

Rosaline and Myer Feinstein Lecture Series, 2005

Explorations in the History of Hebrew Indexes By Bella Hass Weinberg

[Slide 1: Title page] This past winter, David Gilner was playing telephone tag with me. He left me a message saying he had good news. Since David was AJL President when I completed editing volume 10 of *Judaica Librarianship* as well as the cumulative index to the journal, I expected the news to be that volume 11 of *Judaica Librarianship* was about to be published. When David finally reached me and told me I had been elected unanimously to be this year's CARLJS lecturer, I was absolutely stunned.

[Slide 2: Hines Award] This news brought to mind my shocked reaction when I got a call from the American Society of Indexers informing me that I had been selected to receive the Hines Award for distinguished service to the profession. The prior winners had been either octogenarians or dead!

At my acceptance speech for the Hines Award, I said, "False modesty is not one of my character flaws, and I can't say ... it never crossed my mind that some day I might get the ... Award, but in considering the prior winners, I thought that at present I lacked two qualifications: eligibility for Social Security, and a subscription to *Modern Maturity*."

[Slide 3: JL 11] In the introduction to the inaugural CARLJS lecture, Michael Grunberger described Feinstein lecturers as "senior librarian[s]." I've never thought of myself in that category. Scary.

I'm particularly honored to be the second female Feinstein lecturer, following Dina Abramowicz, of blessed memory, with whom I was privileged to work at YIVO for nearly three decades. Yiddish was my mother tongue, and I cataloged my father's Hebrew and Yiddish sheet music collection as a little girl. I was hired by YIVO when I was eight to be its consultant on transliteration.

Well, once David Gilner told me the good news, I had to pick a topic. Unlike prior CARLJS lecturers, who spoke on highly specific topics, I selected the theme "Explorations in the History of Hebrew Indexes," a synthesis of my research in this area, much of which has been published in journals devoted to information science and indexing. A couple of the conference papers were presented in Israel and Holland. Several papers on specific aspects of this research have been presented at AJL conferences, but usually during one of five concurrent sessions, and hence very few *Judaica* librarians heard the papers. It's fun to have a plenary session. I hope you will share my excitement about the topic I have selected. I shall present some technical and structural concepts as well as personal reflections on the implications of this research.

[Slide 4: AJL '91] My explorations into the history of Hebrew indexes actually began with an AJL Convention paper. In 1991 I was asked to present a paper on *Judaica* indexes. I thought it would be nice to have a little historical component, and I asked Brad Sabin Hill to give me some

leads because he had written that early printed Hebrew books contained indexes of great sophistication. I recall doing the research at the Gottesman Library of Yeshiva University.

[Slide 5: Shapiro] In 1992, a paper by Fred Shapiro in *Journal of the American Society for Information Science* described what he believed to be the earliest citation index, a table of cases published in the year 1743 in a British book of law. **[Slide 6: Shepard's]** Prior to his discovery, it had been believed that the citation index was invented by Shepard in the 1800s. *Shepard's Citations* shows which older cases are cited in newer ones. The speaker on copyright this morning mentioned the concept of *precedent*, which is basic to British and American case law.

[Slide 7: Toldot Aharon] From dabbling in the history of Judaica indexes, I knew that Hebrew citation indexes had been published as early as the 1500s. *Toldot Aharon* is a well-known work in this genre, which shows where Biblical verses are cited in the Talmud.

During a Research Leave in 1995-96, I looked for Hebrew citation indexes in printed and manuscript form. **[Slide 8: Mafteah ha-Derashot]** The earliest known one is said to have been compiled by Maimonides in his youth, and hence dates back to the 12th century.

[Slide 9: JASIS 97] My study was published in a special issue of *Journal of the American Society for Information Science*. **[Slide 10: JASIS Chart]** The paper contains a table of the earliest Hebrew citation indexes. **[Slide 11: Historical Studies]** It was reprinted in a book called *Historical Studies in Information Science*, one of whose editors is Prof. Michael Buckland, last night's keynote speaker.

