

Interpreting the image: using advanced computational techniques to read the Vindolanda Texts

Melissa Terras

*School of Library, Archive and Information Studies, University College London,
London, UK*

m.terras@ucl.ac.uk

Abstract

Purpose: This paper provides an overview of the development of a computer system designed to aid historians in the reading of the stylus tablets from the Roman fort of Vindolanda. The different stages in developing the system are outlined, and preliminary results given.

Methodology/Approach: Literature review regarding Vindolanda, stylus tablets, and the process of reading an ancient document. Knowledge Elicitation techniques used to explicitly model expert processes used to read an ancient document. Corpus construction: gathering a corpus of character forms, and lexicostatistics. Implementing an advanced cognitive imaging system utilising Artificial Intelligence techniques which uses the explicit model of how experts read ancient text, and the corpora collated, to produce plausible interpretations of the document.

Findings: This paper describes the developmental stages undertaken to construct a system which can read in images of an ancient document and produce plausible interpretations of the document, to aid the historians in the lengthy process of reading an ancient text. In carrying out the development, an explicit representation of how experts approach and reason about damaged and deteriorated texts was formulated, and a large corpus of letter forms and linguistic data was captured, providing experts with a valuable dataset and understanding of their own tools and processes. Preliminary results from the resulting computer system are presented which demonstrate the usefulness of the technique, although more work is needed to develop this into a stand alone computer system.

Research limitations: Focused on the Roman stylus tablets from Vindolanda, near Hadrian's Wall, although the technique could be extrapolated to cover other types of ancient documents from any period should corpora regarding letter forms and linguistic data be available.

Practical implications: It is demonstrated that using techniques from Artificial Intelligence and Cognitive Psychology can result in further explicit understanding of humanities expert processes, which allow computational systems to be constructed. The resulting computational system is a tool for the humanities expert, which carries out a task in a similar manner, allowing for faster reasoning time and quicker hypotheses development.

Originality/value: Presents the first known system to intakes an image of an ancient text and output a plausible interpretation of the text in a reasonable time frame, assisting the papyrologist in resolving ambiguities in the damaged and abraded text.

Keywords: Vindolanda, Knowledge Elicitation, Artificial Intelligence, Cognitive Imaging System, Expert Processes.

1. Introduction

The ink and stylus texts from Vindolanda are an unparalleled source of information regarding the Roman Army and Roman occupation of Britain for historians, linguists, palaeographers, and archaeologists. The visibility and legibility of the handwriting on the ink texts can be improved through the use of infrared photography. However, due to their physical state, the stylus tablets (one of the forms of official documentation of the Roman Army) have proved almost impossible to read.

This paper describes a system designed to aid historians in reading the stylus texts: in the process developing what appears to be the first system developed to aid experts in

reading an ancient document. This system brings together years of research in image analysis, Artificial Intelligence (AI) architectures, papyrology, and palaeography to make a complete AI system that intakes a signal and outputs a symbol. The system is not an "expert system" that automatically "reads" and provides a transcription of the texts, rather it is a papyrologist's assistant that mobilizes disparate knowledge, such as linguistic and visual clues, and uses these to speed up the process by which an expert can arrive at the most likely interpretation of a text.

Before designing and building any tools to aid papyrologists in the reading of texts, it is a necessary requirement firstly to ask: just what does a papyrologist do when trying to read and understand an ancient text? A proposed model of how experts read ancient documents is presented, and this representation was subsequently used as a basis for the development of the computer system which can aid historians in the reading of the Vindolanda texts by "reading" texts in the same manner that they do themselves. The architecture of the system was developed from a system previously designed as a proof of concept to analyse aerial images, and the system was trained on linguistic and palaeographic corpus data gathered from previously read Vindolanda texts.

It can be demonstrated that the system takes in an image of unknown text as input and outputs plausible interpretations of that text by propagating various solutions to the problem and presenting the expert with the most likely results. This aids the papyrologist in sorting through differing hypotheses whilst trying to read an ancient text, which is often a lengthy and difficult process. The system presented here is currently being developed into a stand-alone desktop system, and it is envisaged that the tools will be applicable to different linguistic systems should there be data sets to train the system on.

Section 2 of this paper provides background information regarding the Vindolanda texts, and Section 3 discusses related research undertaken on the stylus tablets and stresses the need for the development of a cognitive visual system to aid in the reading of the tablets. Section 4 provides an overview of the Knowledge Elicitation techniques utilised in order to construct an explicit model of the processes used to read an ancient text, Section 5 details the gathering of corpus information and the construction of models to train the system, and Section 6 describes the architecture used to construct the system and presents results from the system in action. A summary of this research is given in Section 7 which demonstrates the contribution this research has made towards constructing a stand alone system to aid papyrologists in their task.

2. The Vindolanda texts

The discovery of the tablets at Vindolanda[1], a Roman Fort built in the late 80s AD near Hadrian's Wall at modern day Chesterholm, has provided an unparalleled resource regarding the Roman occupation of northern Britain and the use and development of Latin around the turn of the first century AD. Textual sources for the period in British history from AD 90 to AD 120 are rare, and the ink and stylus tablets are a unique and extensive group of documents providing a personal, immediate, detailed record of the Roman Fort at Vindolanda from around AD 92 onwards (Bowman and Thomas, 1983, 1984, 2003; Bowman, 1999).

The ink tablets, carbon ink written on thin leaves of wood cut from the sapwood of young trees, have proved the easiest to decipher. In most cases, the faded ink can be seen clearly against the wood surface by the use of infrared photography, a technique used frequently in deciphering ancient documents (Bearman and Spiro, 1996). The majority of the six hundred writing tablets that have been transcribed so far contain personal correspondence, accounts and lists, and military documents (Bowman and Thomas, 1983, 1994, 2003).

The two hundred stylus tablets found at Vindolanda appear to follow the form of official documentation of the Roman Army found throughout the Empire (Turner 1968; Fink 1971; Renner 1992). It is suspected that the subject and textual form of the stylus tablets will differ from the writing tablets as similar finds indicate that stylus tablets tended to be used for documentation of a more permanent nature, such as legal papers, records of loans, marriages, contracts of work, sales of slaves, etc. (Renner 1992), although the linguistic aspects of the tablets will be similar as they are contemporaneous documents from the same source, probably written by the same scribes.

Take in Figure 1. Stylus tablet 836, one of the most complete stylus tablets unearthed at Vindolanda

Manufactured from softwood with a recessed central surface, the hollow panel of the stylus tablets was filled with coloured beeswax. Text was recorded by incising this wax with a metal stylus, and tablets could be re-used by melting the wax to form a smooth surface. Unfortunately, in nearly all surviving stylus tablets[2] the wax has perished, leaving a recessed surface showing the scratches made by the stylus as it penetrated the wax[3]. In general, the small incisions are extremely difficult to decipher. Worse, the pronounced wood grain of the fir wood used to make the stylus tablets, staining and damage over the past two thousand years, and the palimpsestic nature of the re-used tablets further complicate the problem; a skilled reader can take several weeks to transcribe one of the more legible tablets, whilst some of the texts defy reading altogether. Prior to this research, the only way for the papyrologists to detect incisions in the texts was to move the text around in a bright, low raking light in the hope that indentations would be highlighted and candidate writing strokes become apparent through the movement of shadows, although this proved frustrating, time consuming, and insufficient in the transcription of the texts.

