



FREDERICK JELINEK

1932–2010

Elected in 2006

“For contributions to statistical language processing with applications to automatic speech recognition.”

BY BIING-HWANG (FRED) JUANG AND LAWRENCE RABINER

FREDERICK JELINEK, an information-theoretic linguist, died September 14, 2010, of a heart attack, at the age of 77. Jelinek was a Czech American researcher in information theory and natural language processing. Born just before the war, on November 18, 1932, in Czechoslovakia to a Jewish father (who was a dentist and died of disease in the Theresienstadt concentration camp in 1945) and a Catholic mother (who converted to Judaism), Fred Jelinek and his family managed to escape being sent to the concentration camps, due to his mother’s background, and to eventually emigrate to the United States in the early years of the communist regime.

Jelinek studied the newly developing field of information theory at the Massachusetts Institute of Technology under Professor Robert Fano. Jelinek had begun to develop an interest in linguistics after the immigration of his wife, who initially enrolled in the linguistics program at MIT; he often accompanied her to linguistics classes. After his graduate degree, he joined Cornell University in anticipation of the opportunity to work with Charles F. Hockett, a prominent linguist there. Although the intended research program in linguistics fell through, Jelinek went on to teach and develop information theory at Cornell in the subsequent years before joining IBM Research in 1972. Jelinek’s affiliation with Cornell was not a long one, but he deserved credit for the creation and

recognized prominence of Cornell's program in information theory.

Interestingly, at IBM, Jelinek got the opportunity to pursue research in the two fields of most interest to him—linguistics and information theory—through the challenge of designing an automatic typewriter that would respond to the human voice. He saw speech recognition as an information theory problem, rather than the traditional view that speech recognition could (or to some, should) be solved using fundamental principles in acoustics and/or linguistics (or more likely a combination of the two technologies). As an admirer of Claude Shannon, the “Father of Information Theory,” Jelinek's approach followed the fundamental teaching of the “Shannon game,” in which a person competes with a computer, which keeps an array of conditional probabilities that define the odds of a given word following a specified sequence of words, in completing a partial sentence.

Information theorists believe that a machine that chooses the most probable word to succeed the three (or more) preceding (already seen) words would be more statistically likely than the guess made by a human, thereby winning the game. The machine need not actually understand or guess the meaning the sentence really intends to express. The computer-selected word is thus based on the so-called n -gram statistical grammar commonly used in speech recognition systems today. With such types of grammar as the discrete source models, Jelinek treated the speech recognition problem, or rather the natural language processing problem, as a noisy-channel discrete decoding problem and advocated the method of maximum likelihood decoding. This statistical approach to linguistics, built on the framework of information theory, stayed in the course of the axiomatic probability theory, avoided the mathematically intractable problem of human intelligence, but clashed with the conventional discipline of linguistics.

Jelinek was well known for his wry humor, most especially this oft-quoted quip: “Every time I fire a linguist, my payroll goes down and the performance of my speech recognizer goes up.” Jelinek's engineering results, as manifested in systems

such as IBM's Tangora, eventually won the respect of many practically minded linguists or linguistic engineers and earned him many distinctions, including induction into the National Academy of Engineering. Jelinek can be considered a pioneering figure in the field of information-theoretic linguistics and its most prominent advocate.

In 1993 Jelinek retired from IBM and joined John Hopkins University as director of the Center for Language and Speech Processing and Julian Sinclair Smith Professor of Electrical and Computer Engineering. Soon after joining Johns Hopkins, Jelinek started to organize the annual summer workshop on spoken-language research, which has benefitted many young-generation research engineers and students, even to this date.

Jelinek won numerous distinctions. He received the Signal Processing Society Award of the Institute of Electrical and Electronics Engineers in 1997, the ESCA Medal for Scientific Achievement in 1999, the European Language Resource Association's first Antonio Zampolli Prize in 2004, the James L. Flanagan Speech and Audio Processing Award in 2005, and the 2009 Lifetime Achievement Award from the Association for Computational Linguistics. He received a *honoris causa* Ph.D. from Charles University in 2001 and was inducted into the National Academy of Engineering in 2006.

Fred Jelinek is survived by his wife, Czech screenwriter Milena Jelinek; a daughter and son; a sister and stepsister; and three grandchildren.