

The Digital Archive of South Indian Inscriptions (DASI) — A First Report

Sascha Ebeling

"Most urgent is the need for comprehensive computer databases of the now unmanageably vast published epigraphic material; very little has been done in this direction, and the need for it is growing constantly."

(Salomon 1998: 224)

1 Introduction

The purpose of the present paper¹ is to briefly describe an international project which was launched in the year 2002 as a cooperation between scholars from Uppsala University (Sweden), Cologne University (Germany) and a number of South Asian scholars.² The aim of the project is to build up a comprehensive computerised archive of South Indian inscriptions which will focus—at a first level—on the digitisation of the entire existing corpus of inscriptions in the Tamil language, published and unpublished. This overall plan will be carried out in several successive steps. According to the particular priorities of the project members, we have decided to begin with first digitising Tamil inscriptions which contain information about Buddhism in South India and Sri Lanka, i.e. a subset of the existing inscriptional corpus. Once this subset has been completed, we will gradually further enlarge the corpus, entering other thematically arranged groups of Tamil inscriptions. After the Tamil material has been completed, we will take up the inscriptions in other Dravidian languages and in Sanskrit.

The choice of Tamil inscriptions pertaining to Buddhism as the first subset to be digitised was made due to the fact that they represent a relatively small and therefore manageable amount of data which has been collected and

¹ An earlier version of this paper was presented at Uppsala University in June 2002. I would like to thank all the members of the extended AKAR (*Arbeitskreis für Asiatische Religionsgeschichte*), Prof. Dr. Annette Wilke and Dr. Brigitte Luchesi, and above all Prof. Dr. Peter Schalk for their comments and suggestions.

² The project team currently consists of Prof. Dr. R. Nagaswamy (former director of the State Department of Archaeology, Govt. of Madras, India), Prof. Dr. A. Velupillai (former Professor of Tamil, Sri Lanka, and Visiting Professor, Virginia, USA), Prof. Dr. P. Schalk (Uppsala University, Sweden), Prof. Dr. Gunilla Gren-Eklund (Uppsala University, Sweden), Dr. Thomas Malten and the author of the present paper (Institute of Indology and Tamil Studies, University of Cologne, Germany). Initial funding was generously provided by the *Margot och Rune Johanssons Stiftelse*, Uppsala, Sweden.

thoroughly (re-)analysed in a recent publication edited by two of the project members (Schalk/Veluppillai 2002). Manageability of the first data is an important factor, since much of the work on DASI involves tasks which are tackled here, as far as we know, for the first time in the history of epigraphy and which pertain to both the technical side (e.g. entering the data into the computer (typing, proof-reading), compiling the software for presenting and handling the data, scanning images of the inscriptions, indexing and cross-referencing the entries, etc.) and the 'scholarly' side (e.g. discussing the questions of which particular inscription to include, how to translate and interpret it, acquiring further unpublished inscriptions and images (digital photographs in the field or copies of estampages)). In preparing the first small number of entries for DASI, all these issues have to be addressed, and the experience we gain during this process may then be applied to the handling of larger epigraphic corpora. The result of the DASI project will hopefully be useful to those who work with epigraphic material: South Asianists, linguists, historians of religion, literature, art, economics, politics and social life.

For those readers who may be less familiar with Tamil epigraphy, a very brief and sketchy review of the field will be given in the next section of this paper. Section 3 will discuss details of the Digital Archive of South Indian Inscriptions.

2 Tamil inscriptions and the current state of Tamil epigraphy

The utility of inscriptions for a variety of purposes has been acknowledged basically since the beginnings of Indian epigraphy fostered by pioneer indologists (such as Charles Wilkins (1749-1836) and others)³ at the end of the 18th century. In the words of James Heitzman, who refers to the body of inscriptions connected to the Cōla period which alone number more than 10,000: "Any study of the early history of south Asia must take into account the world portrayed by this data base" (Heitzman 1997: 2). Moreover, inscriptions may also be studied by linguists as a valuable source for diachronic language analysis. Although a large number of grammars of inscriptional Tamil exist,⁴ the known Tamil inscriptions have not been systematically employed to complement current theories about the historical development of the Tamil language. Neither do we know much about the relationship between inscriptional and literary Tamil, nor have the various findings of the inscriptional grammars been compared and critically discussed.⁵ Therefore, one of the original purposes for developing DASI was to provide the necessary data for such language analysis.

³ See Salomon (1998: 199ff).