3. Research and the Vindolanda texts

Science, University of Oxford, to analyse these texts (Bowman *et al.*, 1997; Schenk, 2001; Molton *et al.*, 2003; Brady *et al.*, 2005; Pan *et al.*, 2004). These techniques involve scrutinising the document surface using image processing techniques to provide new information that the human eye cannot see. However, these techniques do not provide a means of sorting through this information to provide the most likely interpretation of the data. Therefore, it was necessary to develop a computer system to aid the historians in the reading and interpretation of these images themselves, to

speed up the reading process: the system is designed to work in combination with these image processing algorithms and the experts themselves.

4. The process of reading an ancient document

In order to identify the tools that could be built to aid the papyrologists in their transcription of the Vindolanda tablets, it was first necessary to try and gain an understanding of what the papyrology process actually entails. Although the readings generated from ancient documents provide one of the major primary information sources for classicists, linguists, archaeologists, historians, palaeographers, and scholars from associated disciplines, surprisingly little research has been carried out regarding how an expert constructs meaning from deteriorated and damaged texts (Terras, 2002)[4]. Techniques borrowed from the field of Knowledge Elicitation (McGraw and Harbison-Briggs, 1989; Waterman 1986), a structured process used to gain understanding of the behaviour of experts in order to train expert systems, were used to gather quantitative and qualitative information about how papyrologists work, resulting in an in-depth understanding of the ways different experts approach and reason about damaged and abraded texts. Firstly, as with all knowledge acquisition tasks, the domain literature was researched, and any associated literature was collated. Three experts were then identified who were working on the ink and stylus texts, and who were willing to take part in this investigation. The experts were observed whilst going about their tasks, and unstructured interviews were undertaken, where the experts described their domain, and the individual processes and techniques that they preferred. More structured interviews were then undertaken, when the experts were asked to describe particular facets of their task, such as the identification of letter forms and the role of grammar, word lists, and external historical and archaeological resources in the reading of the documents. A series of Think Aloud Protocols (TAPs) were then undertaken (a technique adopted from experimental psychology, where the expert is urged to utter every thought that comes to mind whilst undertaking a specified task) and the experts were given structured tasks to complete. These sessions were recorded, transcribed, and analyzed using Content Analysis techniques (Terras 2002, 2005). These exercises demonstrated that the experts use a recursive reading mechanism which oscillates between different levels, or modules, of reading, and the process was rationalized into defined units, to develop a connectionist model of how papyrologists approach and start to understand ancient texts.

The proposed model of how experts read ancient text is presented in Figure 2. An expert reads an ancient document by identifying visual features, and then incrementally building up knowledge about the document's characters, combinations of characters, words, grammar, phrases, and meaning, continually proposing hypotheses, and checking those against other information, until s/he finds that this process is exhausted. At this point a representation of the text is prepared in the standard publication format. At each level, external resources may be consulted, or be unconsciously compared to the characteristics of the document. Although a simple representation, the model shows the overall scope of the process of reading an ancient text: further, more detailed procedural analysis regarding every level (or "agent") in this system is presented in Terras (2002).

5. Gathering corpus information

In order to train a computational system to function in the way that the model, above, does, it was necessary to collate information regarding the type of letter forms and language used at Vindolanda in such a way that it could be incorporated into the system. The Vindolanda ink tablet corpus is the sole major contemporaneous linguistic resource to the Vindolanda stylus tablets, and this was used to generate a corpus of textual data and letter forms regarding Latin of this period. The palaeographic data was then manipulated to produce a set of "average" characters which the system uses for comparison with unknown stroke input data.

5.1 Textual corpus

The Vindolanda ink corpus used comprised of the 230 Latin texts published in Bowman and Thomas (1994), plus 56 new texts that had been read in preparation for the publication of the next volume of the Vindolanda texts (Bowman and Thomas 2003). There were 27,364 characters (excluding space characters) in total, comprising 6,532 words, or word fragments. In the corpus there were 2,433 unique word tokens (1,801 words appeared only once, the rest were repeated). This provided an adequate corpus on which to base any conclusions about the language used at Vindolanda[5].

Additional linguistic analysis of the Vindolanda ink texts was undertaken (Terras, 2002) using a common Corpus Linguistics program for lexical analysis, WordSmith (<http://www1.oup.co.uk/elt/catalogue/Multimedia/WordSmithTools3.0/>), in order to generate lexicostatistics such as word lists, letter frequency, and bigraph frequency. This provided statistics regarding the language used at Vindolanda, which could then be integrated into the system described here.

5.2 Image corpus

The ink and stylus tablets from Vindolanda use a form of handwriting commonly referred to as Old Roman Cursive. There is a paucity of documents from this period of the Roman Empire, and so the development of the character forms in ORC is the subject of much academic debate (see Bowman and Thomas, 1983, 1994, 2003). In order to train the system, a selection of characters from the Vindolanda ink texts previously read by Bowman and Thomas were encoded using a markup scheme and a tool developed to annotate aerial images to train artificial intelligence systems (Terras and Robertson, 2004). This resulted in an XML representation of each character on a stroke by stroke basis which captured information about stroke length, position, direction, and relation to other strokes. In total, 1,506 individual characters from the ink tablets, and 180 characters from the stylus tablets were annotated, providing a training set for the system.

Characters were annotated firstly by drawing around the outline of a character using a computer mouse, and assigning it a character label. Individual strokes were then traced, and numbered by selecting the option from a drop down menu. Stroke ends were then identified and labelled. Finally, stroke junctions were noted and assigned labels. An example of this is shown in Figure 3, using the letter S from the start of ink tablet 311 as an example.

Take in Figure 3. Steps taken in annotating a letter

These annotations are preserved in an XML file which textually describes the annotated image. The placing of the strokes is preserved by noting the co-ordinates of each feature. The annotation of the corpus represents approximately 300 hours' worth of work, and is the largest palaeographic corpus of Old Roman Cursive in existence. The corpus can also be used to analyse the ORC letter forms (Terras, 2006).

5.3 Generating models from the corpus

The primary aim of building such a corpus was to provide data with which to drive an Artificial Intelligence system, which could input unknown text and output probable interpretations. To match unknown to known characters, it is necessary to have models of those known characters to compare the unknown data to. Annotating the character forms in the way described above meant that vector co-ordinates for all the characters had been captured within the XML data, and so character models for each type of character could be generated.

