Biosemiotics manuscript No.
(will be inserted by the editor)

Qualitative and quantitative examples of natural and artificial phenomena

Antoni Hernández-Fernández [1,2,*]

Received: date / Accepted: date

Abstract The dichotomy between the qualitative and the quantitative has been a classic throughout the history of science. As will be seen, this dichotomy permeates all ontological levels of reality. In this work, phenomenological examples potentially related to semiosis are presented at the different levels established by Mario Bunge and Josep Ferrater Mora, contrasting the qualitative categorizations with the quantifiable physical reality. Likewise, the need to continue in the quantification of the biosemiotic and linguistic studies will be presented, while, in contrast, the need to establish a qualitative framework in the little-addressed study of technosemiotics will be raised, of potential interest given the notable advances that are expected in communication systems for inert artifacts in the next years. In short, in the thesis defended here the qualitative precedes the quantitative in the defining path of science.

Keywords Biosemiotics · Quantitative Linguistics · Technosemiotics · Bungean ontology · Categorization examples · Artificial systems

1 Introduction

In this Special Issue the dichotomy between qualitative and quantitative approaches has been raised in biosemiotics (Faltýnek and Lacková 2021) and, in addition, recently the *sign* character of life and the explanatory power of quantitative approaches have being debated (Faltýnek and Lacková 2020), which is not only an interesting and heated controversy in semiotics but also in linguistics and epistemology. As we will see, the starting question is not whether

Antoni Hernández-Fernández
[1] LARCA Research Group, Complexity and Quantitative Linguistics Lab., Institut de Ciències de l'Educació, Universitat Politècnica de Catalunya, Carrer Jordi Girona 1-3, 08034 Barcelona, Catalonia, Spain. Tel.: +34 649005153
E-mail: * antonio.hernandez@upc.edu
[2] Societat Catalana de Tecnologia, Secció de Ciències i Tecnologia, Institut d'Estudis Catalans, Carrer del Carme 47, 08001 Barcelona, Catalonia, Spain.

This is a post-peer-review, pre-copyedit version of an article published in *Biosemiotics*.
The final authenticated version is available online at: http://dx.doi.org/10.1007/s12304-021-09423-1

quantitative approaches can develop biosemiotics, but when they will. Obviously, in this sense it is assumed that biosemiotics, like linguistics, bets on its scientific character. The examples are many, but in general in the history of science the qualitative interpretation of reality has always preceded quantitative studies in all disciplines. Biosemiotics needs quantification in order to be formalized in a scientific way, like the rest of the sciences. In this sense, despite having followed parallel paths, both within the broad framework of general linguistics, biosemiotics and quantitative linguistics are disciplines doomed to meet.

Another premise is that the agents in biosemiotic processes do not necessarily have to be human, since they can be other living beings, or even artifacts in technosemiotic or cybernetic processes (Sharov 2010). Actually, the need for the distinction between life and everything artificial is at the very root of cybernetics (Wiener 1948) and it has generated a deeper definition of the concepts of agent and agency from a broader perspective (Sharov 2010; Sharov 2018). The need to separate communication processes from semiotic ones was qualitatively clarified by Eco (1979), although Eco insisted on the need for a human addressee to interpret the coded signal in order to be able to speak *properly* of "communication process"(Eco 1979). It is true that Eco extends semiotics to the study of everything that can be considered a "sign", so that all meaningful phenomena are signs (Chandler 2017), but implicitly there is always *a human* behind the semiotic phenomenon (Eco 1979).

To begin with, in a point of view opposite to this anthropocentric one assumed by Eco (1979) and other scholars, we could affirm that the prefix "bio-" is redundant both when talking about (bio)semiotics and when talking about (bio)linguistics, since both semiotic and linguistic systems are inherently linked to life. In fact, paraphrasing Von Uexküll (1942) all living beings are *carriers of meaning* (Von Uexküll 1942) and therefore are semiotic agents, potential interpretants (still a controversial concept, by the way, see (Colapietro 1988)) in the sense of Peirce (Peirce 1894).

Von Uexküll's approach thus eliminated the implicitly 'human' character of the semiotic agent extending it to the various manifestations of life. Each organism has its own meaningful perception of the world (*Umwelt*), far from human (Von Uexküll 2010; Machado and Romanini 2012). We coexist with a myriad of living beings that are subjects of their own perceptual worlds, with functional cycles adapted to their communicative contexts (Von Uexküll, 2010: 49), more or less complex(Von Uexküll 2010), that constitute authentic semiotic universes parallel to ours, as close as often unknown and of which we still know very little or only some of their plural codes (Emmeche 2004).

Biosemiotics studies the production and interpretation of signs and codes in living systems from an interdisciplinary perspective (Hoffmeyer 2008; Favareau 2010; Kull 2016), and as recently pointed out and synthesized by Schult, Preik and Kirschner (2020:2):

> "The basic idea is that biology is based on communication processes on all levels, from the molecular level to ecosystems (Gérard 2014), so

> that life processes can be viewed as sign-mediated interactions. Life phenomena are, in all aspects, natural semiotic systems. Permeating the entire biology, the sign is a useful key and a practically relevant access to the understanding of life. The connection between biology and semiotics is characterized by Pattee (1995) as follows: *Communication is the essence of life.*"

On the other hand, quantitative linguistics, from its origins (Zipf 1949; Grzybek 2012), shares the broad and formalizing view of biosemiotics in the study of communication systems (Köhler et al. 2008), but with the mathematical gaze of science, also converging with other open perspectives from fields such as information theory (Shannon 1948; Shannon and Weaver 1949) or genetics (Tsonis et al. 1997). Thus, there have been no few efforts to extend the statistical regularities found in human language, mathematically formalized and known as linguistic laws (Meyer 2002; Altmann and Gerlach 2016), to the study of human speech (Torre et al. 2019) and other communication systems which range from the vocalizations of other primates (Gustison et al. 2016; Watson et al. 2020) or their gestural communication (Heesen et al. 2019), to chemical communication (Hernández-Fernández and Ferrer-i-Cancho 2016), genomics (Ferrer-i-Cancho et al 2013; Li 2012; Hernández-Fernández et al. 2011) or proteomics (Eroglu 2014; Shahzad et al. 2015; Caetano-Anollés et al. 2017), to name some fields where studies abound in recent years. However, part of theoretical linguistics continue without considering the advances of quantitative linguistics or biosemiotics, without trying to explain them from their theoretical formalisms, while in contrast ethologists or scholars of animal communication or zoosemiotics do try because they have found in these approaches a fruitful way to carry out comparative studies (Sebeok 1965).

