
ORIENTAL MANUSCRIPTS AND NEW INFORMATION TECHNOLOGIES

M. G. Carter

THE PLATONIC EDITION: SOME CONSEQUENCES OF COMPUTER EDITING FOR TEXT-BASED SCHOLARSHIP IN ARABIC GRAMMAR*

In this paper I would like to offer some very personal speculations about the effects of computers on the study of medieval Arabic grammatical texts, in particular the *Kitāb* of Sībawayhi (d. ca. 180/796), though most of what I shall say could be applied to any field of scholarship which relies for its data on manuscript sources. I will first define my position, then look at the specific case of editing the *Kitāb* in hypertext, and will conclude with some general thoughts on the nature and future of research in this area.

We are all in Plato's Cave, still. Although there may be a few individuals who have, over the millennia, escaped from the Cave, and have even come back bravely (and often suicidally) to tell us what it is like outside, we remain as Plato described us, prisoners chained to one wall and condemned to pass our entire lives gazing at the shadows of reality flickering on the opposite wall.

And that is, in effect, as close as we shall ever get to Sībawayhi, whether in manuscript or print: no matter how many versions we consult (and there are at least seventy-eight manuscripts) [1], we will never know what the authentic, original form of the *Kitāb* was. We do not even know that it ever had an authentic, original form, still less whether Sībawayhi himself (who seems to have left the actual task of compilation to his pupil al-Akhfash) ever considered his work in that light.

Nevertheless, we must exploit the new technology to recreate the *Kitāb*, to translate it from the old medium to the new. Without wishing to minimize the enormous difficulty of digitalizing all seventy-eight manuscripts (plus the others that will inevitably emerge), the several printed editions, Jahn's translation and as much secondary literature as possible, this aspect of the work is no more than an essential preliminary, and therefore not important in the context of this paper. Quantity is *per se* uninteresting to a Platonist. What does concern us is the effect of new data-base on both traditional and innovative scholarship.

Let us first dispose of the concept of a "critical edition". I maintain that it would be an absolute misuse of the new technology to try to create one: paradoxically, the electronic version of the *Kitāb* should never become final-

ized in any printed text. Ask yourselves, what authority would it have? Whose "edition" would it be? The Arabs never confined themselves to one *Kitāb*: as Geneviève Humbert has shown, there were two extremely well defined alternative "editions" in circulation and probably a third, minority version which is still under investigation [2]. Abū 'Alī al-Fārisī (d. 371/981) actually consulted five named and several anonymous *Kitāb* manuscripts for his commentary *al-Ta'liq* [3], and Monique Bernards reports that Ibn Wallād (d. 332/943) likewise used to check doubtful readings in more than one copy of the *Kitāb* [4]. Even the Qur'ān has never been subjected to the constraint of a unique and invariable earthly version, indeed the idea would have struck Muḥammad as rather unsound [5].

The eventual emergence of the "Eastern" and "Western" versions of the *Kitāb* is probably a scholarly compromise inspired more by convenience and professional interests than by strictly textual considerations, and it was still open for grammarians to refer to other readings, as for example when Abū Naṣr al-Qurṭubī (d. 401/1011) appeals to the authority of al-Zajjāj to support the reading *yudmaru* against *yazharu* [6], not a trivial difference, you must agree. For what it is worth, the Derenbourg, Būlāq and Hārūn editions all have *yazharu*.

But this last point takes us into an area where the new technology can be useful, for it will help us ask (I do not say it will answer, for it cannot) the question, "Why do all the printed editions have *yazharu*?" Actually we already know why: it is because Derenbourg has it, and the other editions merely follow him, but the next question is inevitably, "Why does Derenbourg have *yazharu*?", to which the answer is probably, "Because he did not find any variants in the manuscripts he consulted", and with that we are entering mysterious and challenging waters, because we know, even if Derenbourg did not, that there is a long-standing alternative *yudmaru* going back to al-Zajjāj. With an electronic *Kitāb* (which I shall resist calling an "e-*Kitāb*") it would be a simple matter to interrogate the data-base and trace the distribution of this variant, with useful results for the history of ideas and the sociology of the text which

* Since this was originally delivered as a paper at the International Medieval Conference in Leeds (July 1997), and also at the Semitic Symposium in Kivik, Sweden (August 1997), the written version retains the oral style of presentation, a quality which should be familiar to readers of medieval Arabic manuscripts.

a conventional critical edition could not easily reveal. In the end, of course, it would still leave the exact reasons for the variant a matter of speculative interpretation which no data-base or software can determine for us.