A character model is defined as a probability field that indicates the likely placing of one or more strokes of a two-dimensional character, producing a general representation of a character type. Unknown characters can then be compared to a series of these models, and the probability that they are an instance of each one calculated, the highest probability indicating a match.

On a conceptual level, the (stroke-based) character model is constructed by taking an image of an individual character, finding its bounding box (identifying the rightmost x co-ordinate, and the leftmost x co-ordinate, and the highest and lowest y co-ordinates), and transforming this into a standardised (21 by 21 pixel) grid. The stroke data is blurred slightly to produce a generalised model which allows specific instances to be compared to it more successfully[6]. Each standardised representation is accumulated onto a generalised matrix for each character type: resulting in a generalised representation of each type of character. These are subsequently used as the models to which unknown characters are compared. An example of how these steps combine to generate a character model is given in Figure 4, where a small corpus which contains three 'S' characters is used to generate a character model of an S.

Take in Figure 4. The generation of a character model from a small corpus

Character models of every character contained within the dataset were generated, as shown in Figure 5:

Take in Figure 5. Character models generated from the training set of the ink text corpus

These models were used to train the Artificial Intelligence system to effectively "read" in stroke data of unknown text and output possible, reliable, interpretations of that text. By combining these stroke models with the linguistic data and advanced

Terras, M. (2006). "Interpreting the image: using advanced computational techniques to read the Vindolanda texts", *Aslib Proceedings*, Volume: 58 Issue: 1/2 Page: 102 - 117.

Artificial Intelligence techniques, a system was constructed that was able to replicate the process the historians went through to output possible interpretations of the Vindolanda texts.

6. Developing an AI based system

AI has repeatedly shown that interpreting our world requires bringing to bear a great deal of world knowledge. The process of reading an ancient document, especially documents that are in a very bad state of disrepair, such as the Vindolanda stylus tablets, is an especially good example of the need to mobilize a great deal of *a-priori* knowledge. It is clear that the remnants of writing on the tablets themselves contain insufficient information to recover the original written text, but by applying sufficient linguistic knowledge of character shapes, word and letter frequency, and grammatical information, the tablets can sometimes be read. Implementing a system that can automate the process requires an architecture that can fuse the many different kinds of knowledge in order to arrive at an estimate of the most probable interpretation.

Numerous architectures have been proposed to solve interpretation problems: the HEARSAY system (Erman, *et al.* 1980) was developed for speech recognition and more recently Hidden Markov Models (HMM) have been used in Natural Language Processing (NLP) (Charniak, 1993) but these approaches have drawbacks. Blackboard systems can be hard to control, and HMMs can be too restrictive to incorporate complex and diverse kinds of knowledge (Robertson, 1999, 2001; Robertson and Laddaga 2002, 2004).

To construct a computational model of the papyrology process, we adopted, and adapted, the GRAVA architecture, developed by Dr Paul Robertson[8], which was built to provide a flexible backbone upon which image interpretations problems can be solved. The GRAVA (Grounded Reflective Adaptive Vision Architecture) system was originally developed to analyse aerial satellite images (Robertson, 1999, 2001). The GRAVA architecture uses Minimum Description Length combined with Monte Carlo Select algorithms to find a fair way of comparing data across semantic levels, and is implemented in YOLAMBDA, a dialect of LISP. To test his system, Robertson constructed a small program which could effectively "read" a nursery rhyme, utilising data regarding letter form, letter frequency, and word frequency, to propagate the most likely interpretation of the image of the text. This system was adapted considerably to be able to work with the Vindolanda data, and the architecture and mathematical underpinnings of the adapted system developed for this project are described in detail in Robertson, *et al.* (2005), Terras and Robertson (2005), and Terras (2006). It is enough, here, to show a model of the system architecture, which demonstrates that it replicates the lower modules, or "agents" of the papyrology process, using AI techniques to sort through the data to propagate the most likely interpretation of the image.

All agents work with data generated from the Vindolanda corpus, as discussed in section 5: the Character Agent uses the set of character models plus data regarding the frequency of letters in the Vindolanda corpus, whilst the Word Agent utilises a list of words generated from the documents read at Vindolanda so far.

The basic schematic of the system is shown in Figure 6. To input an image of an unknown document, this image has to be annotated (either manually or automatically, using the image processing algorithms) to identify the key features, and the stroke data prepared for comparison (for more details see Robertson, *et al.*, 2005). The stroke data is then passed on to the Character Agent, which compares this data to the character models (generated previously from the Annotated Corpus), and calculates a description length based on which character it is most likely to fit, and the frequency of characters which appear in the corpus ("i" is much more likely to appear than "q", for example). One of these characters is selected by a random process, and is passed, with its description length, to the Word Agent. As characters are passed to the Word Agent, they form a string of characters, these resulting "words" are then compared to known words in the corpus, and a description length is calculated for how well they fit known words from Vindolanda. The description length from the Character Agent and the Word Agent is then added, giving a global description length for that run of the system. Subsequent runs of the system generate different global description lengths, as the random process in selecting which letters are passed to the Word Agent assures that different combinations of characters that may fit are put forward as possible interpretations of the problem. When the description lengths of these separate runs of the system are compared, the Minimum Description Length gives the overall most likely interpretation of the test image. The system has effectively sorted through the different permutations of the data, and provided the experts with the most likely solution to the problem.

6.1 The system in action

When implemented in practice, the system outputs the most likely interpretation of the texts, as can be seen in Figures 7 and 8. An annotated image is entered into the system, and after eight iterations, the system outputs the correct interpretation of the text, keeping note of the alternative possible hypothesis it has also thrown up.

Take in Figure 7. A section of ink tablet 255 which reads "ussibus puerorum meorum"[8], used as input.

Take in Figure 8. Output from first successful run on the section of 255, indicating that the system resolves upon the correct interpretation of the image.

More remains to be done to make this a useful tool for the papyrologists. The word lists contained within the system need to be expanded, to allow for fragmentary sections of words which may not appear in the existing corpus. The process of annotating a document and feeding it into the system is time consuming, and requires an understanding of the annotation tools themselves. The system will require development to become a stand-alone desktop application, as it currently depends on the Knowledge Engineer and Engineering Scientist being present to run it (this development is being undertaken at the Department of Engineering Science and the Centre for the Study of Ancient Documents at Oxford University, in collaboration with the School of Library, Archive and Information Science at University College London). Nevertheless, the system is a working tool which can take in images of texts

and output plausible interpretations of those texts, which is speeding up the process of reading the Vindolanda stylus texts by propagating realistic interpretations of complex data. The system is also being developed to work with the image processing algorithms, so that particularly damaged and complex incisions of unknown text can be automatically fed into the system (Robertson *et al.*, 2005). Other types of graphical and linguistic textual systems could also be analysed using this technique, provided the statistical data regarding letter forms, word frequency, and letter frequency could be obtained.

