Natural Language as a Technological Tool

Simone Aurora (✉)

Università degli studi di Padova, dipartimento FISPPA, Piazza Capitaniato 3, Padova, 35139, Italy

simone.aurora@unipd.it

Abstract

Language is a technological tool, since it is the outcome of a process of exteriorisation of a set of intentional practices. This process of exteriorisation is semiotic in nature. Language as technology and technology as language are thus ways of socialize consciousness. On the basis of some recent results in applied linguistics, the paper suggests that language and technology have to be considered as "functioning" when they enable social relations, by collectivising consciousness and producing a sort of social intelligence as well as an increase of complexity; on the other hand, language and technology have to be considered as "non-functioning" when they hinder socialization, privatize consciousness and reduce complexity, as is the case in automatized and algorithmic treatment of languages. This concept of language requires a reconsideration of the ways in which linguistics and philosophy of language understand semiotic practices and demands a shift from an "autonomist view" to a "political view" of language.

Keywords: Natural language; Technological tool; Social intelligence; Algorithm

Аннотация

Язык – это технологический инструмент, поскольку он является результатом процесса экстериоризации набора интенциональных практик. Этот процесс экстериоризации носит семиотический характер. Таким образом, язык как технология и технология как язык являются способами социализации сознания. На основе некоторых недавних результатов в прикладной лингвистике в статье предлагается рассматривать язык и технологию как "функционирующие", когда они обеспечивают возможность социальных отношений путем коллективизации сознания и создания своего рода социального интеллекта, а также увеличение сложности; с другой стороны, язык и технологии следует рассматривать как "нефункционирующие", когда они препятствуют социализации, приватизируют сознание и уменьшают сложность, как в случае с автоматизированной и алгоритмической обработкой языков. Эта концепция языка требует пересмотра способов понимания семиотических практик лингвистами и философами языка, и подразумевает переход от "автономистского" к "политическому" взгляду на язык.

Natural Language as a Technological Tool

1. COMPUTATIONAL CAPITALISM AND THE "DEATH" OF LANGUAGES

As Peter K. Austin and Julia Sallabank (2011) write in the introduction to a recent important book, "it is generally agreed by linguists that today there are about 7,000 languages spoken across the world; and that at least half of these may no longer continue to exist after a few more generations as they are not being learnt by children as first languages. Such languages are said to be endangered languages" (2011, p. 1). This circumstance does not represent a neutral linguistic fact but is rather the effect of a precise political-economic framework, namely what Bernard Stiegler (2016) has called "computational capitalism" – that is an era in which "*calculation prevails over every other criteria of decision-making*, and where algorithmic and mechanical becoming is concretised and materialised as logical automation and automatism, thereby constituting the advent of nihilism, as computational society becomes a society that is automated and remotely controlled" (pp. 8-9) – and Antoinette Rouvroy and Thomas Berns define as "algorithmic governmentality", that is "a certain type of (a)normative or (a)political rationality founded on the automated collection, aggregation and analysis of big data so as to model, anticipate and pre-emptively affect possible behaviours". "Algorithmic governmentality", they add further, "produces no subjectification, it circumvents and avoids reflexive human subjects, feeding on infra-individual data which are meaningless on their own, to build supra-individual models of behaviours or profiles without ever involving the individual, and without ever asking them to themselves describe what they are or what they could become" (Rouvroy & Berns, 2013, p. X). As this paper would like to show, then, the functioning of this new form of capitalism implies a radical reduction of the semantic, syntactic and morphological potentialities of natural languages. As a matter of fact, these "could [even] progressively evolve to seamlessly integrate the linguistic biases of algorithms and the economical constraints of the global linguistic economy" (Kaplan, 2014, p. 62). These claims seem to be coherent with what Austin and Sallabank write, as they observe that "economic, political, social and cultural power tends to be held by speakers of the majority languages, while the many thousands of minority languages are marginalized and their speakers are under pressure to shift to the dominant tongues. In the past sixty years, since around the end of World War II there have been radical reductions in speaker numbers of minority autochthonous languages, especially in Australia, Siberia, Asia and the Americas" (Austin and Sallabank, 2011, p. 1). "Reduction" is thus one of the main key-words of computational capitalism. These assertions remind what Deleuze and Guattari wrote in their brilliant analysis of linguistics in A Thousand Plateuaus (see Aurora, 2017). "The scientific enterprise of extracting constants and constant relations is always coupled", they affirm, "with the political enterprise of imposing them on speakers and transmitting order-words" (Deleuze and Guattari, 1987, p. 101). Indeed, from a purely linguistic point of view, there is no point in considering, for instance, British English as standard English and the so called Black-English merely as a deviation from standard English. This distinction is only a political distinction, by which, so Deleuze and Guattari (1987) write, "language is homogenized, centralized, standardized, becoming a language of power, a major or dominant language"