It may emerge that al-Zajjāj's reading is not part of a formal manuscript tradition but occurs independently in one of the innumerable glosses and marginal comments which cover some of the manuscripts of the *Kitāb*. This is an aspect of the work for which hypertext is a perfect medium: computer graphics can with relative ease produce a screen facsimile of the manuscript page with the complete contents of each folio converted into a standard font, but both the screen display and a physical print-out would be difficult to read. It would be far more convenient to assign all the glosses to hypertext links, where also the variants, translations and references to the secondary literature will reside, not to mention the codicological, historical and geographical information about the manuscript itself, in other words a whole library, with the *Kitāb* as its starting point.

It is obvious that the hypertext edition makes all kinds of traditional scholarship easier. If we have all the manuscripts at our electronic fingertips, so to speak, we can then, with the help of the appropriate software, view the state of the work at any point in its transmission, compare versions across space or time, call up the text in the form it had in copies known to be in the possession of specific grammarians, trace a variant back to its probable origin, follow up a topic in the commentary literature (for as well as the *Kitāb*, this data-base would ideally contain all the commentaries as well), check the translation of Jahn, bring up the relevant work of other scholars on a technical term or topic, in short, we can exploit the passive omniscience of the computer to gather information in support of kind of research endeavour.

Only the scale and complexity distinguish these activities from what academics have been doing since scholarship began. But the prospects are still exciting: two tasks immediately present themselves as perfect projects for our omnivorous and indefatigable amanuensis: one has apparently already started (in France, I am told), namely to publish the glosses, and here I will only make a recommendation from the side of the playing field, being unaware of how far the game has progressed. My suggestion is simply to use the computer to reassemble all the signed glosses as a mini-corpus for each known author: in one very important case, al-Akhfash, we have no other surviving grammatical work, and this might be the only practical means of reconstructing at least the general contents of his grammatical thought.

The other is rather less exciting but no less valuable, and that is to draw a map of the variants in the *Kitāb*, of which there are enough manuscripts to make the exercise genuinely fruitful. There are two kinds of variants, signed and unsigned, and characteristically the signed variants involve major technical disagreements (such as the example from al-Zajjāj above) while the unsigned ones are apparently random and have no consequences for the grammatical issues, e.g. *ḍaraba* 'amrun *zaydan* instead of *ḍaraba* *zaydun* 'amran. My hypothesis [7] is that these variants are not in fact random, but part of a sophisticated identification system to prevent unauthorized use of this rather profitable text: I assume that professionals knew where to look for these seemingly innocent differences and could thus trap those who had no right to be teaching the *Kitāb* because

they had not obtained the proper scholarly *ijāza* [8]. But to test this hypothesis, let alone attempt some sort of proof of it, first requires the registration and mapping of every single variant, after which it might be possible to look for patterns, e.g. that a family of manuscripts (identified by the traditional methods, e.g. by circumstantial evidence such as provenience, scribe or ownership history) would exhibit a consistent distribution of these unsigned variants, which could then be used to identify other manuscripts. In this way a particular set of intrinsically meaningless variants could become as distinctive as the empty holes in those cards we used to store our data on!

The temptation most to be resisted, however, is to undertake purely quantitative studies, at which a true Platonist can only shudder. It would be invidious to single out current work in which statistical methods are applied to features of language which by their nature are not usefully quantifiable: instead I will simply confess to having done something of the kind myself, albeit in a very informal way, and this will perhaps be a suitable penance for having done so in the first place. In a recent article I pointed out that the absolute numbers of quadriliteral verbs in modern colloquial Egyptian seemed significantly higher than those for Classical Arabic, and produced some extremely vague figures from the dictionaries to prove it [9]. An Aristotelian might favour such techniques but, whether the figures are set out in tables, spread sheets, graphs or pie charts, the proof of the pudding is in the eating, and on this question I side with Benjamin Disraeli: "There are three kinds of lies" — he is reported to have said, by Mark Twain, of all people — "lies, damned lies and statistics".