Conclusion

This research has discussed the development of a computer system to aid historians in the reading of ancient texts, from initial demonstration of the problem, work with experts to ascertain their needs, the gathering of relevant information and statistical data, and identifying, adopting and adapting a technology which enables the disparate knowledge used to read such texts to be mobilised. As such it is a proof of concept. More development needs to be undertaken before this system becomes a stand alone application, but its success is grounded on the fact that it has such a rich dataset to draw information from, and that working closely with the papyrologists from the outset has enabled the construction of a system that behaves in a similar manner to the experts themselves, making this a useful and intuitive tool which speeds up the time taken to read an ancient text.

There were further unexpected rewards from this research. The data set of Old Roman Cursive letter forms has proved useful in determining palaeographic differences between the characters which appear on the stylus and ink texts (Terras, 2006). The modelling of the papyrology process is the first detailed analysis of how experts who read ancient texts operate (Terras, 2002, 2006), allowing an insight into how texts which are so important to our understanding of history and society can actually be deciphered and read. From a computer science point of view, the Vindolanda texts provided real world data to test the implementation of a system which had just been applied to a simplistic problem (the nursery rhyme data). The technologies used have demonstrated that information from differing semantic levels, such as image and word data, can be compared in a reasonable time frame utilising our comparison technique and this could be applied to comparing any sort of differing data in computer systems (up until now, it has been time consuming and computationally expensive to compare and contrast differing semantic data computationally).

This research is also a demonstration of fruitful collaboration between differing academic tribes and territories: the use of advanced computational techniques in a humanities discipline benefiting both the scholars in the humanities and the computer and engineering scientists. Future work will entail delivering this system as an application for papyrologists to use, aiding the experts in their complex and time consuming task.

Notes

1. See Bidwell (1985) for further information regarding Vindolanda, Birley (1999) for an account of the discovery of the tablets, and Bowman (2003) for an introduction to the content of the Vindolanda texts.
2. It is estimated that around 2,000 of such tablets exist outside Egypt (Renner 1992).
3. Only one stylus tablet, 836 (see Figure 1), has been found so far with its wax intact. Unfortunately this deteriorated during conservation, but a photographic record of the waxed tablet remains to compare the visible text with that on the re-used tablet. (The figure shown here is of the tablet after the wax was lost).
4. Aalto (1945), Youtie (1963, 1966) and Bowman and Tomlin (2005) are the only discussions found (as yet) which try to describe what the papyrology process actually entails, with some higher level discussion available in Turner (1968).
5. The representativeness of a Corpus has been much discussed in the field of Corpus Linguistics (Kenny, 1982; Biber, 1983; Oostdijk, 1988; Biber, 1990), the consensus being that "small is beautiful" and that "there is every reason to make maximal use of these corpora [of 2,000 word length] for analysis of linguistic variation until larger corpora become readily available" (Biber 1990, p. 269) The major corpus of American English, The Standard Corpus of Present Day Edited American English (known as the Brown Corpus, see Francis and Kuchera (1979)) uses 2,000 word samples to represent its various genres. Zipf's "law" (Zipf, 1935, reprinted 1965), which predicts an "ideal" set of frequencies (and hence probabilities) for lexical items given a particular vocabulary size, maintains that a dictionary size of 1,000 words will give a percentage cover of 56.220 per cent of the language, although this is based on an analysis of English language texts. In the case of this research, the corpus used is 100 per cent of the known corpus of Latin for this period.
6. That is to say, the stroke data is convolved with a Gaussian Blur operator to reduce over-fitting.
7. Formerly of the Robots Group, Department of Engineering Science, University of Oxford. Now Research Scientist of the Computing Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology.
8. Meaning "for the use of my boys", referring to some cloaks and tunics for Clodius Super's *pueri*.

References

- Aalto, P. (1945), "Notes on methods of decipherment of unknown writings and languages", *Studia Orientalia, Edidat Societas Orientalis Fennica*, XI.4.
- Bearman, G.H. and Spiro, S. (1996), "Archaeological applications of advanced imaging techniques", *Biblical Archaeologist*, Vol. 59 No. 1, pp. 56-66.
- Biber, D. (1983), "Representativeness in corpus design", *Literary and Linguistic Computing*, Vol. 8 No. 4, pp. 243-57.
- Biber, D. (1990), "Methodological issues regarding corpus-based analysis of linguistic variation", *Literary and Linguistic Computing*, Vol. 5 No. 4, pp. 257-69.
- Bidwell, P.T. (1985), *The Roman Fort of Vindolanda at Chesterholm, Northumberland*, Historic Buildings and Monuments Commission for England, London.
- Birley, R.E. (1999), *Vindolanda IV: Fasc. IV, Writing Materials*, Northumberland, Roman Army Museum Publications, Greenhead, Carvoran.

Terras, M. (2006). "Interpreting the image: using advanced computational techniques to read the Vindolanda texts", *Aslib Proceedings*, Volume: 58 Issue: 1/2 Page: 102 - 117.

