

# Yannis Haralambous

Yannis Haralambous is well known for his Omega system and his study of non-Latin fonts.

[Interview completed 6 January 2009.]



*Dave Walden, interviewer:* Please tell me a bit about your personal history.

**Yannis Haralambous, interviewee:** I was born in Athens, Greece, in 1962. My father was one of the first Greeks to obtain a Ph.D. in Geology in the fifties, in Bonn (Germany), and that is where he met my mother. Whether scientific or literary, the house was full of books, in many languages. Already as a teenager I got fascinated by the Fraktur script trying to read an old Brockhaus. One of my hobbies at the time was already related to typography and history: I was collecting stamps with errors (you have to detect and explain the error). My first book, at the age of sixteen, was a catalogue of Greek postal stationeries. I always wanted to be a mathematician, and as I went to a French school (the “Lycée léonin”) it was only natural for me to go abroad for studying. So at the age of 17 I left Greece to study mathematics in Lille, France. I graduated in Pure Mathematics in 1985, and in 1990 did my thesis in the field of Algebraic Topology.

*DW:* Please tell me how you first became involved with  $\TeX$ .

**YH:** In 1987, the Math department of Lille University decided to buy two Mac Plus machines to run a piece of software called *Textures*. A few days later my Ph.D. advisor handed me a copy of *The  $\TeX$ book*. I loved it so much that I read it overnight (what I liked the most was the drawing of the computer drinking coffee). It was that event as well as my encounter with Klaus Thull, in 1988, that completely changed my life and made me switch from Mathematics to Computer Science.

*DW:* Yet you still did your thesis in pure math rather than in computer science. Was your decision to change fields in 1988, yet the actual change happened after you finished your degree in math?

**YH:** No, I took the decision to change fields after my thesis (and after the TUG meeting in Cork), when I realized that working on  $\TeX$  was far more exciting. I still taught math for another year, but spent all my free time on studying oriental scripts to model and create them in  $\TeX$  and Metafont.

*DW:* Why did your advisor think you should read *The TeXbook*?

*YH:* To be able to write my publications and ultimately my thesis in TeX. At that time there was no L<sup>A</sup>TeX yet (or let's say: L<sup>A</sup>TeX wasn't very convincing) and one would still write math in plain TeX.

*DW:* Who is Klaus Thull and what was your encounter with him about?

*YH:* Oh god! You don't know Klaus Thull? He's a great figure of early TeX in Germany. He was the first to write a public domain Metafont for MS-DOS, and for that reason he is an honorary DANTE member, for life. He worked on fractals and on Sanskrit and on many other interesting fields. Unfortunately I lost his trace many years ago. You should absolutely interview him.

My first email correspondences over the Atlantic were with Silvio Levy, Barbara Beeton and Dikran Karagueuzian. When I went to my first TUG meeting (in Cork, 1990), two months after my thesis, I had already developed my Gothic fonts and my first Arabic system, and I was discovering a wide and wonderful new world.

*DW:* Did you use Metafont or another system to develop those fonts?

*YH:* Metafont is a fabulous system for making fonts. Instead of drawing contours one would rather use pen strokes with their dynamics, their tension, etc. And since I worked mostly on historical fonts for various oriental writing systems, using a pen was only natural. Not to mention that metaness is a great thing when you have various font sizes for titles, text, footnotes, critical apparati, marginal notes, etc.

*DW:* My impression is that work with fonts is not typically central to careers in either mathematics or computer science. I see your work history on your website (<http://omega.enstb.org/yannis/>). Please tell me a little about these positions and how your work with TeX and fonts fits in.

*YH:* My research is on digital typography, electronic documents and the preservation of the cultural heritage of the book in the digital era. When I left University I started my own business (Atelier Fluxus Virus), together with my better half, Tereza, who is an artist and a font designer. It was hard but we were able to maintain two parallel activities: typesetting for various publishers all around the world; and doing research, organizing conferences, keeping up with the developments in this field. Finally I accepted a position as Computer Science professor at Télécom Bretagne, a very nice place where I can teach and do research.

*DW:* Please tell me about the Omega project and how John Plaice fit into that effort.

*YH:* In the early nineties all the typesetting systems I developed (Arabic, Hebrew, Syriac, Sinhala, Khmer, Mongolian, and more), were based on pre-processors written in C. A pre-processor will read a TeX file with a special syntax, will process stuff properly tagged and leave the rest unchanged. This is cumbersome and adds yet another complexity level to typesetting. John was still in Canada at that time (in 1993?). He wrote me an email asking whether I would be interested in a new TeX where these things would happen inside the engine. He had a lot of experience writing compilers so he naturally thought of a special Lex-like language for describing what my pre-processors were doing, which would be compiled and run at high speed inside this new TeX. I was enthusiastic and a few days later he was in Lille. We spent days and nights discussing and analyzing the various issues involved in typesetting all writing systems of the world, and one of that nights (between 3 and 5 am) we decided to call the new system Omega. It would be based on Unicode (which was just released) and have its own fonts. And so it did.

