# Project Description (1994)

## Background

The project described below had its origins several years ago in some consultations in the Netherlands between scholars who saw a need for a new Thesaurus of Ancient (especially, but not only, Biblical) Hebrew. It became apparent to these scholars that a large computer database had many advantages over traditional methods of publication for such a project. The group was widened to include specialists in Hebrew, related languages and linguistics from several European countries and from Israel, and an outline plan was agreed. In 1991 the European Science Foundation agreed to establish and fund a Scientific Network on "The Semantics of Classical Hebrew" under the chairmanship of Professor J. Hoftijzer (Leiden) for three years, so that both theoretical and practical aspects of the design of a semantic database for an ancient language could be explored. These preliminary investigations were carried out in three workshops (Strasbourg, June-July 1992; Florence, April 1993; Cambridge, September 1993 - see *Zeitschrift für Althebräistik* VI/1 (1993) and VII/1 (1994)) and a Symposium (Leiden, July 1994), and a detailed plan for future work was agreed. In February 1994 the Standing Committee for the Humanities of the European Science Foundation expressed its satisfaction with the activities of the Network and recommended them as being "of great intellectual significance and worthy of being assisted in the next stage of its work". [Further details of the work of the Scientific Network can be found in the ESF publication *Scientific Networks* (1993), pp. 46-47.]

This document was originally prepared and approved by the Co-ordinating Committee of the Scientific Network (which is what is referred to in what follows as "The Committee"), but it has been revised and agreed by the Executive Committee which will in future be responsible for the direction of the project (see the section below on Organisation).

## Aim of the Project

It is not the aim of the project to add a new kind of dictionary to the already existing ones, nor to tackle a semantic study on the basis of a special method or methods. Its purpose is to prepare a tool which can be a useful inducement to further semantic research. It will provide a badly needed survey of the results and arguments found in scholarly literature. Thus it will prevent for everyone who uses it the loss of valuable time. Moreover it is planned in such a way that it will be clear not only what work has already been done, but also which workable way of approach has not been used for the study of the lexeme in question. In this way it will stimulate future research by drawing attention to what needs to be done and it will act as a stimulus to fresh, original research. It goes without saying that particularly research in which modern types of approach are used will be encouraged.

## Organisation

These aims can only be realised by international co-operation. The following European universities (represented by the scholars named) have so far indicated their readiness to establish centres for the work of the project, should the necessary funding be obtained: Bonn (Prof. Dr. H.-J. Fabry), Cambridge (Dr. G.I. Davies), Dublin (Prof. Dr. K.J. Cathcart), Florence (Prof. Dr. I. Zatelli), Leiden (Prof. Dr. T. Muraoka and Dr. K. Jongeling), Leuven

(Prof. Dr. A. Schoors, Prof. Dr. M. Vervenne and Dr. P. Swiggers), Oxford (Prof. Dr. H.G.M. Williamson) and Paris (Prof. Dr. A. Lemaire).

Each of these centres will have responsibility for a lexical field or a number of lexical fields (by 'lexical field' is meant a group of lexemes which have a distinct mutual sense-relation). This approach is preferable because the subdividing of the lexical material in this way lends coherence to the description and study of the individual lexemes.

For these centres two levels of central management will be needed:

a) an Executive Committee consisting of representatives of the associated centres, which will oversee the actual running of the project.
b) an Advisory Group, which will include representatives of the ESF and the other funding bodies and some additional specialists in the field.

One centre will be responsible for the construction and maintenance of the database. It will also undertake the programming necessary for the systematic organisation of the input of the database, make bibliographic provisions and oversee the dating and classification of texts. It will, like the other centres, have responsibility for a lexical field or fields. Leiden has been chosen to be this centre.

Every centre will have one or more salaried co-workers. The work on the database can be completed in 10 to 12 years. This period will be subdivided into phases, each phase with its own objective.