- Bowman, A.K. (1999), "The Vindolanda Writing-Tablets 1991-4", *Atti del XI Congresso Internazionale di Epigrafia Greca e Latina, Roma 18-24 Settembre 1997*, Vol. I, pp. 545-51.
- Bowman, A.K. (2003), *Life and Letters on the Roman Frontier, Vindolanda and its People*, 2nd ed., British Museum Press, London.
- Bowman, A.K. and Thomas, J.D. (1983), *Vindolanda: The Latin Writing Tablets. (Tab. Vindol. I.)*, Society for Promotion of Roman Studies, London.
- Bowman, A.K. and Thomas, J.D. (1994), *The Vindolanda Writing-Tablets (Tab. Vindol. II)*, British Museum Press. London.
- Bowman, A.K. and Thomas, J.D. (2003), *The Vindolanda Writing - Tablets (Tab. Vindol. III)*, British Museum Press, London.
- Bowman, A.K. and Tomlin, R.S.O. (2005), "Wooden stylus tablets from Roman Britain", in Bowman, A.K. and Brady, M. (Eds), *Images and Artefacts of the Ancient World*, Oxford University Press, Oxford.
- Bowman, A.K., Brady, J.M. and Tomlin, R.S.O. (1997), "Imaging incised documents", *Literary and Linguistic Computing*, Vol. 12 No. 3, pp. 169-176.
- Brady, M., Pan, X., Terras, M. and Schenk, V. (2005), "Shadow stereo, image filtering and constraint propagation", in Bowman, A.K. and Brady, M. (Eds), *Images and Artefacts of the Ancient World*, Oxford University Press, Oxford, pp. 15-30.
- Charniak, E. (1993), *Statistical Language Learning*, Massachusetts Institute of Technology Press, Cambridge, Massachusetts.
- Erman, L.D., Hayes-Roth, F., Lesser, V.R. and Reddy, D.R. (1980), "The HEARSAY-II speech understanding system: integrating knowledge to resolve uncertainty", *Computing Surveys*, Vol. 12 No. 2, pp. 213-253.
- Fink, R.O. (1971), *Roman Military Records on Papyrus*, Philological Monographs of the American Philological Association, No. 26.
- Francis, W.N. and Kucera, H. (1979), *Brown Corpus Manual*, Providence, Brown University, available at: <http://helmer.aksis.uib.no/icame/brown/bcm.html> (accessed 25 February 2005).
- Kenny A. (1982), *The Computation of Style, An Introduction to Statistics for Students of Literature and Humanities*, Pergamon Press, Oxford.
- McGraw, K.L. and Harbison-Briggs, K. (1989), *Knowledge Acquisition: Principles and Guidelines*, Prentice-Hall International Editions, London.
- Molton, N., Pan, X., Brady, M., Bowman, A.K., Crowther, C. and Tomlin, R. (2003), "Visual enhancement of incised text", *Pattern Recognition*, Vol. 36, pp. 1031-1043.
- Oostdijk, N. (1988), "A corpus linguistic approach to linguistic variation", *Literary and Linguistic Computing*, Vol. 3 No. 1, pp. 12-25.
- Pan, X., Brady, M., Bowman, A.K., Crowther, C. and Tomlin, R.S.O. (2004), "Enhancement and feature extraction for images of incised and ink texts", *Image and Vision Computing*, Vol. 22 No. 6, pp. 443-451.
- Renner, T. (1992), "The finds of wooden tablets from Campania and Dacia as parallels to archives of documentary papyri from Roman Egypt", *Proceedings of the 20th International Congress of Papyrologists*, Copenhagen, 23-29 August 1992.
- Robertson, P. (1999), "A corpus based approach to the interpretation of aerial images", *Proceedings of 7th International Congress on Image Processing and its Applications (IEE IPA99)*, Manchester, Vol. 2, pp. 527-531.

- Terras, M. (2006). "Interpreting the image: using advanced computational techniques to read the Vindolanda texts", *Aslib Proceedings*, Volume: 58 Issue: 1/2 Page: 102 - 117.
- Robertson, P. (2001), *A Self Adaptive Architecture for Image Understanding*, D.Phil Thesis, Department of Engineering Science, University of Oxford, Oxford.
- Robertson, P. and Laddaga, R. (2002), "Principal component decomposition for automatic context induction", in Ishii, N. (Ed), *Artificial and Computational Intelligence, Proceedings of the IASTED International Conference*, Tokyo, Japan, pp. 243-250.
- Robertson, P. and Laddaga, R. (2004), "The GRAVA self-adaptive architecture; history; design; application; and challenges", *24th International Conference on Distributed Computing Systems Workshops (ICDCS 2004 Workshops)*, Institute of Electrical and Electronics Engineers Computer Society, pp. 298-303.
- Robertson, P., Terras, M., Pan, X., Brady, J.M. and Bowman, A.K. (2005), "Image to Interpretation: An MDL Agent Architecture To Read Ancient Roman Texts", submitted.
- Schenk, V. U. B. (2001). *Visual Identification of Fine Surface Incisions*. D.Phil Thesis, Department of Engineering Science. Oxford, Oxford University.
- Schenk, V.U.B, and Brady M. (2003), "Visual identification of fine surface incisions in incised Roman stylus tablets", *ICAPR 2003, International Conference in Advances in Pattern Recognition*.
- Terras, M. (2002), *Image to Interpretation: Towards an Intelligent System to Aid Historians in the Reading of the Vindolanda Texts*, D.Phil. Thesis, Department of Engineering Science, University of Oxford, Oxford.
- Terras, M. and Robertson, P. (2004), "Downs and acrosses, textual markup on a stroke based level", *Literary and Linguistic Computing*, Vol. 19 No. 3, pp. 397-414.
- Terras, M. and Robertson, P. (2005), "Image and interpretation: using artificial intelligence to read ancient Roman texts", *HumanIT*, Vol. 7, p. 3.
- Terras, M. (2005), "Reading the readers: modelling complex humanities processes to build cognitive systems", *Literary and Linguistic Computing*, Vol. 20 No. 1, pp. 41-59.
- Terras, M. (2006), *Image to Interpretation: An Intelligent System to Aid Historians in Reading the Vindolanda Texts*, Oxford University Press, Oxford, in press.
- Turner, E.G. (1968), *Greek Papyri, An Introduction*, Clarendon Press, Oxford.
- Waterman, D.A. (1986), *A Guide to Expert Systems*, Addison-Wesley, Reading, Massachusetts.
- Youtie, H.C. (1963), "The papyrologist: artificer of fact", *Greek, Roman and Byzantine Studies*, Vol. 4, pp. 19-32.
- Youtie, H.C. (1966), "Text and context in transcribing papyri", *Greek, Roman and Byzantine Studies*, Vol. 7, pp. 251-8.
- Zipf, G.K. (1935), reprinted 1965, *The Psycho-Biology of Language*, Massachusetts Institute of Technology Press, Cambridge, Massachusetts.

Acknowledgements

Thanks must go to the supervisors of my doctoral thesis, from which this paper stems: Professor Alan Bowman (Centre for the Study of Ancient Documents, University of Oxford), and Professor Mike Brady (Department of Engineering Science, University of Oxford). The GRAVA system was adopted and adapted with the consent and help of Dr Paul Robertson (formerly of the Robots Research Group, Department of Engineering Science, University of Oxford, and now Research Scientist of

Terras, M. (2006). "Interpreting the image: using advanced computational techniques to read the Vindolanda texts", *Aslib Proceedings*, Volume: 58 Issue: 1/2 Page: 102 - 117.

Massachusetts Institute of Technology Computing Science and Artificial Intelligence Laboratory) and this research could not have come to fruition without his help and generosity. Thanks must also go to the anonymous papyrologists guinea-pigs who gave considerable time and effort to allow this study to take place.



Figure 1. Stylus tablet 836, one of the most complete stylus tablets unearthed at Vindolanda. The incisions on the surface can be seen to be complex, whilst the wood grain, surface discoloration, warping, and cracking of the physical object demonstrate the difficulty papyrologists have in reading such texts.