(p. 101). Such claims echo the famous 1945 assertion by the prominent linguist Max Weinreich (1945) – remarkably originally expressed in Yiddish, a minority language – according to which "[a] language is a dialect with an army and navy" (p. 13).
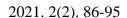
2. LANGUAGE AS A TECHNOLOGICAL TOOL: THE THEORETICAL FRAMEWORK

Within this very general scenario, the two strong claims that will be to made in this paper are 1) that language is one of the primal and most powerful technological tools; 2) that language is the "technological" condition of possibility, both at a transcendental and empirical level, of every process of socialization. Combining these two claims, it is thus possible to affirm that language is the basic technology for socialization of people and things and, accordingly, that the usage of language as a tool have a necessary influence on the models of socialisation that can be produced, created, designed or programmed. As has been shown, though, in the age of computational capitalism linguistic and semiotic technologies catalysing processes of collective and social subjectifications are inhibited, as they are substituted by algorithmic and automatic encoding programs.

Before shifting to the second section of the paper, a clarification is needed. Although in the following references will be mainly addressed to verbal language, this notion needs to be understood in a broad sense, according to which language must be seen as a semiotic system, namely as a system of signs.

Of course one should now provide a definition of sign, but this would require too much space and would lead the argument too far away from the topic. So, one could stay at the common, pre-scientific and intuitive understanding of the nature of sign, according to which, a sign is an object, quality, event, or entity whose presence or occurrence indicates the probable presence or occurrence of something else; however, it is important to keep in mind that signs become elements of a language, only insofar as they form a system, namely as they are structurally organized.

As André Leroi-Gourhan has brilliantly shown, technics and language have evolved in parallel, with each enabling and responding to the other. "As soon as there are prehistoric tools", Leroi-Gourhan (1993) writes, "there is a possibility of a prehistoric language, for tools and language are neurologically linked and cannot be dissociated within the social structure of humankind". "Techniques", Leroi-Gourhan (1993) writes, "involve both gestures and tools, sequentially organized by means of a 'syntax' that imparts both fixity and flexibility to the series of operations involved". "This operating syntax is suggested by the memory and comes into being as a product of the brain and the physical environment" (p. 114). The technical object is nothing but the organized series of the necessary gestures performed to produced it in view of an end; in other terms, it is nothing but the storage of all the neural connections that have been activated to shape matter in a certain way and in view of a specific end. Thus, technical objects can be described as the outcomes of a process of "exteriorization" of memory. This process represents the condition of possibility not only of the technical object, but also of every semiotic-linguistic practice, which accordingly can be described, in this sense, as a technical object. Indeed, the possibility of *selecting* and *combining* a set of hierarchically ordinated operations in view of an end implies the agency of an original discursivity, namely of what one could call a "primal syntax". This primal syntax is what allows to

relate and link together what is stored in memory (*ritention*, in phenomenological terms) and what is designed or projected through imagination (*protention*, in phenomenological terms). Both technical object (an arrow) and sign (for instance the word "arrow") are exteriorized memory, insofar as they are elements in which a specific set of practices, namely a *protocol*, is stored. Thus, there cannot be technics without language – otherwise there would only be objects with a certain shape, but no technical objects in proper sense – and, in turn, no language without technics – otherwise there would only be, for instance, sounds but not words.
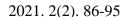
Technical objects and linguistic practices form then what Bernard Stiegler calls *tertiary retentions*, insofar as they allow for a cumulative spatialization and materialisation of neural-individual memory, which constitutes the very condition of possibility of all retentional and protentional capabilities and, accordingly, of consciousness itself.

Given this theoretical framework, it becomes possible to consider technological and linguistic evolution as two intertwined phenomena.