Numbers are by definition meaningless, they are perhaps the purest of Platonic images, inherently abstract in both form and substance. Whatever it is they mean it does not lie and indeed cannot lie in the numbers themselves: ask any Christian neo-Platonist about the significance of the number three! The issue is not quantity but quality: the truth is not decided democratically by votes, and statistics can only apply to phenomena which themselves obey statistical laws (whereby an inherent tautology is likely to result). Reliance on numbers can make a true judgement impossible: I think of the indispensable (at least for the time being) *Lexique-Index* of Troupeau, who decided not to itemize words which occur more than 60 times in the *Kitāb*, thus guaranteeing that the probably most marginal terms (e.g. *isnād*) would be the easiest to find and the probably most central (e.g. *ibtidā'*) also the most difficult. A concrete example: the word *naẓīr* is found 168 times in the *Kitāb*, and is thus too frequent for its individual occurrences to be listed, but the equally important concept of *muḍāra'a*, which is used a total of 111 times in its various forms, is fully itemized because no form of the root occurs more than sixty times. A close examination of Sībawayhi's use of *muḍāra'a* suggests that in fact it is at least as important as *naẓīr*, probably more so, as it has a wider technical range, but if one judged only by the numbers one might conclude the opposite [10].

This is not to say that statistical evidence is to be rejected on principle, but in the humanities it can seldom do more than confirm or contradict a general impression. One of the more convincing uses of computers in Arabic is the work of Bohas and Paoli in metrical studies, a domain which, with its high degree of formalism, seems ideal for such treatment. And the results are truly fascinating, giving

us a detailed knowledge of the structure of Arabic poetry which no modern mind could encompass unaided. Yet the authors do not confuse the computer with the intellect, and fully recognize that the human reaction to poetry (and this only at the metrical level, be it noted) is qualitatively quite different from the processing of poetry by digital means [11].

By raising the issue of quality I intentionally bring into the debate questions of humanism and the nature of the academic life, notions which are none too outstanding in the literature of computer technology. Here, too, great changes lie ahead as scholarship reformulates itself in the new medium, where personal contact is replaced by the Net, research is done for us by algorithms, publication and feedback are both instantaneous and universal, as we display our learning in what I suppose in our field would have to be called a Cyber-Majlis.

For the medieval Arabs the transmission of knowledge was a serious and well controlled business conferring both dignity and profit on the participants. Their medium was, of course, the manuscript, and there was an elaborate and generally effective system of publication which covered the production, distribution and copyright of original works. Although the terms did not exist then, we can be sure that the concepts of "intellectual property", "information management" and "knowledge transfer" were very familiar to them. It is not a little ironic that we still derive our own livelihood today from the same sources as the medieval Arabs, the very manuscripts which we now propose to digitalize.

Digitalization immediately evokes many problems which can only be hinted at here. Permanence: books and manuscripts survive remarkably well considering how they are treated, but what is the reliability of an electromagnetic charge which can be obliterated in a millisecond's loss of current? Access: there was a time when precious books were chained to the desk, but what is now the value of something which can be read by anyone with a computer and a modem? The aesthetic side: when will we abandon typographical conventions which themselves were (and largely still are) a compromise between the physical restraints of hot metal type and the visual criteria of calligraphy? The physical aspect: what difference will it make to read a work on a screen instead of holding it in our hands? And what will it mean in future to "read" a document, given that browsing software will take over most of the functions of reading? Professional matters: who or what will the new academic community be, the ones by whom and for whom these manuscripts are going to be digitalized? And what will be the meaning of the word "erudition" when we all have instant access to the same knowledge base?

