

SOMETHING OLD AND SOMETHING NEW

Victor H. Yngve
The University of Chicago
Home: 28 Crest Drive Dune Acres
Chesterton, IN 46304
v-yngve@chicago.edu
Tel: (219) 787-8340

As part of my research assistantship as a graduate student in physics at the University of Chicago in the late 1940s, I was involved in preparing instrumentation for detecting incoming cosmic radiation in the stratosphere. The electronic circuits we were using were similar to circuits found in the literature on the new automatic electronic computers, the so-called giant brains. It occurred to me, as it did to others, that these machines might be useful for other purposes besides computing—for instance, translating languages to help eliminate some of the barriers between countries in a post-war world.

Around 1950 we had been experimenting with recording our cosmic ray data on the new tape recorders. At one point I toyed with the idea of building a word-for-word translation machine in which a dictionary in digital form would be magnetically recorded. Such a machine would be comparatively simple to design and build. I realized that the translations would be imperfect, but perhaps they would go a long way toward being useful. The linguistic difficulties involved in word-for-word translation even then loomed larger than the mechanical difficulties. I did not actually try to build such a machine, postponing such work for after I had finished my dissertation.

In the meantime I consulted with some University of Chicago linguistics faculty members and entered on an extracurricular program of reading in the field of linguistics. It seemed clear that a solution for many of the multiple meaning problems that showed up could be achieved if the machine could work out a syntactic analysis for each sentence. I started working on this and a pencil-and-paper scheme gradually took shape.

In looking for a job for after graduation to pursue these interests, I visited Claude Shannon at Bell Telephone Labs. I had read Shannon's paper on the entropy of printed English, and thought that perhaps Bell Labs would be interested in my ideas. They weren't interested, but Shannon told me that there was someone at MIT who was organizing a conference on the subject.

I wrote to MIT for information and was privileged to be invited, while yet a graduate student, to attend the first international conference on machine translation, organized by Y. Bar-Hillel, which took place at MIT on June 17-20, 1952. Through the organizing activities surrounding that conference I became acquainted with the early research of others on MT, including the Warren Weaver memorandum.

I interviewed for a job with Léon Dostert at Georgetown University, and was, if I remember rightly, introduced to Paul Garvin and shown his card file for translating some sample Russian sentences into English. I also interviewed at MIT and subsequently took up a full-time research position there on July 1, 1953, Bar-Hillel having returned to Israel.

In my first major paper on MT, I completed the scheme I had been working on for analyzing German sentences syntactically and translating them into English. This scheme was published in 1955 in the Locke and Booth book that came out of the conference. I believe this was the first published table-driven syntactic parsing scheme and perhaps the first published translation scheme that was completely explicit and programmable. It was actually later programmed.

However, problems showed up in efforts to expand this scheme to include a larger dictionary and additional recognized phrase types. There followed a two-step scheme that separated sentence recognition and sentence production. This was published in 1955 in the journal *MT* that Bill Locke, the chairman of the Modern Language Department, and I had started in March 1954. I was disappointed that Garvin's scheme was not published in his name at the time and could only be gleaned from newspaper accounts.

There followed a three-step scheme, which was also published in *MT*, and we developed the COMIT programming language, which allowed the growing group of linguists and logicians to become productive in MT research. The first successful grammar written in COMIT ran on the IBM 704 over the Fourth-of-July weekend of 1960. Our efforts were centered on the scientific and linguistic issues relevant for advancing the possibility of useful machine translations.

In order to test our grammars, we arranged to drive our sentence construction routine with random specifiers so as to produce a random sample of the output of the grammar. Faults in the grammar showed up easily and we could go back and in essence debug the grammar. This was the first testing of grammars by the method of random generation. It worked admirably, and this scheme served us for many years of grammatical research.

Although our grammars grew in size and complexity and were capable of generating a wide variety of sentences at random, this very success pointed up a number of residual problems with which many of you are already familiar. Some of these problems seemed particularly resistant to solution. We felt overwhelmed. Where could we turn for appropriate theory?

Thus, when I moved to The University of Chicago in 1965, it was with the hope of being able to find out what was wrong with linguistic theory that it could not support such an important application as MT.

I can now report, over 33 years later, that it has required nothing less than laying completely new foundations for general linguistics. I should like to tell you more about this exciting new development, but it would take much too long. For those who are interested in learning more, I have prepared a textbook that you can work through at your own pace (Yngve 1996). You might also wish to visit my web page at:

<http://humanities.uchicago.edu/humanities/linguistics/faculty/yingve.html>

or contact me by e-mail at v-yingve@uchicago.edu.

REFERENCE

Yngve, Victor H. 1996. *From Grammar to Science: New Foundations for General Linguistics*. Amsterdam/Philadelphia: John Benjamins.