## Scope

It has been agreed that the project will concentrate on the language of all texts in Ancient Hebrew: the Hebrew Bible, the Ancient Hebrew inscriptions, the Hebrew text of Ben Sira and the Hebrew Qumran texts.

## Semantic Methods

At the first workshop of the Scientific Network a thorough survey was given of the different types of semantic approach which are used (or: can be used) for the study of Ancient Hebrew. Particular attention was given to the question of the problems which arise when these approaches are used for the semantic description of a language (phase), like Ancient Hebrew, which we only know through a relatively small corpus of texts. The presupposition of the work proposed is that each of these approaches offers as such a legitimate way of tackling semantic problems and that each of them can contribute to a smaller or larger extent to their solution.

## Framework

In the database the arguments and results of semantic scholarly work will be collected in one framework, no matter which method or methods were used in the relevant publications or by the co-workers on the project themselves. For such a database a framework is needed in which the relevant data for each lexeme will be collected. This framework will reflect current insights into the relevant methods and their problems. It will be clear that, in view of the varied kinds of scholarly approach, a neutral way of describing the results arrived at by the

application of each of them is needed. In this respect it appeared helpful to subdivide the framework into sections: for each line of approach (or: for a combination of related lines of approach) a separate section. In each of these sections the relevant results of scholarly work can be described according to the neutral descriptional method. During the second and third workshops of the Scientific Network a decision was reached on the definitive form of this framework. In this framework system sections dedicated to lexical fields will also be added.

The division of the framework into sections helps to keep it surveyable and well-organised. In a final section a general survey will be given of the results reached in the other sections.

## Descriptional method

During the second workshop a neutral descriptional method was accepted which formally, to outward appearance, is the same as that used for componential analysis, without the presuppositions of this kind of approach. This componential description of the results found in scholarly literature should be clearly distinguished from componential analysis as a linguistic approach to semantic studies and from the theories embraced by its practitioners. The use of this descriptional method for the aims of the database is possible because it can easily be adapted for the description of the results of other methods than the one it was designed for originally.

## Material included

In view of the relative uncertainty of nearly every solution in our field it was considered wise to be extremely careful before leaving out scholarly proposals and/or arguments as being invalid. This does not mean that results and arguments of scholarly work ought to be inserted indiscriminately into the framework. Each section (except the final one) will be subdivided into two subsections. In the first one those results and arguments are collected which after assessment are considered to be convincing, in the second one the other material.

## Multi-dimensional approach

During the second workshop it was decided that for the database the more sophisticated approach of a multi-dimensional nature was to be preferred, because it offers possibilities for the combination of data from different parts of one framework or from different frameworks, while leaving open the possibility of making printouts of every lexeme entry, which will always remain helpful. It was the opinion of the committee that the database should be maintained in ASCII format to keep it independent of operating systems and programs that might change over the years. During the third workshop there was presented a pilot project, in which demonstrations were given of how such a multi-dimensional approach could work in practice.

## Terminology

For a database system it is of course necessary to use a uniform system of general, basic terms. Such a system was discussed during the third workshop. It was decided to explore the possibility of the use of the descriptive terminology recently developed by a member of the Committee, Prof. Richter from Munich, on the basis of modern research. A list of grammatical and descriptional terms has been formulated, which can be used when the work

on the database starts, also on the basis of the material provided by Prof. Richter. In this way one can ensure a real coherence for the database work and an effective co-ordination of the work of the different centres. The idea is that the input should be done according to the framework system, but the implementation will be such that the output will not be, as far as possible, defined by it. An offer of Dr. Burnard, a specialist in this field from Oxford, to make an analysis of the framework, using the proposed terminological system, in terms of SGML has been gladly accepted by the Committee.

## Hierarchic approach

To guarantee a satisfactory use of the multi-dimensional approach the introduction of a hierarchic code system is desirable. One could present a fixed code system with a hierarchic structure. However, in this way we would commit ourselves to one semantic approach, excluding other approaches, which would be undesirable. Therefore an approach which gives users the opportunity to make their own choices is preferable, because it makes many different approaches possible. In this way future users can approach a problem from e.g. a formal grammatical or a notional standpoint, according to their own wishes. Such a construction, which is technically possible, will be used for the database project.

