Paolo Xella and José Á. Zamora

7 Phoenician Digital Epigraphy: CIP Project, the State of the Art

Abstract: The corpus of Phoenician-Punic inscriptions comprises about 12,000 documents, spread over a very wide area and span of time (all the countries of the Mediterranean region, from the end of the 2nd millennium BCE to the first centuries of the 1st millennium CE). The quantity and nature of the documents have caused considerable difficulties in the knowledge and scientific use of these sources. The project CIP (Corpus Inscriptionum Phoenicarum necnon Poenicarum, also known as the PhDB or Phoenician Data Base) came into being to tackle these problems by producing a collection and a critical edition of all the epigraphic documents in the form of a data bank.

Keywords: epigraphy, North-West Semitics, Phoenician & Punic, corpus, data bank

7.1 Motive of the Project and Institutional Background

According to generally accepted estimates, the corpus of Phoenician-Punic inscriptions comprises about 12,000 inscriptions from all the countries of the Mediterranean. As noted when this project was presented (Cunchillos, Xella, & Zamora, 2005), the sheer quantity and scattered nature of the documents, spread over a very wide span of time, have severely affected research and caused considerable difficulties in the knowledge, availability and use of these sources. In fact, as yet there is not even a simple, complete and reliable list of existing Phoenician inscriptions, still less a critical edition of them. There are only incomplete collections or anthologies (Cunchillos, Xella, & Zamora, 2005, pp. 517-519, with references) most of which need bringing up to date.

This lack of verified documents has had repercussions on the very knowledge of the Phoenician language, making it extremely difficult – and in some cases even impossible – to revise and update basic study tools (such as grammars, dictionaries, concordances, etc.). This state of affairs has seriously restricted the role that epigraphic evidence – the only direct written source of Phoenician culture – should play in general historical information.

In order to tackle these problems and to try to resolve them, by making Phoenician texts – presented with rigorous and uniform criteria – available to the academic...
community, the project CIP – *Corpus Inscriptionum Phoenicarum necnon Poenicarum*, also known as the PhDB or *Phoenician Data Base*, came into being (Figure 7.1).\(^1\) It was born from an Italian initiative that early on made use of collaboration with a CSIC Spanish team (Cunchillos, Xella, & Zamora, 2005; Xella & Zamora, 2007).

![Home menu of the CIP data bank, with basic access links](http://cip.cchs.csic.es/)

*Figure 7.1: Home menu of the CIP data bank, with basic access links*

### 7.2 Aims and General Description of the Project

The project produces a collection and a critical edition of all Phoenician and Punic epigraphic documents in the form of a data bank (realistically speaking, the only form possible). The data bank also aims to include all available information on every Phoenician epigraphic document, presenting the relevant data in an ordered and programmatic form, together with graphic and photographic material.

\(^1\) [http://cip.cchs.csic.es/].
In this electronic edition, the inscriptions are collected and classified according to a uniform and open-ended system, which allows updating of the various regional corpora in real time. This can be done either by inserting new documents or by checking and expanding available information, improving known readings thanks to collations, and by extending data or bibliography in various ways. Each epigraphic document is identified by a unique code, based on a single criterion, with cross-reference to previous editions or collections, together with the most complete and up-to-date bibliography. All the information about the inscriptions can be requested and analysed from various aspects and on different levels (Figure 7.2).

The aim of the CIP is to present a text based on original collations. Whenever it is impossible to collate a document directly, the CIP proposes a textual version resulting from other ways of checking (based on photographs, engravings, copies, etc.), or else based on one or more editions that are expressly indicated in a specific field. As a result, the text presented in the Database is a genuine edition of it, involving the use of a whole set of conventions and critical choices. These parameters are based on those normally used by specialists, with minimal adaptations to the demands of an electronic format. This is in order to accommodate further computerized analysis, such as the generation of automatic segmentation, restorations, concordances, morphological analysis, etc. From a technical point of view, it must be noted that all the textual information is stored and managed in linked “tables” (taking advantage of the use of a relational database).

### 7.3 Basic Technical Data

Such a difficult and wide-ranging project needs the application of new technology (with a precise and suitable methodology, cf. for example Cunchillos, 2000).