It is undoubtedly necessary to establish meeting points between both perspectives: qualitative (classical semiotics) and quantitative, essential in the development of science. Along these lines, the quantitative and qualitative approach to natural phenomena in the systems of the ontological levels of reality will be discussed first (following Bunge and Ferrater Mora proposals (Bunge 1977; Bunge 1979; Ferrater Mora 1979), see figure 1). Subsequently, quantitative linguistics and semiotics will be situated in this general scheme. Next, we will see that humans cannot escape neither from a quantitative physical reality that we perceive, nor from a brain that qualitatively categorizes and processes information. In semiotic systems qualitative categories with which to segment the *continuum* are established, assuming a knowable physical reality that is inescapably quantitative.

Finally, for the peace of mind of the editors of *Biosemiotics* (given the name of the journal) and contrary to what was stated at the beginning, it will be discussed the need to maintain the prefix "bio-" in biosemiotics given the emergence of technologies that, once implemented, could develop their own *technosemiotics*, when they become autonomous and complete agents (Sharov 2010; Sharov 2018). This conceptualization will allow to maintain a contrastive distinction between the semiotics of the natural (*bio*) and that of the artificial (*techno*), because although traditionally inert objects that actively participate in semiosis were unknown (without some living being behind them), the evolution of artificial intelligence systems and synthetic biology pose the future challenge of the emergence of technosemiotics (Fernández 2015; Fernández 2013), or even of the *synthetic* biosemiotics (Joesaar et al. 2019), fields as unknown as once were (and still are) the communication systems of a large number of the living beings on our planet.

2 From matter to mind: an ontological segmentation

The emergence of an apparent conflict between qualitative perception and quantitative scientific inquiry into reality is natural. As we will see below through some examples, our brain organizes and segments reality in a discrete and qualitative way, through concepts and words with which it segments and collapses reality. The use of categories, categorical perception and the subsequent cognitive processes of categorization are universal in human languages (Goldstone and Hendrickson 2010; Harnad 1987), and, as a hypothesis, it should also occur similarly in the *Umwelt* and perceptual mechanisms of other species as a fundamental step in biosemiotic processes (Machado and Romanini 2012). On the other hand, with the birth and development of science, the need to mathematize and quantify that same reality has been established, surely because the success of science derives from its predictive capacity and this, in turn, involves quantitative processes such as counting and calculation. However, mathematical and quantitative capacities are also present in our brain mechanisms and in those of other species, phylogenetically related or not with humans (Nieder 2019).

During the 20th century and so far into the 21st, science has made it possible to discern the material elements that make up different ontological levels with which, traditionally, the continuum of reality has been segmented. Thus, José Ferrater Mora (1979) distinguished four levels in reality *between matter and mind*: Physical, Biological, Social and Cultural (Ferrater Mora 1979). For his part, Mario Bunge extended these levels to five, highlighting their *systemic* character; these are systems to which Bunge gave crucial importance to the transitions between them, in such a way that it also distinguished some intermediate levels (in parentheses in this list), so following Bunge (1979:46): Physical systems, chemosystems (biochemosystems), biosystems (psychosystems), sociosystems and technical systems (Bunge 1979).

Bunge (1998, pp.339-340) later redefined the types of possible systems, establishing among others (Bunge 1998): Natural systems, Social systems, Technical systems, Conceptual systems, Semiotic systems and Artificial systems. Natural systems are those in which all its components, and the bonds among them, are not man-made, whereas technical systems are human creations. The distinction that Bunge (1998) makes between technical and artificial systems is interesting. For Bunge, artificial systems are the result of the union of technical, conceptual and semiotic systems, semiotic systems understood as

