

The Historical Dictionary of the Hebrew Language

Doron Rubinstein



Doron Rubinstein was born in 1964. He studied Hebrew Language and Bible at Tel Aviv University, and taught Hebrew at high school, at Tel Aviv University and at the David Yellin College of Education in Jerusalem. He has been working at the Academy of the Hebrew Language since 1996, and in 2000 became the co-ordinator of the department of modern literature of the Historical Dictionary Project. drubinst@fastmail.fm

Founding the Academy of the Hebrew Language and setting up the Historical Dictionary Project

This article introduces readers to the major research project of the Academy of the Hebrew Language – the Historical Dictionary of the Hebrew Language.

The Academy of the Hebrew Language (AHL) was established following a decree of the *Knesset* [Israeli parliament], “The Supreme Institute of the Hebrew Language Act, 5713-1953”, passed on 27 August 1953. The Academy replaced an earlier institution, ‘The Hebrew Language Council’. The first plenary session was held on 16 November 1953, about three months following the legislation. Its first President was Naphtali Herz Tur-Sinai (Torczyner, 1886-1973, President of the AHL from its foundation until his death). The next Presidents were Professor Ze’ev Ben-Hayyim (born 1907, President 1973-1981), and Professor Joshua Blau (born 1919, President 1981-1993). The current AHL President, since 1993, is Professor Moshe Bar-Asher (born 1939).

About a year after its foundation, on 20 December 1954, at its eighth session, the Academy’s plenum approved an agreement between the AHL and Bialik Institute publishing house of the Jewish Agency, “to publish a historical dictionary of the Hebrew language, containing the lexicon of Hebrew words and their meanings throughout history, from ancient times to our age. This dictionary shall be called ‘The Historical Dictionary of the Hebrew Language of the Academy of the Hebrew Language’” (*Proceedings of the AHL* 1-2.45, 1954-1955). At its next plenary session, on 2 March 1955, the plenum appointed the editorial board of the Dictionary. However, four more years passed before its shape was determined.

The founder and first editor of the Dictionary was Ze’ev Ben-Hayyim. He served as editor until his retirement in 1992, when the current editor, Professor Abraham Tal (born 1931), was appointed.

Ben-Hayyim spent a few months in Europe on his scientific work (end 1957-beginning 1958), and took advantage of this stay to travel and study historical dictionary projects across the continent. In a series of sessions of the editorial board (May-June 1958), he reported his findings and proposed a plan to organize the preparatory work of the Historical

Dictionary (*Proceedings of the AHL* 5.62, 1958). Ben-Hayyim suggested basing work on the Dictionary from its outset on the use of computers, a revolutionary idea in those days.

On 2 January 1959, the editorial board endorsed Ben-Hayyim’s proposal to “decide on one comprehensive historical dictionary to embrace all periods” (*Leshonenu* 23.118, 1959, and *Proceedings of the AHL* 6.87, 1959). This dictionary “may also serve as an excellent basis for the preparation of special Period Dictionaries at some later date.”

This decision determined the shape of the Project from that day onwards, and during the following months, in April-May 1959, the preparatory work for the Dictionary, headed by Ben-Hayyim, started according to that program (*Proceedings of the AHL* 6.87, 1959).

The first attempt to compile such a comprehensive dictionary was made by Eliezer Ben-Yehuda (1858-1922), known as the “reviver of the Hebrew language”: *Millon ha-Lashon ha-Ivrit ha-Yeshana ve-ha-Hadasha* (*A Complete Dictionary of Ancient and Modern Hebrew*, 1908-1959).

The Historical Dictionary Project through the prism of its publications: from the *Source Book* of ancient literature to the *Ma'agarim* CD-ROM

The publications of the Historical Dictionary Project (HDP) offer a fine prism to describe its history: each one of the publications constituted a mid-stage in the process of the Project, and each of them reflected its accomplishments until its publication date. In this section we will review the history of the Project through its publications, practically all of which are publications of the department of ancient literature. One more important publication – the sample pamphlet of the root *'RB* – will be discussed at the end of the article.

The beginning of the Project was devoted to seeking an appropriate way within the framework set by Ben-Hayyim. What did this imply? The HDP had to undertake two tasks before the Dictionary writing could begin: the first – preparing the foundations of the Dictionary and determining its scope, and the second – assembling the sources, that is assembling the texts from which the references would be derived and on which the Dictionary would be based.

