



## Award Ceremony of the Pius XI Medal

Mariano Sigman

Winner of the 2017 Pius XI Medal

Science is about making visible the invisible. The discovery of telescopes allowed scientists in the Renaissance to observe for the first time the most remote and distant corners of the solar system. And with this, solve fundamental questions which for centuries had only been a matter of speculation. In recent years, the instruments and vessels of science have also transformed our capacity to observe human brain activity. This, in turn, has solved controversies about the human mind, which had so far been exclusively a matter of philosophical conversation. For example, we can delve into the dreamer brain, decoding the content of dreams in real time, establishing concisely that dreams occur in real time and are not “just” figments of memory of the awakened mind. Or, from a more medical perspective, to decipher the thoughts of vegetative patients asking if they may have residues or preserved forms of consciousness which might not be expressed overtly. Human thought has never been so transparent.

But all this technology is, of course, useless for investigating one of the most mysterious aspects of human thought. How it came to be what it is. How were the thoughts of our ancestors? We know that their brains were almost identical to ours. But were they conscious in the same way we are today?

There are different intuitions about this longstanding philosophical debate. On one hand, it is natural to think that the deepest aspects of human thought – our ability to be conscious, to form memories, to imagine or to dream – have always been the same. Another possibility is that the social transformations that have so radically shaped our culture may even have forged these structural columns of our thoughts. Without the capacity to use all our technological paraphernalia to investigate brain function, is this question even amenable to science?

The solution to this conundrum is in the traces that our ancestors have left, not only of what they did, of how they lived or fought, but also of how they thought. In the same way that we can reconstruct how the ancient Greek cities looked like, just based on a few bricks, the writings of a culture are the archeological records, the fossils of human thought.

And in fact, doing some form of psychological analysis on some of the most ancient books of human culture, in the seventies Julian Jaynes came with a wild and radical hypothesis: that in about 3000 BC the world was a “garden of schizophrenics”.

Jaynes read exhaustively the human writings in the axial-age – a period approximately between 800 and 200 BC, which marked a radical transformation in Chinese, Indian and Western civilizations. It was during this period that the religions and philosophies that form many of the pillars of modern culture were produced. Studying these foundational texts, Julian Jaynes argued that during this period, human consciousness also went through a radical transformation.

He argued that the first humans described in these books behaved – in different traditions, in different places of the world – as if they were hearing and obeying voices that they perceived as coming from gods or muses, which today we would call hallucinations. And then, only as time went on, they progressively began to understand that they were the creators and owners of these inner voices. And with this, they acquired introspection: the ability to think about their own thoughts.

This may seem quite strange and paradoxical at first, but as with most human abilities that we take for granted, upon reflection it becomes clear that they ought to be forged in a learning process, either through the span of life, culture, or evolution. We all produce inner voices, inner thoughts. Most of us understand that we are the authors of these inner voices. But what if, as often happens during dreams, we do not? Chris Frith, a brilliant cognitive neuroscientist, has argued that this inability may be at the heart of schizophrenia. In this view, hallucinations do not result so much from a vast excess in fueling mental creations, but more in the incapacity to recognize the authorship of such creations.

According to Jaynes, consciousness, prior to Homer, lived in the present and lacked introspection, what now we call primary consciousness and is characteristic of schizophrenia or dreams (except for lucid ones). With the proliferation of texts, consciousness transformed into what we now recognize, a form of consciousness in

which we perceive we are the authors, protagonists and people responsible of our inner thoughts and actions, a consciousness which, in turn, has the richness to interweave with what we know of the past and what we predict or hope for the future.

Jaynes and others have argued that the appearance of written texts was at the heart of this psychological revolution, because it allowed thought to be consolidated on paper instead of being entrusted to the more volatile memory. We should remind those who now reflect so much on how the Internet, tablets, cell phones and the unceasing flow of information can change the way we think and feel, that the information age is not the first material revolution to radically change the way we express ourselves, communicate and, almost certainly, think.

*In summary, Jaynes' theory is that consciousness, at least in the way we perceive it today, when we feel we are the pilots of our own existence, is a quite recent cultural development. This conjecture remains as one of the most polemical, controversial and speculative in cognitive neuroscience, among other things because it relies on very specific examples and subjective interpretations. It was in fact difficult, if not impossible, for Jaynes (and others, at that time) to establish these claims in a quantitative and objective manner. The reason is quite simple: the word introspection (or self-consciousness) that Jaynes argued appeared throughout the axial-age, is not mentioned a single time in the books he studied.*

Things would have been much easier for Jaynes' theory if Plato woke up one day writing: *"Hello, I'm Plato, as of today I have a fully introspective consciousness"*. But this, of course, did not happen.

The development of introspection must be read in between the lines. And here is where psychology meets philology, computer science and linguistics. Doing so requires a more sophisticated mathematical entity than simply counting words: the construction of a "space of words". This is a huge, complex and high dimensional space that contains all words in such a way that the distance between any two of them is indicative of how closely related we feel they are. So one would want in this space the words 'dog' and 'cat' to be close together, but the words grapefruit and logarithm to be very far apart. *And this has to be true for any two words in this space.*

Computational linguistics has shown us that this space can be built quite easily using a simple but effective premise: when two words are related, they tend to appear in the same sentences, paragraph or page, when aggregated over a very large corpus of text. One can then formally define the proximity of any two given words as their likelihood of appearing together in a vast summary of all human linguistic expressions, compared to how often they would appear together simply by pure chance.

Once this space has been constructed, establishing which of two texts is closer to a given concept, is not just a qualitative argumentation but instead a concrete, objective and quantifiable argument. And with this one can inquire about the history of any concept, including introspection, asking how it grows, fades up or down, vanishes and changes in time. A text becomes a stream of words, a stream of words becomes a trajectory in the space of meanings. And then using simple geometry one can ask whether this trajectory approaches any given concept.

And when this is done after digitalizing the books of the Ancient Greek Tradition, ordering them by time, measuring the proximity of each word of this stack of books to introspection, one finds that, as Jaynes had conjectured, these texts become closer with the passage of time to the concept of introspection. There is a slow progression for the older Homeric texts: the Iliad and the Odyssey. And then, about 600 years before Christ, throughout the development of the ancient Greek culture, it begins to ramp up very rapidly to an almost five-fold increase with writings becoming closer and closer to introspection.

A fundamental virtue of an objective and quantitative analysis is that one can replicate the exact same results in a different case, asking whether these results are also true in a different and independent tradition. And, indeed, when this analysis is repeated for the Judeo-Christian books, or for a whole set of traditions throughout the axial-age, one obtains essentially the same pattern.

With this exercise in quantitative philology, using computational tools in humanity's historic archives, Carlos Diuk, Guillermo Cecchi, Diego Slezak and I were able to test Jaynes' "soft" hypothesis that there is a change in the narration of Homeric and Biblical texts that reflects an introspective discourse. In our view, however, it is not possible to examine the "hard hypothesis", settling whether this change reflects the filter of written language, censorship, narrative trends and styles, or whether, as Jaynes conjectures, it expresses the way our ancestors thought. Resolving this dilemma requires ideas and tools that we have yet to even imagine.