



<https://doi.org/10.48417/technolang.2021.04.03>

Research article

Which Theory of Language for Deep Neural Networks? Speech and Cognition in Humans and Machines

Luca Capone (✉)

University Campus Bio-Medico of Rome, 21 Via Álvaro del Portillo, 00128 Rome, Italy

l.capone@unicampus.it

Abstract

The paper explores the relationship between technology and semiosis from the perspective of natural language processing, i.e. signs systems automated learning by deep neural networks. Two theoretical approaches to the artificial intelligence problem are compared: the internalist paradigm, which conceives the link between cognition and language as extrinsic, and the externalist paradigm, which understands cognitive human activity as constitutively linguistic. The basic assumptions of internalism are widely discussed. After witnessing its incompatibility with neural network implementations of verbal thinking, the paper goes on exploring the externalist paradigm and its consistency with neural network language modeling. After a thorough illustration of the Saussurian conception of the mechanism of language systems, and some insights into the functioning of verbal thinking according to Vygotsky, the externalist paradigm is established as the best verbal thinking representation to be implemented on deep neural networks. Afterwards, the functioning of deep neural networks for language modeling is illustrated. Firstly, a basic explanation of the multilayer perceptron is provided, then, the Word2Vec model is introduced, and finally the Transformer model, the current state-of-the-art architecture for natural language processing, is illustrated. The consistency between the externalist representation of language systems and the vector representation employed by the transformer model, prove that only the externalist approach can provide an answer to the problem of modeling and replicating human cognition.

Keywords: Natural Language Processing; Deep Neural Networks; Artificial Intelligence; Philosophy of Language; Philosophy of Science; Psycholinguistics; Linguistics, Philosophy of Technology

Citation: Capone, L. (2021). Which Theory of Language for Deep Neural Networks? Speech and Cognition in Humans and Machines. *Technology and Language*, 2(4), 29-60. <https://doi.org/10.48417/technolang.2021.04.03>



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)

Special Topic:

Technology as Language – Understanding Action in a Technical Condition

Спецвыпуск

“Техника как язык: понимание и действие в техническом мировоззрении”



УДК 004.032.26:81

<https://doi.org/10.48417/technolang.2021.04.03>

Научная статья

Теория языка для глубоких нейронных сетей: Речь и познание у людей и машин

Лука Капоне (✉)

Римский биомедицинский университет, 21 Альваро дель Портильо, 00128 Рим, Италия

l.capone@unicampus.it

Аннотация

В статье исследуется взаимосвязь между технологией и семиозисом с точки зрения обработки естественного языка, т. е. автоматизированного машинного обучения с помощью глубоких нейронных сетей. Сравниваются два теоретических подхода к проблеме искусственного интеллекта: интерналистская парадигма, которая рассматривает связь между познанием и языком как внешнюю, и экстерналистская парадигма, которая понимает когнитивную деятельность человека как конститутивно лингвистическую. Основные предположения интернализма широко обсуждаются. Убедившись в его несовместимости с нейросетевыми реализациями вербального мышления, в статье продолжается исследование экстерналистской парадигмы и ее согласованности с языковым моделированием нейронных сетей. После тщательной иллюстрации сосюрсовской концепции механизма языковых систем и некоторого понимания функционирования вербального мышления в соответствии с Л. С. Выготским, экстерналистская парадигма устанавливается как лучшая репрезентация, которая может быть реализована в глубоких нейронных сетях. Далее проиллюстрировано функционирование глубоких нейронных сетей для языкового моделирования. Сначала дается базовое объяснение многослойного перцептрона, затем вводится модель Word2Vec и, наконец, проиллюстрирована модель Transformer, современная архитектура для обработки естественного языка. Согласованность между экстерналистским представлением языковых систем и векторным представлением, используемым в модели преобразователя, доказывает, что только экстерналистский подход может дать ответ на проблему моделирования и воспроизведения человеческого познания.