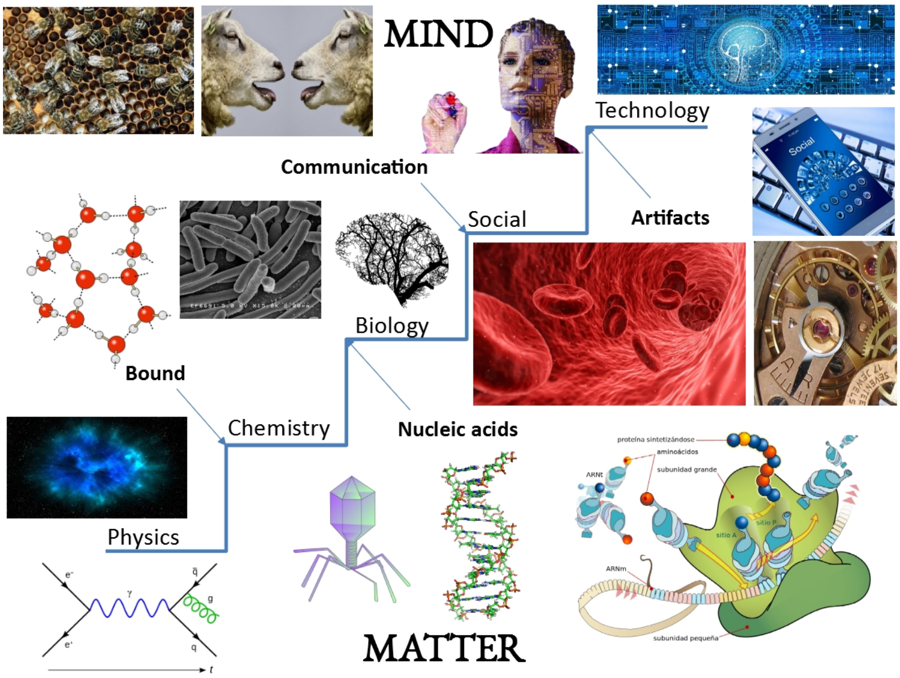

Fig. 1 Artistic interpretation of the scale of systemic levels of the continuum of reality, between the ontological poles of "Matter" to "Mind", posed by Bunge (1979) and Ferrater Mora (1979) (Bunge 1979; Ferrater Mora 1979). Adapted figure by Hernández-Fernández (2019), with permission (Hernández-Fernández 2019).

those composed of signs (Bunge 1998). Artificial languages, designed by human beings, would thus be an example of an artificial system. Bunge points out that languages are semiotic systems that can be natural, artificial or mixed (Bunge 1998). Therefore, the distinction between artificial and technical systems allows Bunge to locate programming languages as artificial systems with a semiotic component, and the same would be applicable to communication systems that are designed for machines.

Some Bunge scholars have already revealed that this asymmetry between the ontological levels and the types of systems is deliberate, so that there is no one-to-one correspondence between them. Thus, Barceló (2021) points out that the existing systems in a bungean sense are countless and are dynamic systems that respond to both external influences and internal changes. In this way, to understand how a semiotic system works, it is necessary to discover the mechanisms that modulate its evolution and its transformations (Barceló 2021). Certainly the ontological levels mentioned above could then be obvious and are nothing more than a qualitative segmentation, with fuzzy limits, of a more complex reality composed of systems and subsystems.

A free and expanded interpretation of these ontologies established between Bunge and Ferrater Mora has been proposed in figure 1 (Hernández-Fernández 2019).

In fact, Bunge and Ferrater Mora during their lives exchanged abundant correspondence in this regard (i.e. see the introductory study by Carla Velásquez in (Ferrater Mora 2018)). In this way, in these transitions (figure 1) we find the chemical bond between the physical and the chemical levels, which establishes unions between atoms and constitutes molecules and compounds; between the chemical and the biological we have the genetic information, encapsulated in DNA and RNA; between the biological and the social, communication systems and psychology; and between the social and cultural levels, artifacts and technologies. Bunge thus includes apart the technical systems that Ferrater Mora integrated at the cultural level (Bunge 1979; Ferrater Mora 1979).

Potentially, semiotic and communication processes can occur at all these levels and, interestingly, in the corresponding transitions. Thus, in life there are organisms that perceive, interpret and act according to the incidence of photons or radiation (Beloussov et al. 2000; Brosche and Strid 2003), and even it has been speculated with quantum approximations (Bischof and Del Giudice 2013); chemical communication is key in microscopic life (Waters and Bassler 2005), inside organisms, in taste and smell, and in ecosystems, such as pheromones and other volatile compounds (see many examples in (Wyatt 2003)); intraspecific and interspecific communication is obvious and well-known at a biological and social level (Leonhardt et al. 2016; Guerrero 2009); and, finally, the emergence and ubiquity of data science is leading scientists to propose to speak of semiotics at the (bio)technological level, using DNA based molecular communication (Sun et al. 2019) or proposing DNA as an information storage system (Church et al. 2012). Likewise, it has been discussed what type of relationship is established between the concepts of information and meaning (Shannon and Weaver 1949), within the framework of biosemiotics in its different manifestations in living beings (Queiroz et al. 2011). In short, epistemologically it is questioning whether there is a correspondence between "data" and "signs" (Compagno and Treleani 2019; Compagno 2018), between the quantitative and the qualitative.

It is this last debate within the development of new technologies that is stirring up the need to establish a dialogue between disciplines and quantitative and qualitative approaches. Whatever the chosen ontological segmentation, the semiotic study of all these transitions represents a challenge. Or many, since they can be given to living species and possible artificial life systems that will be created in the future (with their corresponding artificial languages), and that will arise from human creation as a result of the same knowledge of these transition elements (genetics, communication and technology). Technological artifacts then will permeate all levels and affect the transitions between them.

3 Qualitative and quantitative phenomena

Qualitative phenomena are fundamental in semiotics, but science cannot escape from physical quantities. Table 1 contains some examples of phenomena in which some quantitative correlate has been established in a more or less solid way. The strength of the quantitative approach lies in its predictive power and its tendency to objectively approach the object of study, so in my opinion it will progressively spread in the biosemiotics. The examples are intended to be suggestive for communication scholars to venture what could be happening, in a broad way, in any discipline that claims to aspire to have a scientific character, as is the case of linguistics (Bunge 1983; Hernández-Fernández and Ferrer-i-Cancho 2019). Quantification will then be inevitable.

A well-researched classic case is the perception of color: faced with a *continuum* of wavelengths in the visible, we categorize and name colors according to our language and cultural biases (Baronchelli et al. 2015; Hardin et al. 1997). In the color labels, as in the phonetic inventory or in the use of words, our brain uses mechanisms of categorical perception (Goldstone and Hendrickson 2010; Harnad 1987). Although it is an old problem of correlation between qualitative and quantitative, with clear semiotic and communicative implications in comparative studies (Sandell et al. 1979), there are still very interesting approaches from computational theories (Vaina and Passingham 2017).