Those were all rhetorical questions which I have no intention of trying to answer, mainly because the answer will have to be found on the Web itself, in the Cyber-Majlis, where there will doubtless also be a new rhetoric, and probably, given the perversity of human nature, even rhetorical questions will find an answer from some latter-day Jāhīz or Tawhīdī who just happens to log on. Already the electronic interchange of ideas is well established, and "the way electronic intertextuality changes our relation to the institutions of scholarship" is now earnestly discussed, along with many of my rhetorical questions above [12]. The electronic symposium is now in full growth, and

henceforth scholars will have to reckon with the fact that their ideas no longer belong exclusively on a printed page but will be exposed to a global "readership". I have put readership in quotation marks because I am not sure whether it still applies to staring at a screen, and also because I wonder how many of my listeners noticed the unconscious survival of print habits when I referred to something I had said previously as "above". The slip was genuine and has not been edited out, though as you know, there is never anything final about a document in a word processor!

This brings us back to Platonic editions, which differ from word processor files in that they can never be printed out: they remain always *in potentia*, even after (or perhaps especially after) the manuscripts have been transmuted from "physical marks on a surface" [13] into invisible states of electrons. What will probably happen in the Cyber-Majlis is that a variety of *Kitāb* editions will emerge, as individual scholars base their research on their own choice of readings, and these will compete for authority until perhaps a generally accepted version evolves. In this way we shall replicate electronically exactly what happened with the original manuscripts, and here I could rest my case and say that this is the strongest argument in support of my original assertion that no final "critical" edition should even be contemplated.

I will conclude by speculating about some of the possibilities of the new scholarship which will be facilitated as soon as the software designers can turn their attention away from the more profitable computer games.

One development which seems inevitable is the thesis programme, an algorithm which will skim through a given corpus and abstract from it a well-formed and convincing piece of research which can be uploaded (or as we used to say, submitted) for a doctorate. Of course it will be read and evaluated by another piece of software and the degree will be awarded on the Web, perhaps as a downloadable multimedia diploma. This is only a slight exaggeration, whose aim is simply to emphasize the inevitability of change in the academic process at the same time as asking you not to forget that the relationship between a scholar, his students, his discipline and his material is essentially a personal one, where "face to face" contact is about to be replaced by the electronic "interface".

My tax software suggests another programme which could stimulate (or should I say simulate?) useful enquiry, namely the "What if?" button, which will explore for me the consequences any whimsical speculation, no matter how implausible. For example, what if Ferdinand de Saussure had read Jahn's translation of the *Kitāb* [14], could we then argue that his concept of language as a social phenomenon came (indirectly) from Sībawayhi rather than, according to certain historians of linguistics at least, directly from Durkheim? [15] Our software would obligingly search through all the relevant data and with any luck might discover in the records of the Geneva University library (which doubtless still exist and will eventually be digitalized!) that de Saussure had indeed borrowed and presumably therefore read Jahn's work, whereupon a whole new chapter of the history of linguistics could be written. But behind this facetiousness there is a serious point: software cannot initiate, it can only serve as a tool for the curiosity of the researcher. It has been said of al-Khalīl ibn Aḥmad that the proof of his genius lay in the questions he

asked [16], and it will certainly be a challenge to future scholars to find questions which live up to the power of their research tools.

Unfortunately the Internet, which is where the electronic symposium will meet, is itself a rather indiscriminating entity, with a mesh so tight that it catches far too many small fish. Bulletin Boards, for example, are intellectually just another kind of graffiti, either indecipherable or trivial. The main purpose of computers seems to be playing games, and this has spilled over into educational practice, where learning and playing have become fatally confused. Believe it or not I have seen some years ago an advertisement for a programme for learners of Arabic called "Fun with verbs". I hope you are as appalled as I am that this infantile exercise was aimed at university students. Will future generations of students be taught by a greedy and impatient Tamaguchi which will die if they do not learn their verbs properly? The answer, unfortunately, is almost certainly yes.

There may indeed be technical analogies between the editing of manuscripts and other computer applications such as games, business programmes and the creation of original works: the conventions of graphic presentation and the strategies of manoeuvring through the programme are similar in all of these. But the reading of texts is qualitatively unique, a mediated contact of mind with mind, in our case with Sībawayhi's, and therefore it is neither a peda-

gogical nor a creative experience, still less a game. The ideas come to us merely through a different medium, which should be as transparent as possible, with minimal interference from the presenter (the ideals of typography, for example, were to combine beauty and efficiency in a form which did not distract from the content: computer graphics has a very long way to go in this regard). We should bear in mind Plato's suspicion of everything creative: creativity leads away from the truth, and artists (I paraphrase Plato) are merely wizards who play with images, thus twice removed from reality because their "creations" are in themselves only images of images.