Ключевые слова: Обработка естественного языка, Глубокие нейронные сети, Искусственный интеллект, Философия языка, Философия науки, Психолингвистика, Лингвистика, Философия технологий

Для цитирования: Capone L. Which Theory of Language for Deep Neural Networks? Speech and Cognition in Humans and Machines // *Technology and Language*. 2021. № 2(4) P. 29-60.
<https://doi.org/10.48417/technolang.2021.04.03>



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



LANGUAGE AND COGNITION

Since the beginning of artificial intelligence (AI) history, intelligent systems have been associated with the ability to reproduce human language (Turing, 1950; Russell and Norvig, 2021, chapter 1.3.5). Whether it's scientific research, the latest news or cultural and entertainment products, nothing catches the public's ear more than a machine that can talk "like us". The verbal activity has always inspired feelings of pride in human beings, but also curiosity, especially when compared to other forms of life, not endowed with this faculty.

Soon, not only philosophy, but other disciplines such as linguistics and psychology, realized the cognitive relevance of language, breaking the exclusive association between language and communication. The connection between thinking and speech turns out to be reinforced by these theories, and the idea that human cognition can only come to be through words has become commonplace for many thinkers. These theories are referred to as externalist, as for these authors, human cognition is developed through techniques and tools external to the organism. Orality, and in particular the functional use of signs (see Vygotsky and Luria 1993, pp. 118, 202), is one (and probably the most important) of these techniques.

Nevertheless, the first steps of AI did not take place within this theoretical framework; instead, they passed through information theory, neurosciences, and philosophy of mind (Dreyfus, 2007; see also Russell and Norvig, 2021, paragraph 1.2). It is possible to group these interpretations under the label of internalist theories. These theories consider thought as an internal process expressed by language, but independent from it.

In contact with engineering and computer sciences, the internalists ignored verbal activity and sought to model and reproduce human cognition starting from what they considered the logical structure of thought, only expressed in the grammar of languages (Chomsky, 1988, p. 134), developing rule-based models and ontologies. Most of these projects have had below-expectations outcomes (Ceccato, 1961; see also Dreyfus, 1992, pp. 130–152). Moreover, the recent advances in AI research confirms that for the replication of human intelligence the study of verbal activity cannot be disregarded (Brown et al., 2020; see also Capone and Bertolaso, 2020). The hypothesis of the paper is that an externalist paradigm may prove to be the most effective theoretical approach to reproduce verbal thought on a machine.

In the chapter, the internalist and externalist paradigms are compared by evaluating their respective solutions to the question of the relationship between thinking and speech. This step will be paramount to understand what kind of phenomenon the discipline of artificial intelligence should try to replicate.

Internalism

The internalist solution to the relationship between cognition and language has ancient roots; this position considers language essentially as a nomenclature, necessary for the communication of speech-independent contents (De Mauro 1967, p. 8).

"Now spoken sounds are symbols of affections in the soul, and written marks symbols of spoken sounds. And just as written marks are not the same for all men, neither



are spoken sounds. But what these are in the first place signs of – affections of the soul – are the same for all; and what these affections are likenesses of – actual things – are also the same” (Aristotle, 2014, p. 72 – *De Int.* 16a 3).

Aristotle’s position stems from two theoretical necessities. The first is to provide a better explanation of the linguistic phenomenon than those of his predecessors. Even in Heraclitus there are traces of a mentality that is defined as oral or pre-alphabetic (cf. Ong, 2004; Havelock, 2010). Following what must have been a widespread opinion, Heraclitus appears to believe that the name is an objectively inherent attribute of the object (De Mauro, 1999, p. 43). The second reason is to be found in the dispute with the skeptics. Aristotle needed a theoretical foundation for the principle of identity, which is itself indemonstrable. Tying the meaning of a word to internal concepts, “the same for all”, acts as a guarantee of this principle (De Mauro, 1967, p. 8) ensuring comprehension and communication between human beings.

The conception according to which words of different languages stand for concepts and referents that are the same for everyone, has long been the hegemonic theory in philosophy (fig. 1). The next section illustrates how this idea has survived until today and continues to live on in the theories of mind that came to the fore with the mentalist turn of the 1960s led by Chomsky (1959).

