-1370). The *Diver* Project, developed at Stanford University is another tool to support video annotation (Pea et al, 2004: 54-61). Designed to work with a single video, it nevertheless has a unique feature, that of so-called 'dives', where users can manipulate the viewpoint of a video using a virtual camera viewfinder, allowing zooming, panning and rotation of the original video data. Also developed at Stanford University *VACA* provides a toolkit for annotating or coding several simultaneous videos on a timeline representation (Burr, 2006: 622-627). Now in its third version ELAN was developed at the Max-Planck-Institute for Psycholinguistics (Brugman & Russel, 2004: 2065-2068). It is a fairly comprehensive tool for the annotation of video data, primarily in the field of linguistic research. It supports annotation in tiers, what other projects might call tracks, so several simultaneous annotations can be applied to a single piece of media.

Unlike any of these others, however, DRS has an concordancing tool integrated directly into the environment which allows users to not only search text-based records of language-in-use, but also the existence of gesture (by searching specific gesture codes) within and across the individual records of supervision sessions, and provides an account of concurrent language at the point of gesticulation. This provides a better utility for the exploration of a range of lexical, prosodic and gestural features of conversation, and for investigations of how such interact in real, everyday speech. Therefore, this novel concordancing application makes the DRS more finely attuned with the needs of corpus linguists than its contemporaries.

3. Introducing the DRS

As concordance lines are such a powerful research tool, it was a requirement of the applied linguists in the project that DRS should have the ability to generate them. There was a question as to whether this model could or should be applied to multiple modes of data and, if not, how should those modes be best represented in a re-usable operational interface. It was decided not to try to replicate concordance lines with video data due to the fact that, ultimately, multiple videos would mean smaller images for each event and it would not be possible to distinguish the subtle features of gesture or head movement, for example, with videos of this size.

Figure 1, below, shows the NMMC concordance tool operating within the DRS environment. This concordance tool allows the linguist to search across a large database of multimodal data (250,000 words) utilising specific tokens, phrases, patterns of language or gesture codes as a 'search term'. Once presented with a list of occurrences, and their surrounding context, the analyst may jump directly to the temporal location of each occurrence within the video or audio clip to which the annotation pertains. This concordancer will allow the linguist to research statistical or probabilistic characteristics of corpora, as well as to explore specific tokens, phrases and patterns of both verbal and non-verbal language usage in different file formats:

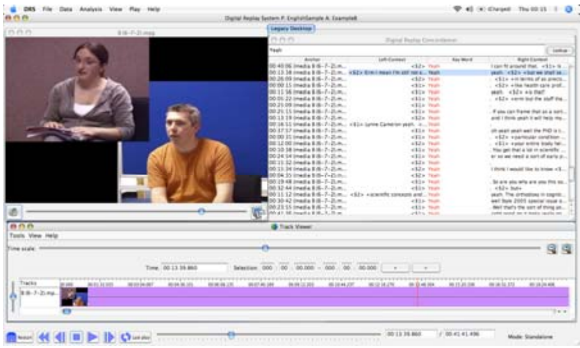

Figure 1: The concordance tool in use within the Digital Replay System environment.

Perhaps the most obvious difference between a standard text-based corpus and a MM corpus is the need to use a timeline as a means of aligning all the data streams. This may have originally been included as a logistical necessity, but in practice it allows a degree of flexibility that standard corpus software tools do not have.

The trend in corpus linguistics has been towards having all the data and metadata together in one file but DRS is much more flexible in this regard. Because it uses a timeline as an anchor, the user can attach as many transcripts or annotations to that timeline as is desired. This means that data and metadata can be stored in separate files, the text can be read by the user unobscured by the metadata and vice versa. The user can therefore use DRS as an analysis tool rather than just a read-only tool like existing software, making DRS a useful tool for a wide variety of users. This is obviously important in work such as ours which cannot be fully automated.

Data in DRS is therefore presented in a 'textured' way with windows of specific and relevant information being integrated and layered behind main frames that display the key search features in a similar way to current textual concordances. The user can organise the loaded data, which can include various different files formats, and subsequent analyses in the 'project browser' functionality, as well as store metadata information, citing where, when, under what conditions and with which participants and equipment the corpus data was collected.

A related advantage of the DRS concordancer is the fact that the input files can be in the rich text format version of Microsoft Word. This may be of limited benefit to professional corpus linguists but from our experience of working with students who are trying to build their own corpora they often find working in unfamiliar environments such as text files or xml off-putting. There may be the possibility to develop a corpus tool which is much more user-friendly for novice corpus linguistics studies or for researchers who wish to use corpus software but are not familiar with existing software such as WordSmith Tools.

