

Lifetime Achievement Award

Martha Palmer and Barbara Di Eugenio Interview Martha Evens

Martha Evens

Martha Palmer preamble: Several of us can remember Martha Evens clearly as a calm, smiling presence at ACL conferences in the 1990s and early 2000s. We knew she was an expert on lexicons and lexical databases—and some of us knew she had been a long serving SIGLEX officer from 1992 to 2004 in various capacities and that she was doing something in the esoteric area of medical informatics, which was pretty much uncharted territory in those days. But mostly we knew she was always interested in what we were doing, asking us about what we were presenting, complimenting us on recent publications, inquiring about our families, and making us feel welcomed and part of the community. Few of us knew that when she graduated Bryn Mawr summa cum laude with a degree in Mathematics in 1955 along with three other women, she had also added German and Greek to the Latin and French she had learned in high school. She was clearly destined for computational linguistics, although the first ACL conference was still 7 years away. After studying more mathematics in Paris as a Fulbright Scholar, she enrolled in a master's program in Mathematics at Harvard/Radcliffe and graduated in 1957. Her eventual husband Len Evens pointed her to an opening as a programmer working with Oliver Selfridge at MIT Lincoln Labs. There she found herself contributing to the first spelling correction program and was immediately hooked on the challenges of using computers to process language. She drove the two boxes of cards containing the first LISP interpreter from MIT to Lincoln Labs as a favor to a friend, not realizing that she would later use LISP extensively herself. She still speaks reverently of her time there with Oliver Selfridge and his continuing influence. Eleven years of marriage and three children later, with her husband settled in a tenured position at Northwestern, she decided to go back to graduate school to study Computer Science at Northwestern. When she received her Ph.D. in 1975, she joined the Illinois Institute of Technology. She had already worked on a Mandarin Chinese parser at Berkeley in the 1960s, so her commitment to natural language processing was long established. In addition to her work on lexicons, she pioneered the development of intelligent tutoring systems for medical topics, and her tutoring system focusing on the circulatory system, called CIRCSIM-Tutor, was used by hundreds of medical students over several years at three different medical schools. She was a co-editor of

<https://doi.org/10.1162/coli.a.00453>

the precursor of *Computational Linguistics* from 1981 to 1984, president of ACL in 1984, and on the Cambridge University Press editorial board for the NLP series from 1982 to 1990. Yet at conferences, she never referred to her professional roles or in any way threw her weight around. She just stayed focused on how best to teach and serve her students and the community. She organized a number of conferences in artificial intelligence and went along to many others with her students, to ensure that as many of them as possible could present papers—and she had a lot of students. In her 25 plus years at IIT, she supervised over 100 Ph.D. students and taught every computer science course on the books except for hardware. In the words of the former Illinois Tech Computer Science chair Eunice Santos, “Martha is very well-known as someone who put her heart into being there to help students. She is very much loved. She is also an incredibly humble person. You’ll learn more about what Martha has done from everybody else than you will from her.” Today we have a rare opportunity to learn a little bit from her. Enjoy.

Interview

Martha Palmer (MP): This is so richly deserved and definitely should have happened sooner but we are really glad we are still able to talk to you about your research and your whole career and the experiences you have had.

Martha Evens (MWE): Thank you for the opportunity. I am so delighted to receive this award. I never thought this would happen, especially since I’ve been out of research for so long.

Barbara Di Eugenio (BDE): Martha, we would really like to know how you got started in natural language processing.

MWE: Right. I majored in Mathematics at Bryn Mawr and graduated in 1955. I got a Fulbright and spent a year in Paris learning more mathematics and went home to study even more mathematics at Harvard. I got a Master’s in Mathematics from Radcliffe. My first day at Harvard, a man named Leonard Evens was sitting on the front steps, greeting the new Mathematics graduate students and showing them around the department. Over that year, we saw each other a lot, and he suggested I try to get a summer job at MIT Lincoln Laboratory, where he had worked in the summer of 1956. I was hired there as a mathematician and, by incredible luck, my supervisor for the summer was Oliver Selfridge. That was the summer when the first copy of the Fortran compiler that left IBM arrived at MIT, so he had me learn to program. He asked me to write a spelling correction program for the Morse code messages, which were full of errors, that the Navy was getting from all over the Pacific. Several years later our program became the first widely available spelling program. Oliver hired me for the next few summers, too. He was very supportive and helpful when I finally applied to graduate school in Computer Science. In the meantime, I got married to Len Evens in 1958 and we produced three wonderful children, and my husband became a tenured Professor of Mathematics at Northwestern after spending some time at University of Chicago and Berkeley. Berkeley was one of the first universities to have a strong Linguistics Department, and while we lived there, I went to their colloquia and read some of the textbooks they used in their courses. Len encouraged me to go back to graduate school in computer science when our youngest child started school. Although I’d been going to listen to some people at the University of Chicago talk about computer science, we thought it would be cheaper and more manageable for me to go to Northwestern.