⁴ The monographs that deal with grammatical issues are Agesthalingom/Shanmugam (1970), Amalanathan (1972), Bhagirathi (1961), Jayakumari (1959), Kanapathi Pillai (1935), Karthikeyani (1981), Panneerselvam (1969), Shanmugam (1969), Veluppillai (1962), Veluppillai (1963-64), Veluppillai (1972), Veluppillai (1976), and Zvelebil (1964). Unfortunately, most of them are available only as unpublished manuscripts, so that access to them has hitherto been difficult if not impossible. DASI will therefore incorporate grammars and linguistic analyses within its framework.

⁵ An exception to this is Veluppillai (1980), chapter 6.

As for inscriptions in the Tamil language, there are literally tens of thousands of written documents on a variety of materials, such as stone, metal plates or pottery ranging from the last centuries before the beginning of the Christian era to the present times.⁶ They are different kinds of what one may call official documents, such as royal proclamations, land transactions and other gifts by kings and ministers to temples, learned Brahmins and other institutions, documents related to taxation, education, irrigation, elections of candidates to village assemblies, memorials to dead persons erected by devotees or descendants, known as hero-stones (Tam. *vīrakkal*, *naṭukal*), etc. The earliest known records in what scholars now agree to be an early form of Tamil, the so-called Old Tamil Brāhmī inscriptions, are a series of about 83 brief dedicatory inscriptions found in natural caves at 26 sites in Tamilnadu. They date from approximately the second and first centuries B.C.⁷ During the following second to the sixth centuries A.D. there are merely a few records which have so far been found and dated on palaeographic grounds.⁸

A much more widespread use of Tamil as an epigraphic language can be seen from the times of the Pallava kings, i.e. from the 7th century A.D. onwards, when bilingual Tamil and Sanskrit copper plate inscriptions began to be issued.⁹ The succeeding Cōla and Pāṇṭiya dynasties continued this practice, and it is from this time onwards that we have a vast corpus of inscriptions, particularly on temple walls and copper plates, which number in the many thousands. According to an estimate made in 1973 by R. Nagaswamy, also a member of the DASI team, one may say that out of a total of 25,000 inscriptions which have so far been discovered in Tamil Nadu, only 6,000 have been published.¹⁰ Unfortunately, this situation does not seem to have changed much during the past thirty years. On the one hand, we have an enormous amount of unpublished material, and on the other hand, whatever has been published is scattered in many different journals and publication series, so that even for specialists of Tamil epigraphy the situation is confusing. In order to put the further study of Tamil inscriptions on a firm footing, so that systematic research becomes possible, it was decided to build up DASI.

⁶ A general overview which divides the available Tamil inscriptions into six consecutive periods (ending in ca. 13th c. A.D.) is given by Krishnan (1990). A comprehensive bibliography on Tamil epigraphy which lists the published Tamil inscriptions as well as secondary literature is currently being prepared by the present author.

⁷ Mahadevan (1995: 174), who also records findings of Tamil Brāhmī inscriptions from Sri Lanka. For the latest list of these inscriptions, see Mahadevan (1992). They record the provision of stone beds to Jaina (or Buddhist or Ājivaka?) ascetics for religious observances like fasting, etc. (see the discussion by Velupillai in Schalk/Velupillai (2002: 168)). Followed by these early cave inscriptions are the graffiti on potsherds from Arikamedu (Begley (1996); Subrahmanian/Venkatraman (1980: 7); Krishnan (1990: 206)).

⁸ Subrahmanian/Venkatraman (1980: 7); Krishnan (1990: 206f.).

⁹ See Salomon (1998: 105) and Brocquet (1993-4). The first bilingual Sanskrit and Tamil inscription is found on the Paḷḷaṅkōyil Plates issued by Siṃhavarman III (Brocquet 1993-4: 90). Zvelebil (1964) has provided a linguistic analysis of this important inscription.

¹⁰ Nagaswamy (1973: 67).

3 The Structure of DASI

As has been stated above the aim of DASI is to provide a comprehensive archive of South Indian inscriptions and related information, so that one may obtain all the information that is currently available for a particular inscription. Thus, the database will contain the text of a particular inscription in both Tamil script and in transliteration, an image of the inscription, a translation of the text, a short summary of the contents, information on date, findspot, and whatever additional information may be available. At present, the project work is being carried out simultaneously on four different sub-projects: 1) the digitisation of the Tamil inscriptions pertaining to Buddhism (with the material taken from Schalk/Vēluppillai (2002), 2) the digitisation of grammars of inscriptions and further linguistic material, 3) the digitisation of the entire *Annual Report on Indian Epigraphy* (ARE) series published by the Archaeological Survey of India (ASI), and 4) the digitisation of the Tamil inscriptions contained in the two series *South Indian Inscriptions* (SII) and *Epigraphia Indica* (EI) series also published by the ASI. The term ARE series here refers to the series of annual reports published under various names since 1887 by the Government of Madras up to 1945 and to its successor published by the ASI.¹¹ The ARE volumes contain lists of inscriptions, published and unpublished, that were reported or examined by the officers of the Archaeological Survey. Some (but by far not all) of these inscriptions may be found published in the major series SII and EI. In that case the information given in the ARE entries in DASI will be linked to the respective published inscription, so that it becomes possible to see, when browsing the ARE entries, whether a particular inscription has been published or not. This may best be illustrated by an example:

