

## Yaacov Choueka: A Profile

The Editors<sup>1</sup>

Yaacov Choueka was born in 1936 in Cairo, Egypt to a family of renowned rabbis who had emigrated from Aleppo to Cairo in the early years of the twentieth century. Yaacov's great grandfather, Hakham Aharon, had been Chief Rabbi of Aleppo, Syria, in the 1870).

In Egypt, Choueka absorbed – and learned to cherish – the Arabic language, on the literary as well as the colloquial level, and developed an appreciation of Egyptian cultural heritage, especially musical heritage. He is quite a *maven* in classical Arabic music, with its unique modalities of *maqamat*, and in the Egyptian musical scene of the 1920s through the 1970s, in particular the musical world of the genial composer Mohammed Abdel Wahab and the inimitable diva Om Kolsum.

At the same time, Choueka's social milieu and his studies at the French Lycée Français opened the world of French history, culture, literature, philosophy and science to him, and in particular the French insistence on clarity, precision and elegance of thought. It is there that he discovered his affinity for mathematical thinking and there that his talents were sharpened and honed.

After the Sinai campaign, in January 1957, Choueka and his relatives followed in the footsteps of other family members who had already left Egypt for Israel.

In October 1957, Choueka started his studies at the Department of Mathematics at the Hebrew University in Jerusalem, where he studied numerous mathematical subjects sitting at the feet of masters of international fame. He found himself, however, mostly attracted to the areas of foundations of mathematics, of mathematical logic and of computability, as sharpened by the use of precise mathematical tools. In 1962, Choueka received his M.Sc. (with distinction), submitting a thesis, supervised by Michael Rabin, on "Ultra-products and non-Standard Analysis", and in 1971 he received his Ph.D., submitting a thesis, also supervised by Michael Rabin, on "Finite Automata on Finite and Infinite Structures".

From the summer of 1962 until the end of 1964, Choueka did his military service in the Israel Defense Forces, ultimately attaining the rank of Captain. During the course

---

<sup>1</sup>Based on material provided by, and discussions with, Yaacov Choueka.

of his service, he attended the first course on programming ever given in Israel and wrote programs in Assembler Language on the first and only computer (the "Philco") then operating in Israel (first, except for an innovative scientific computer – the WEIZAC – built by the Weizmann Institute of Science). In the summer of 1965, he married Sarah Shweka, and together they were blessed with eight children and (at this writing) 20 grandchildren; they hope to celebrate their golden wedding anniversary in a couple of years.

\*\*\*

In October 1964, Choueka joined the Department of Mathematics at Bar-Ilan University, attaining the rank of Full Professor and joining the University Senate in 1990. With the establishment of the Department of Computer Science at Bar Ilan in 2002 (for the creation of which he had struggled hard and stubbornly since the late 1970s), Choueka joined it, remaining there until his retirement, in 2005, as Professor Emeritus of Computer Science.

In the forty years of his academic career, Choueka balanced – in his unique way, which reflects well his personality and his profile – the various aspects of such a career: research, publications, conferences, teaching, guiding students, and academic and public activities. "Cultivating a brilliant student is most certainly more important than publishing a couple of papers which oftentimes have only a short-lived impact", he frequently claimed. Choueka supervised sixteen students in their M.Sc. and Ph.D. studies and theses.

During his early years at Bar-Ilan University, Choueka taught a large variety of subjects in mathematics and computer science, including calculus, set theory, number systems, classical analysis, complex functions, differential equations, linear spaces, and more, concentrating, from the mid-eighties, both in lectures and in research seminars, on mathematical logic, computability theory and automata and formal languages, and, somewhat later, on information retrieval, text algorithms and, finally, on natural language processing.

Until the late eighties, Choueka also taught at several other institutions of higher learning in Israel, offering, for example, a graduate course on automata theory at the Weizmann Institute of Science and a course in computing at the Hebrew University, as well as teaching part-time for a number of years at the Jerusalem College of Technology and at the Jerusalem College for Women. One of his aims in these varied teaching posts was to attract brilliant students to these subjects and encourage them to

continue their graduate studies in these domains. Indeed, this is how some of his former star students, now researchers of international fame, started their career.

\*\*\*

In the mid-sixties, Choueka joined the Responsa Project, a research project in full-text information retrieval systems, initiated a couple of years earlier by Aviezri Fraenkel of the Weizmann Institute of Science. This proved to be a crucial step, one which profoundly influenced Choueka's academic activities, since he became heavily involved in that project and remained involved in it for twenty years. Soon after joining this joint project of the Weizmann Institute and Bar Ilan University, Choueka became the *de facto* leader of the Bar-Ilan team (1967-1974), and in the summer of 1974, when it was decided to transfer the project to the sole trusteeship of Bar-Ilan University, he was appointed Head of the project, a position he held for twelve years.

The Responsa System, a highly successful and sophisticated full-text system for Rabbinic literature, whose corpus spans more than a thousand years of responsa and other works of Jewish heritage, is the first such system ever developed in Israel and one of the very first to be developed for humanities corpora worldwide. With its unique embedded linguistic component, developed by Choueka in the early seventies (!) and never changed or even touched since then, it is still the most influential and most widely-used system of its kind in Israel.

In September 1998, the Responsa Project was selected, along with 49 additional projects, to receive the "Quality Initiative Citation" in a contest for creative, high-quality and visionary projects, in celebration of the 50th anniversary of the establishment of the State of Israel. In 2007, it won the most prestigious prize awarded in Israel, the Israel Prize.

In order to stabilize the Project's activities, and to give it an adequate academic and administrative home, Choueka, with the approval of the University Senate, in 1974, created the Institute of Information Retrieval and Computational Linguistics (two quite novel domains at that time), and chaired it for thirty years, until his retirement in 2005. In the Institute, housed in the framework of the Faculty of Exact Sciences, Choueka initiated and built, in the mid-eighties, *Maba'* ("Bar-Ilan corpus of Modern Hebrew"), a large structured corpus of various strata of Modern Hebrew, a necessary tool for the computerized research of modern Hebrew.

At the height of the Responsa Project activities, a Center for Computers and Jewish Heritage was established at Bar-Ilan, which Choueka also chaired for three years (1983–1986).

Choueka's research interests, which shifted over the years, and the many publications that resulted, detailed in the annexed list of publications, are well represented by the three sabbatical years he spent outside of Israel. His first sabbatical (1974-1975) was at the Department of Mathematics of the University of Illinois at Urbana-Champaign, where he worked and published mainly on automata theory and formal languages. His second sabbatical (1979-1980) was at the Department of Computer Science and the *Institut d'études Médiévales* of the University of Montreal, to which he was invited by the Committee of the Program for Distinguished Visitors. His time there was spent working and publishing on information retrieval and full-text systems (architecture, compression, feedback and linguistic components). His third sabbatical (1986-1987) was at the Bell Communications Research (Bellcore) Laboratories in Morristown, NJ, where he worked and thereafter published on natural language processing (computerized lexicography, dictionaries and grammars; morphology, lemmatization and tagging; word-sense disambiguation; collocation discovery and extraction; corpus linguistics), a year that was followed by two summer visits to that same research laboratory.