Northwestern did not yet have a Department of Computer Science, but I knew I wanted to work with Gilbert Krulee, who was then chair of the Engineering Management Department, so I filled out an application for Engineering Management and started that program part-time in 1969. In 1971, Northwestern created a Department of Computer Science with Krulee as department chair, and I moved over along with him. At the time, Northwestern had a policy of not supporting married female graduate students, but I taught courses in computer science for Northwestern and was able to pay my tuition that way. For my Ph.D. thesis, I wrote a program that was able to read and interpret children's stories, and could answer multiple choice test questions about the stories correctly. The program also could tell students whether their answers to the test were correct, and if their answers were incorrect, it could explain to them why the right answer was correct. Based on this work, I got my Ph.D. at Northwestern in Computer Science in 1975. Robert Dewar was leaving the Illinois Institute of Technology (IIT) to chair the NYU Department of Computer Science that year, and I applied for a position at IIT and was hired. Computer Science was also relatively new at IIT, and I was the first professor hired at IIT with a Ph.D. in Computer Science. Around 1973, Krulee had moved to head Northwestern's new Department of Linguistics, and this led me to start talking more to people in linguistics, in particular, Professors Raoul Smith and Oswald Werner (who was also in Anthropology) but also Judith Markowitz and Bonnie Litowitz, who were graduate students in Linguistics. The five of us wrote a book together,¹ published in 1980, and I wrote other papers in NLP with members of this group. This research group helped me develop as a researcher in the years immediately after my Ph.D. My wonderful husband was always very supportive of my work, and he came home early to watch our children so I could go to seminars in the Department of Linguistics at Northwestern. I ended up staying at IIT until I retired in 2001, but they let me keep my laboratory until 2010, and I continued to work with students until then.

BDE: I'm really interested in asking you about CIRCSIM-Tutor, but I also want to point out to the NLP/ACL crowd that you have been a pioneer in AI in education as well and in using natural language processing in that context of education. You were certainly one of the main movers behind CIRCSIM-Tutor, which is one of the first, if not the first, intelligent tutoring system that modeled conversations between the teacher and the students, and tried to look at various features like the student taking the initiative, and what kinds of questions they asked, and so on. So I wanted to ask you how you started on that project, and what challenges you faced. These days AI is widely accepted, especially in health care. Now they want people in NLP to make progress, but when you started this work in the late 1980s and early 1990s, this was not the state of research. So we'd like to hear what you have to say about how you got started on CIRCSIM-Tutor, the challenges you faced, and the satisfactions you found.

[CIRCSIM-Tutor was a pioneer Intelligent Tutoring System to teach cardiovascular physiology; most relevant, it conducted a natural language interaction with the student, using tutoring strategies employed by expert human tutors.² Evens and her group collected and carefully analyzed human-human tutoring interactions in this domain, and investigated a variety of tutoring

1 Evens, M., B. Litowitz, J. Markowitz, R. Smith, and O. Werner. 1980. *Lexical-Semantic Relations: A Comparative Survey*, Linguistic Research, Inc., Edmonton, Alberta.

2 Evens, M. and J. Michael. 2006. *One-on-One Tutoring by Humans and Machines*. Erlbaum, Mahwah, NJ.

strategies and student behaviors, including differences between face-to-face and computer-mediated tutoring sessions; the usage of hinting and of analogies on the part of the tutor; taking initiative on the part of the students; and several domain-based teaching techniques, for example, at which level of knowledge to teach. All of these strategies were implemented, and several were evaluated in careful experiments. CIRCSIM-Tutor was shown to engender significant learning gains, and was used in actual classes, which is even more striking since the NLP technologies available at the time were severely limited. For further details, please see Di Eugenio et al.^{3]}