*DW:* My understanding is that John Plaice eventually split off from the Omega project. And from the *TUGboat* archives I see you are now working on something called Omega<sub>2</sub> with Gábor Bella. I also think I remember seeing that Aleph was somehow related to Omega and its functionality is being subsumed by Lua<sub>TeX</sub>. Please help me understand what you are working on now and how it relates to those other systems or, for instance, to X<sub>Y</sub>TeX.

*YH:* When John left the Omega project I continued working on it as an experimental platform. My student Gábor Bella did his Ph.D. on a new atomic unit of text, called *texteme*, to solve the problems of the artificial character/glyph duality introduced by Unicode. He partially implemented textemes into Omega, and this is what we called Omega<sub>2</sub>.

Aleph is the typical example of what can happen in the world of free software: a group of people took the code of Omega, added some minor functionalities to it (the extra features of  $\varepsilon$ -TeX, except for bidirectional typesetting, which was already handled by Omega in a different way), and gave it an entirely new name. They did fix some bugs and Aleph is indeed more stable for production than Omega, but does this deserve a new name?

LuaTeX is a great project and I am very glad that the main features of Omega survived in it (although Taco claims he will not be supporting Omega Translations Processes in the long run). A Byzantine music project I'm working on will be entirely based on LuaTeX.

X<sub>Y</sub>TeX is a nice compromise for those who trust their operating system to do part of the typesetting process. What I always liked with TeX is its independence from the surrounding operating system: one is able to fix his/her own rules and re-invent the wheel, if necessary. Not so with X<sub>Y</sub>TeX, which relies on the operating system to do part of the job. Of course, if the end user simply wants a stable and friendly tool to get his/her documents typeset, he/she will find that in X<sub>Y</sub>TeX. But all the adventure of breaking the rules and “boldly go beyond” is lost.

Most of my current projects are not related to building typesetting engines. Whenever typesetting is involved, I use LuaTeX.

*DW:* Do you see yourself producing a production level engine, or do you see yourself more as doing experiments and prototyping that others may pick up for production engines?

*YH:* When Omega started we were rather heading for the first goal, but in the last years we switched to the second one.

*DW:* You directed me to your website for the polytonic system for Greek (<http://www.polytoniko.gr/index.php?newlang=en>). Your website says it must remain neutral but as a lay reader it seems to me that you personally favor the system. What was your motivation for this involvement and what do you hope to accomplish?

*YH:* It is my conviction that the monotonic reform of 1982 has been a monumental mistake of the Greek government, and that it is our moral duty to return to the polytonic system, in order to save the Greek language. The arguments of monotonic linguists are based on the assumption that written language exists *only* to represent oral language, and since accents and breathings have no effect on the latter, they are useless. This assumption is clearly wrong, since, as we all know, written text carries information different than oral—but nevertheless very important. This information is lost when the language is written monotonically.

You misunderstood what I meant by “neutrality”. This Web page is clearly not neutral on the issue of accents and breathings. But it has to remain *politically* neutral;

indeed, among the people requesting the return of the polytonic system there are—unfortunately—some who belong to rather obscure groups (pagans, nationalists, and others). Their presence can give false impressions to people who are discovering our initiative and our arguments.

What I hope to accomplish is to persuade people that they have been fooled by unscrupulous politicians and linguists, and that their cultural heritage has been stolen from them, under their noses. Many people realize at last that the “simplification” of this reform breaks the historical continuity of the language. Please visit our Web site to learn more about this very important issue.

*DW:* Is Atelier Fluxus Virus continuing business? If so, what is the connection between your research projects and your business?

*YH:* Yes, of course. Tereza has taken over Atelier Fluxus Virus, with clients in France, the US and Greece. There is practically no connection between my research projects and the business, at least for the moment.

*DW:* More generally, your research and other involvements cover a broad area. To what extent do you see T<sub>E</sub>X being relevant (and how) to these domains in the future?

*YH:* I can’t imagine myself starting a project which involves document production without using T<sub>E</sub>X. It is my conviction that an electronic document should be optimally presented, including on the text level. Building typesetting engines other than T<sub>E</sub>X would be re-inventing the wheel, and I think that Don has invented a very nice wheel for us. Of course, when we speak of the “T<sub>E</sub>X invention”, we refer to many things which are not all equally important. I think that the most fundamental aspect of T<sub>E</sub>X is the node model: to consider a text as chain of characters, penalties, glue and other types of nodes. This model reflects the rationale of traditional typography and is much more powerful than the character/glyph model of the Web. Of course, it can become even more powerful if we consider additional properties of nodes or new node types. This text model, together with the algorithms used in T<sub>E</sub>X, is, IMHO, the heart of T<sub>E</sub>X, and this deserves to survive.