[Slide 12: Science Citation Index] The inventor of *Science Citation Index*, Dr. Eugene Garfield, was so pleased to learn how far back citation indexes went that he provided me with a travel grant for further research at the Vatican Library. Because Italy is the cradle of Hebrew printing, I expected the Vatican Library to have early Hebrew indexes. In eight days in July 1998 I examined all of its 800 Hebrew manuscripts that were likely to have indexes as well as all the incunabula (books printed before 1500) held by the Vatican library. **[Slide 13: Moreh Nevukhim]** I found more citation indexes, including one to the Hebrew translation of Maimonides' *Moreh Nevukhim, Guide to the Perplexed*, but no subject indexes.

The Vatican holds more than 40 Hebrew incunabula, but the Jewish Theological Seminary (JTS) has a nearly complete collection. After examining all of them, I inferred that there are no subject indexes in Hebrew incunabula.

[Slide 14: Indexes and Religion] My research in the Vatican Library and at JTS stimulated a think-piece called "Indexes and Religion," which observed that the earliest indexes were for the most part not included in medical manuscripts and printed books, but in theological works. This statement is true for Latin works as well as Hebrew ones. The article also examines the question of why Jews compiled citation indexes but not subject indexes.

[Slide 15: Rouse] In documenting that paper, I encountered the claim of Richard Rouse that subject indexes were invented in France. **[Slide 16: Latin concordance]** In the summer of 1999 I did research in that country, primarily on 13th-century Latin Biblical concordances and subject

indexes. A concordance is a word index; a subject index includes concepts that may not be expressed in the text. In every library I visited, I examined Hebrew manuscripts as well, including Elias Levita's Biblical concordance at the Bibliothèque Nationale in Paris. **[Slide 17: Structure of Masoretic List]** In the gift shop of that library, I bought the catalogue of an exhibition of Hebrew manuscripts called *A la Main Forte*, and in browsing through it one Shabbos morning, it struck me that the lists in Masoretic Bibles were concordance-like structures, with the Hebrew word as the heading, and the Biblical phrase as the locator, since Hebrew Bibles had no chapter and verse numbers in the Middle Ages. This observation stimulated a new direction of research, on the Mesorah, the apparatus designed to standardize the text of the Hebrew Bible and to ensure that it would be transmitted accurately.

In the summer of 2000 I interviewed many experts on the Mesorah in Israel and also examined microfilms or photostats of the relevant manuscripts at the Institute of Microfilmed Hebrew Manuscripts at the Jewish National and University Library in Jerusalem. **[Slide 18: Okhlah]** I spent hours with the microfilm of the Halle manuscript of *Okhlah ve-Okhlah*, a 10th-century work of the independent *Masorah*, that is, a Masoretic compilation not included in or with a Biblical text. *Okhlah ve-Okhlah* consists of numerous thematic lists of words from the Bible, many arranged alphabetically. I did not visit the city of Halle, which is in Germany, but another manuscript of *Okhlah ve-Okhlah* is held in Paris, and a facsimile of it has been published.

[Slide 19: Indexer '01] After doing further research on Masoretic Bibles in England, in the summer of 2001, I described them in a paper published in *The Indexer* in October 2001 **[Slide 20: AJL 2004]** and at last year's AJL Convention.

I could not believe the sophisticated index structures found in the lists compiled by the Masoretes, or Ba'alei Mesorah, in the 10th century. **[Slide 21: Positional]** Among the amazing lists produced by the Ba'alei Mesorah are those that identify words at the beginning of a verse (*rosh pasuk*); **[Slide 22: Compound words]** combinations of adjacent words, such as *tel 'olam*; **[Slide 23: Permuted]** permutations of words, for example, *mishpat va-tsedek, tsedek u-mishpat*; **[Slide 24: Truncation]** and words ending in certain letters, such as double *nun*. All of these indexing structures were previously believed to have been invented in the online era, some 40 years ago.

The Masoretes provided frequency counts for Biblical words, something we might think could have been done only with the aid of computers. At the exhibits someone is selling flash cards of the most frequent words in the Bible. The company should have credited the Ba'alei Mesorah.

[Reflections]

Up to this point I have summarized research on the earliest Hebrew citation indexes and on early Hebrew word indexes compiled by the Masoretes. I shall now share some reflections on this research.