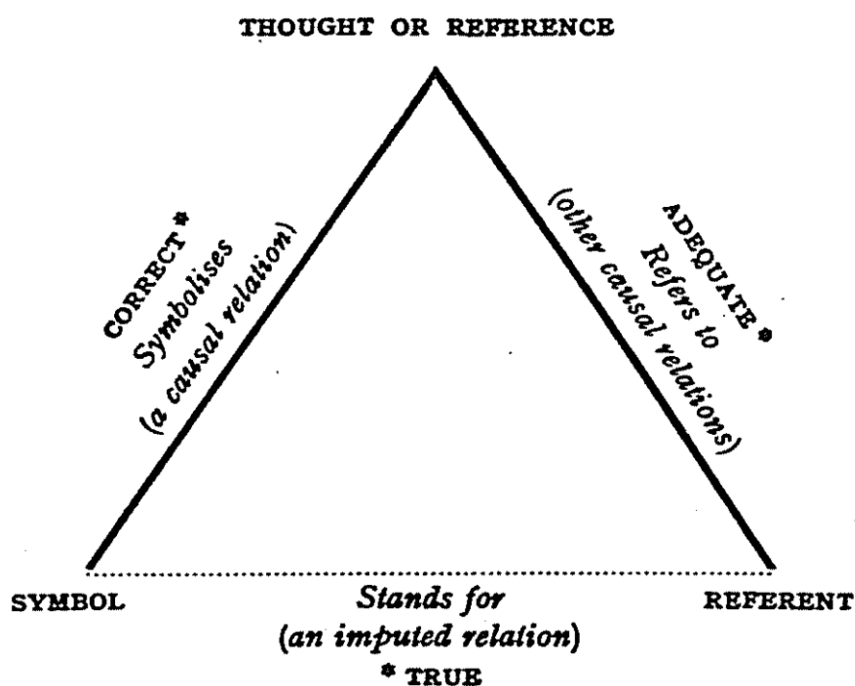


Figure 1. Ogden and Richard’s (1946) semiotic triangle (p. 11). A well-known graphical representation of the internalist conception of the relationship between language and thought.



The Mentalist Turn

Just like Aristotle, Chomsky tries to anchor language to universal structures, responsible for the functioning of every possible language. What he proposes as a theory of syntax is actually a theory of mind, with assumptions that even trespass into the physiological level. The brain is described in analogy to the hardware of a computer, on which various independent modules function as software. The ensemble of these modules constitutes the mind. The module in charge of language acquisition is the LAD (language acquisition device) (Chomsky, 2015, p. 58).

Language learning is not really something that the child does; it is something that happens to the child placed in an appropriate environment, much as the child's body grows and matures in a predetermined way when provided with appropriate nutrition and environmental stimulation. (Chomsky, 1988, p. 134)

The main argument in favour of this thesis is that of the *poverty of the stimulus*. The external stimulus is not sufficient for the acquisition of language. The subject requires a biological endowment that provides the data and structures for the development of a grammar (Chomsky, 1988, p. 153).

We may think of the language faculty as a complex and intricate network of some sort associated with a switch box consisting of an array of switches that can be in one of two positions. [...] When these switches are set, the child has command of a particular language and knows the facts of that language: that a particular expression has a particular meaning, and so on." (Chomsky, 1988, pp. 62–63)

In the Chomskyan proposal, syntax occupies a leading role, while semantics has practically no place, since concepts are innate (actualized according to the parameters of the LAD) and not the result of learning.

After Chomsky, the cognitive and mind studies have diversified a lot, but all of them followed the internalist paradigm, according to which concepts, representations and in general every meaning, is predetermined by internal structures, and language is nothing but the tool to manifest these inner states.

It is possible to summarize the general assumptions of the internalist approach as follows:

- Universalism of cognitive structures;
- Cognitive structures determine mental contents (concepts), and the rules for their relations (grammar);
- Language is an encoding/decoding tool for communicating internal states.

Internalism and Automation

Following the postulate that thought consists of concepts (internal states) plus grammar (structure of rules), the internalist project encouraged approaches to the modeling and reproduction of human cognition based on the definition of rule sets, instructions, and concept structures designed to imitate internal thought processes.



The earliest automatic systems appeared in the field of problem solving and logic. There were systems for proving theorems and chess-playing programs. These systems operated in controlled environments, far removed from the context in which human thought works.

The first attempts to approach reasoning and communication in natural language date back to the late 1960s. Marvin Minsky, head of the AI lab at MIT declared that within a generation there would have been intelligent computers like HAL, from Kubrick’s famous film (Dreyfus, 2007). A few years earlier, Alan Newell and Herbert Simon claimed to have solved the problem of the mind-body relationship. According to them, the mind could be conceived as a system of physical symbols. In their opinion, bits and symbols can be used to represent the human mental world of concept (Russell and Norvig, 2021, paragraph 1.3.1).

