Reading Ancient Scripts: Investigating the Human Visual System for Artificial Intelligence in Palaeography

Ségolène Tarte
segolene.tarte@oerc.ox.ac.uk
University of Oxford, United Kingdom

Rachel Mairs
r.mairs@reading.ac.uk
University of Reading, United Kingdom

Mihaela Duta
mihaela.duta@psy.ox.ac.uk
University of Oxford, United Kingdom

Chrystalina Antoniades
chrystalina.antoniades@ndcn.ox.ac.uk
University of Oxford, United Kingdom

With the renewed vitality of research in Artificial Intelligence, thanks in particular to the continued development of neural network techniques for machine learning, computer vision technologies developed for handwriting recognition offer innovative ways of conducting research in palaeography (Brusuelas, 2016; Hassner et al 2014; Muzurelle, 2011; Stutzmann, 2015)

In this context where artificial intelligence often endeavours to replace human intelligence, or at least to emulate it, we are undertaking to understand better what it is that human intelligence does when reading ancient handwritten scripts. Ultimately, our ambition is to nudge artificial intelligence for palaeography to be intelligent enough to identify where human intelligence is still superior to machine intelligence (and therefore better left the upper hand) and where researchers can benefit from algorithmic support.

Handwriting recognition is a challenging task – both for humans and machines – because handwritten scripts inherently straddle two interlinked spheres of intelligence: that of visual shapes, and that of semantics.

This work builds on previous research (Terras, 2006; Youtie, 1963) that has identified six strands of human strategies in palaeography through ethnographic analysis, the results of which were cross-referenced with the cognitive sciences literature (Tarte, 2014). These strands were: visual perception, aural feedback, motor feedback, semantic memory, structural knowledge acquisition, and creativity; all continuously interacting with and feeding back into each other to some degree. In this project, we concentrate on the task of reading ancient handwriting – one of the many aspects of palaeographical research, whether it is concerned with mediaeval scripts or with more ancient scripts.

In this paper, we present some findings and observations made in the process of designing experiments to investigate some of the mechanisms underlying handwriting recognition in a palaeographical context; preliminary results from the experiments themselves are forthcoming.

To explore in depth how humans handle the variability of the shapes of signs in a given script, our experiments aim to bridge between traditional ethnographic methodologies, geared towards the gathering of qualitative data, and cognitive sciences methodologies, geared towards the gathering of quantitative data. The script of choice was Demotic, a script of the Late Egyptian language and whose use spanned ten centuries, therefore displaying a large variability in shapes. The team of scholars involved in designing and conducting our experiments was made of: an Egyptologist/Classicist, an Ethnographer, a Neuroscientist, and a Computer Scientist. Many of the observations reported here stem from the epistemological encounters of very different traditions of research; they emerged through the interdisciplinary conversations that took place in the process of designing the experiments.

The outcome of these conversations was the following experimental setup, building on the principles of the protocols outlined by Althaus and Plunkett (2015) and Longcamp et al (2008).

Experiment

Volunteers are invited to two experimental sessions that take place in a library setting, where they interact with a tablet computer using a stylus. The first session is a learning session and the second is a delayed recognition session. The sessions take place at least one week apart.

During the learning session volunteers learn to recognise 5 Demotic signs (target signs). This session

comprises of a familiarisation phase followed by an immediate recognition phase. The familiarisation phase can comprise of one of the following three familiarisation conditions:

- static passive familiarisation – 3-second repeated presentation of each sign
- static active familiarisation – repeated presentation of each sign with time for the volunteer to draw the sign on the tablet using the stylus
- dynamic familiarisation – 3-second repeated presentation of movies depicting the drawing of each sign

During the familiarisation phase each sign is presented 8 times, twice in each of 4 distinct hands. The presentation order is pseudo-randomize to ensure that signs don't appear twice in a row. Each volunteer is assigned randomly to one of the three familiarisation conditions. The familiarisation phase is followed by a 2-step immediate recognition phase. In the first recognition phase, pairs of Demotic signs are presented; each pair is made of one target sign and one distractor sign (the distractors are also Demotic signs), and in the second recognition phase words containing the target sign are presented.

The delayed recognition session comprises of three phases: a delayed recognition, a repeated familiarisation and a re-enforced delayed recognition. The two delayed recognition phases are similar to the immediate recognition phase, while the repeated familiarisation is similar to the familiarisation of the learning session.

At the first session, all volunteers are given a short video introduction that aims to prime them towards scripts and writing before starting the experiment; the second session ends with a debrief where volunteers are given the freedom to ask questions about the tasks or scripts to the experimenter.

Negotiations

In the process of designing this experimental setup, a number of fascinating questions emerged due to the intrinsic multidisciplinary nature of the project. The main interdisciplinary negotiations that took place can be summarised thematically as revolving around: the definition of an alphabet; the nature of script and materiality; and biases and mixed methodologies.

Alphabet

As the researchers and the volunteers evolve in an environment where the default script is alphabetic, we decided to choose an alphabetic script. The choice of Demotic can therefore be questioned: Demotic is not an alphabetic script, even if most of its signs have a phonetic value. We wanted however to use real data – as opposed to an invented alphabet – so was there any context in which Demotic might have been used as an alphabet? From the Ptolemaic period onwards, the frequency of Greek names in Egypt is higher, and therefore documents bearing Greek names resort to transliterating them into Demotic. Although there was no received orthography for those transliterations, it becomes more acceptable, in this specific context, to consider Demotic signs as alphabetic signs (for the purpose of this experiment, determinatives, when present, were regarded as not part of the word).