In this respect, a reference of paramount interest can be made to what French linguist Sylvain Auroux (1994) has named, with regard to the history of the evolution of language, the "automatisation", in which, according to Auroux, we find ourselves and which sees the rise of computer based computational treatments of language. In this phase, language seems to tend to become more and more *automatic*. More specifically, the automatisation of language takes place, as this paper suggests, at three different, although closely intertwined, levels, that one can define as *simplification*, *standardisation* and *anticipation.* This process of automatisation of language seems to play a pivotal role in the transformation, which is typical of the current political-economic scenario, of knowledge into mere information and of speakers into mere customers and consumers.

By *simplification* of language, it is meant the decrease of language complexity, that is the limitation of the range of possible syntactic, morphological and semantic forms; consider, for instance, the increasing impoverishment of vocabulary or the vanishing of abstract terms or complex syntactic constructs that many languages are being subjected to due the increasing and all the more pervasive presence of automatic encoding systems in everyday life. This is particularly apparent in the domain of *machine translation*, as various researches have recently demonstrated. Beata Beigman Klebanov and Michael Flor, for instance, have shown that "the lexical tightness of human-composed texts is tighter than that of the machine translated materials; human references are tighter than machine translations" (2013, p. 27), while another recent study allows to quantify "the loss of lexical richness in Machine Translation (MT) systems compared to Human Translation (HT)" (Vanmassenhove et al., 2019, p. 222).

By *standardisation*, it is meant a process which tends to remove linguistic differences, like for instance dialects or idiolects or minority languages, and to reduce the possibilities of linguistic variation. As already suggested, "language standardisation, central to language planning and policy, is inherently ideological […] [More specifically], standard language ideology encompasses assumptions about language correctness; belief in 'the one best variety'; and a demotion of all (non-standard) varieties" (McLelland, 2020, p. 1-2). Standardisation leads to a gradual *reduction* of linguistic variety, of what can be called *Semio-diversity*, which usually takes the form of the death of (non-standard) languages: "the most common process leading to language death" – writes Isaac Muhte

(2016) in a paper analysing the case of Ndau dialect in the linguistic context of Shona, a Bantu language, "is one in which a community of speakers of one language becomes bilingual in another language, and gradually shifts allegiance to the second language until they cease to use their original (or heritage) language. This is a process of assimilation which may be voluntary or may be forced upon a population. Speakers of some languages, particularly regional or minority languages may decide to abandon them based on economic or utilitarian grounds, in favour of languages regarded as having greater utility or prestige" (Mhute, 2016, p. 63).
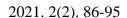
By *anticipation* it is finally meant the result of the application of technical devices which anticipate and, to some extent, prescribe the linguistic choices of the speakers. A very common example of these kind of technologies is represented by what are significantly called "predictive keyboards", that is keyboards that suggest upcoming words for fast typing. Indeed, as a recent study has clearly shown, "writers are sensitive to these differences: when presented with predictive text suggestions, people wrote shorter and more predictable language. In short, predictive text suggestions – even when presented as single words – are taken as suggestions of what to write" (Arnold et al., p. 136). This is due to the fact that "the suggestions are, by construction, the words that are the most predictable in their context. Thus, writers who follow these suggestions may create writing that is more predictable than they would create without such suggestions" (Arnold et al., 2020, p. 128).

These three levels, namely *simplification, standardisation* and *anticipation* are closely intertwined. Standardisation implies and, at the same time, requires simplification and both are improved by linguistic predictive technologies; *anticipation* demands, in turn, a simplified and standardised language in order to be effective. The result of this complex linguistic device could lead to what can be called the "entropic death" of language and languages. It is not by chance that the titles of two important books by prominent linguists such as Claude Hagège and Nicholas Evans bear the title *Halte à la mort des langues* (Hagége, 2002) and *Dying words* (Evans, 2009). Indeed, the automatisation of language seems to result in the tendency to create a unique, oversimplified, predictable and concrete language, a language that serves the purposes of global consumerist capitalism, in so far as it limits itself to the exchange of data and information and is made incapable of producing social collective meaning, namely those socialization processes which should represent the main goals of language as technology, and, accordingly, to serve as both the source and the repository of shared knowledge and common political agency.

3. LANGUAGE AND SOCIALIZATION: A RESEARCH PROGRAM

Understanding language as a technological tool for socialization clearly implies a general reconsideration of the very nature of linguistic agency and accordingly an epistemological shift both in the philosophy of language and in the scientific treatment of language, namely linguistics.