The internalist approach can be found in the work of these scholars. The elements of thought are imagined as copies of the objects of the world that inhabit the mind, designated by the words of the different languages. Concepts can be related according to various rules (generally imagined on the model of logical relations). It was just a matter of matching symbols to concepts and then calculate according to defined rules. Dreyfus describes the situation as follows:

far from replacing philosophy, the pioneers in CS [Cognitive Simulation] had learned a lot, directly and indirectly from the philosophers. They had taken over Hobbes’ claim that reasoning was calculating, Descartes’ mental representations, Leibniz’s idea of a “universal characteristic”—a set of primitives in which all knowledge could be expressed,—Kant’s claim that concepts were rules, Frege’s formalization of such rules, and Russell’s postulation of logical atoms as the building blocks of reality. In short, without realizing it, AI researchers were hard at work turning rationalist philosophy into a research program. (Dreyfus, 2007)

Following these assumptions, developers began to build “microworlds” (Russell and Norvig, 2021, paragraph 1.3.3), within which the programs they implemented could attempt to solve tasks. The use of controlled and ideal environments was necessary to avoid the *frame problem*. A controlled environment requires relatively little basic knowledge for an agent to be able to operate appropriately within it. An open system, on the other hand, requires a lot of information and prior knowledge and will present the agent with several choices that cannot be unambiguously decided, and which are unmanageable by such systems. A program, following specific instructions, can easily identify a cube in a world populated only by cubes, spheres and pyramids. The same rules can hardly be used unambiguously in an open environment. The frame problem is closely related to that of the meaning of concepts and their representation, and will be of fundamental importance later in the paper. Suffice it for now to say that this problem has never allowed these approaches to move from controlled to open (or natural) environments. Similarly, it has never allowed them to move from a simplified language to real natural language commands.

Attempts to represent language and make computers capable of understanding natural language have a long history within the discipline of AI. Scholars such as Ceccato



(2003) tried to represent language according to categories and semantic frames, attempting to create a sort of taxonomy of concepts, hoping to overcome the frame problem. This approach has long been used, as in the case of the OWL (Ontology Web Language) project¹, or in the creation of word graphs and networks, such as Wordnet (Fellbaum, 2006) and ConceptNet (Liu and Singh, 2004)². Although useful and interesting, these attempts continue to persist in an internalist and, in the words of Dreyfus, rationalist conception of thought³. In this view, concepts are represented according to features and categories, i.e. other concepts, entering an infinite regression from which there is no exit⁴.

Another eloquent example of the inadequacy of the internalist approach is provided by the way in which this paradigm has dealt with the machine translation task. If languages are nothing but by-products of the interaction between innate cognitive structures and the environment, it is to the universal grammar and therefore to the mind that one must look in order to understand the functioning of thought. In addition, according to Aristotle, if concepts and external referents are the same for all languages, it follows that there is a fundamental commensurability between each language. This is what the principal investigators of the various research groups that arose from the 1960s onwards thought.

The first year [...] revealed that in two respects we had somewhat underestimated the extent of the research [...] the degree of exact correspondence between the expressions of different languages is much less than anyone ever suspected [...] Two languages, for example, may present the same observational objects, but the relation in which they are presented is different; or there is difference in the elements into which the named thing is divided". (Ceccato, 1960 as cited in De Mauro 1999, pp. 168–169, trans. by the author)

Silvio Ceccato, head of a research group, personally verifies that the Aristotelian foundation of the principle of identity does not hold. Each language carves out the plane of content (the concepts, fig. 2) and the plane of expression (the signs and the way they combine with each other, fig 3) in different and not always corresponding ways. These differences are not merely expressive alternatives. The way a language carves out and articulates its concepts is something essential to thought. The externalist approach stems from this assumption. The internalist project of modeling cognition by tracing the basic rules and elements that constitute thought failed before it even began.

¹ <https://www.w3.org/OWL/>

² For a further review of alternative NLP (Natural Language Processing) approaches see Cambria and White (2014)

³ ConceptNet, as well as other knowledge graphs and ontologies, can be used together with other AI models (based on distributional semantics and not involved with internalism), but at this level of the paper it is important to understand the ineffectiveness of these approaches when used exclusively.