Ontological level	**Phenomena**		**Formula/ Method**	**Refs**
	Qualitative	**Quantitative**		
Physical	Color	Wavelenght (λ)	$\lambda \cdot f = c$	(Baronchelli et al. 2015; Hardin et al. 1997)
Chemical	Acidity / Basicity	pH	$pH = -log_{10}[H_3O^+_{(aq)}]$	(Bates 1948; Lim 2006)
Biological	Classification of species (Linnean system)	Genome code assignement	DNA barcodes	(Marakeby et al. 2014; Blaxter 2016)
Linguistical	Speech sounds	Intensity, duration, frequency	$f = c/\lambda$; $\beta(dB) = 10\log_{10}(\frac{I}{I_0})$	(Goldstone and Hendrickson 2010; Schatz et al. 2021)
	Classification of populations (town, village, city, metropolis...)	Urban Quantitative Indicators	Power laws in urban indicators	(Bettencourt et al. 2007; Gabaix 1999)
Technological	Chip Scale Integration (SSI,MSI,LSI,VLSI,ULSI)	Transistor Count (Number/Area)	Moore's law	(Kaeslin 2008; Noyce 1977)
	Programming Languages	Quantitative Classification and Phylogeny	$f(r) = \frac{A}{r}(N+1-r)^b$	(Valverde 2016; Smith and Szathmary 1997)

Table 1 Some examples of phenomena at different bungean ontological levels and their associated physical quantities or quantitative approach. The chosen examples have implications in some communication systems of living beings or in technological communication. The last column gives some relevant references regarding the chosen example.

pH is a scale used to specify the acidity or basicity of a substance and is used as a measure of the hydronium ion concentration in chemistry (Bates 1948), with values that are often erroneously bounded between 0 and 14, but which can be negative (Lim 2006). In chemistry textbooks, qualitative interpretations used to be given to pH, establishing classifications for substances (i.e. very acidic, acidic, neutral, alkaline or very alkaline), classifications that some public organisms have standardized, as is the case of the classification of soils from the United States Department of Agriculture (Kellogg 1993). pH variations in ecosystems, often anthropogenic, can affect the communication of organisms (Heuschele and Candolin 2007) or their reproduction (Räty and Huhta 2003). The study of the influence of pH in the communication of plants or other living beings is undoubtedly still a challenge for science.

Both the wavelength of radiation and the pH are clear examples in which a *continuum* it is qualitatively segmented. However, can we do the same and quantify discrete elements of nature (beyond simply counting them)? In this sense, scientific advances have been allowing quantitative proposals, unsuspected years ago. The discovery of DNA and the vastness of microscopic species that overspread our planet have shown that the current Linnaean classification system of organisms is inadequate to name and order all that immense genetic diversity, and for this reason new quantitative proposals have been made that go beyond the Linnaean qualitative system (Marakeby et al. 2014; Blaxter 2016): before studying the biosemiotics of an organism, we should be able to name it correctly.

Of course Genome Code Assignement, or DNA barcodes, could *complement* current biological classification of organisms (Marakeby et al. 2014; Blaxter 2016), in the same way that quantum mechanics or relativity did not replace Newtonian mechanics, very useful in most studies of everyday physics. There is, however, an important nuance in this example: the assignment of genetic codes to organisms implies an important quantitative leap in the naming of species, to date based on fundamentally qualitative observations and in the concept of *species*. Meanwhile, the physical formulations of quantum mechanics or relativity are also quantitative in its majority of formulations, like Newtonian mechanics.

The acoustic identification of speech sounds is another type of categorical perception in which we organize the phonological inventory of the languages we know, after a learning period. We do it automatically in our mother tongue(s), to which we are massively early exposed (Roy et al. 2015) and later we can do it in other languages that we learn. Statistical learning mechanisms (Saffran et al. 1996) could explain the phenomena of phonological space segmentation, in addition to another wide range of basic and higherorder cognitive functions (Bogaerts et al. 2020). However, as recent research shows, the problem of learning and segmentation of phonetic categories is not yet fully resolved (Schatz et al. 2021).

As an example of the verbal segmentation of social reality, the denomination of human populations has been chosen in Table 1. The development of civilizations will require having clear quantitative parameters to define populations, something that has been speculated on since the pioneering work of Malthus (Malthus 1798; Brown et al. 2014). Thus, we usually classify urban population nuclei according to their size (hamlet, town, village, city, metropolis, megalopolis...) but, as with all non-scientific classification and segmentation of a perceptual continuum, it is problematic, particularly at the boundaries between segments and at their extremes. In this sense, starting from the quantitative study of the main urban indicators (population, extension, energy consumption,...) a more objective classification can be made, relevant in a global context in which the environmental impact of cities begins to be worrying (Brown et al. 2014). There have already been quantitative attempts in this line (Bettencourt et al. 2007), such as the suggestive study of Zipf's law in cities by Gabaix (1999) (Gabaix 1999).

Finally, two examples of proposals for quantitative modeling of technological phenomena that we usually discern qualitatively have been chosen. The first is the study of the evolution of the miniaturization scale of transistors, related to computational power. Regarding its evolution, the famous Moore's law in 1965 established that the number of transistors on a chip should double in each technology generation (Moore 1965). We will not enter here into the problems of this quantitative law of technological development (see (Lundstrom 2003) in this regard), but in this case the names of the integrated circuits (SSI, MSI,...) have had to be industrially standardized quantitatively, simply by counting the number of transistors per unit area (Kaeslin 2008; Noyce 1977). However, this count per unit area should consider the type of technology used if more accurate predictions of technological evolution are to be made (Lundstrom 2003), showing us the difficulties and limitations of modeling in the technological field despite this *artificial* quantitative approach.