The approach to setting out data is clear and easy to use. The user is able to view not only the concordancer but also a number of original transcripts and video and audio representations of the texts under investigation. This could potentially have a significant impact on the way corpus linguistics is conducted and mean that there could be a much more even balance between quantitative and qualitative methodologies as DRS allows the user to switch between these two approaches much more easily than second generation corpus software.

It is important to note that the concordancer is still being developed to provide frequency counts of the data. The integration of this utility will eventually allow the linguist to research statistical or probabilistic characteristics of corpora, as well as to explore specific tokens, phrases and patterns of language usage (both verbal and non-verbal) in more detail. The current version of the DRS concordancer allows the analyst to search across texts as well as within texts, and provides a reference to the text from which specific concordances were derived. This proves very useful because with this reference, the relevant cited text and attributing video can be opened within the analysis viewer, and specific words or encoded features of interest can be investigated in more detail. It is hoped that when the tracker is integrated within DRS, this feature can be used to allow the linguist to search for key terms and then the tracker on the associated video(s) in order to start to map relationships between language and gesticulation.

More generally, DRS provides support for a variety of different techniques for qualitative social science analysis. At its most basic level it allows a user to transcribe or annotate video and audio files, then play back those annotations synchronously with the reference media. DRS provides powerful tools for the synchronization of diverse temporal media types, such as video, audio, system logs etc. The number of simultaneous media files that can be played back is limited only by processor power and screen real estate, meaning several different views can be used to create a holistic approach to one's data analysis. DRS also provides facility for a complex in-line, hierarchical coding system, where media files can be coded either in real time by assigning pre-defined codes to key presses, or simply selecting areas of the timeline where a particular code should apply. When coupled with the output from the computer vision gesture recognition system, these 'coding tracks' can be automatically generated – thus saving the labour involved in coding by hand. DRS also provides a method of constructing multimedia *DRSDocuments* which contain references to both supporting media and corresponding points in the timeline allowing references within text to be used to link to particular associated media.

4. Defining gestures of interest

As a 'point of entry' for detecting and marking up gesture-in-talk in the NMMC to-date, preliminary linguistic analyses and classifications of each stream of data have been undertaken, and findings were cross-compared in order to determine patterns that may occur both within and across each data stream. This process is outlined below (adapted from Gu's 2-step approach to 'segmenting the MM text into various units', and 'annotating the units', 2006: 134):

- Defining and classifying verbal behaviour
- Defining and classifying non-verbal behaviour
- Combining the verbal and non-verbal, and highlighting the potential for exploring patterns and relationships between the two

A key question when applying MM mark-up, annotation and coding schemes is whether a MM approach will allow us to gain a better understanding of the relationship between the different modes. Whether or not the verbal replaces or supports the gestural stream in terms of discourse function requires a close examination of a relatively large data set, as well as further iterations of the development of categories of the coding scheme. In a further step it will then be possible to develop an integrated coding scheme which includes both verbal and gestural properties. This research is on-going (for preliminary results please refer to Carter & Adolphs, 2008; Adolphs, 2008; Knight et al., 2008).

In an attempt to mark-up the visual 'mode' of the data within DRS we have used a novel technique by which gestures are recognised and annotated by the system in order to automate the creation of these coded annotation tracks (this can later be encoded following rigorous analysis of the patterns between certain movement sequences). This is achieved by way of a tracking algorithm (developed by computer vision experts at the University of Nottingham) which can be applied to a video of a speaker and reports in each frame the position of, for example, the speaker's hands in relation to their torso, within a pre-defined granularity (as an Excel based output, defining the specific location of the tracker body part and the exact frame in which this occurs). Examples of the tracker are seen in figures 2 and 3.

The tracking algorithm reports the position of, for example, the speaker's mouth in relation to their eyes, in each frame (as seen in figure 2, for more information see Evans & Naeem, 2007). The circular nodes seen in figure 2 are the tracking targets (with a pre-defined granularity), which have the flexibility to allow the user to adjust the size of the tracked locations in relation to the specific size of, for example, the eyes and mouth of the participant (this has proved particularly useful when using close-up images in which participants have larger eyes and mouths). These targets are manually positioned at the start of the video and subsequently, as the tracking is initiated, a horizontal line is automatically drawn in the centre of these three nodes, marking an initial y-axis location (with position 0). Consequently, subsequent vertical head movements are denoted as causing a marked change in the y-axis in a + or – direction (+ being a head up movement and – being a head down movement).