Indeed, as John E. Joseph (2006) observes in the preface of his book *Language and Politics*, "in the last two decades, applied linguistics has abandoned the structuralist view of language as a self-contained, neutral system, in favour of a conception of language as political from top to bottom, in its structure as well as its use (p. ix). The consequences of
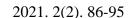
this conceptual shift can be seen in the growing attention paid by linguistic scholars to topics like "language choice, linguistic correctness, (self-) censorship and hate speech, the performance of ethnic and national identity in language, gender politics and 'powerful' language, rhetoric and propaganda, and changing conceptions of written language, driven in part by technological advances (Joseph, 2007, p. ix). If we look at the *canon* of 20th-century philosophy of language though, we discover that these interconnections between language and politics have been largely neglected and this conceptual shift has been generally overlooked. If we explore the philosophical reflection on language outside the disciplinary field of classic philosophy of language instead, it is possible to recover a conceptual shift analogous to that occurred within the language sciences, that Gerald Posselt has named the "ethico-political turn" in the philosophical study of language. This ethico-political turn results, to a large extent, from the criticisms directed towards structural linguistics and its *autonomist position*, as Ranko Bugarski (1999) named it, and is thus to be found especially within the so-called post-structuralist tradition. By autonomist position, Bugarski means the view according to which linguistics should not be subordinated or linked to philology, philosophy, sociology or some other discipline, but rather considered as a self-contained science. In fact, the object of linguistics, namely the system of language, can be studied, according to this view, solely by means of a description of the differential relations between the elements of the system, that is to say linguistic signs.

It is not by chance that Saussure's (1959) definition of language 'presupposes the exclusion of everything that is outside its organism or system' (p. 20). In order to clarify this point, Saussure uses a well-known analogy, comparing language to the game of chess: The fact that the game passed from Persia to Europe is external; against that, everything having to do with its system and rules is internal. If I use ivory chessmen instead of wooden ones, the change has no effect on the system, but if I decrease or increase the number of chessmen, this change has a profound effect on the 'grammar' of the game. (Saussure, 1959, pp. 22-23) Therefore 'one must always distinguish between what is internal and what is external. In each instance one can determine the nature of the phenomenon by applying this rule: everything that changes the system in any way is internal' (Saussure, 1959, p. 23). Accordingly, the science of language must necessarily exclude all the non-linguistic variables that can somehow be related to language, like ethnological, political, institutional or geographical variables, since they are 'external' variables and, as such, they are completely ineffective for a definition of the "forces that are permanently and universally at work in all languages" and for a deduction of the "general laws to which all specific historical phenomena can be reduced" (Saussure, 1959, p. 6).

However, although sharing a similar conceptual shift, according to which language is no more construed as autonomous but rather "political from top to bottom", philosophers studying language usually disregard linguistic analyses and professional linguists consider philosophical reflections just as unreliable and unscientific speculations. As a result, professional linguists often combine accurate applied studies with naive epistemological and metatheoretical assumptions, whereas philosophers often conjecture the nature of language without being acquainted with the technical tools required to describe it.
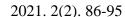
This paper, by completely assuming the ethico-political turn to language defined by Gerald Posselt, aims to provide a contribution to understand "language, along with its cognitive-communicative function, in its subjectivizing and community-grounding dimension, as well as to envisage the different modes and practices by which we are constituted as epistemic, ethical and political subjects" (from the description of the research project "Language and Violence. The ethico-political turn to language after the *linguistic turn*" http://language.univie.ac.at/about/).