⁴ Trying to describe the logical form of the proposition and therefore of thought, Wittgenstein in his *Tractatus Logico-Philosophicus* speaks of the primitive signs as follows: "The meanings of primitive signs can only be explained by means of elucidations. Elucidations are propositions that contain the primitive signs. So they can only be understood if the meanings of those signs are already known." (Wittgenstein, 2002, 3.263). To be sure, the later Wittgenstein of the *Philosophical Investigations* can be said to follow the externalist paradigm.

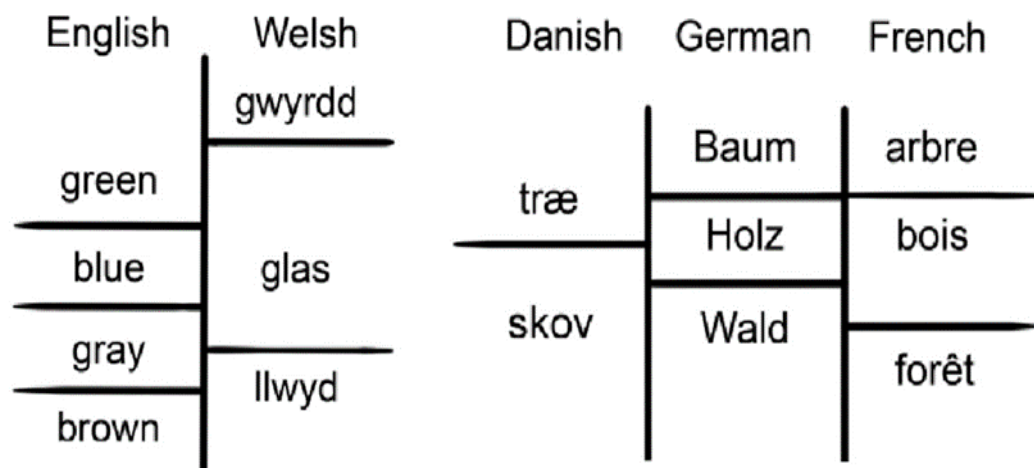


Figure 2. Scheme of differentiations of some meanings in different languages (Hjelmslev, 1969, p. 53–54)

Uygarlaştıramadıklarımızdanmışsınızcasına

(behaving) as if you are among those whom we could not civilize

uygar: civilized
 _laş: become
 _tır: cause somebody to do something
 _ama: not able
 _dik: past participle
 _lar: plural
 _ımız: 1st person plural possessive (our)
 _dan: among (ablative case)
 _mış: past
 _sınız: 2nd person plural (you)
 _casına: as if (forms an adverb from a verb)

Figure 3. Structure of a Turkish sentence (example from Jurafsky and Martin, 2009, p. 46)

In summary, internalism is concerned with the search for the internal structures of the mind, the same for all individuals. This affects its approach to the representation of cognition, leading it to isolate two elements: the concepts (lexicon) and the structure of rules that relate them (grammar) to produce thought.

Nevertheless, the above examples have shown that neither of these two elements, isolated by internalism, is a constant measure among languages. Different languages, and the ways in which they structure cognition, cannot be sidelined in addressing the problem of replicating human thinking.



Externalism

According to externalism, what is commonly called human thought is not an entirely internal phenomenon, ascribable to an innate genetic heritage, but relies on external material elements, and develops throughout the history of the relationship between the human being and technical artifacts. Among its exponents are linguists, philosophers, psychologists, anthropologists, archeologists and scholars from many other disciplines (see Wittgenstein, 1968; Vygotsky, 1987; Leroi-Gourhan, 1993; Whorf, 2012; Ihde and Malafouris, 2019).

From this point of view, the rational, discursive, systematic and almost syllogistic character attributed to thought, far from being something universal, is the result of the relationship of human beings with articulated language, syllabic writing, printing press (book diffusion) and many other technologies.

Although not unique, language, in its material (phonic) aspect, is considered by many thinkers to be the most important technical supplement for the phylogenetic and ontogenetic development of the human being. For example, Vygotsky and Luria (1993) showed how the child's preverbal thinking functions in ways that are structurally different from those of the adult. The preverbal cognition modalities, observed in some primates and in some people with language disorders, are based on perceptual, tactile and mainly operational ways of processing experience, which have little to do with the reflective, detached and to some extent more analytical thinking offered by articulate language. In particular, some externalists conceive the relationship between thinking and speech as a functional and structural unit.