This article was translated from Hebrew by IJK, the translation was edited by Raphael Gefen, then additions to the English version have been introduced by DR and IJK. The Hebrew version is available online: <http://kictionaries.com/kdn/kdn12-3-7-heb.html>

Preparing the foundations of the Dictionary – setting the precise working procedure and determining the database structure – lasted several years. The results in their different stages appeared in the first two publications of the HDP: *Barayta di-Melekhet ha-Mishkan* and *Megillat 'Ahima'az*, as described below.

Compilation of sources for the Dictionary was completed in 1963, with the publication of *Sefer ha-Meqorot* [Source Book] for sources “from the canonization of the Bible until the conclusion of the Geonic period”. A second edition was published in 1970. The essence of the book was the HDP’s intention to process all Hebrew writings of the time, which are known to us. The book provided a chronological list of sources that were candidates for processing in the Dictionary, indicating the literary genre of each: inscriptions, documents of the Judean Desert, Rabbinical literature (Talmudic and Midrashic), etc. Further details were given in a separate list: a list of the *mesivot* [transmissions], that is the “good” manuscripts of each work, a list of its important printings, a list of important studies dealing with it, etc. In due course these data were concentrated in the bibliographical records of the Academy’s database. Each source in the HDP database, without exception, has a bibliographical record that includes the above-mentioned data.

The introduction to the *Source Book* served Ben-Hayyim as a forum to raise some of the Project’s problems, two in particular: the problem of chronology and the problem of manuscript selection.

The problem of chronology concerns the time gap between the source itself and the surviving manuscripts in which it was transmitted, i.e. appeared in writing. The transmission of the *Mishnah* is a good example of this problem: the *Mishnah* was compiled in the beginning of the third century CE, but its best manuscript, the Kaufmann Manuscript, dates from the twelfth century CE, nearly a thousand years later! In any assertion whatsoever about Mishnaic Hebrew (and about the language of any other such source) the researcher must take into account the huge gap in time between the *Mishnah* and the Kaufmann Manuscript.

The second problem concerns determining the choice of manuscripts that will be utilized for the purposes of the Dictionary. The Project did not wish to exploit in full all the manuscripts of every given source. Instead, it wanted to utilize manuscripts that met certain chosen criteria, the principal being that only the

earliest manuscripts of each source were to be utilized for the purposes of the Dictionary. A single manuscript out of these was chosen to serve as the “main transmission”, the manuscript according to which the text of the said source was installed in the Academy’s database.

At the same time that the *Source Book* was being completed, first attempts were made to process full sources by computer. In 1961, before the third World Congress of Jewish Studies, there appeared the booklet *Samples of a Concordance and Word Collections of Barayta di-Melekhet ha-Mishkan – an automation attempt in Hebrew research with IBM computers*. This booklet presents the first attempt to process a Hebrew text with the aid of computers, although the processing procedure that is presented in it is not the one that the Project has finally adopted.

Four years later, in 1965, before the fourth World Congress of Jewish Studies, there appeared the booklet *Megillat 'Ahima'az – text, concordance and lexical analysis*. The editorial board wanted thereby to show its method of preparing the material for the Historical Dictionary by means of the computer and of the technological means at its disposal then – from setting up the source to setting up its concordance. This booklet already presents the method of the Dictionary, as was destined to be applied with very few changes almost until the present time.

This method was described in full detail in the booklet *The HDP and the Ways of its Making*, published in 1969, before the fifth World Congress of Jewish Studies. An earlier version of this method was described already in the *Source Book*, pp.232-235 (reprinted in *Leshonenu* 27-28.171-175, 1964).

It is important to emphasize one feature of the concordance. Its structure and its code system enable a uniform treatment of texts from all layers of the Hebrew language: Biblical Hebrew, Mishnaic and Talmudic Hebrew, Medieval Hebrew and Modern Hebrew. The mere existence of this possibility – the possibility to treat texts from all the language layers in one system – is a remarkable illustration of the historical uniformity of Hebrew morphology. Despite all the – very many! – differences between the historical layers of the language, all of them together and each and every one of them on its own represent aspects of **one** language – Hebrew.

Eight years after the publication of *Megillat 'Ahima'az*, in 1973, there appeared *The Book of Ben Sira: text, concordance and analysis of the vocabulary*. The book contained the Hebrew source of the *Book*



Ma'agarim CD-ROM

The Academy of the Hebrew Language, Jerusalem, 1998
Second Edition, 2001
Databases of the *Historical Dictionary of the Hebrew Language*, from the second century BCE to the eleventh century CE

HEBREW SPELLING

Hebrew is a Semitic language, related to Aramaic and more distantly to Arabic. The Hebrew alphabet (*aleph-bet*) consists of 22 letters, all of which represent, primarily, consonants. Four of these – *Aleph, He, Waw* and *Yod* – serve also as vowel letters [*matres lectionis*].