A history of technological progress can be described as a list of significant major transitions of evolution (Smith and Szathmary 1997) that require the emergence of elements or agents resulting in a novel property in the former that was not present in the latter (Valverde 2016). They are often qualitative transitions but can be quantitatively modeled from the perspective of complex systems, as Valverde (2016) showed for the case of the evolution of programming languages (Valverde 2016). His interesting proposal shows us the way forward for the modeling of semiotic phenomena in cultural transmission using complex networks. Let us recall here Bunge's definition of an artificial system, as a confluence of technical, conceptual and semiotic systems, applicable to programming languages (Bunge 1998). In Table 1 the frequency-rank $f(r)$ beta distribution of programming languages popularity is showed as an example (Valverde 2016), where r is the programming language rank, N is the maximum rank, A is a normalization constant and a and b the empirically fitted parameters.

4 Conclusions

As has been seen, the continuum of ontological reality – classified here following the structural levels proposed by Ferrater Mora and Bunge (Bunge 1979; Ferrater Mora 1979; Ferrater Mora 2018) – it is usually categorized and conceptualized through language in a discrete, qualitative way, at the cost of segmenting a continuous perceptual reality. The fact that our systemic interpretations and discourses are based on these qualitative linguistic labels (mainly *words*) should not make us renounce the quantitative approach, necessary to scientifically study any natural phenomenon. However, for a quantitative approach to be given, it needs to be based on clear qualitative conceptions. In the history of science it has always been like this: the qualitative has preceded the quantitative.

Otherwise, going back to the examples in Table 1, in the case of light, it would be impossible for us to understand the communicative systems of many organisms, without exactly measuring the wavelengths of the radiation involved; nor would we ever know how many plants react to pH variations, if we stick only to qualitative considerations on the acidity or basicity of the soils;

nor will we be able to rigorously study the enormity of living beings whose genetic variations escape the classical Linnaean denominations of species nor, finally, will we be able to scrutinize the linguistic laws and the universals of language, if we impose qualitative limitations on ourselves.

In this sense, in some previous works on speech it has been seen that the statistical laws of language are found with greater precision if their study is attended based on quantitative physical magnitudes of voice (frequency, duration and amplitude), than merely attending to symbolic simplifications (like texts) that, in many cases, only reproduce the linguistic laws to the extent that they are more or less faithful to the physical reality of which they pretend to be a mirror (Torre et al. 2019; Hernández-Fernández and Ferrer-i-Cancho 2019). Another controversial example has been the categorization of emotions in humans (Ekman 1992), which can be seen as the establishment of labels regarding the perception of the continuum of facial features, something fundamental in facial recognition systems (Mellouk and Handouzi 2020).

We are perhaps in a transition in linguistics and semiotics towards the quantitative (Köhler et al. 2008; Barbieri 2007), which is the result of advances in computing that allow us to easily analyze data that was unapproachable a century ago. We could say that the physical, chemical and biological levels have been successfully quantified, the linguistic level is on the way, but what about the cultural and technological one?

To date, research in *technosemiotics* has been posed as an emerging challenge or as a bridge between 'natural', or biological, semiotics (*biosemiotics*) and artifacts of increasing communicative complexity (Fernández 2015; Fernández 2013). However, while in biology centuries have passed since the Linnaean tradition of systematization in the study and qualitative classification of living beings, which now allows novel and solid quantitative approaches as has been seen (Marakeby et al. 2014; Blaxter 2016), there is a lack of qualitative tradition in the evolutionary study of technology and artifacts, as the limitations of quantitative approaches at the technological level reflect (Moore 1965; Valverde 2016): Technology lacks a clear definition of qualitative traits on which, *a posteriori*, quantitative approaches can be applied.

If reviewed in detail, most quantitative studies in technology are reduced to analyzing the *evolutions in the time scale* of the different technologies, with an underlying concept of 'progress' or 'growth' typical of the 19th century (Valverde 2016; Smil 2019), while the most successful efforts of evolutionary studies are related to technology effects in the so-called *anthropocene*, and have come from ecology approaches where qualitative and quantitative efforts converge (Smith and Szathmary 1997; Ellis 2019; Ellis 2015). An example worthy of review is the defining effort of Ellis (2015) and his clear distinction between the qualitative and the quantitative (Ellis 2015). In other words, while biosemiotics has arrived at a time when biology is mature enough to consider quantitative studies in its midst, technosemiotics still needs for qualitative prior work.

Technosemiotics extends Wiener's classic cybernetics (Wiener 1948) and therefore the following work is pending to support this subdiscipline:

i) the establishment of the traits that define artifacts, analogously to how living beings are defined, considering the existence of hybrid entities between the biological and the technological;
ii) the extension to the technological elements of all those parameters of biological communication systems that are applicable to technology and technological agents (Sharov 2018).

Regarding this second point, it would be good not to grant limitations to the technology that living beings possess (*biocentrism*): thus for example an artifact could potentially produce and *understand* (process) several acoustic signals simultaneously in parallel, something very difficult for a mammal. On the other hand, as a hypothesis, it seems a priori easier to know the *new technological umwelt* of the complex artifacts that we will design and create, than those of living beings that are modified by biotechnologies, especially when we are completely unaware of the communicative mechanisms and *umwelt* of the starting organisms. However, we will once again be lost with respect to the new technological agents that we are capable of creating if we are left alone with the data and ignore the qualitative mechanisms underlying its operational functioning. Controlling future artificial intelligence systems implies a good understanding of their internal functioning and technosemiosis, as well as delving into the conception of a technosemiotic agent (Sharov 2010; Sharov 2018).