MWE: Yes, in 1976 a cardiologist named Daniel Hier, who was working at Michael Reese Hospital a few blocks from the IIT campus, called me to ask if there was some way we could write a program that would input a patient's cardiac symptoms and suggest diagnoses; we did that, and wrote several papers based on that program. After Michael Reese closed down, Dan moved to University of Illinois Medical Center and hired one of my students to work there. Some doctors at Chicago Medical School including David Trace contacted Daniel and asked him about using computers in medicine, and he suggested they talk to me. My students and I collaborated with David Trace and others at Chicago Medical School to write a software system called MEDAS that explained diagnoses to a patient and suggested possible treatments, gave discharge instructions, and created a patient database. Chicago Medical School is now called Rosalind Franklin University of Medicine, and is in the far northern suburbs of Chicago, so it was a bit of a drive to go there, but well worth it. When I went there, I met Frank Naeymi-Rad, who was their head computer administrator, and Tim Koschman, a computer scientist there. Frank and Tim got interested in the work we were doing and later got Ph.D.s at IIT working with me. Tim played a key role in developing MEDAS and Frank got doctors interested in medical informatics and later started a company in this field that has done very well. I wrote a number of programs with Frank and Tim and with other students to expand and improve MEDAS, and I think of MEDAS as a kind of great grandmother to the EPIC medical informatics system.

MP: And I understand you had a lot of Ph.D. students—a hundred? How did you manage that? You must have been supervising 10 students at a time or 15, or even 20 all at once. How in the world did you do that?

MWE: Yes, how did I manage? Remember that these students were spread out over 30 years or more. I had several group meetings every week, and when I went home, I often took a thesis draft or a paper home to edit, and spent the evening making comments or suggesting changes. You know, just teaching computer science has been a lot of fun because the students are so excited. In the 1960s, I spent several years teaching mathematics at National College of Education (now Lewis University) to students training to be elementary school teachers, many of whom really hated mathematics. I enjoyed that, but in a very different way. In contrast, my computer science students were full of enthusiasm and very talented. They did excellent work and figured out how to talk to doctors and did most of the real work, and I got to put my name on their papers.

³ Di Eugenio, B., D. Fossati, and N. Green. *Intelligent Support for Computer Science Education: Pedagogy Enhanced by Artificial Intelligence*. CRC Press, 2021.

MP: A lot of your students got their Ph.D.s in things other than natural language processing or medical applications of natural language processing.

MWE: My first few Ph.D. students were students left behind by Robert Dewar, and their theses were related to compilers or databases. However, almost all the rest wrote their theses on natural language processing or lexical databases. In fact, I had 17 students write theses adapting some of what we had learned about natural language processing in English to Arabic. I learned a little Arabic in the process, but not at all well. We learned a lot about Arabic linguistics, and then the same thing happened in Korean and Chinese. So we would have CIRCSIM-Tutor meetings once a week and Arabic and Korean meetings also.

MP: Yeah, I know what that's like. I've done that too. You seemed to go to a lot of conferences. For conferences that were not too far away, did you drive? Did you take a bunch of students together to the conferences? Did you have a van?

MWE: No, but one of my students did. In fact, one of my students drove a taxi to make ends meet.

MP: But I think none of them are taxi drivers now! A lot of your students are professors at universities, or CEOs of companies, or deans, or even one I think is the president of a university. Is that right?

MWE: Yes, they worked hard and did good work. They learned a lot and were very excited about their research. It was just a whole lot of fun.

MP: So, how many of them worked on CIRCSIM-Tutor?

MWE: Oh, dear. I'm not sure.

MP: Just roughly, 25 or 50?

MWE: Well, Michael Glass did a lot of that work. Michael is now teaching and helping to run computer science at Valparaiso University. He played a key role in building the Computer Science department there. Frank Naeymi-Rad added a lot of the medical vocabulary.

BDE: Can I say something about Michael? He was my postdoc for two years, and my most cited paper is a joint paper with Michael, about the squib on the intercoder agreement on Kappa.⁴ Sorry about bringing up my own work but it's to Martha's credit that her graduate students were so good.