Finally, shortly after his retirement, Choueka was asked to head a complex project for the digitization of the large, unique and critically important Cairo Genizah collection of more than 300,000 Jewish manuscripts, and since then, his research and publications shifted to the area of automatic processing of large collections of historical manuscripts and the computer analysis of digital images of such manuscripts.

\*\*\*

Three international trips, the first two taken by Choueka in the early stages of his career and all sponsored by relevant organizations, are worth mentioning.

In the summer of 1975, as a preparatory step for his assumption of the position of Head of the Responsa Project, Choueka made a tour of all the major full-text systems in the United States and Canada, educating himself in the advanced experimental hardware and software that was being developed in this field and establishing professional contacts with all the major players in this - then novel - technology.

In 1980, the Council of Higher Learning in Israel was examining ways for advancing the status of Israel's university libraries in the digital sphere, and Choueka was asked to visit all the libraries in the United States and Europe in which digitized library systems were being installed (there were about ten of them) and present his recommendations. In his report, Choueka detailed the functioning of the various systems, pointed out their grave deficiencies, and recommended against adopting any of them, suggesting instead the fostering of a then-small Israeli company that was developing a library system by the name of Aleph. His recommendation was endorsed, and today, more than thirty years later, Aleph is the implemented system in all Israeli universities, and is, by far, the most common computerized system of library files in Western university and public libraries.

In the summer of 1990, Choueka was invited by the USSR Academy of Sciences for a two-week visit to Moscow and Leningrad (now St-Petersburg), to lecture at, and exchange information with the researchers of, VINITI, the USSR National Information Retrieval Institute in Moscow. The main task of this institute, housed in a gigantic twelve-story building, was to summarize and translate all scientific and technical publications of the West, disseminating them throughout the USSR scientific and technical communities, producing along the way more than a *million* abstracts a year, and yet not one operational projector could be found there to assist Choueka in his lectures. These were the tumultuous final years of the strict communist regime, still holding tight a few short months before the fall of the Berlin Wall. Choueka's visit was colored with events, adventures, and ways of life that imprinted unforgettable images and sensations on his mind and much affected some of his later attitudes in the scientific, cultural and personal spheres.

\*\*\*

Starting in the mid-seventies, Choueka was heavily involved in international and national conference activities, contributing papers, chairing sessions, participating in panels or serving as a member of the organizing or program committees. In fact, many of his contributions, including some of the most frequently cited ones (such as item [43] in the list of publications appearing at the end of this chapter), appeared in conference proceedings and were not subsequently published in journals.

In particular, Choueka organized two international conferences and one national one, all in Jerusalem: the Fifteenth International Conference of the Association for Literary and Linguistic Computing (ALLC, 1988), which he chaired and at which he gave the

opening address, the Second International Conference of the Association Internationale Bible et Informatique (AIBI, 1988), and the Eighteenth National Conference of the Information Processing Association of Israel (IPA, 1983), which he also chaired. The ALLC conference, in particular, was distinguished by three "firsts": it was the first conference in computing and the humanities at which free email was available to all attendants, the first at which remote access was available so that speakers could connect to their systems at home and display them at the conference, and the first time a special panel was organized on "Corpus Linguistics", then an emerging discipline.

A selected list of conferences to which Choueka was invited to give a keynote speech or a plenary talk is included at the end of this chapter.

Two remarkable incidents – probably unique in the world of conferences and speakers - from this quite active conference career are worth mentioning here.

In 1977, Choueka was invited by the organizers of ICCH 3 -- The Third International Conference on Computers and the Humanities (Waterloo, Canada) -- to give an opening plenary lecture on the Responsa Project, one of the first examples at that time of the application of computer technologies to Humanities' domains (incidentally, the other plenary speaker was the Iranian Ambassador to Canada...). While he was in the middle of his talk, Serge Lusignan, one of the conference organizers, came running through the lecture hall, approached the podium, asked Choueka to stop for a minute, and announced through the microphone: "I just got a telephone call from Jerusalem that Choueka's wife, Sarah, gave birth twenty minutes ago to a Choueka boy! Mother and child are feeling well". The standing ovation of the audience (really intended to applaud Sarah) was the last thing Choueka heard at that conference, since he hurried immediately after the announcement to catch the first plane back to Tel-Aviv....

In 1982 Choueka was invited by the Norwegian Research Center for Computers and Law of the University of Oslo, to spend two weeks in Oslo, visiting the Center and giving a talk at the Faculty of Law Auditorium on the Responsa Project, as a prototype of a computerized system for storing and searching legal precedents.

Choueka visited the Center for a couple of days, after which he went touring in Norway with his wife, returning to Oslo to find that the Oslo media was in a frenzy about his planned lecture. Not being able to read Norwegian, he asked the then Chief Rabbi of Norway, Michael Melkior, to explain the situation to him and to keep him updated daily about any developments. These were the days of the Lebanon war, with

Arik Sharon as Minister of Defense and Menachem Begin as Prime Minister, and the newspapers were full of vicious allegations about the Israeli Defense Forces committing war crimes against civilians in Lebanon. Journalists were writing vehemently against the planned talk, questioning how the university could invite a scholar from Israel under such circumstances, demanding that the university cancel the talk, and urging the guest to return immediately to his country. Others wrote that politics should not be mixed with science, that the visit had been arranged a long time ago, that the guest had, personally, nothing to do with what was happening in Lebanon, etc. The balance shifted when the University Students' Organization took sides and sent a formal letter to the university asking it to immediately cancel the talk, threatening the organization of large protests and the use of force, if necessary, to prevent anyone from attending the lecture. Violence was in the air, and Choueka was asked by the university if he would prefer to cancel his talk. Despite the advice of the Israel embassy officials that he should indeed cancel the talk, given the delicate situation, and the fact that they could not provide him with the necessary security measures, he told the university that he would give the lecture, unless the university itself cancelled it.

A few minutes before the talk was to start, he went to the Center offices, from where he was escorted, hands tightly coupled with those of the Center Head, for a ten-minute walk to the auditorium through a very narrow, empty, strip in the midst of a tumultuous sea of thousands of students shouting and brandishing their menacing fists and barricading the entry to the lecture hall. Somehow they got through the door, and there they found a completely empty lecture hall, but for the dean of the university, who had come directly from his office and was waiting on the presidential platform.