Technically speaking and without going into detail, it should be noted that the CIP project uses customized applications for processing data. All the data banks created rely upon well-known commercial software: relational databases with client-server architecture, programmed *ad hoc*. In this way, it is possible to generate new data from the data already entered, to organize the records functionally (in linked “tables”) and to share them, having available a single block of data brought up-to-date in real time. Members of the team can access this main block of data on-line thanks to a simple web browser, independently of the type of hardware or system-software used (Zamora, 1997, 2007).
Figure 7.2: Top: Searching by countries layout. Below: List of inscriptions layout
7.4 Organization and Structure of the Corpus

In the last few years, the decision to use various kinds of information technology has prevented the enormous amount of material accumulated during the various phases of the project from generating dramatic problems in managing the documentation. Instead, it has turned into a genuine “open-ended catalogue”. It is a sort of working edition that has proved to be very useful, right from the start, not only for the development of the project itself, but also, in general, for epigraphic, linguistic and historical research.

As noted above, the organization of the material has been achieved by adapting the tools of information technology to the methodology and aims of epigraphy, as the very structure of the corpus demonstrates. Each document has a main file (a “record”) in a primary table, which contains a set of information arranged (in “fields”) and standardized and normalized where possible and useful: this basic information concerns both the inscribed object (from the date and place where it was found to its formal and material characteristics) and the text (from its transcription to the relevant bibliography), and is linked with more information arranged and distributed in other files (organized in separate related tables). Different layouts allow the user to integrate all the available information in several practical ways (Figure 7.3).

The normalization of part of the information is by no means arbitrary. By distinguishing and standardizing specific data, we have tried to set up reliable and effective criteria for selecting the documents. In this way, uniform research on groups of inscriptions has been possible based, for example: on their find-spot or date, type of inscribed object, technique, material, etc. Research can be carried out on a single topic or on a combination of topics. Research on any other type of information (material, bibliography, text, etc.) or by selecting other criteria is always possible as well.

The dominant criterion for organising these documents is based on the find-spot (conventionally called “locality”) of every inscribed object. As for other data, a linked table has been created (so allowing it to be managed separately) with records for each locality containing more information (together with photographs and maps). This find-spot is part of a geographical area (field: “nation” or “country”) corresponding to the modern country to which it belongs. Although this criterion is only remotely based on historical and cultural reasons (and thus just some meaningful areas of the past vaguely coincide with nations of the present), it allows for a simple and objective classification. Within each nation, when considered useful or necessary, a further regional subdivision is used (field: “region”) identified by a number, which is not based on historical criteria but on clear geographical and/or administrative criteria (Figure 7.4).
This threefold classification (“nation”, “region”, “locality”), completed by a progressively increasing number, forms the set of initials and numbers, or coding sequence, tagging every single inscription, which thus receives a (unique) alphanumeric code that identifies it, producing a “code” or “siglum” of the type lb1byb0001 (= Lebanon/region number 1/Byblos/inscription number 1). In addition, the Database envisages the insertion of cross-references to other collections (for example: CIS I 158 = ICO Sard. 24 = KAI 67 = KI 62). For this purpose, another linked secondary table has been created. The inventory number of the inscribed object in a museum or excavation is also entered into the appropriate field (to facilitate working with groups of inscriptions kept in the same place, for example) linked to another secondary table. Generally speaking, all the standardized fields (material on, and technique by which the inscription is written, type of inscribed object, etc.) are linked to a separate group of files, arranged in a secondary table, with more information.

Even though the CIP does not claim to become a database of epigraphic images (for which there are other projects), it also allows for the insertion of photographs, drawings and various other graphic materials intended especially for rapid identification of the inscribed object (Figure 7.5). In a similar way, the CIP takes advantage of including all the bibliography judged to be relevant for each document to construct a separate (but always linked) bibliographical table.
Figure 7.4: Top: Line of text file (Basic layout). Below: List of find-spots (example of layout integrating data from secondary tables)
7.5 State of the Database and Future Outlook

So far, the CIP has collected a total of more than 9,500 inscriptions, and some regional corpora have been catalogued almost completely. However, there are various levels of work: the corpora with fewer documents have made it easier to produce satisfactory critical editions, as is also the case of corpora that have benefited from research projects with epigraphic implications. Where the mass of documents and the resulting difficulties of collection are greater, instead, the project has indicated setting up a base of information to be improved critically at a later phase.

The future outcome of the project includes not only on-line consultation of the whole corpus but, in addition, conventional and electronic publication of catalogues, regional corpora and other research tools extracted from the Database or derived from the project that will be considered useful for research purposes. In the mainframe of the project, a collective monograph in two volumes, aiming to offer a wide overview of current knowledge on Phoenician epigraphy, is about to be published (Amadasi, Xella & Zamora, forthcoming).
Bibliography