Throughout history, two systems of Hebrew spelling have evolved: vocalized and unvocalized, with many variations within each system. In vocalized spelling, called also ‘defective’ or ‘grammatical’ spelling, all of the vowels are indicated by diacritical vowel points (*niqqud*), and some of them also by vowel letters. In unvocalized spelling, called also ‘plene’ spelling, the vowel points are omitted, but some of them are substituted by additional vowel letters (*Waw* and *Yod*). The latter is the spelling system commonly used in Israel today.

Rules regulating unvocalized spelling were first issued by the Hebrew Language Council and later by the Academy of the Hebrew Language, most recently in 1993. However, some of these rules are often disputed, and many writers and publishers do not fully apply them. Moreover, since the very existence of two spelling systems within the same language is quite confusing, it is gradually being realized that the whole subject should be reconsidered. A concrete proposal for a modest reform was raised recently by Mordechai Mishor in a special session of the Academy of the Hebrew Language (May 2004).

of *Ben Sira* according to the manuscripts that served the HDP, the concordance of the book, as it was set up in the Project, as well as numerous lexicological lists; an index of the *tevot* and their lexicographic entries, a list of the entries in descending frequency order, etc. As pointed out by Ben-Hayyim in the preface, “the editorial board of the Historical Dictionary sought to present this work in the usual form of a printed book *without its being typeset*. The reason for this was that typesetting and everything connected with it were liable to impair the accuracy of the material and its fidelity to the MSS, which had been achieved with such great effort by the mechanographical method. [...] This is perhaps the first time that an attempt has been made to apply this form of printing to a Hebrew book of so extremely complicated [...] nature, both because of the way the contents are arranged and the range of letters and symbols used [...]” (p.viii).

The next two publications of sources of the Historical Dictionary and their concordances were no longer printed on paper. In 1988 a part of the database appeared on microfiche: *Materials for the Dictionary – Series I – 200 BCE-300 CE*. These microfiches were edited in a similar manner as the *Book of Ben Sira* from 1973.

In 1998 there appeared the *Ma'agarim [Databases]* CD-ROM, including the Historical Dictionary’s database from the period of the second century BCE to the first quarter of the fifth century CE. A second edition of this CD was published in 2001, where the database was expanded and included texts up to the first half of the eleventh century.

The department of modern literature and the department of medieval literature

Until 1969 the activity of the HDP focused on ancient literature. That year the department of modern literature was established. This department was intended to process sources from the years 1750-1947. Two events demarcate this 200-year period: beginning in the sixth decade of the eighteenth century, with the publication of the journal *Qohelet Musar* (edited by Moses Mendelssohn; only two issues of which appeared), among the heralds of the *Haskalah* period in Europe, and ending in the year 1947, before the establishment of the State of Israel. This date, which is some decades ago now, offers a better perspective for examining the state of the language in that period.

Two historical facts are at the essence of the work of the department of modern

literature: one – the multitude of Hebrew sources in the said period, and the other – the printed book (or journal) being the major means for disseminating the writings of that period.

Thus, the editorial board made two decisions that differentiated between the work of the department of modern literature and that of the department of ancient literature: first – to process only a selection of sources from the relevant period, not all of them, and second – to process each source according to its first (and sometimes only) publication in print. Excluded from the first decision were three great authors, all the work of whom it was decided to process: Mendeli Mokher Sefarim, Hayyim Nahman Bialik and Shmuel Yosef Agnon. The practical meaning of the second decision made it necessary to find the place of first publication of, for example, each of Agnon’s hundreds of stories – and to locate the journal or the book where it was printed.

In 1977 the first pamphlet of the *Source Book* appeared “for the period from 1750 onwards [...] a selection of writings from the Hebrew belles-lettres (1860-1920)”.

The department of medieval literature was established at the end of 1999, with the aim of treating literature of the 700-year period between 1050-1750 CE. At present this department deals with the Geonic literature.

Intermediate summary

As of April 2004, the HDP databases have been supplied with about 3,500 sources ranging from one or a few words (inscriptions and old coins) to tens of thousands or hundreds of thousands of words (the *Mishnah* and the *Tosefta*, books from the eighteenth and nineteenth centuries, etc). The total number of word tokens in these sources is about 18,500,000, about 8,500,000 of which have received lexicographical entries.