First, in the field of librarianship you get to use everything you've learned. Having majored in French in college certainly helped me negotiate the libraries of France, but the most dramatic example of using what I'd learned occurred at the Vatican Library. During my first few days there, I worked with the catalog of the Vatican's Hebrew manuscripts that was compiled in

Israel; its language of cataloging was, of course, Hebrew. One day I arrived and did not find the catalog in its place. I asked the librarian, and he informed me, “The Cardinal took it!” Even as a full professor, I could not pull rank over the Cardinal. My only alternative was to work with the 18th-century catalogs of Vatican manuscripts, whose language was Latin. I had studied that language for two years in high school, and this turned out to be essential for my research on the history of indexes.

Closely related to the theme of *learning* is that of *knowledge*. [Slide 25: AJL 99] At the AJL Convention held in 1999, I discussed the types of knowledge assumed on the part of users of Hebrew reference works: [Slide 26: Canonical order] knowledge of the order of the books of the Tanakh, of the *parshiyot* (portions of the Pentateuch), of the tractates of the Talmud, and the like. I recall that CARLJS lecturer Jacob Kabakoff was in the audience and seemed to enjoy the discussion.

The idea for that paper emanated from my analysis of Hebrew indexes. Basic to the concept of index is a *known order* of arrangement, but many Judaica librarians admit that they do not know the order of the *parshiyot* or of the tractates of the Talmud by heart.

As my AJL colleagues know, although I occasionally dabble in reference, I am at heart a cataloger. The importance of accurate cataloging hit home again and again as I was doing this research. In each country I visited, I discovered errors in the catalogs of both books and manuscripts. Upon returning from the research trips, I prepared letters explaining the errors and suggesting corrections. The librarians and manuscript curators were always grateful for these error reports.

[Slide 27: *Moreh Tsedek*] It gave me particular pleasure to have made some bibliographic discoveries in the course of this research. The Biblical citation index *Moreh Tsedek* was described by most bibliographers as part 2 of ‘*Avodat ha-Levi*, a citation index to the commandments. YIVO has a separately bound copy of *Moreh Tsedek*, and Yeshiva University’s copy is bound with a different work. I theorized that Hebraica bibliographers had worked with a volume containing two related books bound together, but ‘*Avodat ha-Levi* and *Moreh Tsedek* are two independent works. I convinced Isaac Yudlov of this, and he changed the record in the database of the Institute for Hebrew Bibliography.

[Slide 28: AJL 1998] At the 1998 AJL Convention I described this research in detail in a paper called “How Accurate are Hebraica Catalogs and Bibliographies?: Tracing the Case of *Moreh Tsedek*.” This also relates to the theme of knowledge: How do we know that anything we record in a bibliographic record is correct?

[Slide 29: Genizah] In the summer of 2000 at the Institute of Microfilmed Hebrew Manuscripts, Dr. Ezra Chwat pointed me to items from the Cairo Genizah that might be citation indexes.

[Slide 30: Structure of fragment] I analyzed their structure and presented my findings to him. Some were essentially lists of references to older sources; others were citation indexes. As I describes a fragment in the latter genre, a light bulb went off in Dr. Chwat’s head; he ran to find a photostat of a Genizah fragment that was a match to the one I had in hand. The two fragments

were held by different libraries.

[Slide 31: EAJS] At the Congress of the European Association for Jewish Studies held in Amsterdam in July 2002, I put forth the opinion that when such a case occurs, the library with the smaller fragment should donate it to the library with the larger one. My paper touched on several other subjects, but this controversial opinion generated the most comment from the audience. I am an apolitical person, and this seems to be a political issue. As a researcher I need comprehensive collections, but libraries are possessive of their treasures, even fragments. As the third speaker in this morning's digitization session said, "Libraries compete viciously."

Previously I spoke of using what one has learned. Research on the history of Hebrew indexes has stimulated me to delve into subjects that I had not previously studied. For example, the all-girls' yeshivas I attended did not teach *Torah shebe-'al peh* (Oral Law, that is, Mishnah and Gemara). **[Slide 32: 'En Mishpat]** After analyzing the citation index embedded in the Talmud as well as its other reference apparatus, I was interested in learning more about the structure of the Talmud and started attending classes on this subject. I have since learned so much and am able to contribute to the discussions on the basis of findings from my research.