The three men waited the customary fifteen minutes, but, faithful to their warning, the students didn't allow anyone to enter the hall. The Center Head asked Choueka if he would like, nonetheless, to give the talk to an empty audience, and Choueka said: certainly, that was why he had come to Oslo. The dean then stood up, arranged his tie, and in a solemn voice congratulated "our distinguished visitor from Israel, etc.", then translated his words into Norwegian, and finally invited Choueka to the podium. Using the projector and his colorful transparencies, Choueka gave an enthusiastic one-hour talk to a large and splendid empty hall, at the end of which the dean stood up and asked if there were any questions, and there being none, since no one was there to ask, and after some hand-clapping, he grabbed Choueka by the elbow, hurried

him through a back door and dark corridors to the rear exit of the university, where a taxi was waiting for him, engine running, which took him at full speed to the airport....

\*\*\*

A striking feature of Choueka's many-faceted academic activities is his talent at initializing, heading, steering and implementing large and complex projects, both in the academic and the national or public spheres, projects that involve a delicate blend of research, technology and applications, on the one hand, and, on the other, require a solid grip on computer technologies, on the subject-matter (texts, Jewish heritage, linguistics) and on project management in general (leadership, budgets and timetables). One such project has to do with his long involvement with Hebrew computerized processing, which began in 1964 when he developed a fully operational, accurate, and complete morphological analyzer for Hebrew, one of the very first for any language. In the early seventies, he developed the linguistic Hebrew/Aramaic component embedded in the Responsa system mentioned above, a component that is still operational today and is still unique among all operational Jewish heritage full-text systems. In 1989, Choueka initiated and acted as Director and Principal Investigator (1989–1997) of *Rav-Milim*, a broad, robust, comprehensive and integrated computerized infrastructure, with the tools and technologies for the advanced processing of contemporary Hebrew texts, at the Center for Educational Technology in Ramat-Aviv. The outcome of this project consisted of several modules, the most important ones being the following: *Milim*, a complete and accurate morphological analyzer for Modern Hebrew (1989), which is still the only such program fully operational and currently embedded in many major governmental, intelligence and commercial full-text systems in Israel (the spelling-checker derived from *Milim* was chosen by Microsoft to be the standard Hebrew spelling-checker for Word); *Nakdan*, a computerized program for the automatic vocalization of Hebrew words, quite a daunting task, and its companion *Nakdan-text* for vocalizing running Hebrew texts, using a large assortment of word-disambiguation techniques; and *Rav-Milim*, a new dictionary of Modern Hebrew (6 vols., CET, Steimatzky and Miskal, 1997), which introduced new and revolutionary ideas and standards in Hebrew dictionary-making and was the first dictionary to adapt and embed the modern international approach and methodologies in dictionary-making inspired by the computerized processing of textual corpora. *Rav-Milim* was also the first dictionary of



Hebrew with an online electronic version enriched with a number of sophisticated grammatical and other modules, available and continuously updated on the Internet ([www.melingo.com](http://www.melingo.com)). *Nakdan-text* won Choueka and the Rav-Milim team the Israel Prime Minister Prize for Computing in 1997, and, earlier, the 1992 Annual Prize of the Information Processing Association of Israel.

Additional projects that Choueka headed and brought to completion, as well as consulting positions of importance that he held, are detailed at the end of this chapter. Those projects that were developed at Bar-Ilan were covered by some 15 grants received from competitive international, as well as national, funding agencies.

\*\*\*

The active participation of Choueka in international meetings was accompanied by international and binational activities, of which we will mention here his position as a member of the Advisory Committee of Domestic, a joint Israeli-German project for full-text retrieval systems with mini-computers; his participation by invitation in the crucial meeting of the Text Encoding Initiative (TEI) that took place at Vassar College (Poughkeepsie, NY) in 1987, an initiative meant to set-up standards for the international exchange of tagged text-corpora, which ultimately gave birth to the XML (and many other xML) languages; his being a member, and then a co-chairman, of the Joint Committee for France-Israel Cooperation in Computer Science Research Projects, established by the Ministry of Science and Technology (1989-1999); and finally his participation by invitation in the NEH-sponsored meeting that took place in Princeton, NJ (1990) on the future of the Princeton-Rutgers National Center for Machine-Readable Texts.

From the early eighties and throughout his career, Choueka was noticeably involved in academic and public activities related to the framework of higher learning in Israel, on one hand, and to computer-related education in the Israeli educational system on the other.

In the sphere of higher education, he served as a member or chairman of the following bodies, all appointed by the Council for Higher Learning in Israel: the Universities' Computing Resources Committee (1983–1986), the High Committee for Appointment of Associate and Full Professors in Colleges in Israel (1985–2013), the Super-Computing Resources Committee (1988–1991), the Committee for Accrediting Haifa University for a Master of Science program in Computer Science (2001–2002), and later (2005) Status Reviewer of this program, the Committee for Accrediting Colleges

in Israel for First Degree (B.Sc.) Programs in Computer Science and Software Engineering (2003-2005), and recently (2013), the Committee for the Evaluation of the Associate/Full Adjunct Professor Degrees in the Institutes of Higher learning in Israel and of the title of Assistant Professor at the Technion.

At the other tip of the educational scale, Choueka played a key role, from the early seventies and until the early nineties, in the design, planning and implementation of the Ministry of Education's preliminary plans to introduce computer studies in all elementary and high schools in Israel. He was a co-organizer and the main lecturer of a first-of-a-kind program in Israel to help teachers of mathematics in schools adapt to teaching computers. These were the days of IBM punch-cards, when programs had to be punched on these cards and fed into the computer, and responses were sometimes received only a week later, and the days of Newton, a small hand-held "computer" with 32K (!) of memory, designed exclusively for writing simple programs in BASIC. His activities in this domain included being a member or chairman of the following bodies, all appointed by the Ministry of Education: the Committee on Computer Education in the Israeli Educational System (1975–1985), the Curriculum Committee for Courses in Computing in the Intermediate and High Schools in Israel (1983–1986), the National Steering Committee for Telematics (Computers and Communications) and chairman of its sub-committee on Computers and Education (1983-1986), the Steering Committee for Computers in Education (1990–1992), the National Planning Committee for Science and Technology in the Educational System in Israel ("Tomorrow 2000") and chairman of its sub-committee on Computers and Education (1990-1992), and a position on the Board of Directors of the Centre for Science and Technology Education and of its sub-committee for Grants (1994-2000). Additionally, Choueka was a member of the following academic or national bodies: the National Council for Information and Communications (1989-1990); the National Broadcasting Authority (1980-1984) and the Committee on Regulating Advertising in Cable Television (Ministry of Communications, 1992); the Committee for Landau Research Prizes of the National Lottery Organization (1989-1992); the Committee for "Chlore" Prizes (1995-1996); the National Committee for Infrastructures in Information Technologies (Ministry of Science and Technology, 1999-2002); and the Advisory Board for Equal Opportunities for Digital Access - *Lehava* (Ministry of Finances, 2002). Choueka also served on the Board of Trustees of *Mivhar*, the College for Science Education, in Bnei-Brak.