MWE: Well, Michael was an undergraduate student at the University of Chicago and he started doing computational linguistics there, so he knew a lot before he came to me. Since then, he has been experimenting with mathematics tutoring programs and has

4 The Kappa coefficient of intercoder agreement [Cohen, J. 1960. A coefficient of agreement for nominal scales. *Educational & Psychological Measurement*, 20: 37–46]; [Di Eugenio, Barbara, and Michael Glass. 2004. The kappa statistic: A second look. *Computational Linguistics*, 30(1):95–101].

done a lot of good work along with Jung Hee Kim, who is now a Professor at North Carolina State.

MP: So you've also done a lot of work with lexical databases. You wrote a book on this subject that just got reprinted not that long ago.⁵ Did you find that the work with lexical databases was useful for the intelligent tutoring systems like CIRCSIM-Tutor?

MWE: Actually, I edited two books on lexical databases and wrote parts of them also. Yes, an intelligent tutoring system needs to know a lot about language. It needs to understand a lot of words to handle questions that students ask and to try to encourage them to keep going, or whatever is needed, and that takes a lot of vocabulary. CIRCSIM was actually invented not by me but by two professors at Rush School of Medicine, Joel Michael and Allen Rovick. Joel and Allen had difficulties teaching medical students in a required first year physiology course how to solve problems in medicine, and it took a lot of time to help each student individually, so they asked me if we could write a tutoring system which would help students work through these difficulties. We took recordings of a lot of one-on-one tutoring sessions between Joel and Allen and their students. Based on these, we did a lot of complicated planning to address communication problems between students and the tutoring system, and we tried a number of different ways to address these issues. In the process, we learned that usually having two medical students work together with the tutoring system worked best in practice. Reva Freedman, who is now a Computer Science Professor at Northern Illinois University, did a lot of the planning and Michael Glass did most of the parsing.

MP: And the system was actually used at Rush Medical School? For how long?

MWE: Well, as long as Joel and Allen were still teaching there at least. There were others at Rush who started using CIRCSIM-Tutor after Joel and Allen retired, too. One of my students, Tim Koschman, taught at Southern Illinois Medical School using CIRCSIM-Tutor, so it was used some there, but I don't know for how long, or how many students used it. After I retired from teaching at IIT, Joel and I wrote a book about the system.

MP: But at Rush, though, there must have been hundreds of students who used the system over several years? That's quite a record for a natural language processing system, especially a tutoring system. Barbara, do you know of any comparable applications?

BDE: The only ones that come to mind are the database tutoring systems from the Tanya Mitrovic group in New Zealand, but they came much later; likewise for the Auto-Tutor series of ITSs on introductory Computer Science (University of Memphis) or the Atlas-Andes ITS on physics (University of Pittsburgh)—interestingly, Reva Freedman contributed to the NLP component of the latter. So that's why I think CIRCSIM-Tutor was really the first tutoring system that took natural language processing seriously.

MP: And that was actually used successfully over a long period of time.

⁵ Evens, M., editor. 1988/2009. *Relational Models of the Lexicon: Representing Knowledge in Semantic Networks*. Cambridge University Press.

MWE: Yes, we also used it for summer students studying physiology.

MP: Were there any moments when you were working on that system or another system when things just seemed to go horribly wrong, and you despaired of whether or not you'd ever be able to get it to work?

MWE: Yes, indeed. However, thanks largely to Frank Naeymi-Rad, we knew that the systems we had written earlier worked. Joel and Allen's students got really interested in the program and helped us by telling us about problems and making suggestions about adding vocabulary and plans.

MP: So they helped you beta test?

MWE: Yes, they told us when there were bugs in the program, so yes, it was very helpful.

MP: Yes, that's right. Getting somebody to use a tutor is always a big advantage.

MWE: Well, the students who used the system liked it and that helped. A lot.

MP: It sounded like you've really enjoyed your Ph.D. students. Was that the most satisfying part of being a computer science professor?

MWE: Oh, well, I don't know whether it's still true but at that point the people who were studying computer science were really excited about it. I enjoyed teaching compiler courses especially.

MP: Yes, you seemed to teach just about everything. I mean a lot of us teach artificial intelligence and natural language processing, but you also taught compilers and programming languages, and I think algorithms. Was there anything you didn't teach? Did you just teach most of the computer science curriculum?

MWE: Oh, I didn't teach any hardware courses. My husband knew hardware and played a major role in setting up the computer network at the Northwestern Department of Mathematics. My husband handled all the hardware at home until he died a year and a half ago.