A typical example of a research project involving T<sub>E</sub>X is an ongoing project with the Voltaire Museum of Geneva on building a computer-driven Monotype machine. T<sub>E</sub>X is producing a DVI file using special fonts and then we post-process the result to adapt it to the mechanical constraints of the Monotype machine.

*DW:* I have thought before about buying a copy of your book on Fonts and Encodings (the English edition), and now have a copy in hand—very impressive. Please tell me about your effort to publish that book.

*YH:* I have been dealing with encodings and with font formats ever since the late eighties. At some time I realized that I had gathered a lot of contacts and documentation and that this subject was not treated in any published book. Everybody knows what PANOSE is, but it is very hard to find its specification. There is a description of TrueType hints on the Microsoft Web site, but it is rather obscure and can certainly not be used as a primer. So I proposed to Xavier Cazin, a good friend working at the French division of O’Reilly, to write a book on these issues. He immediately encouraged me and we agreed on the basis of a 600 page book. As it is often the case, while working on the book I discovered more and more subjects, and we ended up with a bit more than a thousand pages (most of it typeset in 9 points!). It was a very intense experience and I am very proud of it. The book is dedicated to my father, who passed away while I was writing it.

As Xavier explained to me, when O’Reilly US is interested in a book published by some branch outside the US, they will ask a US author to write on the same topic. In my case

and for the first time, they decided to translate the book. I was very lucky since I found the ideal translator: Scott Horne is a native English speaker with an outstanding talent for learning an uncountable number of languages. He is a computer scientist, a  $\TeX$  user and a musician. And a perfectionist in everything he does. He did a first translation of the book and as I was reading it I rewrote many parts which needed to be updated, so that the English version was up-to-date when it came out (in the fall of 2007). Not only did he find and correct all the errors of the original edition, but he even translated jokes and *jeux-de-mots* so that I got very positive critics saying “The result is a book that seems to have been written by a human not a droid” (Rick Jelliffe) and “Thus this book becomes the first technical reference text we’re aware that actually contains a running gag” (Designorati).

Currently (early 2009), the book is being translated into Russian. My knowledge of this language is probably not sufficient to interact with the translator in the same way I did with Scott. But I will rewrite parts of the book which need to be updated (latest versions of FontLab and FontForge, font management in Lua $\TeX$ , photofonts, etc.).

DW: In one of our email exchanges, you mentioned “third-world” writing systems. Please tell me more about this.

YH: Computing has been invented and developed in the West. The way we deal with text is not compatible with oriental writing systems. For example, Chinese ideograms are either encoded using Unicode code points (which means that adding new characters is lengthy and can certainly not be spontaneous) or by using higher-level markup like CDL (which needs special software and uses dozens, if not hundreds, of bytes for a single character): neither approach is optimal. Another example: the Arabic writing system is fluid; typography (both hot lead and electronic) segments Arabic into individual glyphs, but this segmentation is absurd: in Arabic calligraphy letters are naturally connected, when drawing one letter you already think and prepare yourself for the next one. Arabic produced by the method invented by Gutenberg (segmenting and re-assembling) is bound to be artificial, unnatural, degenerate. For these and many other reasons, it is quite difficult to adapt computers to oriental writing systems, unless of course one makes many compromises and gets doubtful results.

What I call “third world” writing systems are those which have suffered so much from technical compromises that it is too late (or very hard) to return to their authentic form, that is, the form they had before the computer. Unfortunately it happens quite often that the complexity of a script is inversely proportional to the size of its market; that is the case of Mongolian, Khmer, etc. Although both Microsoft and the Unicode consortium have spent a lot of time and effort on this kind of scripts, they will always remain bounded to technical insufficiencies.  $\TeX$  has always been an open alternative to legacy systems, and this openness is a source of hope for a better future for these scripts. But for this to happen,  $\TeX$ -based systems must remain independent of OS resources such as Uniscribe or Pango; this is the case for Lua $\TeX$  but not necessarily for X $\TeX$ .

As for your last question, knowledge of a language and of its script are orthogonal. To be competent in a script you need to know its history and evolution, the esthetics surrounding it, its relation with other scripts, the differences between handwritten, calligraphic, and printed, etc. None of these is trivial, and one must be constantly in contact with written material to get the “touch” of a script. A lot of harm has been done to many writing systems by incompetent font designers . . . .

DW: Thank you, Yannis, for taking the time to do this interesting and (for me) educational interview.