\*\*\*

The mid-sixties were the beginning years of computer activities in Israel, preparing the ground for the creation of the Information Processing Association (IPA) of Israel, and for Choueka's close involvement with its operation. Choueka was Chairman of its Working Group on Computational Linguistics (1966-1973), a member of its Executive Board of Directors (1975), a member of its General Council (1975-1988, with some short interruptions), Chairman of its Committee for Annual Awards in Computer Science and Data Processing Applications (1982), the Organizer and Chairman of its Eighteenth National Conference (1983), the Chairman of its Committee for Grants to graduate students in Computer Science (1989), and presented papers at its Annual National Conferences in 1966, 1969, 1972, 1981, 1990, and 1995.

In 1981, Choueka was the recipient (with Paul Bratley from Montreal) of the IPA Annual Award for the Best Research Paper in Computer Science (item [14] in the List of Publications), and he (together with Yoni Neeman and the *Rav-Milim* team) was awarded its Annual Prize for the Best Scientific and Technological Applications of Computers in 1992. In 1990, on the occasion of the 25th anniversary of the IPA, he was awarded a recognition certificate as one of the veterans of computing in Israel, and in 1992, his many services to the Association were recognized by a special award. In terms of public activities, a mention should be made of his participation in the activities of the World Center for Aleppo Jewish Heritage (the Aleppo Jewish Community being one of the oldest Jewish communities in the world, with written records more than a thousand years old). Choueka served as a member of the Board of Directors and Chairman of its Research Committee for thirty years (1985–2005). He initiated and supervised, among others, the publication of a comprehensive bibliography of all published books authored by Aleppo rabbis (numbering about 400), designing each book description as a double spread: a page of biographical and bibliographical information, accompanied by an image of the book's frontispiece (such a written monument apparently has never been erected for any other Jewish community), and of a lavishly-published Pentateuch with the popular and time-honored translation to Arabic (the *Sharh*), a translation that any Aleppine child (boy or girl) was supposed to learn and know by heart. For his many contributions to preserve Aleppine Jewish heritage he was awarded a special recognition plaque in 2004.

To celebrate Choueka's 35 years of teaching and research, the Seventh International Symposium on the Foundations of Artificial Intelligence (BISFAI) was convened in his honor (Ramat Gan, 2001), and then on the occasion of his retirement, the Annual Meeting of the Israeli Society of Computational Linguistics (ISCOL) was convened in his honor (Ramat Gan, 2004).

### **Major Projects and Consulting Positions**

1962-1965: Development of the first operational **Morphological Analyzer** of Hebrew – see above.

1965–1986: The **Responsa** Project – see above.

1968-1975: Chief Advisor for the computerization of the activities of a large **Intelligence Agency**, in particular the development of a sophisticated system for names matching and retrieval. The basic algorithms, considered then to be among the best in the world, were still operational in 2006.

1977–1980: Chief Advisor for the computerization of the comprehensive **Arabic-Hebrew Dictionary**, (A. Sharoni (Ed.) 3 vols., IDF Publishing House, 1987).

1982–1985: Chief Advisor for text processing, information retrieval and electronic publishing of the **Encyclopedia Hebraica**; responsible for the semi-automatic production and automatic publishing of the Encyclopedia Index volume, the first complex book ever to be published in Israel entirely by computerized technologies, with data transfer from the Encyclopedia offices to the printer's offices through regular telephone lines.

1988–1990: **Maba'** - Bar-Ilan Corpus of Modern Hebrew – see above.

1986–1989: Supervising the building (by a graduate student) of **Hyper-Talmud**, a hypertext system for the Babylonian Talmud, one of the first large ('non-toys') hypertext systems ever developed.

1988–1997: **Rav-Milim** – see above.

1997–2002: Chief Advisor to **Yedioth Ahronoth**, the largest daily newspaper in Israel, for "LateNews: a Full-Text and Images Knowledge Base for 60 years of Yedioth Ahronoth", a project in which the full one million pages of the newspaper (from 1939 on) were digitized and made available both as images and as full-text for search and retrieval needs.

- 1999: Responsible for one of the very first projects in Israel for data mining, data verification and data correction concerned with the largest insurance company in Israel databank of millions of data items.
- 2000–2005: Chief Scientist of Celebros, an Israeli start-up concerned with information retrieval, natural language processing and commercial entities.
- 2004–2005: Chief Advisor for the algorithmic aspects of the name-matching module of the Israel Borders Control System, currently the official operational system for controlling passengers' movements in all entry/exit points of Israel.
- 2006–2013: Head of the Cairo Genizah Computerization Project, part of the Friedberg Genizah Project. The related Genizah website the central user interface to the system, currently contains 450,000 (!) high-quality digital images covering all Genizah fragments collections in more than 60 libraries all over the world, about a million data items related to these fragments, and a rich, innovative and highly sophisticated set of tools for manipulating these images and metadata. This is most probably the largest and most advanced website currently operational for dealing with large collections of historical manuscripts.
- 2008–2013: Head of a research group (a cooperation between Tel Aviv University Faculty of Computer Science and the Friedberg Genizah Project) for computer analysis of digital images of historical manuscripts, that developed many innovative and world-first systems for such an analysis, systems that were presented (and published) in more than ten international conferences (winning the best paper award in one of them), which were inspired by - and then applied to - the Genizah case.
- 2012: Head of a project to build a computerized comprehensive corpus of Judeo-Arabic (basically a Jewish dialect of Arabic written in Hebrew characters), a language of critical importance to the study of Jewish history and culture of the Middle Ages, equipped with a full-text research software, and intended to help scholars better understand the language and its structure, and designed so as to apply on it the newest techniques of natural language processing to reveal the various inter-relationships of words and expressions in this language.
- 2012: Head of the Friedberg Talmud Bavli Variants Project, whose aim is to create a website in which will be displayed high-quality digital images of all Talmud manuscripts and first printings all over the world, together with their accurate transcriptions and a novel, very flexible and dynamic way of displaying the

variant readings of any Talmudic text, by presenting a synopsis that can be locally tailored to the needs of every user.