From a functional point of view, verbal thinking emancipates language from the communicative function that internalism has assigned to it. Thought and speech, two processes with separate genetic roots, come together during ontogenetic development to give rise to verbal thinking (Vygotsky, 1987, p. 109). Thus, speech comes to be shaped as a process in which thought is not expressed but formed (Vygotsky, 1987, p. 110; Wittgenstein, 1968, 244).

From a structural perspective, signs, in their very materiality, are constitutive elements of human cognitive activity, an essential technical supplement to the differentiation of meanings in the indistinct continuum of preverbal thought (Vygotsky, 1987, p. 250; see also Saussure, 2011, p. 112). No wonder that for externalists, the investigation of language is a constitutive part of the investigation of thought.

In the next chapter it is shown how articulated language structures verbal thinking.

Identity of Concepts

The principle of identity (related to the frame problem) is an issue for externalists too. What guarantees communication and how can people understand their own and others' thoughts? Since the externalist approach cannot rely on a priori concepts and universal grammars, it tries to solve the problem starting from signs.

Describing linguistic entities, Saussure defines concept (or meaning) as "a quality of its phonic substance", at the same time he states that "a particular slice of sound is a quality of the concept" (Saussure, 2011, p. 103). The meaning is a property inseparable from the *signifier* side of the sign (the acoustic or graphic image) and vice versa. This



duality is constitutive of the linguistic sign. Words are not labels applied to objects and concepts already available all along. A sign without a meaning is not a sign, but an empty sound, which is distinguishable only up to a point (Vygotsky, 1987, p. 49), a meaning without a signifier is not a meaning, literally it is nothing.

Without signs articulation, the plane of content (thought) as well as the plane of expression (vocal or graphic) remain two indistinct continuums in which discrete units (words and concepts) cannot be identified.

However, in order for “an issuer or receiver to establish a semiotic relationship between two entities, it is evidently necessary that he or she can operate with each entity as that particular, determinate entity” (De Mauro 2019, p. 7 trans. by the author). Aristotle’s problem is not solved yet. Linguistics in particular has long been confronted with the problem of identifying the units of speech. According to the intrinsic duality of the linguistic sign, meanings and signifiers cannot but determine themselves by delimiting each other, through participation in a common system of signs, i.e. the *langue*⁵ (fig 4).

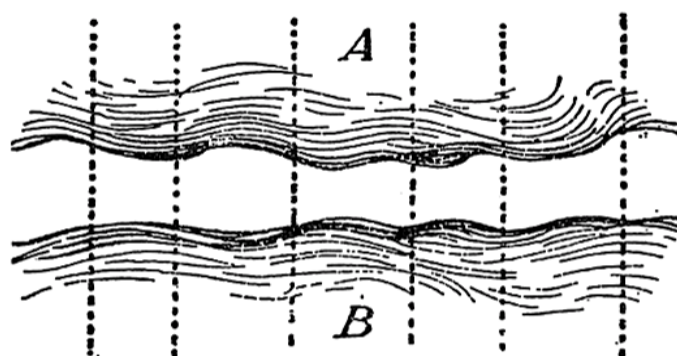


Figure 4. Divisions in the continuous and parallel planes of content and expression (Saussure, 2011, p. 112)

The picture shows the two parallel planes of content and expression (A and B); the *langue* serves as an intermediary between preverbal thought and sound. It consists of a system of reciprocal boundaries within the planes of content and expression. Thus, units are secured by differentiation. The chaotic and indistinct preverbal thought is led to specify and articulate itself according to the differentiations imposed by the *langue*.