The real *Kitāb* (if we are lucky) will lie hidden somewhere in the huge electronic library constructed and administered by the computer, which will act purely as a custodian for Sībawayhi's ideas. The work as such can never escape from this library in any finite form, nor can we come closer to it than the contemplation of the versions recreated on our screens. We remain trapped in Plato's Cave, and while others happily avoid reality by playing in their Multi-User Dungeons (how Plato would have enjoyed that image!), we shall browse for the reality of the *Kitāb* in our computers.

However, everybody is still gazing at the same screens, which have become the new walls of our Cave on which the shadows of the truth are still flickering.

Notes

1. The manuscript history of the *Kitāb* has been exhaustively traced by Geneviève Humbert in *Les voies de transmission du Kitāb de Sībawayhi* (Leiden, New York, Köln, 1995), supplemented by the same author's "Remarques sur les éditions du *Kitāb* de Sībawayhi et leur base manuscrite", in *Studies in the History of Arabic Grammar II, Proceedings of the 2nd Symposium on the History of Arabic Grammar, Nijmegen, 27 April—1 May 1987*, eds. C. H. M. Versteegh and M. G. Carter (Amsterdam, 1990), and "Un témoignage fossile du *Kitāb* de Sībawayhi", in *Développements récents en linguistique arabe et sémitique*, ed. G. Bohas (Damascus 1993).

2. See chapters 3—8 of Humbert's *Les voies de transmission du Kitāb de Sībawayhi*.

3. See *al-Ta'liqa*, ed. 'Awad ibn Aḥmad al-Qūzī (Cairo/Riyadh, 1990—96), i, introduction, p. 52f.

4. *Changing Traditions: al-Mubarrad's Refutation of Sībawayh and the Subsequent Reception of the Kitāb* (Leiden, New York, Köln, 1997), p. 44.

5. The perfect version, of course, is in Heaven with its author. Of the various *hadīths* discouraging dispute about Qur'anic variants perhaps the most vivid is one in which Ibn Mas'ūd hears a man reciting the text differently from the way he had heard it himself from Muḥammad. An outraged Ibn Mas'ūd leads the man by the hand to confront the Prophet, only to be told, "You've both got it right, don't argue about it. The ones before you who argued are all dead now". Bukhārī, *Ṣaḥīh*, p. 44 (= *Khuṣūmāt*) 1. The doctrine of the "Seven Readings" formalizes this situation and allows the user to construct a potentially infinite number of Qur'āns (in this light it seems unnecessary to limit the number of possible interpretations of each verse to 60,000, cf. I. Goldziher, *Die Richtungen der islamischen Koranlegung*, Leiden, 1920, p. 257). The theme has been brilliantly taken up by Italo Calvino, *If on a Winter's Night a Traveller*, describing the error of 'Abdullāh ibn Abī Sarḥ in trying to hold Muḥammad to an unreasonable standard of accuracy (p. 182f in the San Diego edition of 1981).

6. *Sharḥ 'uyūn Kitāb Sībawayhi*, ed. 'A. 'A. 'Abd Rabbih (Cairo, 1984), p. 121.

7. See M. G. Carter, "Arabic literature", in *Scholarly Editing, a Guide to Research*, ed. D. C. Greetham (New York, 1995), p. 555f.

8. Humbert again gives a lively account of the way in which a copy of the *Kitāb* was obtained illegally and the culprit thrown in jail for it, see her "Copie "a la pecia" à Bagdad au IX^e siècle?", *Gazette du livre médiéval*, 12 (1988), pp. 12—5.

9. "Signs of change in Egyptian Arabic", in *Understanding Arabic, Essays in Contemporary Arabic Linguistics in Honor of El-Said Badawi*, ed. Alaa Elgibali (Cairo 1996), pp. 137—43, esp. 140.