### **Topical List of Publications of Yaacov Choueka**

#### **A - Automata and Computability Theories**

1. Y. Choueka *Bounded Sets of Nodes in Transfinite Trees*, Israel J. of Mathematics, Vol. 11, 1972, 254-257.
2. Y. Choueka, *Structure-automata*, IEEE Trans. on Computers, Vol. C-23, 1974, 1218-1227.
3. Y. Choueka, *Theories of Automata on  $\omega$ -tapes -- a Simplified Approach*, J. of Computer and Systems Sciences, Vol. 8, 1974, 117-141.
4. Y. Choueka, *Book Review of "Computability and Decidability" by J. C. Loekx (Lecture Notes in Economics and Mathematical Systems, No. 68, Springer-Verlag, 1972)*, Philosophia, Vol. 4, 1974, 213-215.
5. Y. Choueka, *Finite Automata, Definable Sets and Regular Expressions over  $\omega^n$  tapes*, J. of Computer and Systems Sciences, Vol. 12, 1978, 81-97.
6. A. Amir and Y. Choueka, *Loop-programs and Polynomially Computable Functions*, Int. J. of Computer Mathematics, Vol. 9, 1981, 195-205.
7. Y. Choueka and D. Peleg, *An Essentially Finite Characterization of omega-regular Languages*, Bulletin of the European Assoc. of Theoretical Computer Science (EATCS), no. 21, Oct. 1983, 21-24.
8. A. Amir and Y. Choueka, *Polynomial Computations in Non-deterministic Loop-programs and PL-programs*, Int. J. of Computer Mathematics, Vol. 4, 1983, 209-221.
9. A. Amir and Y. Choueka, *A Syntactical Definition of the P=NP Problem*, Int. J. of Computer Mathematics, Vol. 17, 1985, 217-228.

#### **B - Full-Text Systems and Information Retrieval**

10. Y. Choueka, *Itur Piskei-din beShitat haText haMale* (Retrieval of Precedents by the Full-text Method), Institute of Criminology Publications, No. 18, The Hebrew University, Jerusalem, 1971, 39-44 (Hebrew).

11. R. Attar, Y. Choueka, D. Schindler and A. Fraenkel, *Kvazim Leshoniyim beMa'arakhot liDlyat Meida'* (Linguistic Files in Document Retrieval Systems), Proc. of the 8th National Conf. of the Information Processing Assoc. of Israel (IPA) (Tel-Aviv, 1972), Y. Moneita (ed.), Information Processing Assoc. of Israel, Jerusalem, 1972, 218-247 (Hebrew).
12. R. Attar, Y. Choueka, N. Dershowitz and A. S. Fraenkel, *KEDMA - Linguistic Tools in Retrieval Systems*, J. of ACM, Vol. 25, 1978, 52-66.
13. Y. Choueka, *Computerized Full-text Systems and Research in the Humanities*, Computers and the Humanities, Vol. 14, 1980, 153-169.
14. P. Bratley and Y. Choueka, *Processing Truncated Terms in Document Retrieval Systems*, Information Processing and Management, Vol.18, 1982, 257-266. (A preliminary version appeared in the Proc. of the 16th National Conf. of the Information Processing Assoc. of Israel (IPA) (Jerusalem, 1981), E. Attar and A. Amit (eds.), The Information Processing Assoc. of Israel, Jerusalem, 1981, 67–93.)
15. Y. Choueka, *Linguistic and Word-manipulation Components in Textual Information Systems*, in *The Application of Mini- and Micro-Computers in Information, Documentation and Libraries*, C. Keren and L. Perlmutter (eds), Elsevier, Amsterdam, 1983, 405–417.
16. Y. Choueka, *Conversational ou Concordances Imprimées: Le Problème de l'Exploitation d'un Gros Corpus*, Informatique et Sciences Humaines, Institut des Sciences Humaines Appliquées, Université Paris-Sorbonne, no. 61-62, 1984, 93-105 (French).
17. Y. Choueka, *Statique et Dynamique dans le Traitement Automatique des Banques Textuelles*, Proc. of the 11th Int. Conf. of the Assoc. for Literary and Linguistic Computing (ALLC) (Louvain-La-Neuve, 1984), J. Hamesse (ed.), Editions Slatkine, Geneva, 1985, 115-120 (French).
18. Y. Choueka, A. S. Fraenkel, S. T. Klein and E. Segal, *Improved Techniques for Processing Queries in Full-text Systems*, Proc. of the 10th Annual Int. ACM-SIGIR Conf. on Research and Development in Information Retrieval (New Orleans, 1987), C. T. Yu and C. J. van Rijsbergen (eds.), Assoc. Computing Machinery, New York, 1987, 306-315.
19. Y. Choueka and A. Aviad, *Yissum Hypertext Ivri beTahanat Avodah* (Hypertalmud - A Hypertext System [for the Talmud] on a Workstation), Proc. of

the 25th National Conf. of the Information Processing Assoc. of Israel (IPA) (Jerusalem, 1990), The Information Processing Assoc. of Israel, Jerusalem, 1990, 281-290 (Hebrew).

20. Y. Choueka, N. Dershowitz and L. Tal, *Matching with Hierarchical Ontology*, Language, Culture, Computation: Studies in Honor of Yaacov Choueka, volume I, N. Dershowitz and E. Nissan (eds.), Lecture Notes in Computer Science, vol. 8000, Springer-Verlag, Berlin [this volume].

### **C - The Responsa Project**

21. Y. Choueka, M. Cohen, J. Dueck, A. S. Fraenkel and M. Slae, *Full-text Document Retrieval: Hebrew Legal Texts (Report on the First phase of the Responsa Retrieval Project)*, Proc. of the ACM Symposium on Information Storage and Retrieval (Maryland, 1971), M. Minker and S. Rosenfeld (eds.), Assoc. Computing Machinery, New York, 1971, 61-79.
22. Y. Choueka, M. Cohen, J. Dueck, A. S. Fraenkel and M. Slae (a solicited expanded form of the preceding paper) *Full-text Law Retrieval - The Responsa Project*, Pamphlet No. 3, Working Papers on Legal Information Processing Series, J. Schweitzer Verlag, Berlin, 1972, 64 pp.
23. Y. Choueka, *The Responsa Project: A Status Report*, Newsletter of Medieval Data Processing, Institut d'Etudes Médiévales, Université de Montréal, Montréal, Vol. VII, 1977, 35-42.
24. Y. Choueka, M. Slae and S. Spero, *The Responsa Project - Revisited*, Proc. of the Assoc. of the Orthodox Jewish Scientists, Vol. 5, Feldheim Publications, New York, 1979, 19-66.
25. Y. Choueka, *Proyekt haShe'elot heHatshuvot (The Responsa Project)*, in Mavo leTikhnut ulHaqarat haMahshev (An Introduction to Programming and Computers), The Open University, Vol.4, 1979, 64-69 (Hebrew).
26. Y. Choueka, M. Slae and A. Schreiber, *The Responsa Project: Computerization of Traditional Jewish Law*, in Legal and Legislative Information Processing, B. Erez (ed.), Greenwood Press, 1980, ch. 18, 261-286.
27. Y. Choueka, *RESPONSA: An Operational Full-Text Retrieval System with Linguistic Components for Large Corpora*, Law in Multicultural Societies, Proc.