We must remember that the artificial systems we design have, in addition to a technological component, their corresponding conceptual and semiotic components (Bunge 1998). However, we could add a certain *social* component, understanding that machines and other artificial systems already communicate with each other as *relatively* autonomous agents, since they are still under human supervision. Technosemiotics should be especially concerned with the study of these artificial systems in case, in the future, the autonomy of the machines were total and real. Another aspect open to reflection is whether our current confusion about the approach we should give to the study of artificial communication systems is due to the fact that, for the first time, the quantitative has been put before the qualitative, so that we are flooded with data that we do not know how to categorize or we are unable to study, simply because we had not considered a systematic definition or categorization of technological systems and their parts.

In conclusion, both biosemiotics and technosemiotics still have a long pending roadmap. While biosemiotics will require greater quantitative approximations in the years to come, technosemiotics still needs a solid qualitative foundation.

Acknowledgements Thanks to Alfons Barceló and Ramon Ferrer-i-Cancho for the discussions that have largely led to the philosophical reflections contained here, and to Ľudmila Lacková for her valuable comments on the draft of the article.

Funding

This work has been funded by the project PRO2021-S03 HERNANDEZ of Institut d'Estudis Catalans. AHF is also funded by the grant TIN2017-89244-R from Ministerio de Economía, Industria y Competitividad (Gobierno de

España) and the recognition 2017 SGR-856 (MACDA) from AGAUR (Generalitat de Catalunya).

Conflicts of interest/Competing interests. None.

References

Altmann, E. G., & Gerlach, M. (2016). Statistical laws in linguistics. In *Creativity and universality in language* , 7-26. Cham: Springer.

Barbieri, M. (Ed.). (2007). *Introduction to biosemiotics: The new biological synthesis.* Dordrecht: Springer Science & Business Media.

Barceló, A. (2021). *Larga vida a la economía crítica: el enfoque de la reproducción.* Ediciones Asociación de Economía Crítica. Available in: http://www.asociacioneconomiacritica.org/larga-vida-a-la-economia-critica-el-enfoque-de-la-reproduccion/

Baronchelli, A., Loreto, V., & Puglisi, A. (2015). Individual biases, cultural evolution, and the statistical nature of language universals: The case of colour naming systems. *PloS one*, *10*(5), e0125019.

Bates, R. G. (1948). Definitions of pH Scales. *Chemical reviews*, *42*(1), 1-61.

Beloussov, L., Popp, F. A., Voeikov, V., & Van Wijk, R. (Eds.). (2000). *Biophotonics and coherent systems.* Moscow: Faculty of Biology, Lomonosov State University.

Bettencourt, L. M., Lobo, J., Helbing, D., Kühnert, C., & West, G. B. (2007). Growth, innovation, scaling, and the pace of life in cities. *Proceedings of the National Academy of Sciences*, *104*(17), 7301-7306.

Bischof, M.,& Del Giudice, E. (2013). Communication and the emergence of collective behavior in living organisms: a quantum approach. *Molecular biology international*, *2013*, ID:987549.

Blaxter, M. (2016). Imagining Sisyphus happy: DNA barcoding and the unnamed majority. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1702), 20150329.

Bogaerts, L., Frost, R., & Christiansen, M. H. (2020). Integrating statistical learning into cognitive science. *Journal of Memory and Language*, *115*, 104167.

Brosché, M., & Strid, Å. (2003). Molecular events following perception of ultraviolet-B radiation by plants. *Physiologia Plantarum*, *117*(1), 1-10.

Brown, L. R., Gardner, G., & Halweil, B. (2014). *Beyond Malthus: The Nineteen Dimensions of the Population Challenge.*. New York: Routledge.

Bunge, M. (1977). *Treatise en Basic Philosophy III. Ontology I: the Furniture of the World.* Dordrecht-Boston: Reidel.

Bunge, M. (1979). *Treatise en Basic Philosophy IV. Ontology II: a World of Systems.* Dordrecht-Boston: Reidel.

Bunge, M. (1983). Lingüística y filosofía. Barcelona: Ariel.

Bunge, M. (1998). Semiotic systems. In: Altmann & Koch (Eds.) *Systems: new paradigms for the human sciences*, 337-356. Berlin: De Gruyter.

Caetano-Anollés, G., Minhas, B. F., Aziz, F., Mughal, F., Shahzad, K., Tal, G., ... & Kim, K. M. (2017). The compressed vocabulary of the proteins of archaea. In *Biocommunication of Archaea*, 147-174. Cham: Springer.

Chandler, D. (2017). *Semiotics: the basics.* New York: Taylor & Francis.

Church, G. M., Gao, Y., & Kosuri, S. (2012). Next-generation digital information storage in DNA. *Science*, *337*(6102), 1628. https://doi.org/10.1126/science.1226355

Colapietro, V. M. (1988). *Peirce's approach to the self: A semiotic perspective on human subjectivity.* Suny Press.

Compagno, D. (Ed.). (2018). *Quantitative semiotic analysis.* Berlin: Springer.

Compagno, D., & Treleani, M. (2019). Introduction to Meaningful data/Données signifiantes. *Semiotica*, *2019*(230), 1-17.

Eco, U. (1979). *A theory of semiotics.* Bloomington Indiana: Indiana University Press.

Ekman, P. (1992). Are there basic emotions? *Psychological Review*, *99*(3), 550–553. https://doi.org/10.1037/0033-295X.99.3.550

Ellis, E. C. (2019). Evolution: biodiversity in the Anthropocene. *Current Biology*, *29*(17), R831-R833.

Ellis, E. C. (2015). Ecology in an anthropogenic biosphere. *Ecological Monographs*, *85*(3), 287-331.

Emmeche, C. (2004). A-life, organism, and body: The semiotics of emergent levels. In: Bedeau, M., Husbands, P.; Hutton, T.; Kumar, S. and Hideaki, S. (eds.), *Workshop and Tutorial Proceedings: Ninth International Conference on the Simulation and Synthesis of Living Systems (Alife IX).* Boston, 117–124.