At present, the computer department of the Academy is developing a new software for the HDP, and the three Project departments are gradually proceeding to use it. The transformation to the use of the new software is accompanied by a real revolution in the structure of the Dictionary’s databases, and it necessitates their comprehensive updating. The task of updating the ancient literature database has now been completed, and serves as the base for updating the modern literature database, which is currently being done.

There are several aspects to this updating, the principal ones of which are: completing the vowel points that

for economy reasons was omitted before (indicating the first two vowel points in each word sufficed to distinguish between most of the homographic words in the language); setting standard pointing to those entries, which for various reasons had different spelling from the Hebrew standard today; attributing the Dictionary entries to their roots. If the word's root was not determined, it is attributed to another selected form – the “*neta*”.

**The form of the Historical Dictionary:
The Root 'RB – Specimen Pamphlet
and a thought about the future**

We have so far dealt with describing the work of the HDP, with its source books and its databases. But what will be the form of the Dictionary itself?

The editorial board tried to answer this question in 1982 and published – on the pages of the journal *Leshonenu* – a specimen pamphlet, containing the complete lexicographic treatment of one root of the Hebrew language – the root 'RB (*Leshonenu* 46.3-4.165-267). The root 'RB was chosen as a sample because of the many difficulties it poses to the lexicographer in determining its branches of meanings, whether because of homonymy or because of polysemy. As Ben-Hayyim, who edited the pamphlet, wanted to present in it the continuity of use of the words derived from the root 'RB, he included references not only from the Dictionary's databases – the ancient literature database and the modern literature database – but also from sources from historical layers of the language not yet processed in the HDP.

However, this pamphlet was written and printed when the personal computer and information networks such as the Internet were still in their infancy. The enormous development that has occurred in information technology in recent years requires the Project to adapt continuously, and it will naturally have an influence on the design of the Dictionary. Nevertheless, the infrastructure work that has been done, is being done and will be done on the Project, offers a solid base for the compilation of the Historical Dictionary of the Hebrew Language, whatever end form it may have.

GLOSSARY

Historical terms

Amora *pl.* **Amora'im** אמוראים a sage of the Talmudic period (3rd-5th centuries CE); hence **Amoraic**.

barayta (more commonly **baraita**) *pl.* **baraytot** ברייתא a Talmudic quotation from a Tannaitic source outside the Mishnah; in a wider sense: any piece of Tannaitic material not incorporated in the Mishnah.

Barayta di-Melekhet ha-Mishkan ברייתא דמלאכת המשכן (**barayta**) “on the Building of the Tabernacle”.

Book of Ben Sira (The) ספר בן סירא a book of the Apocrypha; composed by Shim'on ben Jeshua ben 'El'azar ben Sira (2nd century BCE); a.k.a. The Wisdom of Ben Sirach, Ecclesiasticus, etc.

Gaon *pl.* **Geonim** גאונים a post-Talmudic sage, mainly in Babylonia (6th-11th centuries CE); hence **Geonic**.

Haskalah השכלה the Jewish ‘Enlightenment’ movement in Europe (c. 1770s-1880s) that promoted the adoption of secular European culture.

Megillat 'Ahima'az מגילת אהימעץ The Scroll of 'Ahima'az; originally: ספר יוהסין **Sefer Yuhasin** (The Book of Genealogy); composed in Italy by 'Ahima'az ben Palti'el (11th century CE).

Midrash מדרש a legal and legendary commentary on the Bible often characterized by non-literal interpretation; also: an anthology of pieces of this; hence **Midrashic**.

Mishnah משנה the collection of oral law compiled by Rabbi Judah ha-Nasi (beginning of the 3rd century CE); also: a single paragraph of this; hence **Mishnaic**.

Talmud תלמוד a compilation of the Mishnah and its Amoraic commentary (**Gemara** גמרא); there are the Jerusalem (or Palestinian) Talmud (beginning of the 5th century CE) and the more authoritative Babylonian Talmud (end of the 5th century CE); hence **Talmudic**.

Tanna *pl.* **Tanna'im** תנאים a sage of the Mishnaic period (1st-2nd centuries CE); hence **Tannaitic**.

Tosefta תוספתא a compilation of baraytot, arranged according to the order of the Mishnah (end of the 3rd century CE).

Linguistic terms

millon מילון a dictionary; derived from **milla** מילה a word; said to be the first new word introduced into Modern Hebrew by Eliezer Ben-Yehuda (1880).

neta נטע a stem.

niqqud ניקוד (diacritical) vowel points, vowel marks.

shoresh שורש a root; a morpheme that consists of a sequence of (usually) three consonants and carries the basic meaning of a word.

teva *pl.* **tevoth** תיבות a word token.