- of the Int. Assoc. of Law Libraries (IALL) Meeting (Jerusalem, 1985), E. I. Cuomo (ed.) The Hebrew University, Jerusalem, 1989, 47-82.
28. Y. Choueka, *Responsa, A Full-Text Retrieval System with Linguistic Processing for a 65 million-word Corpus of Jewish Heritage in Hebrew*, Data Engineering, IEEE Computer Society, 12 (4) (Special Issue on non-English Interfaces to Databases), 1989, 22-31.
  29. Y. Choueka, *RESPONSA - A Full-Text System with Linguistic Components for Large Corpora*, in Computational Lexicology and Lexicography, a volume in honor of B. Quemada, A. Zampolli (ed.), Giardini Editions, Pisa, 1990, 181-217.
  30. Y. Choueka, *RESPONSA - A Full-Text System with Linguistic Components for Large Corpora*, Proc. of the 15th Int. Conf. of the Assoc. for Literary and Linguistic Computing (ALLC) (Jerusalem, 1989), Y. Choueka (ed.), Champion-Slatkine, Paris-Geneve, 1990, 51-92.

#### **D - Text and File Compression**

31. Y. Choueka, *Tekhnikot Mehirot le'Itur uShlifa bcMilon ubeKonkordanzia be'Ivrit* (Fast Searching and Retrieval Techniques for Large Dictionaries and Concordances in Hebrew), Hebrew Computational Linguistics, Bar-Ilan University, Ramat-Gan, no. 6, 1972, 12-32 (Hebrew; English abstract p. E-33).
32. Y. Choueka, A. Fraenkel and Y. Perl, *Polynomial Construction of Optimal Prefix Tables for Text Compression*, Proc. of the 19th Allerton Conf. on Communication, Control and Computing, Allerton House (Monticello, Ill., 1981), 1981, 762-768.
33. Y. Choueka, S. T. Klein and Y. Perl, *Efficient Variants of Huffman Codes in High-level Languages*, Proc. of 8th Annual Int. ACM-SIGIR Conf. on Research and Development in Information Retrieval (Montreal, 1985), Assoc. Computing Machinery, New York, 1985, 122-130.
34. Y. Choueka, A. Fraenkel, S.T. Klein and E. Segal, *Improved Hierarchical Bit-vector Compression in Document Retrieval Systems*, Proc. of the 9th Annual Int. ACM-SIGIR Conf. on Research and Development in Information Retrieval (Pisa, 1986), F. Rabitti (ed.), Assoc. Computing Machinery, New York, 1986, 88-96.
35. Y. Choueka, A. S. Fraenkel and S. T. Klein, *Compression of Concordances in Full-text Retrieval Systems*, Proc. of the 11th Annual Int. ACM-SIGIR Conf. on

Research and Development in Information Retrieval (Grenoble, 1988), Y. Chiramiella (ed.), Presses Universitaires, Grenoble, 1988, 597-612.

## **E – Computational Linguistics and Natural language Processing**

36. Y. Choueka and A. Kasher (eds.), *Balshanut Hishuvit* (Computational Linguistics), Proc. of the Conf. of the Information Processing Assoc. of Israel (IPA) Working Group on Computational Linguistics (Ramat-Gan, 1969), IPA Publications, Jerusalem, Vol. 2, 1969 (Hebrew).
37. Y. Choueka, *Balshanut Hishuvit - haYisumim* (Computational Linguistics -The Applications), in *Computational Linguistics*, Y. Choueka and A. Kasher (eds.), IPA Publications, Jerusalem, Vol. 2, 1969, 93-115 (Hebrew).
38. Y. Choueka, *Kvuzat ha'Avoda le'Balshanut Hishuvit – Seker Pe'ilut – Razuy uMazuy* (The [IPA] Working Group on Computational Linguistics - a Report), Hebrew Computational Linguistics, Bar-Ilan University, Ramat-Gan, no. 1, 1969, 55-58 (Hebrew).
39. Y. Choueka and F. Dreizin, *Mechanical Resolution of Lexical Ambiguity in a Coherent Text*, Proc. of the Int. Conf. on Computational Linguistics - COLING 76 (Ottawa, 1976), paper no. 4.
40. Y. Choueka and D. Goldberg, *Mechanical Resolution of Lexical Ambiguity - A Combinatorial Approach*, Proc. of the Int. Conf. on Literary and Linguistic Computing (Tel-Aviv, 1979), Z. Malachi (ed.), Tel Aviv University, Tel-Aviv, 1979, 149-165.
41. Y. Choueka, S.T. Klein and E. Neuwitz, *Automatic Retrieval of Frequent Idiomatic and Collocational Expressions in a Large Corpus*, ALLC J., Vol. 4, 1983, 34-38.
42. Y. Choueka and S. Lusignan, *Disambiguation by Short Contexts*, Computers and the Humanities, Vol. 19, 1985, 147-157.
43. Y. Choueka, *Looking for Needles in a Haystack or: Locating Interesting Expressions in Large Textual Databases*, Proc. of the Int. Conf. on User-Oriented Content-Based Text and Image Handling (RIAO) (Cambridge, Mass., 1988), 1988, 609-623.

44. Y. Choueka (ed.), *Literary and Linguistic Computing: Proc. of the Fifteenth Annual Int. Conf. of the Assoc. for Literary and Linguistics Computing (ALLC), Jerusalem, 1988*, Editions Champion-Slatkine, Paris-Geneve, 1990.
45. E. Gur and Y. Choueka, *A Natural Language Interface in Hebrew for Relational Databases*, Proc. of the 30th National Conf. of the Information Processing Assoc. of Israel (Jerusalem, 1995), The Information Processing Assoc. of Israel, Jerusalem, 1995 (Hebrew).
46. Y. Choueka, E. Conley and I. Dagan, *A Comprehensive Bilingual Word Alignment System: Application to Disparate Languages - Hebrew and English*, in *Parallel Text Processing*, J. Veronis (ed.), Kluwer Publications, 2000, 69-96.
47. Y. Choueka, *Natural Language Processing: Challenges, Achievements and Problems - A Sober Personal Perspective*, Language, Culture, Computation: Studies in Honor of Yaacov Choueka, volume III, N. Dershowitz and E. Nissan (eds.), Lecture Notes in Computer Science, vol. 8003, Springer-Verlag, Berlin, 2013.

## **F – Hebrew Linguistics and Computational Linguistics**

48. Y. Choueka and M. Shapiro, *Nitu'ah Mekhanografi shel haMorphologia ha'Ivrit: Efsharouyot veHesegim* (Machine Analysis of Hebrew Morphology: Potentialities and Achievements), *Lěšonénu* (A Journal for the Study of the Hebrew Language and Cognate Subjects), Vol. 26-27, 1964, 354 -372 (Hebrew).
49. Y. Choueka and M. Shapiro, *Machine Analysis of Hebrew Morphology: Potentialities and Achievements*, *The Finite String*, Vol. 3 (5), 1966.
50. Y. Choueka, *Mahshevim veDikduk: Nituah Mekhani shel haPo'al ha'Ivri* (Automatic Grammatical Analysis of the Hebrew Verb), Proc. of the 2nd National Conf. of the Information Processing Assoc. of Israel (IPA) (Rehovot, 1966), A. Kasher (ed.), Information Processing Assoc. of Israel, Jerusalem, 1966, 49-67 (Hebrew).
51. Y. Choueka, *Automatic Grammatical Analysis of the Hebrew Verb*, *The Finite String*, Vol. 4(4), 1967.
52. Y. Choueka, *Automatic Derivation and Inflection in Hebrew Morphology*, Newsletter of ICRH – Institute for Computer Research in the Humanities, New York University, Vol. IV (11), 1969.