Eroglu, S. (2014). Language-like behavior of protein length distribution in proteomes. *Complexity*, *20*(2), 12-21.

Faltýnek, D., & Lacková, Ľ. (2020). In the case of protosemiosis: Indexicality vs. iconicity of proteins. *Biosemiotics*, 1-18.

Faltýnek, D., & Lacková, Ľ. (2021). Quantity versus Quality. *Biosemiotics*, ***pending, in this special issue***.

Favareau, D. (2010). *Essential readings in biosemiotics. Anthology and commentary.* Dordrecht: Springer.

Fernández, E. (2015). Evolution of signs, organisms and artifacts as phases of concrete generalization. *Biosemiotics*, *8*(1), 91-102.

Fernández, E. (2013). Semiosis and Control—from biosemiotics to technosemiotics and back. *Thirteenth annual gathering in biosemiotics.* Castiglioncello. Available at $https : //www.lindahall.org/services/reference/papers/fernandez/semiosis_{a}nd_{c}ontrol.pdf$.

Ferrater Mora, J. (1979). *De la materia a la razón.* Madrid: Alianza.

Ferrater Mora, J. (2018). *Escritos sobre ciencia.* Pamplona: Editorial Laetoli.

Ferrer-I-Cancho, R., Forns, N., Hernández-Fernández, A., Bel-enguix, G., & Baixeries, J. (2013). The challenges of statistical patterns of language: The case of Menzerath's law in genomes. Complexity, 18(3), 11-17.

Gabaix, X. (1999). Zipf's law for cities: an explanation. *The Quarterly Journal of Economics*, *114*(3), 739-767.

Goldstone, R. L., & Hendrickson, A. T. (2010). Categorical perception. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(1), 69-78.

Grzybek, P. (2012). History of quantitative linguistics. *Glottometrics*,*23*, 70–80.

Guerrero, A. (2009). Inter and intraspecificity of chemical communication. *Chemical Ecology*, *1*, 415-437.

Gustison, M. L., Semple, S., Ferrer-i-Cancho, R., & Bergman, T. J. (2016). Gelada vocal sequences follow Menzerath's linguistic law. *Proceedings of the National Academy of Sciences*, *113*(19), E2750-E2758.

Hardin, C. L., Hardin, C. L., & Maffi, L. (Eds.). (1997). *Color categories in thought and language.* Cambridge: Cambridge University Press.

Harnad, S.A. (ed.) (1987). *Categorical perception: The Groundwork of Cognition.* Cambridge: Cambridge University Press.

Heesen, R., Hobaiter, C., Ferrer-i-Cancho, R., & Semple, S. (2019). Linguistic laws in chimpanzee gestural communication. *Proceedings of the Royal Society B*, *286*(1896), 20182900.

Hernández-Fernández, A., Baixeries, J., Forns, N., & Ferrer-i-Cancho, R. (2011). Size of the whole versus number of parts in genomes. *Entropy*, *13*(8), 1465-1480.

Hernández-Fernández, A. (2019). De Bunge a l'educació tecnològica al segle XXI, 145-177. In: Bunge, M. (2019). *Filosofia de la tecnologia.* Barcelona: Societat Catalana de Tecnologia, Institut d'Estudis Catalans.

Hernández-Fernández, A., & Ferrer-i-Cancho, R. (2016). The infochemical core. *Journal of Quantitative Linguistics*, *23*(2), 133-153.

Hernández-Fernández, A. & Ferrer-i-Cancho, R. (2019). *Lingüística cuantitativa. La estadística de las palabras.* Barcelona: EMSE EDAPP / Prisanoticias

Heuschele, J., & Candolin, U. (2007). An increase in pH boosts olfactory communication in sticklebacks. *Biology Letters*, *3*(4), 411-413.

Hoffmeyer, J. (2008). *Biosemiotics: An examination into the signs of life and the life of signs.* Scranton, PA: University of Scranton Press.

Joesaar, A., Yang, S., Bögels, B., van der Linden, A., Pieters, P., Kumar, B., Dalchau, N., Phillips, A., Mann, S., & de Greef, T. (2019). DNA-based communication in populations of synthetic protocells. *Nature nanotechnology*, *14*(4), 369–378. https://doi.org/10.1038/s41565-019-0399-9

Kaeslin, H. (2008). *Digital integrated circuit design: from VLSI architectures to CMOS fabrication.* Cambridge University Press.

Kellogg, C. E. (1993). *Soil survey division staff: soil survey manual.* Washington: United States Department of Agriculture.

Köhler, R., Altmann, G., & Piotrowski, R.G.(Eds.). (2008). *Quantitative Linguistik/Quantitative Linguistics: Ein internationales Handbuch/An In-*

ternational Handbook (Vol. 27). Berlin: Walter de Gruyter.

Kull, K. (2016). The biosemiotic concept of the species. *Biosemiotics*, *9*, 61–71.

Leonhardt, S. D., Menzel, F., Nehring, V., & Schmitt, T. (2016). Ecology and evolution of communication in social insects. *Cell*, *164*(6), 1277-1287.

Li, W. (2012). Menzerath's law at the gene-exon level in the human genome. *Complexity*, *17*(4), 49–53.

Lim, K. F. (2006). Negative pH does exist. *Journal of chemical education*, *83*(10), 1465.

Lundstrom, M. (2003). Moore's law forever?. *Science*, *299*(5604), 210-211.

Machado, I., & Romanini, V. (2012). Semiotics of Communication: From Semiosis of Nature to Culture. *Biosemiotics*, *5*(1), 47-60.

Malthus, T. R. (1798). *An essay on the principle of population, as it affects the future improvement of society, with remarks on the speculations of Mr. Godwin, M. Condorcet, and other writers.* London: J.Johnson's in St. Paul's Church-Yard.