10. To make matters worse, Troupeau simply equates *naẓīr* with the terms *akh*, *ukht* and *akhawāt*, but since these occur so frequently we cannot easily follow them up.

11. G. Bohas, B. Pauli, *Aspects formels de la poésie arabe* (Toulouse 1997), esp. pp. 31f and 181ff.

12. The notions here are abstracted from *Hyper/Text/Theory*, ed. George P. Landow (Baltimore and London, 1994), esp. the editor's introductory chapter.

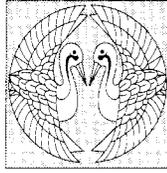
13. *Ibid.*, p. 13.

14. This far from whimsical question was originally posed by J. Fück in 1895 but did not appear in print until 1981, see M. G. Carter, "Arab Linguistics and Arabic Linguistics", *Zeitschrift für die Geschichte der arabisch-islamischen Wissenschaften*, 4 (1987/8), p. 211, n. 20.

15. Cf. K. Koerner, "On the Problem of "Influence" in Linguistic Historiography", in *Papers in the History of Linguistics, Proceedings of the Third International Conference on the History of the Language Sciences*, eds. H. Aarslef, L. G. Kelly, H.-J. Niederehe (Amsterdam and Philadelphia, 1987), pp. 13—28, esp. 19—22.

16. E. Bräunlich, "Al-Ḥalīl und das Kitāb al-'Ain", *Islamica*, 2 (1926), p. 61.

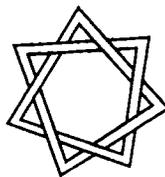
RUSSIAN ACADEMY OF SCIENCES
THE INSTITUTE OF ORIENTAL STUDIES
ST. PETERSBURG BRANCH



Manuscripta Orientalia

International Journal for Oriental Manuscript Research

Vol. 3 No. 4 December 1997



75ESA
St. Petersburg-Helsinki

CONTENTS

<i>TEXTS AND MANUSCRIPTS: DESCRIPTION AND RESEARCH</i>	3
S. Levitt. Sinhalese Painted Wooden Bookcovers	3
T. Sultanov. Turkic Versions of the <i>Tārīkh-i Rashīdī</i> in the Manuscript Collection of the St. Petersburg Branch of the Institute of Oriental Studies	17
Vladimir Polosin. Two Late Eighteenth-Century Ottoman Fiscal Documents from the Manuscript Collection of the St. Petersburg Branch of the Institute of Oriental Studies	30
<i>TEXT AND ITS CULTURAL INTERPRETATION</i>	35
E. Rezvan. The Qur'ān and Its World: IV. "Raise Not Your Voices above the Prophet's Voice" (Society, Power and Etiquette Norms)	35
<i>PRESENTING THE COLLECTIONS</i>	45
T. Deryagina, O. Frolova. Antoni Muchliński and His Collection of Arabic Manuscripts in the St. Petersburg University Library	45
<i>ORIENTAL MANUSCRIPTS AND NEW INFORMATION TECHNOLOGIES</i>	54
M. Carter. The Platonic Edition: Some Consequences of Computer Editing for Text-Based Scholarship in Arabic Grammar	54
<i>PRESENTING THE MANUSCRIPT</i>	59
O. Bolshakov. The St. Petersburg Manuscript of the <i>Maqāmāt</i> by al-Ḥarīrī and Its Place in the History of Arab Painting	59
<i>BOOK REVIEWS</i>	67
<i>Manuscripta Orientalia</i> in 1997, vol. 3, Nos. 1–4 (list of contributions)	70

Front cover:

"Abū Zayd and al-Ḥārith talking". Fragment of a miniature from manuscript C 23 of the *Maqāmāt* by al-Ḥarīrī in the collection of the St. Petersburg Branch of the Institute of Oriental Studies.
Illustration to *maqāma* 50, p. 349, 17.5 × 9.0 cm.

Back cover:

"Abū Zayd as a teacher in a school at Ḥimṣ". A miniature from manuscript C 23 of the *Maqāmāt* by al-Ḥarīrī in the collection of the St. Petersburg Branch of the Institute of Oriental Studies.
Illustration to *maqāma* 46, p. 318, 18.0 × 19.5 cm.