Figure 2: The Head tracker in action

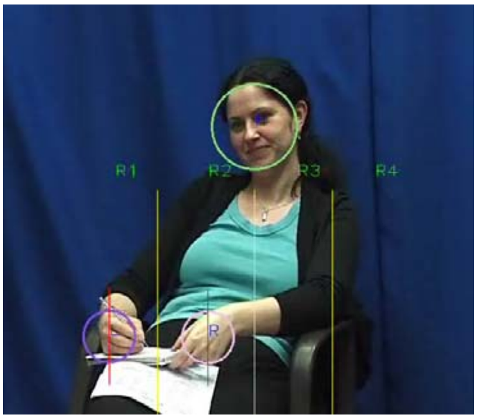

Figure 3: The hand tracker in action

The horizontal line also rotates to the left and right depending upon the position of the eyes, monitoring the angle of motion around the y-axis (when tracking head movements). The observation of the head angle from one tracked frame to the next proves invaluable to the analyst as such can help to reveal the characteristics of specific types of head movement (a feature that was integrated into the tracker as a result of various consultations with us), so for example such information can help to identify head shakes or head rotations as being distinct from a basic up-down head nod, should marked changes in the head angle be observed simultaneously with a marked change in the y-axis.

When tracking hand and body movement (figure 3), instead of the single horizontal line used as the point of movement reference in the head tracker, we are presented with three vertically positioned lines marking four zones on the image, R1 to R4 (R2 and R3 mark the area within shoulder width of the participant, acting as a perceived natural resting point for the arms, hence R1 and R4 mark regions beyond shoulder width). In consequence, the algorithm tracks the video denoting in which region the left hand (labeled as **R** by the tracker, since it is located to the right of the video image) and right hand (labeled as **R** by the tracker, since it is located to the right of the video image) are located in each frame. This provides an output similar to that given in figure 4.

	A	B	C
1	Number Of Tracked Objects = 3		
2	************************************		
3			
4	frame	LHand	RHand
5	1	3	3
6	2	3	3
7	3	3	3
8	4	3	3
9	5	3	3
10	6	3	3
11	7	3	3
12	8	3	4
13	9	3	4
14	10	3	4
15	11	3	4
16	12	3	4
17	13	3	4
18	14	3	4
19	15	3	4
20	16	3	4

Figure 4: An example of the output rendered when using the hand tracker.

In figure 4 the movement of each hand can therefore be denoted as a change in region location of the hand, so for example for Rhand (the left hand), we see a sequence of outputted zone 3 for frames 1 to 7, which changes to a sequence of zone 4 for frames 8 to 16. Ergo this notifies the analyst that the RHand has moved across one zone boundary to the right during these frames.

In theory, in order to track larger hand movements, the analyst can pre-determine a specific sequence of movements which can be searched and coded in the output data. So if, for example, the analyst had an interest in exploring a specific pattern of movement, considered to be of an *iconic* nature, i.e. a specific combination of the spontaneous hand movements which complement or somehow enhance the semantic information conveyed within a conversation, it would be possible to use the hand tracker to facilitate the definition of such gestures across the corpus (for in-depth discussions on iconics and other forms of gesticulation, also see studies by Ekman & Friesen, 1969; Kendon, 1979, 1982, 1990, 1994; Argyle, 1975; McNeill, 1985, 1992; Chalwa & Krauss, 1994 and Beattie & Shovelton, 2002). Obviously the analyst would be required to 'teach' the tracking system be means of pre-defining the combination of movements to be coded as 'iconic gesture 1', for example (so perhaps a sequence of RHand or LHand movements into from R1 to R4 and back to R1 across x amounts of frames), in order to convert the raw output into data which is both more meaningful and useable.

Further to this, it is viable to note that in order to further enhance the efficacy of the hand tracker, the current prototype not only outputs the hand locations across individual frames, but also provides an 'average' location of each hand across the span of one second. Whereas the head tracker was designed to deal with the most subtle of head movements, some of which may last for less than one second, the hand tracker is designed to deal with more emphatic, 'large' hand signals which may last 3,4 or 5 seconds, in addition to more subtle movements, as required.

Each potential gesture sequence (signified by the tracking output) is then labeled with a suitable code, and these codes are presented as a track of annotations anchored to the original video. Due to the fact the concordance tool treats textual and coded annotations in the same way internally, they may be presented simultaneously to the analyst and thus allow her to interact with both types of annotation in the same manner, applying the same skills and techniques appropriate to the use of traditional corpora when performing an analysis (for examples see Scott, 1999; Scott & Johns, 1993).