Let us take a random entry from the lists in ARE, e.g. inscription number 98 of appendix C (stone inscriptions copied in 1920) in the ARE volume for the years 1919/1920. The entry found in this volume will look like this in the (machine-readable) DASI format:

¹¹ The various titles of these reports and their publication history are given in Salomon (1998: 253).

```

<NUM>0001</NUM>

<SOU>ARE-1919/20-C98</SOU>

<REF>SII-13-55</REF>

<LOC>On the east wall of the ruined man2d2apa
to the north of the shrine of the goddess
in the Vaidyanaltha temple at Tirumal6ava1d2i
(Ud2aya1rpall2ayam taluk,
Trichinopoly district)</LOC>

<DYN>Choll2a</DYN>

<KIN>Ra1jakelsarivarman</KIN>

<DAT>3rd year</DAT>

<LAN>Tamil</LAN>

<REM>Incomplete. Registers a gift of gold
for offering crushed rice (aval)
on the day of Alrdral in the month of Malrgal6i
to Gan2apati by Ke1ndala Vel2l2a1lla Pil2l2aiyalr
of Kumbal2l2umbulr in Olyma1-na1d2u,
a district of Ton2d2ai-na1d2u</REM>

```

fig. 1: Sample data-set for ARE entries in the DASI format.

Note that the information given in individual columns in a table in the original ARE volume occurs in the DASI format with tags (in an SGML-based encoding), i.e. three-letter abbreviations to mark a particular category of information. <NUM> is an internal serial number. <SOU> denotes the original source of the information, i.e. here inscription No. 98 of appendix C (stone inscriptions copied in 1920) in the ARE volume for the years 1919/20. <REF> is a cross-reference to the published inscription in a volume of SII which we shall discuss below. <LOC> refers to the place of the inscription, <DYN> to the dynasty, <KIN> to the king mentioned in the inscription, <DAT> to the date of the inscription, <LAN> to language and alphabet, and <REM> refers to further remarks. Instead of diacritic marks, which are always problematic in machine-readable text, we have used a simple system of adding numbers to the character that is to be marked with a diacritic: 1 = bar above (a1 = ā), 2 = dot below (n2 = ṅ), 3 = dot above (n3 = ṅ̇), 4 = acute accent (s4 = ś), 5 = tilde above (n5 = ñ), and 6 = bar below (l6 = ḷ). These transliterations may look odd at first glance, but they are easy to handle in a digital environment and may be instantly converted into any font to suit the user's purpose.

Now, since the inscription given in the above sample was published, the tag <REF> contains a cross-reference to the published inscription which is number 55 in volume 13 of SII. The DASI entry for this inscription looks like this:

```

<NUM>0001</NUM>
<SOU>SII-13-55</SOU>
<REF>ARE-1919/20-C98</REF>
<LOC>Tirumal6avaId2i, Udaiyarpalaiyam Taluk, Trichinopoly
District</LOC>
<POS>On the east wall of the ruined man2d2apa to the north of
the shrine of the goddess in the Vaidyanaltha temple</POS>
<DAT>3rd regnal year</DAT>
<DES>This records an undertaking backed with a security in land,
given by two servants of the temple to make an offering of
flaked rice to god Gan2apati in the temple at Punalvalyil on the
day of Alrudral in Malrgal6i month, with the money endowed for
the purpose and left in their charge
by a certain S4eIndan6 Vell2aIl2a-Pil2l2aiyalr of
Kumbal2l2umbulr in Olymal-nald2u,
a subdivision of Ton2d2ai-nald2u</DES>
<TXT>
<Z=001><g>svasti SRI [|*]
<t>kovirAjake<g>ca<t>ripa<g>nma<t>[R*]kku <b>yANTu 3-Avatu</b>
pu-
<Z=002>n_alvAyil SRIkoyil <g>gaNapati bhaTTAra<t>rkku
<g>caGrAditta-
<Z=003>val toNTainATTu oymA(n_)n_ATTu kumpaLLumpUr ce-
<Z=004>[n]tan_ veLALA piLLaiyAr <g>gaNa<t>patiyArkku mArkazit
ti-
<Z=005>ruvAtirai nAL aval amutu ceyya tUNi nAnAzi aval
<Z=006>aTTuvataKa kONTa pon_ ezu mañcATiyum kun_Ri ivvUr kAma
<Z=007>kkANi aiyARan_ <g>devasvA<t>miy[yu]m aiyARan_ mA-
<Z=008>... <g>deva<t>n_um iviruvom iv[va*]va(l)lukku nAgkaL
<Z=009>oRRivaytta [nilamAvatu melUr egkaL] //...//
</TXT>
<ILL> no illustration available </ILL>
<TRA> no translation available </TRA>