Marakeby, H., Badr, E., Torkey, H., Song, Y., Monteil, C. L., Heath, L. S., & Vinatzer, B. A. (2014). A system to automatically classify and name any individual genome-sequenced organism independently of current biological classification and nomenclature. *PLoS One*, *9*(2), e89142.

Mellouk, W., & Handouzi, W. (2020). Facial emotion recognition using deep learning: review and insights. *Procedia Computer Science*, *175*, 689-694.

Meyer, P. (2002). Laws and theories in quantitative linguistics. *Glottometrics*, *5*, 62–80.

Moore, G. (1965). Moore's law. *Electronics Magazine*, *38*(8), 114.

Nieder, A. (2019). *A brain for numbers: the biology of the number instinct.* Cambridge: MIT press.

Noyce, R. N. (1977). Large-scale integration: What is yet to come?. *Science*, *195*(4283), 1102-1106.

Peirce, C.S. (1894). What is a sign, in: Peirce, C. S. 1998. *Charles S. Peirce: The Essential Writings.* (Moore, Edward C., ed.) Amherst: Prometheus Books.

Queiroz, J., Emmeche, C., Kull, K., & El-Hani, C. (2011). The biosemiotic approach in biology: Theoretical bases and applied models. In: Terzis, G. & Arp, R. (Eds.) *Information and living systems: Philosophical and scientific perspectives*, 91-130.

Räty, M., & Huhta, V. (2003). Earthworms and pH affect communities of nematodes and enchytraeids in forest soil. *Biology and Fertility of Soils*, *38*(1), 52-58.

Roy, B. C., Frank, M. C., DeCamp, P., Miller, M., & Roy, D. (2015). Predicting the birth of a spoken word. *Proceedings of the National Academy of Sciences*, *112*(41), 12663-12668.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926-1928.

Sandell, J. H., Gross, C. G., & Bornstein, M. H. (1979). Color categories in macaques. *Journal of comparative and physiological psychology*, *93*(4), 626-635.

Schatz, T., Feldman, N. H., Goldwater, S., Cao, X. N., & Dupoux, E. (2021). Early phonetic learning without phonetic categories: Insights from large-scale simulations on realistic input. *Proceedings of the National Academy of Sciences*, *118*(7).

Sebeok, T. A. (1965). Animal Communication: A communication network model for languages is applied to signaling behavior in animals. *Science*, *147*(3661), 1006-1014.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, *27*(3), 379-423.

Shannon, C. E. & Weaver, W. (1949). *The Mathematical Theory of Communication.* Urbana: University of Illinois Press.

Sharov, A. A. (2010). Functional information: Towards synthesis of biosemiotics and cybernetics. *Entropy*, *12*(5), 1050-1070.

Sharov, A. A. (2018). Mind, agency, and biosemiotics. *Journal of Cognitive science*, *19*(2), 195-228.

Shahzad, K., Mittenthal, J. E., & Caetano-Anollés, G. (2015). The organization of domains in proteins obeys Menzerath-Altmann's law of language. *BMC systems biology*, *9*(1), 1-13.

Schult, J., Preik, O., & Kirschner, S. (2020). The Importance of Biosemiotics for Morphology. *Biosemiotics*, 1-13.

Smil, V. (2019). *Growth: from microorganisms to megacities.* Cambridge, MA: Mit Press.

Smith, J.M.& Szathmary, E. (1997). *The major transitions in evolution.* Oxford: Oxford University Press.

Sun, Y., Ito, M., & Sezaki, K. (2019). Channel Capacity Analysis of Diffusive DNA based Molecular Communication. (2019). *IEEE Wireless Communications and Networking Conference*, doi:10.1109/wcnc.2019.8885693

Torre, I. G., Luque, B., Lacasa, L., Kello, C. T., & Hernández-Fernández, A. (2019). On the physical origin of linguistic laws and lognormality in speech. *Royal Society open science*, *6*(8), 191023.

Tsonis, A. A., Elsner, J. B., & Tsonis, P. A. (1997). Is DNA a language? *Journal of theoretical Biology*, *184*(1), 25-29.

Vaina, L. M., & Passingham, R. E. (Eds.). (2017). *Computational Theories and Their Implementation in the Brain: The Legacy of David Marr.* Oxford: Oxford University Press.

Valverde, S. (2016). Major transitions in information technology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1701), 20150450.

Von Uexküll, J. (1942). *Meditaciones biológicas. La teoría de la significación.* Madrid: Revista de Occidente.

Von Uexküll, J. (2010). *A foray into the worlds of animals and humans: With a theory of meaning (Vol. 12).* Minneapolis: Univ. of Minnesota Press. Translation of: *Streifzüge durch die Umwelten von Tieren und Men-*

schen, 1934 Verlag von Julius Springer; and *Bedeutungslehre*, 1940 Verlag von J. A. Barth.

Waters, C.M. & Bassler, B.L. (2005). Quorum sensing: cell-to-cell communication in bacteria. *Annual Review of Cell and Developmental Biology*, *21*, 319-46. doi: 10.1146/annurev.cellbio.21.012704.131001. PMID: 16212498.

Watson, S. K., Heesen, R., Hedwig, D., Robbins, M. M., & Townsend, S. W. (2020). An exploration of Menzerath's law in wild mountain gorilla vocal sequences. *Biology Letters*, *16*(10), 20200380.

Wiener, N. (1948). *Cybernetics or Control and Communication in the Animal and the Machine.* Cambridge, MA: MIT press.

Wyatt, T.D. (2003). *Pheromones and animal behaviour: communication by smell and taste.* Cambridge: Cambridge University Press.

Zipf, G.K. (1949). *Human behavior and the principle of least effort.* Cambridge, MA: Addison-Wesley.