This automated approach to tracking hand gestures has two significant benefits over the manual analysis undertaken by McNeill and colleagues. Firstly, there is the potential to save a great deal of time through automation. The tracker can be run much faster than many other image tracking approaches, working at close to real time. It is also possible for the tracker to produce a variety of different outputs at the same time. Secondly, tracking techniques should be able to more accurately recognise the intensity of gestural movement than a human observer can.

Mark-up specific gesture sequences without the tracker necessitates a more labour intensive approach, similar to that taken by systems such as anvil and studiocode. A hierarchical 'coding scheme' must be defined in advance with necessary codes bound to particular keys – then as a media file (typically, but not necessarily, a video) is played back, these keys can be depressed to signify an instance of a particular gesture code. These codes are then stored in a 'coding track' which can be exported to packages such as SPSS for statistical analysis, or used in conjunction with the text in the concordancer within DRS, allowing analysis of instances of co-occurrence between codes and utterances. This process of gesture mark-up within the NMMC is still on-going.

5. Computational requirements and ethical concerns

A MM corpus does of course come with certain computational requirements. The automated gesture recognition algorithm requires considerable computing power to apply, and the project will look at means of making this realistically usable for a range of computing systems, in the coming few months. Additionally, the storage requirements for multi-media data are obviously significantly higher than those of text alone. With the NMMC aiming to cover 250,000 words, a complete copy would require circa 35 gigabytes to store. DRS provides network streaming facilities to allow this data to be stored on a central server, however keeping video in a network accessible location necessitates certain ethical questions.

The key ethical concern involves the question of the anonymity of participants and externals (i.e. people mentioned in the recordings). Anonymity is a relatively easy to address if using spoken language data to build a corpus, due to the fact that it is generally represented in a text based format and thus can be altered at the transcription phase of corpus development.
However, in MM corpora audio data is 'raw' in that it captures speech patterns of each participant which, as they are specific to an individual, exist almost an 'audio fingerprint', potentially making those involved easily identifiable. In order to allow the files to be adequately used for, for example, the exploration of phonetic patterns associated with particular word usage, any 'editing' of audio streams can result in data that is misleading or misrepresentative.

When handling video data the problem is compounded, since the nature of the recording techniques, which include up-close images of specific participants, makes it difficult to conceal the identities of participants. One traditional method would be to pixellate the video data, obscuring the faces of those involved using a technique such as the one proposed in (Newton et al., 2005). However, anything which obscures the features of the individual consequently reduces the usefulness of the video for understanding non-verbal communication. A more heavyweight approach might be to apply a selective encryption algorithm such as (Rodrigues et al., 2006) across the faces which can be decrypted to show the original video by those with the appropriate decryption key, while those without it will see only the obscured video. This would allow the videos to be freely available but still maintain the anonymity to those external to an accepted group of analysts.

Further to this, it is also necessary that there is a consistency between all three modes of data, as one of the fundamental aims of this project stresses, it is essential is that all modes should be equally accessible to corpus searches, allowing not only text-based linguists but also researchers investigating the use of gesture to access data. Therefore, for example, it would be counterproductive to exhaustively omit or alter details in the written transcript when the corresponding audio files remain unchanged. In effect, there is a need for the developer to strike a balance between the quest for anonymity in the data, and its usability and accuracy for research, a balance that is difficult that is far from straightforward to sustain.

6. Summary

The paper has outlined some of the key technical, ethical and practical problems and considerations faced in the development and exploration of MM corpora for the future of linguistic enquiry. It presents a novel MM corpus UI (user-interface), the DRS, which unlike it's contemporaries, allows users to organize, store, annotate, code and search (using the concordancer) audio, textual and video data between and across the different data streams. DRS therefore provides the linguist with an easy-to-use corpus tool-bench for the exploration of relationships between the linguistic characteristics and context of specific gestures, and the physically descriptive representations of those gestures extracted from video data. The study of this relationship leads 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 us 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 seminal work of McNeill, 1992 and Kendon, 1990, 1994).

7. Acknowledgements

The research on which this paper is based is funded by the UK Economic and Social Research Council (ESRC), e-Social Science Research Node DReSS (www.ncess.ac.uk/nodes/digitalrecord), and the ESRC e-Social Science small grants project HeadTalk (Grant No. RES-149-25-1016).

8. References

Abercrombie D. (1963). *Studies in phonetics and linguistics*. London: Oxford University Press.