My all-girls' yeshivas *did* teach Hebrew grammar, but before delving into the Mesorah, my knowledge of it had been limited to *keri* and *ketiv*, notes on variant readings of the Biblical text. I had not known that the Masoretes focused on Hebrew vowels as well as letters. Masoretic notes on vocalization of Hebrew will inform my forthcoming paper for the World Congress of Jewish Studies, on *sheva merahef*.

Prior to conducting historical research I had focused on the newest methods of information organization and retrieval. Exploring the history of Hebrew indexes made me appreciate the auxiliary sciences of history. Dr. Seth Jerchow of the Jewish Theological Seminary identified the watermark on the manuscript of the citation index *Bet Zevul* for me to establish the approximate date when it was written. I was exposed to the disciplines of codicology (the study of manuscript books) and paleography (the study of old writing), which make it possible to date and localize manuscripts.

[Slide 33: Beit-Arié] In Jerusalem I was privileged to consult the codicological database Sfar-Data, which was established by Prof. Malachi Beit-Arié, the world's expert on Hebrew manuscripts, to find out whether Masoretic Bibles were written in 13th-century France. Quite a few were, which lent support to my theory that the Latin concordances got the idea from Hebrew Masoretic Bibles.

[Slide 34: Sotheby's] One of the modern areas of information science in which I have worked is the development of controlled vocabularies. The lack of vocabulary control in the field of indexing is a serious problem for the researcher in this domain. Tables of contents have been called *indexes* in the earliest catalogs of manuscripts, regardless of the language of cataloging - Latin or Hebrew. Last night Michael Buckland described the mistranslation of the term *Register*, which often meant *index*. The translation of the term as "registration" hid Goldberg's invention.

The terminological problem persists to date, as is evident in the recent Sotheby's catalogue: An

item described as having an *index* turned out to have just a table of contents when I examined it. A table of contents follows the order of a book; an index is arranged in a different order, usually alphabetically.

[Slide 35: IFLA] Through a colleague who works for the Rare Book and Manuscript Library of Columbia University, Dr. Consuelo Dutschke, I learned of efforts to prepare a standard for the cataloging of manuscripts. I studied the section on paratext, structures outside the body of the text, including tables of contents and indexes, and discussed desirable additions to the standard at the IFLA conference held in Jerusalem in Summer 2000. The paper was called “Who Invented the Index? An Agenda for Research.” My position was that the presence of indexes in codexes has to be noted consistently to facilitate research in this area.

[Slide 36a: Function words] During that summer, in interviewing experts on the Mesorah who resided in Israel, I found that they hold diametrically opposed opinions on the methods used by the Ba’alei Mesorah to compile lists of Biblical words. **[Slide 36b]** Some of them think that the masters of the Mesorah knew the entire Tanakh by heart and could recall all occurrences of function words such as *‘al* and *el*; others believe that the Ba’alei Mesorah must have had an indexing system because no one could recall such function words.

[Slide 37: Wellisch] Hans Wellisch wrote that Jews did not need concordances because they knew the Bible by heart. Former Librarian of Congress Daniel Boorstin observed in his book *The Discoverers* that indexes begin when memory declines. We today are certainly reliant on indexes to recall all types of information, but the memory power of the Ba’alei Mesorah remains a question. There are relics of the indexing techniques of the compilers of the first Latin Biblical concordance, but unfortunately we have no card files from the Masoretes.

[Slide 38: AJL Abstract] In a recent lecture on Jewish intellectual history, Rabbi Adam Mintz observed, “You’re not a scholar unless someone disagrees with you.” Contemporary experts on the Mesorah are thus true scholars: they disagree on the indexing methods of the Masoretes.