53. Y. Choueka, *Balshanut Hishuvit Ivrit – Lamma veKeizad?* (Hebrew Computational Linguistics, Why and How?) in *Computational Linguistics*, Proc. Conf. of the IPA Working Group on Computational Linguistics (Ramat-Gan, 1969), IPA Publications, Jerusalem, 1969, Vol. 2, 133-138 (Hebrew).
54. Y. Choueka, *Dikduk Yezira Formali laMila haShemanit be'Ivrit* (A Formal Generative Grammar for Hebrew Noun-words), Bar-Ilan Decennial, Vol. 11, M.Z. Kaddari (ed.), Bar-Ilan University Press, Ramat-Gan, 1969, 106-128 (Hebrew).
55. Y. Choueka, *Hipalgut haZurot lefi haOrekh – Ba'ya beBalshanut Statistit* (Length-distribution of Words in Modern Hebrew – a Problem in Statistical Linguistics), Hebrew Computational Linguistics, Bar-Ilan University, Ramat-Gan, no. 1, 1969, 33-34 (Hebrew).
56. Y. Choueka and S. Yeshurun, *Prakim beBalshanut Statistit Ivrit* (Statistical Aspects of Modern Hebrew Prose), Proc. of the 5th National Conf. of the Information Processing Assoc. of Israel (IPA) (Jerusalem, 1969), A. Kasher and U. Schild (eds.), The Information Processing Assoc. of Israel, 1969, 295-319 (Hebrew; English abstract: 26-27).
57. Y. Choueka, *Milhemet haMilon haMilhamti* (The War about the “War Dictionary” [of Isaiah]), Maase Hoshev (Bulletin of the Information Processing Assoc. of Israel, - IPA), 2, 1972, 25-26 (Hebrew).
58. Y. Choueka (ed.), *Hebrew Computational Linguistics*, Bar-Ilan University, no. 5, 1971 and no. 6, 1972.
59. Y. Choueka, *An Exhaustive Computerized Thesaurus of the Hebrew Language*, Abstracts of the Sixth World Congress of Jewish Studies (Jerusalem, 1973), The Int. Assoc. of Jewish Studies, Jerusalem, 1973, abstract E-1.
60. Y. Choueka, *Milim – Ma'arekhet leNituah Dikduki Male, Meduyak uMekhuvan laMorfologia shel ha'Ivrit bat-Zemanenu* (Milim, a System for the Complete, Accurate and On-line Analysis of the Morphology of Current Hebrew), Abstracts of the 7th Conf. for Computers in Education, Ramat-Gan, 1990, The Israeli Assoc. for Computers in Education, p. 62.
61. Y. Choueka, *Rav-Milim, Ma'aerket Kolelet leMihshuv Ha'ivrit veHora'atah* (Rav-Milim, a Comprehensive System for Hebrew Computerization and its Teaching), Abstracts of the 10th Conf. for Computers in Education, Tel Aviv, 1993, The Israeli Assoc. for Computers in Education, p. 116.

62. Y. Choueka (Chief Editor), *Rav-Milim Hatza'ir* (Young Rav-Milim [A Dictionary of Contemporary Hebrew for the Young]), Steimatzki, C.E.T. and Miskal, Tel-Aviv, 2 vols., 1996.
63. Y. Choueka (Chief Editor), *Rav-Milim – haMilon haShalem la'Ivrit haHadasha* (Rav-Milim - the Complete Dictionary of Contemporary Hebrew), Steimatzki, C.E.T. and Miskal, Tel-Aviv, 6 vols., 1997.
64. Y. Choueka, *Rav Milim – A Modern Dictionary for an Ancient but Thriving Language*, Kernerman Dictionary News, July 2004, 27–31.
65. H. Cohen and Y. Choueka, *Hebrew Lexicography*, Encyclopedia of Languages and Linguistics. 2nd ed., Keith Brown (ed.), Elsevier, Oxford, 2006, 268–272.

### **G – The Cairo Genizah, Image Processing of Historical Manuscripts**

66. L. Wolf, R. Littman, N. Mayer, T. German, N. Dershowitz, R. Shweka and Y. Choueka, *Automatically Identifying Join Candidates in the Cairo Genizah*. Int. Journal of Computer Vision, 94(1), 118 – 135, 2011. (Extended abstract in Proc. of IEEE Workshop on eHeritage and Digital Art Preservation, 2009, in conjunction with ICCV 2009 (Tokyo, Japan, 2009), Best Paper Award.)
67. R. Shweka, Y. Choueka, L. Wolf and N. Dershowitz, “*Veqarev otam Ehad el Ehad*”: *Zihuy Ktav-yad veZeruf Kit'ei haGenizah be'Emtsa'ut Mahshev* (Identifying Handwriting and Joining Genizah Fragments by Computer), *Ginzei Qedem* (Genizah Research Annual) Vol. 7, 2011, 171-207 (Hebrew).
68. L. Wolf, N. Dershowitz, L. Potikha, T. German, R. Shweka and Y. Choueka, *Automatic Palaeographic Exploration of Genizah Manuscripts*. In *Codicology and Palaeography in the Digital Age 2*, F. Fischer, C. Fritze and G. Vogeler (eds., with B. Assmann, M. Rehbein and P. Sahle), vol. 2, Books on Demand, Norderstedt, Germany, 2011, 157-179.
69. L. Wolf, N. Dershowitz, L. Potikha, R. Shweka and Y. Choueka, *Computerized Paleography: Tools for Historical Manuscript*, Proc. of the 18th IEEE Int. Conf. on Image Processing (ICIP 2011), B. Macq and P. Schelkens (eds.), Brussels, Belgium, 2011, 3545-3548.
70. L. Wolf, L. Litwak, N. Dershowitz, R. Shweka and Y. Choueka, *Active Clustering of Document Fragments using Information Derived from Both Images*

- and Catalogs*, Proc. Int. Conf. on Computer Vision (ICCV), D. N. Metaxas, L. Quan, A. Sanfeliu and L. J. Van Gool (eds.), Barcelona, Spain, 2011, 1661-1667.
71. Y. Choueka, *Computerizing the Cairo Genizah: Aims, Methodologies and Achievements*, Ginzei Qedem (Genizah Research Annual), Vol. 8, 2012, 9\*-30\*.
72. R. Shweka, Y. Choueka, L. Wolf, N. Dershowitz and M. Zeldin, *Automatic Extraction of Catalog Data from Genizah Fragments' Images*, Literary and Linguistic Computing (LLC) Journal, Special issue on Digital Humanities, 2013. (Extended abstract in Proc. of the Digital Humanities Conf. (DH 2011), Palo Alto, CA, 2011.)
73. Y. Choueka, *miSimta'ot Qahir laInternet haMahir: Genizat Qahir uMihshuvah* (From the Alleys of Cairo to the Internet Highway: The Cairo Genizah and its Computerization), to appear in *Matmonim uGenizot KeParashyot baHistoria*, O. Rimon (ed.), Hecht Museum, Haifa University, Haifa, 2013, 84-96.
74. Y. Choueka, *1,000-year Manuscripts meet Computer Technology the Cairo Genizah Case*, to appear in the Bulletin of the Israeli Academic Center in Cairo, 28, 2013.