MP: It's very handy to have a built-in hardware tech at the house. Everybody in my house keeps thinking I'm supposed to do that. So, let's go back to lexical databases. One of the issues with using pre-existing lexical databases is that quite often the system users will come up with new terminology or new phrases for something and it's hard to match those with the existing entries. Do you have any ideas about that? How did you handle that issue in your tutoring system?

MWE: Sometimes we had to ask the students what unfamiliar words meant.

MP: Get the users to repeat what they said or phrase it a little differently? And then did you continually try to update the system? Were you constantly putting in new phrases, taking phrases that initially the system had not understood and then adding them in?

MWE: Well, yes, and some dictionaries try to do that, too. We would buy a new edition of Webster's and the Cambridge dictionaries every time they were available. If we could get them to send us something before it was printed, that was great, because they were starting to build their dictionaries on the computer, too, so they were interested in what we were doing.

MP: So then while your system was being used, were you and your students in a constant state of updating and maintaining it?

MWE: I was very fortunate to get to know several professional lexicographers from around the world, and they invited me to their conferences, which were mostly held in the area around Toronto or in other places in Canada, and that was great for helping us update our systems. These interactions led me to edit the book *Relational Models of the Lexicon: Representing Knowledge in Semantic Networks*.

BDE: Martha, you were a pioneer as a woman in mathematics and computer science. You already told us a little bit at the beginning how you got started in the field. But as one of the very few women in the field, especially at the beginning, do you have any thoughts on the experience? What was the most frustrating part for you? What was the funniest moment that you encountered in the 1980s?

MWE: In the 1980s in the IIT Department of Computer Science we had at one time as many women taking the master's exam in computer science as men. The exact same numbers, as it happened. It only happened once, but I think at that point a lot of early work on compiling lists and building databases was done by women, and IIT encouraged me to have programs every semester for women who might want to do computer science. The Department of Computer Science was in Engineering some of the time and at other times in Arts and Sciences. Computer science was a kind of engineering that people could do without needing much physical strength or doing something dangerous, and I think that attracted a lot of women to computer science. Fortunately, at the time there was a broad push to get more women into computer science.

MP: What was the hardest part? Did you ever have a moment when you felt like a door was shut in your face because you were a woman instead of a man?

MWE: In the other engineering departments, it happened sometimes and mechanical engineering people were particularly unenthusiastic, I think. They actively told women students interested in mechanical engineering to go away and the administration complained when they heard about it.

MP: I bet they did.

MWE: I just kept telling the people in that department that they were losing a lot of good students that way.

MP: That's a good response. Did anything funny ever happen, like somebody knew you were Professor Evens but they didn't realize you were a woman?

MWE: Oh, yes, that certainly happened. People often assumed that I was a secretary and then there were people who had to be told several times that I had a Ph.D. When

I was hired, I was the only woman faculty member in the Department of Computer Science at IIT, but after about five years happily there were more. In the 1980s, we persuaded a lot of women to study computer science, but unfortunately since then the percentage of women students in computer science at IIT has declined to some degree.

MP: What makes you happiest about your career and the work you've done?

MWE: Happiest? Well, all the students from all over the world, getting to know them, and seeing their excitement, and helping them develop their ideas. Their excitement made me feel good and helped keep me going.

BDE: OK, Martha, as a final question for you, would you have words of wisdom for young researchers in natural language processing today?

MWE: I had a lot of luck. When I was starting, a lot of people were excited about what they were doing and passed some of that excitement on to me. I have learned a lot from linguists as well as from computer scientists. I've had a lot of luck, beginning with getting connected to Oliver Selfridge.

MP: That was a good start.

MWE: When my husband was teaching at Berkeley, Oliver visited several times to give talks in linguistics, and I was able to talk to him during these visits. That helped me connect to linguistics research there, and the same thing happened later in the Chicago area. Oliver made a point of getting coffee with me when he came, and that was really encouraging.

MP: So maybe your advice would be to find a great mentor.

BDE: And always be enthusiastic and excited about your research.

MWE: Find something that you can get excited about and then think about ways to communicate that excitement to others.

MP: Well, Martha, thank you so much. Thank you for spending time with us today. But thank you even more for your long illustrious career, including your 100 Ph.D. students, your fielded systems, your 300-odd publications and all of the contributions you've made to the field, both as a researcher and also as a sterling example of what it means to dedicate your life to your teaching and your students.

BDE: I couldn't say it better.

MWE: Thank you so much. Thank you both.