With reference to linguistics – apart from a clear and general link to pragmatics, as the subfield of linguistics which studies the way in which context influences the meaning and the structure of statements and that "requires a consideration of how speakers organize what they want to say in accordance with who they're talking to, where, when, and under what circumstances" (Yule, 1996, p. 3) – various methodologies and perspectives need to be taken into account, which include, without being limited to: 1) *Systemic functional linguistics*, according to which "[t]he internal organization of natural language can best be explained in the light of the social functions which language has evolved to serve. Language is as it is because of what it has to do" (Halliday, 2004, p. 309), with a special focus on the notion of "regulatory function", using Halliday's wording, that is "the use of language to control the behaviour of others, to manipulate the persons in the environment – the 'do as I tell you' function" (Halliday, 2004, p. 306); 2) *Sociolinguistics* and, more specifically, the notion of "sociolinguistic variable", that is each and every linguistic variable "which is correlated with some non-linguistic variable of the social context: of the speaker, the addressee, the audience, the setting, etc." (Labov, 1972, p. 237); 3) *Language and gender*, with particular attention to the "dominance approach". As Penelope Eckert and Sally McConnell-Ginet summarize, "analysts associated with a dominance framework generally argue […] that differences between women's and men's speech arise because of male dominance over women and persist in order to keep women subordinated to men" (Eckert & McConnell-Ginet, 2003, p. 2); 4) *Critical Discourse Analysis*, in so far as it involves a "discourse analysis which aims to systematically explore often opaque relationships of causality and determination between (a) discursive practices, events and texts, and (b) wider social and cultural structures, relations and processes; to investigate how such practices, events and texts arise out of and are ideologically shaped by relations of power and struggles over power […] In referring to opacity, [we are] suggesting that such linkages between discourse, ideology and power may well be unclear to those involved, and more generally that our social practice is bound up with causes and effects which may not be at all apparent" (Fairclough, 1995, pp. 132-133); 5) *Dialectology*, especially with reference to the notion of "linguistic variety" which should basically substitute that of "language", since "[t]he term 'language' […] is from a linguistic point of view a relatively non-technical term" and should therefore be replaced by the notion of "variety" in order to be able to refer "to any particular kind of language", like for instance "Yorkshire English", "Leeds English" or even "middle-class Leeds English" (Chambers and Trudgill, 2004, p. 5).

For what concerns philosophy, I think that useful references should be, without being limited to: 1) Michel Foucault's archaeological approach and especially his notion of "discursive formation", defined as a set of scientific statements, among which subsists a system of regularities governed by what Foucault names *rules of formation* which are always context-dependent. As Foucault (2002) writes, "the rules of formation are

conditions of existence (but also of coexistence, maintenance, modification, and disappearance) in a given discursive division" (p. 42). More specifically, Foucault identifies four rules of formation: 1) *Rules of formation of objects*; 2) *Rules of formation of enunciative modalities*; 3) *Rules of formation of concepts*; 4) *Rules of formation of strategies* (see Foucault, 2002); 2) Deleuze and Guattari's (1987) philosophy of linguistics, as developed in the fourth chapter of *A Thousand Plateaus*, where they try to define a sort of prolegomena to a new linguistics, grounded on the refusal of what they call, with a certain degree of arbitrariness, the four "postulates of linguistics", and on the replacement of the notion of "structure" with that of "machine", which is clearly a technological notion, and on the primacy of Pragmatics and, moreover, on the replacement of the notions of "subject of the statement" and "subject of enunciation" with the concept of "collective assemblage of enunciation", which presupposes the fact that a fully individual – not social in nature – statement is in the end impossible; 3) Derrida's deconstructive approach, which "questions whether the reasons for imposing a theoretical division between the normal case and the deviation are theoretically justified". Indeed, as Johan Blomberg (2016) observes "a general semiotic theory cannot", according to Derrida, "merely dismiss different kinds of discourses as marginal cases and by extension demote them as less relevant. The deconstructive approach would prefer to understand why such distinctions are *imposed* and also to trace the consequences of avoiding them" (Blomberg, 2016, p. 51, italics mine).

4. CONCLUDING REMARKS

In order to conclude, some brief remarks are needed, that basically sum up the core ideas expressed in the paper. Language is a technological tool, since it is the outcome of a process of exteriorisation of a set of intentional practices; indeed, this process of exteriorisation of a protocollar syntax is semiotic in nature, insofar as it socialize intentionality; thus, language as technology and technology as language are ways of socialize consciousness, namely of inserting it in a network of relations with people and things which are external to it; language and technology are really "functioning", then, when they enable social relations, by collectivising consciousness and producing a sort of social intelligence as well as an increase of complexity; on the other hand, language and technology are not "functioning" when they hinder socialization, privatise consciousness and reduce complexity; this concept of language requires a reconsideration of the ways in which linguistics and philosophy of language understand semiotic practices and demands a shift from an "autonomist view" to a "political view" of language, a shift which should allow to interpret the contemporary political scenario and to react to the risk of an "entropic death" of human symbolic capabilities.