In particular, the child’s thinking is initially a speech-independent process, not structured into meanings (or concepts) yet. Around the second year of age, the

⁵ Saussure distinguishes *langage*, the faculty of articulating signs of all kinds, from *langue*, the particular sign system (i.e., Italian, English, Russian etc.) that determines speech (see in this paper, *Langage, Langue* and *Speech*). The idea is that *langage* is a universal faculty of human beings. Every human can develop a functional use of signs and structure preverbal thought according to a particular signs system (the mother tongue). These systems differ in the way they structure preverbal thought (as seen in fig. 2 and 3). Thus, the *langue* is neither universal, nor individual, it is a social phenomenon. Saussure says: “For language is not complete in any speaker; it exists perfectly only within a collectivity” (Saussure 2011, p. 14).



developmental lines of language and thought intersect, giving rise to verbal thinking, articulating preverbal thought through external signs (Vygotsky, 1987, p. 257). Even internal speech, which is often entirely identified with thought, is actually belated compared to the external manifestations of verbal thinking. According to Vygotsky, when children learn to speak, they go through an egocentric speech stage. Egocentric speech plays an essential cognitive and operational guiding function in child behavior (Vygotsky, 1987, p. 114). Over time, the manifestations of egocentric speech fade away, but its cognitive function, migrating inward, is progressively internalized. Egocentric speech becomes internal speech (verbal thinking).

To summarize, each *langue* processes its units, the signs (each composed of concept and acoustic image), establishing itself as a form between two amorphous masses (Saussure, 2011, p. 112). Thus, the identification of concepts and phonic images is guaranteed by the solidarities between signs. The *langue* system is formed of two-sided signs which oppose each other along the two planes of expression and content (fig. 5).

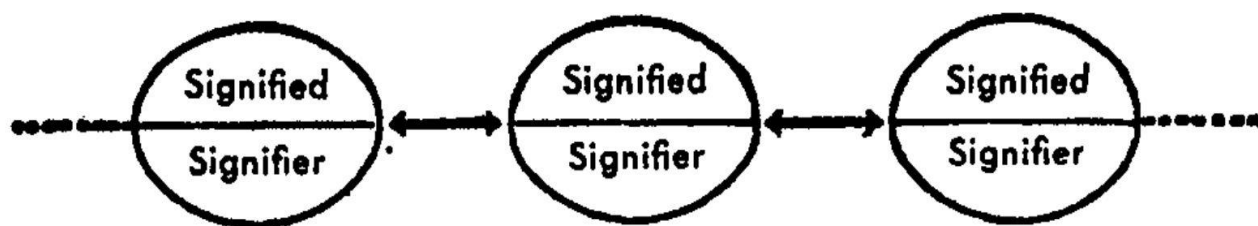


Figure 5. Bifacial signs opposing each other (Saussure, 2011, p. 115)

Signs Differentiation

Each different language (each *langue*) carves out the two planes in its own way (see fig. 2 and 3), which makes languages structurally different. Concepts occupy different positions within different systems, and the meaning (or value) of a sign varies according to the space the sign occupies within the system to which it belongs.

In Figure 2 the meanings of *blue* and *glas* partially overlap, but *glas* has a broader value since its meaning is not subdivided with other signs that delimit it, as with *blue*. A similar example can be made with grammatical morphemes. The value that in English falls on the sign *the*, in languages like Italian, is divided into many signs that articulate the semantic plane of gender and number, opposing each other: *il, lo, la, i, gli, le*.

The concept underlying this way of understanding thinking and speech is that of *differentiation*. The differentiations of the planes of expression and content have been studied by many disciplines and perspectives, below is an example of how meanings differentiate in language learning.

Children begin to speak by first uttering isolated words, gradually coming to articulate longer and longer chains. In mastering the signifying aspect of the *langue* they go from the part to the whole. Instead, in terms of their meaning, first words are to be considered as whole sentences. In semantic terms, children begin with the whole, with the word-phrase, and only later they break down their thinking into a series of separate and connected verbal meanings. By analogy with the plane of expression before the



arrival of language, Vygotsky (1987) describes preverbal thought as a fused, indistinct continuum (p. 250).

The child who utters “mom” for the first time, is not pointing to an object in the world, the referent of the uttered sign. The uttered word has a global and idiosyncratic meaning. The child’s verbal activity has not gained the articulate and representative character (Malafouris, 2007; Montani, 2020), that is too often attributed to any activity with symbols and signs (Capone, 2020), yet. In Heidegger’s (1996) words, it can be said that the children do not have a world of objects yet. Their word-phrase is uttered in conjunction with certain experiences, emotional impulses, external circumstances that are all but defined.