## The instructions for the co-workers

The future co-workers on the database project need to be given a set of basic rules. For fruitful co-operation within each centre and among the centres it is of vital importance to have at one's disposal a set of rules which guarantees a uniform approach on a number of central points. Such a set of rules was presented at the third workshop, and agreement was reached about the definitive text. The co-workers will also be provided (by the organizing centre) with a bibliography containing basic lists of commentaries, handbooks to be consulted, agreed abbreviations and, wherever necessary, advice on the use of existing bibliographical reference works. Works not included in the range defined could be found and be referred to if they are cited in the primary selection of literature. The co-workers will also be provided with the framework text (with short introductory remarks), with the list of descriptional terms mentioned above and with a number of sample entries. These sample entries will be provided to avoid unnecessary and undesirable variation. There will be some scope for co-workers to add the results of their own research to what is found in the existing literature. But it is not their first task to supply answers to open questions. It is important that non-experts should not try to add material in specialised fields if they do not know them well (one could, for instance, think here of certain types of etymology and of the ancient translation of the Hebrew Bible, the Versions). In appropriate cases means for publication have already been offered by the *Zeitschrift fr Althebräistik* and by the series Arbeiten zu Text und Sprache im Alten Testament.

## Tagged Bible text

The Committee wants to avoid overlapping with work already done by others. So, for instance, morphologically tagged Bible texts are in existence. During the third workshop Prof. Richter suggested the use of the tagged Bible text prepared by his institute in Munich. The Committee has considered this text acceptable for use by the co-workers on the database. It goes without saying that legal advice has to be sought on the copyright issues. A one-to-one rendering of the codes of this Bible text in English (which will be the language of the project) will be necessary. Moreover, the use of the Munich Bible text will be very helpful,

because it was based on a manuscript text with Tiberian vocalisation, but was coded in such a way that it could also be viewed without the vocalisation and punctuation (added in the second half of the first millennium A.D.) and even without so-called *matres lectionis* (consonant signs used for indication of vowels already in the first millennium B.C. and partly added into the text later on). So comments in the database could be based on all these forms of the texts.

## Specialists

It goes without saying that for cases where help from specialists is needed (for etymology, for instance, from Egyptologists, Akkadian scholars or Iranists), every centre (or group of centres) will look for their own specialists. In some cases it may be possible for a university which is not a main centre of the project to be associated with it for the purpose of giving specialist advice on such matters.

## Relations with other fields

Although the project deals with the semantics of Ancient Hebrew, this does not mean that it will remain out of touch with the work of scholars in fields with a related character and related problems. Some such scholars were guests at the second and third workshops, and one of them gave a lecture at the latter workshop. We have learned from their experiences. The system designed by us is designed in such a way that, with some slight adaptations, it can also be applied to at least a number of their fields. We have considered this to be a second, very important, task. The system proved to be suitable not only for the semantic presentation of older Semitic languages (especially those known only through a relatively small corpus of texts), but also for the semantic presentation of a non-Semitic language like Old Armenian.

## Costs

The main costs to be borne by each centre will be the salary costs of the research worker(s) there. Some provision will also need to be made for expenditure on equipment, reference works, consumables, travel (to meetings of the Executive Committee) etc. Each centre will calculate its requirements in accordance with local circumstances. In addition there will be some central costs to which it is hoped that each centre will make some contribution. These central costs will include the expenses of occasional meetings of the Advisory Group and the cost of services provided by the Leiden centre to the project as a whole, such as computer programming, handling and checking of material sent in by co-workers elsewhere, and the classification of biblical (and other) texts. There is a possibility that much of these central costs will be funded by a special grant.

---

Drafted by J. Hoftijzer (Chair) and G. I. Davies (Secretary)
August 1994