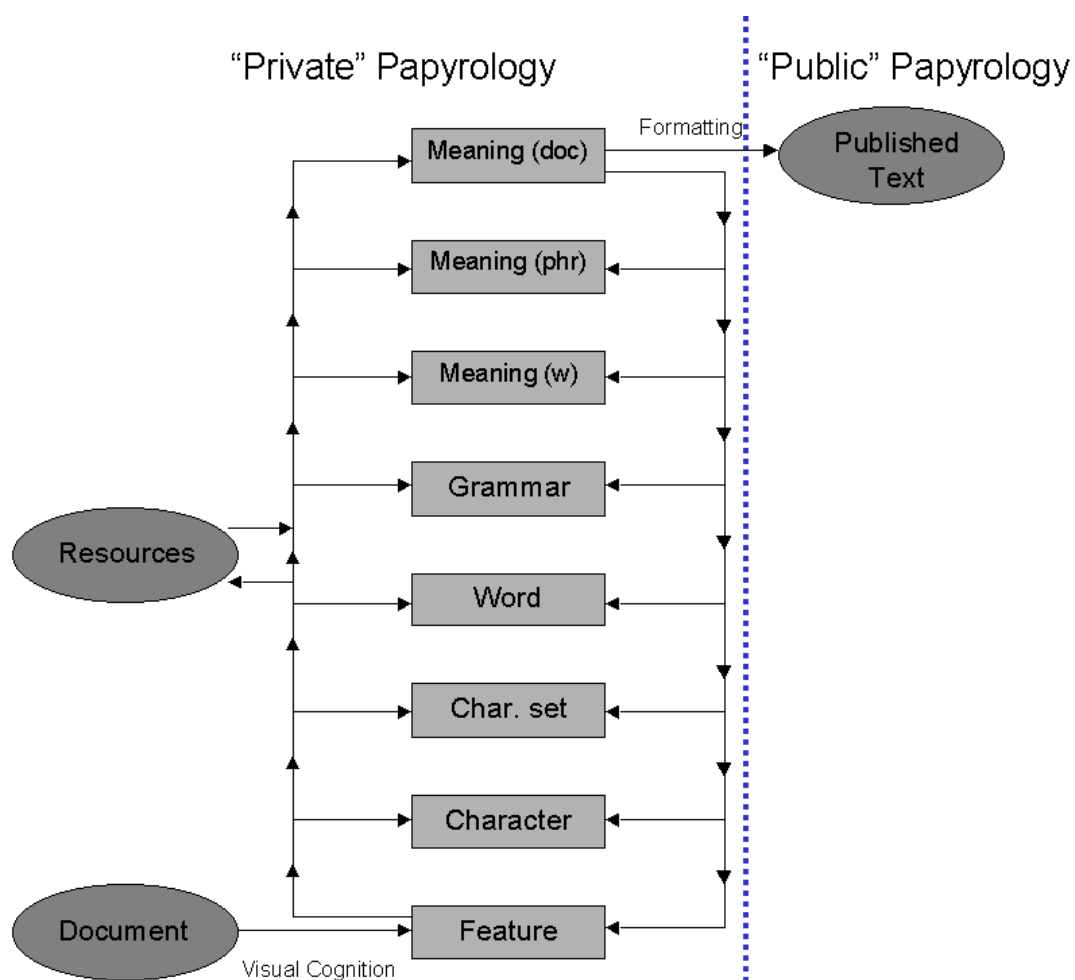


Figure 2. The proposed model of how experts read an ancient text. Public papyrology (Youtie, 1963) refers to the published reading of texts in a common format, after the reading has taken place. Private papyrology is the implicit process the experts undertake when reading a text.

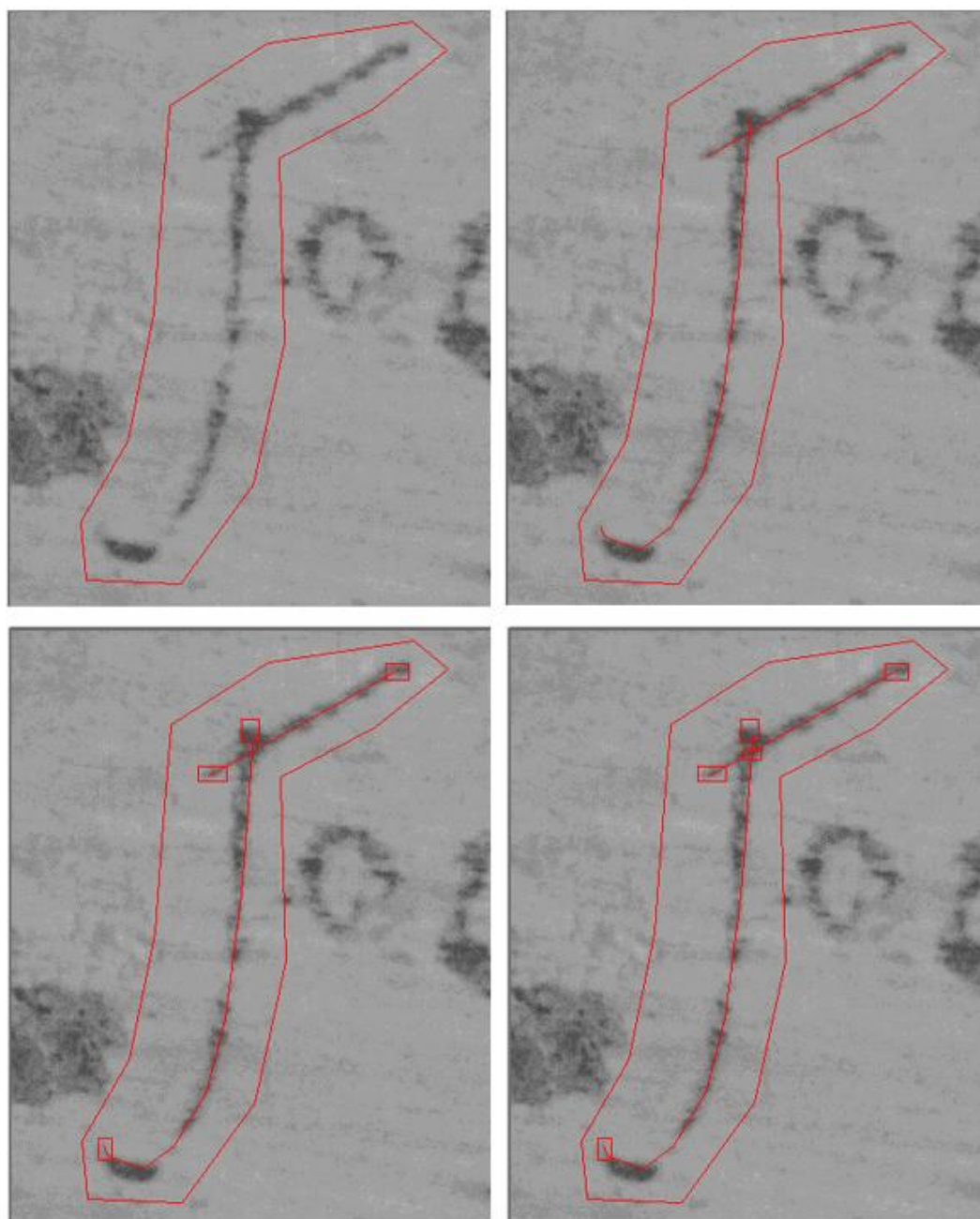


Figure 3. Steps taken in annotating a letter. The outline is traced at first, followed by individual strokes, stroke endings, and stroke meeting junctions. Further information about the markup used to annotate these characters can be found in Terras and Robertson (2004).