```

fig. 2: Sample data-set for SII entries in the DASI format.

Here, the same tags as for the ARE entries have been used, while <SOU> here gives the source in SII and <REF> is of course the cross-reference back to the ARE entry. Additionally, the following tags have been introduced: <POS> provides specific details about the physical position of the inscription, <DES> gives a brief description of the contents, <TXT> refers to the text of inscription proper, <Z> contains line numbers for the inscriptional text, <ILL> contains a hyperlink to a digital image of the original inscription, and <TRA> contains an English translation of the inscription. Translations given in the original sources (i.e. SII or EI) will be included. For the many cases in which the inscription has been published without an English translation, a translation will be made for DASI.

Converting the text of the inscription itself into a machine-readable format is fraught with particular difficulties. Firstly, the published text contains both Tamil and Grantha characters which distinguish Tamil and Sanskrit words or individual characters, so that we need to include this distinction in the digital text. This is done by using the tags <t> and <g> to 'switch on' Tamil and Grantha characters respectively, so that line 1 in the above sample "<g>svasti SRI [| |*] <t>kovirAjake<g>ca<t>ripa" means "svasti śrī [| |*]" in the original is written in Grantha script, while "kovirājake" is written in Tamil script, followed by a "ca" in Grantha and "ripa" in Tamil, etc. The tag indicates bold print in the original source. Secondly, here again one should refrain from using diacritics, so we have employed an ASCII-based transliteration that uses capital letters instead of diacritics for the Tamil (and occasionally Sanskrit) text.¹² The other marks and special characters, e.g. "[| |*]" and the various brackets, are those commonly used in Indian epigraphy.¹³ From the above examples, it will become clear that the digital text is a faithful rendering of the original printed publication. This will be supplemented by images of the inscriptions, so that the specialist may check the published text with the original without having to travel physically to the findspot to read the inscription *in situ*. Also, grammars and dictionaries of inscriptional Tamil will be contained in DASi, so that the vast amount of inscriptional data becomes more manageable, and that also students of Indian epigraphy or non-specialist users may be enabled to work with DASi.

4 Conclusion

What are the advantages of having a computerised archive of South Indian inscriptions? DASi hopes to provide a powerful tool for making the vast amount of published and unpublished epigraphic material accessible and manageable. In collating inscriptions which are often published in series, journals or commemorative volumes which are not easily available even to the specialist, and in trying to assemble, step by step, all the available information on Tamil epigraphy, DASi seeks to assist scholars in their work with the rich and—despite major efforts—still very much understudied material. The data of DASi will be fully searchable and convertible into any format required by the user, including printouts. Indices may be automatically compiled. Lasting data storage does not pose a problem, since all the data are entered in simple encoding formats so that new CD-ROMs can be prepared any time. One of the uses envisaged for DASi is as a mobile resource on a laptop computer in the field, so that epigraphists and archaeologists may check new findings directly with existing material. Furthermore, the archive is modular, so that it can be extended or modified at any time to suit anybody's interests. As has been stated above, the project will be carried out in several successive steps in each of

¹² This transliteration scheme has for the past decade been the standard for all the digital Tamil text corpora developed at Cologne University. It is explained on the Cologne website at <<http://www.uni-koeln.de/phil-fak/indologie/tamil/otl.html>>.

¹³ For a description of these conventions, see Salomon (1998: 162-4) and the literature cited there.

which the archive will be expanded to comprise more inscriptions. Let me stress, finally, that DASI sees itself very much as an open effort and that we would welcome suggestions or further collaborators. As DASI develops further, we hope to provide stimuli to epigraphists who specialise in other areas, so that Richard Salomon's observation that "very little has been done in this direction" (Salomon 1998: 224) may soon become outdated, and that similar endeavours may be taken up for North Indian inscriptions or for the rich epigraphic material elsewhere.