Adolphs, S. (In press, 2008) *Corpus and Context: investigating pragmatic functions in spoken discourse.* Amsterdam: John Benjamins.

Adolphs, S. & Carter, R. (2007). Beyond the word: New challenges in analysing corpora of spoken English. *European Journal of English Studies 11*(2).

Argyle, M. (1975). *Bodily Communication*. London: Methuen.

Badre, A., Guzdial, M., Hudson, S. and Santos, P. (1995). A user interface evaluation using synchronized video, visualizations and event trace data. *Software Quality Journal, 4*(2):101–113.

Beattie, G. & Shovelton, H. (2002). What properties of talk are associated with the generation of spontaneous iconic hand gestures? *British Journal of Social Psychology 41*, 3: 403-417.

Brugman, H and Russel, A. (2004). Annotating multi-media / multi-modal resources with elan. In *LREC 2004*, pages 2065–2068.

Burr, B. (2006). Vaca: a tool for qualitative video analysis. In CHI '06: CHI '06 extended abstracts on *Human factors in computing systems*, pages 622–627, New York, NY, USA, 2006. ACM.

Carter, R. and Adolphs, S. (2008). Linking the verbal and visual: New directions for Corpus Linguistics. *Language and Computers special issue 'Language, People, Numbers', 64*; 275-291.

Chawla, P. & Krauss, R. M. (1994). Gesture and speech in spontaneous and rehearsed narratives. *Journal of Experimental Social Psychology 30*: 580-601.

Conrad, S. (2002). Corpus linguistic approaches for discourse analysis. *Annual Review of Applied Linguistics 22*: 75-95.

Ekman, P. & Friesen, W. (1969). The repertoire of non-verbal behavior: Categories, origins, usage and coding. *Semiotica 1*, 1: 49-98.

Evans, D. and Naeem, A. (2007). "Using visual tracking to link text and gesture in studies of natural discourse", Online Proceedings of the *Cross Disciplinary Research Group Conference 'Exploring Avenues to Cross-Disciplinary Research',* November 7, University of Nottingham.

Gu, Y. (2006). Multimodal text analysis: A corpus linguistic approach to situated discourse. *Text and Talk 26*, 2: 127-167.

Kendon, A. (1979). Some emerging features of face-to-face interaction studies. *Sign Language Studies 22*: 7-22.

Kendon, A. (1982). The organisation of behaviour in face-to-face interaction: observations on the development of a methodology. In Scherer, K.R. & Ekman, P. (eds) *Handbook of Methods in Nonverbal Behaviour Research*. Cambridge: Cambridge University Press.

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.

Kipp, M. (2001) Anvil - a generic annotation tool for multimodal dialogue. *In Proceedings of the 7th European Conference on Speech Communication and Technology (Eurospeech),* pages 1367–1370, Aalborg.

Knight, D. (2006). 'Corpora: The Next Generation', Part of the AHRC funded online *Introduction to Corpus Investigative Techniques*, The University of Birmingham. http://www.humcorp.bham.ac.uk/

Knight, D., Evans, D., Carter, R. and Adolphs, S. (Forthcoming, 2008). Redrafting corpus development methodologies: Blueprints for 3rd generation "multimodal, multimedia'' corpora. *Corpora Journal*.

McNeill, D. (1985). So you think gestures are nonverbal? *Psychological Review 92*, 3: 350-371.

McNeill, D. (1992). *Hand and Mind.* Chicago: The University of Chicago Press.

Newton, E.M., Sweeney, L. and Malin, B. (2005). 'Preserving Privacy by De-Identifying Face Images', *IEEE Transactions on Knowledge and Data Engineering , 17* (2), pp. 232-243.

Pea, R., Mills, M., Rosen, J., Dauber, K., Effelsberg, W., and Hoffert, E. (2004). The diver project: Interactive digital video repurposing. *IEEE MultiMedia, 11*(1):54–61, 2004.

Rodrigues, J. M., Puech, W., Meuel, P., Bajard, J.C. and Chaumont, M. (2006). 'Face protection by fast selective encryption in a video', *The Institution of Engineering and Technology Conference on Crime and Security 2006*.

Saferstein, B. (2004). 'Digital technology and methodological adaptation: Text on video as a resource for analytical reflexivity', *Journal of Applied Linguistics 1* (2), pp 197–223.

Scott, M. & Johns, T. (1993) *MicroConcord.* Oxford: Oxford University Press.

Scott, M. (1999). *Wordsmith Tools*. Oxford: Oxford University Press.