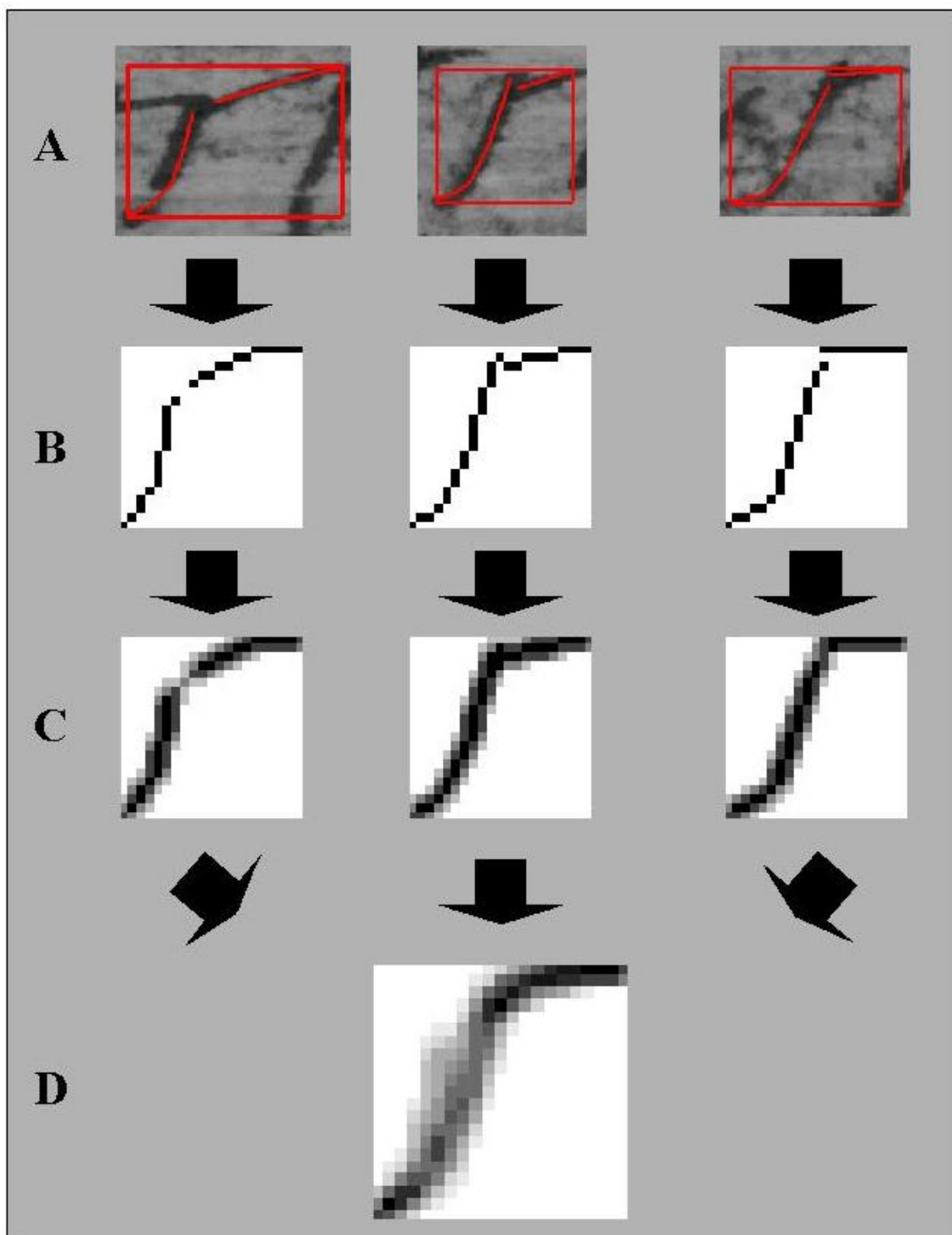


Figure 4. The generation of a character model from a small corpus. Three letter S's are identified, and bounding boxes drawn around them (A). The stroke data is then transformed into a 21 by 21 pixel grid (B). A Gaussian Blur is applied (C). The composite images generate a character model (D). The darker the area, the higher the probability of the stroke passing through that pixel. In this way, the probabilities of the stroke data occurring are preserved implicitly in the character models.