Bibliography

- Agesthalingom, S. and S.V. Shanmugam. 1970. *The language of Tamil inscriptions (1250-1350 A.D.)*. Annamalainagar: Annamalai University.
- Amalanathan, C. J. 1972. *The language of Pandya inscriptions*. unpublished M. Litt. Thesis. Annamalainagar: Annamalai University.
- Begley, Vimala. 1996. *The Ancient Port of Arikamedu. New Excavations and Researches 1989-1992*. vol. 1. Pondichéry: EFEO.
- Bhagirathi, S. 1961. *The language of the Tamil inscriptions: 1050-1250 A.D.* unpublished M. Litt. Thesis. Annamalainagar: Annamalai University.
- Brocquet, Sylvain. 1993-4. "Le kāvya épigraphique du Tamil Nāḍu". *Bulletin d'Etudes Indiennes* 11-12 (1993-4): 85-110.
- Heitzman, James. 1997. *Gifts of Power. Lordship in an Early Indian State*. New Delhi: OUP.
- Jayakumari, V. 1959. *A study of the language of Tamil inscriptions from the 7th century up to the middle of the 11th century A.D.* unpublished M. Litt. Thesis. Madras: University of Madras.
- Kanapathi Pillai, K. 1935. *A Study of the Language of the Tamil Inscriptions of the 7th and 8th Centuries A.D.* unpublished Ph.D. thesis. London: University of London.
- Karthikeyani, G. 1981. *The linguistic study of Tamil inscriptions (900-1050 A.D.)*. Ph.D. Thesis. Annamalainagar: Annamalai University.
- Krishnan, K.G. 1990. "Characteristic Features of Tamil Inscriptions". in: Nayak, H.M. and B.R. Gopal (eds.). *South Indian Studies*. Mysore: Geetha Book House.
- Mahadevan, Iravatham. 1968. "Tamil-Brahmi Inscriptions of the Sangam Age". *Second International Conference-Seminar of Tamil Studies, January 3 to 10, 1968*. Madras. 1-41.
- Mahadevan, Iravatham. 1992. "Occurrence of the *pulli* in the Tamil-Brahmi Script". in: Lockwood, Michael (ed.). *Indological Essays. Commemorative Volume for Gift Siromoney*. Madras. 141-167.
- Mahadevan, Iravatham. 1995. "From Orality to Literacy: The Case of the Tamil Society" *Studies in History* 11.2 (1995): 173-188.
- Nagaswamy, R. 1973. "Archaeology and Epigraphy in Tamil Nadu". in: Thani Nayagam, X.S. and François Gros (eds.). *International Association of Tamil Research. Proceedings of the Third International Conference Seminar, Paris 1970*. Pondichéry: Institut Français d'Indologie. 56-71.
- Panneerselvam, R. 1969. *Tamil inscriptions of the 7th, 8th and 9th centuries A.D.* unpublished Ph.D. Thesis. Trivandrum: University of Kerala.
- Salomon, Richard. 1998. *Indian Epigraphy. A Guide to the Study of Inscriptions in Sanskrit, Prakrit, and the other Indo-Aryan Languages*. New York: Oxford University Press.
- Schalk, Peter and Āṅṅappiḷḷai Vēluppīḷḷai (eds.). 2002. *Buddhism among Tamils in Pre-Colonial Tamilakam and Īlam*. 2 volumes (= Acta Universitatis Upsaliensis. Historia religionum 19 + 20). Uppsala: Uppsala University.
- Shanmugam, S.V. 1969. *The Language of Tamil inscriptions 1350-1700 A.D.* unpublished Ph.D. Thesis. Annamalainagar: Annamalai University.
- Subrahmanian, N. and R. Venkatraman. 1980. *Tamil Epigraphy*. Madurai: Ennes Publications.

- Velupillai, A. 1962. *A Study of the Language of the Tamil Inscriptions of the reigns of Caṭavarman Cuntara Pāṇṭiya and Māravarmān Kulacēkara*. unpublished Ph.D. thesis. Peradeniya: University of Ceylon.
- Velupillai, A. 1963-64. *A study of the language of Tamil inscriptions of the period 800-950 A.D.* unpublished D. Phil. Thesis. Oxford: University of Oxford.
- Velupillai, A. 1972. *The Pandya Inscriptions. A Language Study*. Jaffna: Jaffna Archaeological Society.
- Velupillai, A. 1976. *Study of the Dialects in Inscriptional Tamil*. Trivandrum: Dravidian Linguistics Association.
- Velupillai, A. 1980. *Epigraphical Evidences for Tamil Studies*. Madras: International Institute of Tamil Studies.
- Zvelebil, Kamil V. 1964. *Tamil in 550 A.D. An Interpretation of Early Inscriptional Tamil*. Praha: Czechoslovak Academy of Sciences.