## **H – Judaica**

75. Y. Choueka, *al haFormalism shel Hayahdut haHilkhtit* (On the Formalism of Halakhic Judaism), *Deoth*, 21, 1962, 12-28 (Hebrew).
76. Y. Choueka and H. Sabato (eds.), *Minhat Aharon*, a collection of essays in memory of Rabbi Aharon Choueka, Jerusalem, 1980, 357 pp. (Hebrew).
77. Y. Choueka, *Pirkei Haym*, in *Minhat Aharon*, Y. Choueka and H. Sabato (eds.), Jerusalem, 1980, 15-32 (Hebrew).
78. Y. Choueka, *ha'Ona'ah Shtut*, in *Sefer Higayon*, Studies in Rabbinical Logic, M. Koppel and E. Merzbach (eds.), Zomet Institute, Alon Shevut, 1995, 153-163 (Hebrew).
79. Y. Choueka, M. Sabato, *Beit Arba'a Qabbin - leHeshbonah shel Mishna AhatbeMassekhet Baba Batra bidei Mefaresheiha*, to appear, *Sidra*, 2013 (Hebrew).

## **I – Other**

80. Y. Choueka, *haMahshevim baUniversit'aot beIsrael – Skira* (Computers in the Universities in Israel - a Survey), Proc. of the 2nd National Conf. of the Information Processing Assoc. of Israel (IPA) (Jerusalem, 1966), A. Kasher (ed.), Jerusalem, 1966, 93-115 (Hebrew).
81. Y. Choueka and G. Rahaman, *Nitzul haMahshev beMehkarim Hilkhatiim ubeMehkarim beMada'ei hRruah* (The Use of Computers in Halakhic and Humanistic Research), M.T.T., no. 9-10, 1971, 151-154 (Hebrew).
82. Y. Choueka, *Law in the Age of Electronic Computers - Technological Aspects*, Institute of Criminology Publications, no. 18, The Hebrew University, Jerusalem, 1971, 3-8 (Hebrew).
83. Y. Choueka, *The Multi-faceted Contributions of Aviezri Fraenkel to Information Retrieval and the Related Areas*, The Electronic Journal of Combinatorics, 8(2), 2001, paper # I4.
84. Y. Choueka, *From the Alleys of Cairo to the Internet Highways, An Intellectual Journey: Languages, Information and Intelligence*, Language, Culture, Computation: Studies in Honor of Yaacov Choueka, volume I, N. Dershowitz and E. Nissan (eds.), Lecture Notes in Computer Science, vol. 8000, Springer-Verlag, Berlin, 2013 [this volume].

#### **J – Unpublished (Circulated)**

85. Y. Choueka, *On Regressive Functions and a Problem of Erdos and Hajnal*, 1967, 17 pp.
86. Y. Choueka, *Structure Automata*, 1972, 54 pp. (selected sections appeared in Item 2).
87. Y. Choueka, *The Responsa Project – What, How and Why, 1976–1981 – a Status Report*, Institute for Information Retrieval and Computational Linguistics, Bar Ilan University, Ramat-Gan, 1981, 59 pp.
88. Y. Choueka, *Le Traitement Automatique du Gros Corpus: Méthodologie et Techniques*, Invited talk, 9eme Colloque sur l'Ordinateur et le Texte Médiéval, Institut d'Etudes Médiévales, Université de Montréal (Montréal, 1982) (French).
89. Y. Choueka, *Two Conferences on Optical Storage: An Informal Report and Some Reflections*, Technical Memorandum TM-ARH-008691, Bell Communications Research, Morristown, NJ, 1987, 16 pp.

90. Y. Choueka and Y. Neeman, *Nakdan-Text, an In-Context Text-Vocalizer for Modern Hebrew*, The Fourth Bar-Ilan Symposium on Foundations of Artificial Intelligence (BISFAI-95) (Ramat-Gan, 1995).

### Invited Talks

- The Third International Conference on Computing in the Humanities (ICCH 3) (Waterloo, Canada, 1977)
- International Colloquium on the Logic Analysis of Tense and Quantifiers (Stuttgart, 1979)
- Ninth colloquium of the *Institut d'Etudes médiévales* (University of Montreal) on The Computer and the Medieval Text (Montreal, 1982)
- International Conference on the Application of Mini- and Micro- Computers in Information, Documentation, and Libraries (Tel-Aviv, 1983)
- National Seminar of the Information Processing Association of Israel (IPA) on Computers and Natural Languages (Tel-Aviv, 1984)
- Eleventh International Conference of the Association for Literary and Linguistic Computing (ALLC) (Louvain-La-Neuve, 1984)
- Second Congress of the International Association of Law Libraries (IALL) on Law in Multi-Cultural Societies (Jerusalem, 1985)
- International Conference on Value of Research Data for Government and Business (IFDO/ASSISST) (Jerusalem, 1989)
- National Seminar on Hebrew Computational Linguistics (Tel-Aviv, 1989)
- Eighth National Conference of the Israeli Association for Computers and Education (Tel-Aviv, 1990)
- The Sixth National Annual Meeting on Information (INFO 91), Seminar on Full-text Systems (Tel-Aviv, 1991)
- The Seventh National Annual Meeting on Information (INFO 92), Seminar on Full-text Systems (Tel-Aviv 1992)
- International Conference on Cultural Resources in the Electronic Era (Tel-Aviv, 1995)
- Third Mediterranean Exhibition of Technological Innovation (MEDITERINTEC III) (Naples, 1998)



- International Symposium on Semitic Linguistics (Tel-Aviv, 1999)
- Seventh Bar-Ilan International Symposium on the Foundations of Artificial Intelligence (BISFAI 07) (Ramat-Gan, 2001)
- Arabida International Seminar on Complexity (Lisbon, 2004)
- Judaeo-Arabic Culture and Arabic Speaking World: Linguistic, Textual and social Crosspollination (Society for Judaeo-Arabic Studies Conference) (Cambridge, 2011)
- International Colloquium on Written Arabic, Writing Arabic; Corpora and Lexica – WAWA 3 (Roma 2012)