In turn, this deliberate choice also made a search for images of signs written in different hands easier. By querying http://www.trismegistos.org/ (Brouz and Depauw, 2015) for a list of Greek personal names in Demotic documents, and cross-referencing the results of this search with those from a query on papyri.info for all papyri in Demotic that have accompanying digital images, it was possible to identify and isolate images of Greek personal names written in Demotic on papyri.

Nature of script and materiality

The question of the homogeneity of the signals presented to the volunteers is important in the cognitive sciences. However, isolating the signs from their support means possibly removing information (e.g. faint ink on more salient fibres of a papyrus, degraded or stained papyri). So in an effort to present realistic data, we have decided to use greyscale images, to crop the images of words/signs and to simply uniformize the overall look of the images of words by aligning their histograms; we have also endeavoured to present all signs in such a way that they have a similar size (so some scaling was performed).

A further question was that of the phonetic dimension of the script. From a cognitive sciences perspective, as the focus is on the visual, it didn't make sense to add a phonetic element to the script at this stage. Only when it is better understood how the visual processes handle variability in shape, can it be envisaged to add a multisensory layer of complexity.

This raises some intriguing questions with regard to the nature of a script. In particular, isn't the essence of an alphabetic script to be a notation system for phonetic word utterances? What does the removal of its phonetic dimension entail for the script (regardless

of whether the existing phonetic dimension is a modern convention or an actuality)? Are we denaturing the script by presenting its signs stripped of their phonetic values?

Biases and mixed methods

Negotiating between the highly-controlled design of experiments in the cognitive sciences and the naturalistic settings and exchanges of ethnography proved complex. The main concern in such endeavours is to not compromise the validity of the gathered data with respect to the frameworks of the traditional disciplines. Free exchanges between volunteer and experimenter have the potential to bias the overall results for the cognitive sciences, as the bias is uncontrolled; it was therefore important for any such exchanges to happen only after all data of the controlled variables are collected.

Even then, leaving space for free exchange within the confines of the experiment worried the cognitive scientists who feared to open the door to highly irrelevant questions or even somewhat leftfield questions (e.g. "Can you see my gift of clairvoyance?") It appeared that such questions might be favoured by the running of such experiments in medicalised environments (hospital, psychology department), so we ran our experiments at the library instead, thereby establishing an environment resonant with our overarching theme of reading.

The process of designing these experiments as well as the forthcoming results have some bearing on the understanding of expertise of course, but also on how one might proceed when faced with a large corpus of unedited textual artefacts: Can crowdsourcing approaches be specifically geared towards the strength of the human visual system? Can algorithmic approaches palliate the weaknesses of the human visual system?

Acknowledgments

Bibliography

Althaus, N., and Plunkett, K. (2015). "Timing matters: The impact of label synchrony on infant categorization." *Cognition*, 139:1- 9.

Broux, Y., and Depauw, M. (2015). "Developing onomastic gazetteers and prosopographies for the ancient world through named entity recognition and graph visualization: Some examples from trismegistos people."
In L. Aiello and D. McFarland (eds), *Social Informatics. SocInfo 2014 International Workshops, GMC and Histinformatics*, Lecture Notes in Computer Science 8852, Springer, pp 304-13.

Brusuelas, J. (2016). "Engaging Greek: Ancient lives". In G. Bodard and M. Romanello, (eds), *Digital Classics Outside the Echo-Chamber: Teaching, Knowledge Exchange & Public Engagement*, Ubiquity Press, London, pp 187-204. .

Hassner, T., Sablatnig, R., Stutzmann, D., and Tarte, S. (2014). *Digital Palaeography: New Machines and Old Texts* (Dagstuhl Seminar 14302). Dagstuhl Reports, 4(7):112-34.

Longcamp, M., Boucard, C., Gilhodes, J.-C., Anton, J.-L., Roth, M., Nazarian, B., and Velay, J.-L. (2008). "Learning through hand- or typewriting influences visual recognition of new graphic shapes: Behavioral and functional imaging evidence." *Journal of Cognitive Neuroscience*, 20(5):802–815.

Muzerelle, D. (2011). "À la recherche d'algorithmes experts en écritures médiévales." *Gazette du livre médiéval*, 56–57:5–20.

Stutzmann, D. (2015). "Clustering of medieval scripts through computer image analysis: Towards an evaluation protocol." *Digital Medievalist*, 10.

Terras, M. (2006) *Image to Interpretation. An Intelligent System to Aid Historians in Reading the Vindolanda Texts*. Oxford University Press, Oxford.

Tarte, S. (2014). "Interpreting textual artefacts: Cognitive insights into expert practices." In C. Mills, M. Pidd, and E. Ward, (eds), *Proceedings of the Digital Humanities Congress 2012, Studies in the Digital Humanities*. Sheffield: HRI Online Publications.

Youtie, H. C. (1963). "The papyrologist: artificer of fact." *Greek, Roman and Byzantine Studies*, 4(1):19-33.