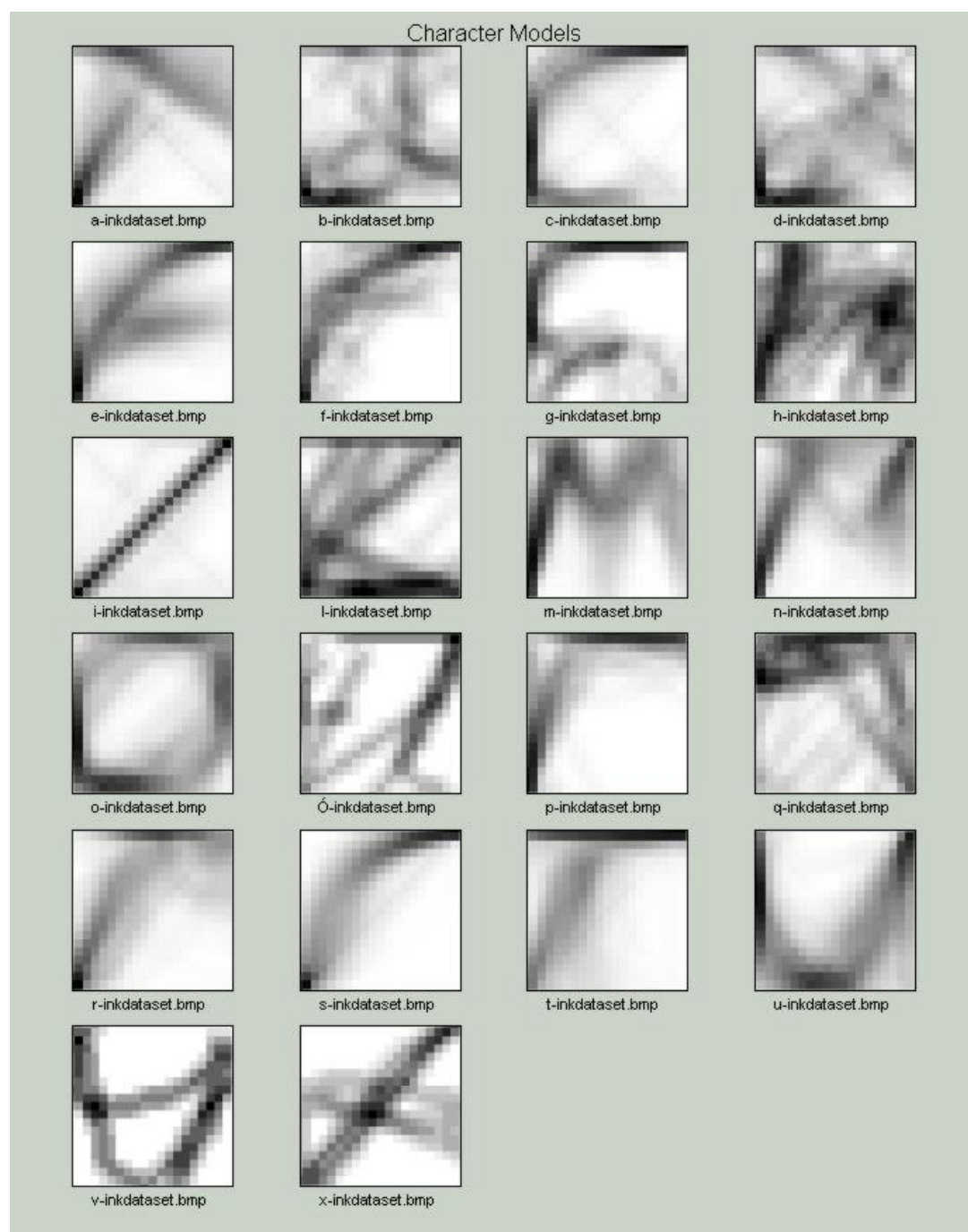


Figure 5. Character models generated from the training set of the ink text corpus. The darker the area, the higher the certainty that a stroke will pass through that individual section of the character box. The letter M, for example, can be clearly seen, suggesting that it is made in a similar fashion each time it is used. The letter H is messier, but this is partly due to the number of letter Hs available to make the character models: the more letters available, the higher the accuracy of the character model. There is no letter j, k, w, y or z, as they were not part of the Old Roman Cursive alphabet.

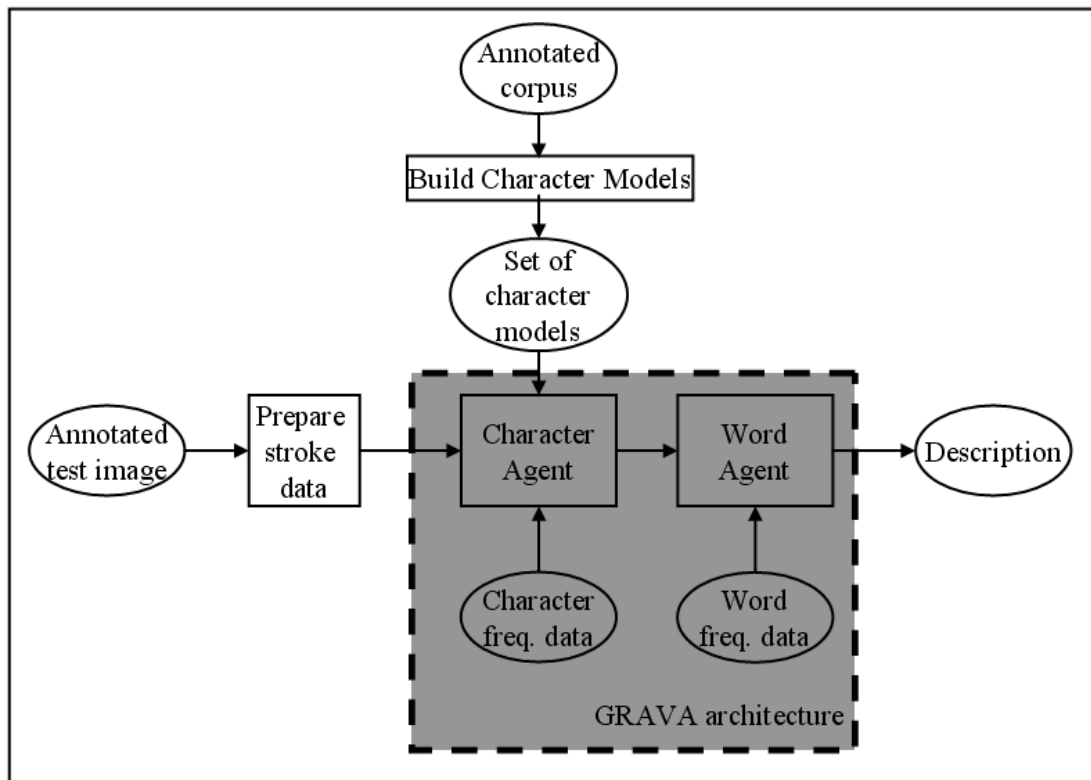


Figure 6. Basic schematic of system. Robertson's GRAVA architecture is highlighted to indicate the processes which are carried out as part of the final run of the system.

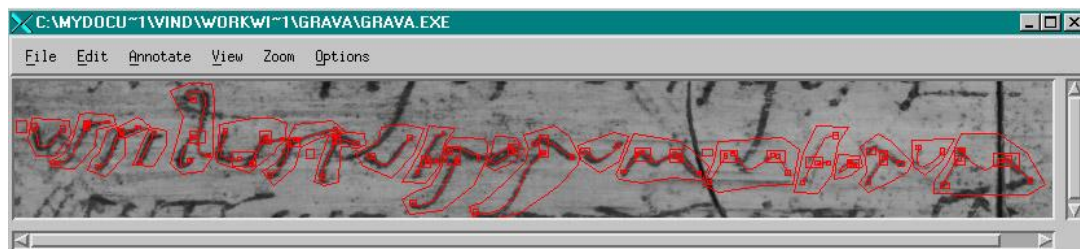


Figure 7. A section of ink tablet 255, which reads "ussibus puerorum meorum"[8], used as input.

```
Grounded Reflective Agent Vision Architecture (GRAVA) Version 2.0.  
Yolambda listener pushed. Type :exit to return to GRAVA.  
  
=> ... load the system and the data ...  
  
=> (runCycles 25)  
iteration 0 DL=440.220794 interpretation=( ... ((2482 252) (2517 250))) ... )  
iteration 1 DL=64.085075 interpretation=( u r s i b u s p u e r o r u m m n o a u m )  
iteration 2 DL=49.374412 interpretation=(ussibus puerorum m n o r u m )  
iteration 3 DL=48.831413 interpretation=(ussibus puerorum m n o a u m )  
iteration 5 DL=47.816696 interpretation=(ussibus puerorum m e o a u m )  
iteration 8 DL=36.863136 interpretation=(ussibus puerorum meorum )  
iteration 25  
  
=> :exit
```

Figure 8. Output from first successful run on the section of 255, indicating that the system resolves upon the correct interpretation of the image. The system outputs each interpretation which is deemed to be more successful than the last, and so builds up a list of plausible interpretations. The last solution matches the correct solution to the problem.