REFERENCES

Arnold, K. C., Chauncey, K., & Gajos, K. Z. (2020). Predictive Text Encourages Predictable Writing. In *25th International Conference on Intelligent User Interfaces* (pp. 128-138). ACM. https://doi.org/10.1145/3377325.3377523

Aurora, S. (2017). From Structure to Machine. Deleuze and Guattari's Philosophy of Linguistics. *Deleuze and Guattari Studies, 11*(3), 405-428.

Auroux, S. (1994). *La révolution technologique de la grammatisation*. Mardaga.

Austin, P. K., & Sallabank, J. (eds.), (2011). *The Cambridge Handbook of Endangered Languages*. Cambridge University Press.

Blomberg, J. (2016). How can Linguistic Meaning be Grounded – in a Deconstructionist Semiotics?. *Public Journal of Semiotics, 7*(1), 43-58. https://doi.org/10.37693/pjos.2016.7.15966

Bugarski, R. (1999). The Autonomy of Linguistics: Saussure to Chomsky and Beyond. In S. Embleton, J. E. Joseph and H.-J. Niederehe (Eds.), *The Emergence of the Modern Language Sciences. Volume 2: Methodological Perspectives and Applications* (pp. 29-41). John Benjamins.

Chambers, J. K., & Trudgill, P. (2004). *Dialectology*. Cambridge University Press.

Deleuze, G., & Guattari, F. (1987). *A Thousand Plateaus. Capitalism and Schizophrenia*. University of Minnesota Press.

Eckert, P., & McConnell-Ginet, S. (2003). *Language and Gender*, Cambridge University Press.

Evans, N. (2009). *Dying Words . Endangered Languages and What They Have to Tell Us.* Wiley-Blackwell.

Fairclough, N. (1995). *Critical Discourse Analysis: The Critical Study of Language*. Longman.

Foucault, M. (2002). *The Archaeology of Knowledge*. Routledge.

Hagège, C. (2002). *Halte à la mort des langues*. Odile Jacob.

Halliday, M. A. K. (2004). The Functional Basis of Language. In *On Language and Linguistics* (pp. 298-323). Continuum.

Kaplan, F. (2014). Linguistic Capitalism and Algorithmic Mediation. *Representations, 127*(1), 57-63. https://doi.org/10.1525/rep.2014.127.1.57

Klemanov, B. B., &. Flor, M. (2013). Associative Texture is Lost in Translation. In *Proceedings of the Workshop on Discourse in Machine Translation (DiscoMT)* (pp. 27-32). Association for Computational Linguistics. http://www.mt-archive.info/10/DiscoMT-2013-Klebanov.pdf

Joseph, J. E. (2006). *Language and Politics*. Edinburgh University Press.

Labov, W. (1972). *Sociolinguistic Patterns*. University of Pennsylvania Press.

Lerou-Ghouran, A. (1993). *Gesture and Speech*. The MIT Press.

McLelland, N. (2020). Language standards, standardisation and standard ideologies in multilingual contexts. *Journal of Multilingual and Multicultural Development*, 1-16. https://doi.org/10.1080/01434632.2019.1708918

Mhute, I. (2016). Standardisation a Considerable Force behind Language Death: A Case of Shona. *Journal of Education and Practice*, *7*(9), 62-65.

Rouvroy, A., & Berns, T. (2013). Algorithmic governmentality and prospects of emancipation. Disparateness as a precondition for individuation through relationships? *Reseaux, 177*(1), 163-196. (E. Libbrecht, Trans.).

Saussure, F. de (1959). *Course in General Linguistics*, The Philosophical Library.

Stiegler, B. (2016). *Automatic Society. Volume 1. The Future of Work*, Polity Press.

Vanmassenhove, E., Shterionov, D., & Way, A. (2019). Lost in Translation: Loss and Decay of Linguistic Richness in Machine Translation. In *Proceedings of Machine Translation Summit XVII, Volume 1: Research Track* (pp. 222-232). European Association for Machine Translation.

Weinreich, M. (1945). Der YIVO un di problemen fun undzer tsayt. *Yivo-bleter*, *25*(1), pp. 3-18.

Yule, G. (1996). *Pragmatics*. Oxford University Press.