I was always honored when Dina Abramowicz would say to me, “Bella, you are a scholar.” Some might substitute the word *perfectionist* or *fanatic*. **[Slide 39: Me’ir Nativ]** Several years ago in the Lowy Collection in Ottawa, I observed that there was a stoplist in the first Hebrew concordance, *Me’ir Nativ*, which was compiled in the 1400s and first printed in 1524. **[Slide 40: English stoplist]** A stoplist is a list of words not indexed; **[Slide 41: Luhn]** it was until recently believed to have been invented approximately 50 years ago, in the computer era, by Hans Peter Luhn of IBM. **[Slide 42: Hebrew stoplist]** I have mentioned the 15th-century Hebrew stoplist at a couple of conferences, but I would like to write a full paper about it. I’ve collected substantial literature about stoplists in general and about the one in *Me’ir Nativ* in particular, but there’s one document I have to check before I can claim that this Hebrew stoplist was the first one.

The OCLC record for *Me’ir Nativ* says that this concordance was modeled on the Latin one of Arlottus di Prato, compiled in the year 1290. There is one Latin article about Arlottus, which I got, and according to a Classics professor at St. John’s University, who helped me get through it, the article does not mention the stoplist. Although I infer from the compiler’s introduction to *Me’ir Nativ* that he invented the stoplist to save time in compiling the concordance, I shall not be

satisfied until I see Arlottus's concordance and can determine that it does not contain a stoplist.

I do not expect any repository to send me a 13th-century manuscript on interlibrary loan. If I find out where Arlottus's manuscript is, I shall have to travel to see it – unless the entire document has been digitized.

Given that I have no sense of direction – my husband says I can't find my way out of a paper bag – I'm amazed at how much I have traveled to explore the history of indexes, and I hope G-d will give me the strength to travel further to add to this body of research.

[Slide 43: Merton] An important book by sociologist Robert Merton that is related to the concept of citation indexing is entitled *On the Shoulders of Giants*. The book contains Hebrew text documenting the earliest occurrences of the phrase “a dwarf on the shoulders of a giant.” While I am of above-average height for a woman, in all this research I have stood on the shoulders of giants, such as Hans Wellisch, who explored the history of indexes in general, and Shimeon Brisman, who recorded the history of Hebrew indexes.

Descriptions of early Hebrew citation indexes existed before I started this research, but they were published in the literature of Jewish studies, not information science. While I believe I contributed original analyses of these indexes, my contribution was primarily creating a link between Jewish studies and information science by combining my knowledge of Hebrew and indexing, and showing the equivalence of terms in the literatures of the two disciplines.

Besides the giants who preceded me, and whom I cited in my publications, I am indebted to many librarians and manuscript curators who assisted me in a variety of ways. Some have been acknowledged by name in my publications; quite a few are present in this audience.

Information scientists have been pleased to learn that some of their indexing techniques go back so far in time. **[Slide 44: CHF]** In 2002 I was invited to speak at a conference on the history of scientific information systems. **[Slide 45: Call for papers]** The call for papers said that these systems postdate the Second World War. Prof. Buckland made a similar point last night in describing Vannevar Bush's Memex, envisioned in 1945. At the 2002 conference, I described indexes compiled as much as a millennium earlier than World War II. I hope that the members of this audience are proud that quite a few of those structures were invented by those who spoke and read Hebrew, just as Prof. Buckland made us proud that someone name Goldberg invented the search engine.

Anyone who has ever written a dissertation knows that you have to describe the limitations of your research. While my Hebrew is pretty good, I do not know Arabic, and it is possible that those who standardized the Koran in the 7th century had index structures like those of the Mesorah. And as Hebrew grammarians learned much from their Arab colleagues, early Hebrew dictionaries, which were concordance-like reference tools in that they cited the Bible, may have had models in Arabic script.

[Slide 46: Eshet Hayil] As I noted previously, the concept of *known order* is basic to indexes, and alphabetical order is the primary sequence used. Alphabetical order is found in the Tanakh,

for example, in the poem *Eshet Hayil* (Woman of Valor). The concept of citation, or acknowledging an intellectual debt, also has a feminine source in the Tanakh: *Va-taged Esther be-shem Mordekhai* (And Esther said in the name of Mordekhai). In the published version of this paper, I hope to acknowledge all of my intellectual debts, and to arrange the bibliographic references in alphabetical order.

Thank you.