The subsequent meanings acquired by children constitute general designations. They are likely to learn the particular word *rose* at a very early age and to use it in a general way for any situation that evokes it, in the way *flower* is used (Vygotsky, 1987, p. 163), but also for many other situations related to the experiences that surrounded that word (that provided the background for the emergence of that meaning), such as a scent or a day outdoors, not necessarily involved with flowers.

Experiences, feelings, objects, and sensations are held together by words in changing and heterogeneous ways. With time and practice, the child will learn new words, begin to differentiate their concepts, and order the semantic plane accordingly.

Preliminary Conclusions

The literature on the conflicting relationship between externalism and internalism is extensive. The purpose of this paper touches this issue only tangentially. The goal is to investigate, from a practical perspective, whether an operationalization of language based on externalist assumptions can prove better than the internalist approaches in replicating verbal thinking.

The first chapter ends with the following conclusions:

- Cognition is closely related to semiotic activity (articulation of signals and meanings);
- Semiotic activity is based on the differential relations between the signs of a language system (*langue*) and not on the formulation of a deep structure of rules;
- The modeling and reproduction of cognition cannot be separated from a thorough understanding of the language system in which it develops.

In the next chapters, through an examination of the most recent advances in NLP (Natural Language Processing), the legitimacy of the externalist paradigm will be confirmed.

HOW TO REPRESENT LANGUAGES

The current state-of-the-art NLP systems are the gateway for the modeling and reproduction of verbal thinking, by implementing networks capable of solving heterogeneous assignments. In order to do so, they exploit the semantics of the language systems (*langue*). The hypothesis of this chapter is that the implementation, or rather the training of these networks, is carried out consistently with the externalist paradigm.

Before illustrating the relationship between AI and externalism, it is necessary to clarify the terms of the issue at hand.



Deep Neural Networks

AI is a broad discipline with a very ramified, albeit short, history. The following pages focus solely on deep neural networks (DNNs), the most common subset of artificial intelligence systems currently in use. DNNs are machine learning systems that spread after the failures of rule-based systems (Russell and Norvig, 2021, paragraph 1.3.3).

Although the name may suggest a closeness to cognitive science and neuroscience, the comparison between computational and biological neurons is improper. Neural networks are in no way brain models; it is more correct to imagine a neural network as a large parametric function in the form of a computational graph (fig. 6).

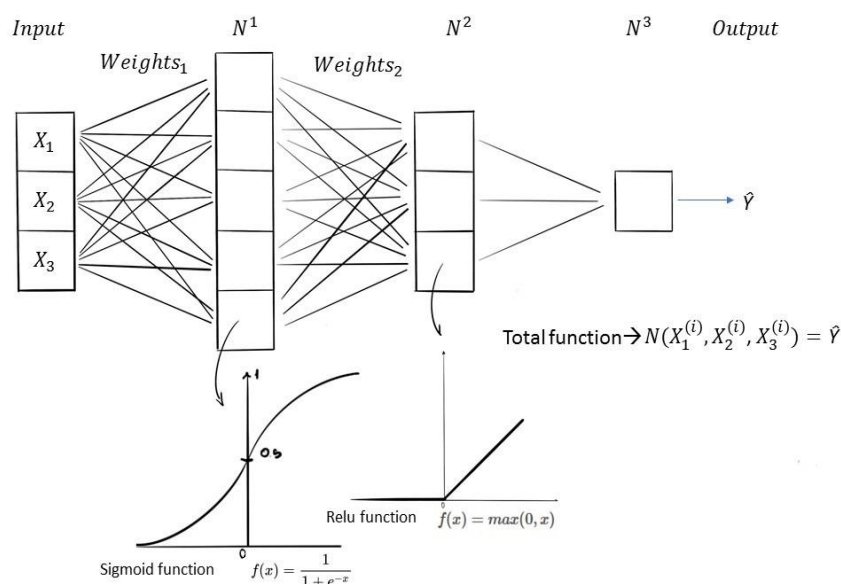


Figure 6. Graphical representation of a neural network with two examples of activation functions

The network is formed by Input, Output and inner layers. The inner layers are composed by neurons and weights (or parameters). The neurons are basically activation functions, they are the invariable part which composes the function that is the network. The parameters, placed before each layer of neurons, are the variable part of the function, these are the values that are learned by the network during training. The idea behind DNNs is that a large enough network (function) is likely to be able to represent any distribution of data.

DNNs can perform two types of tasks: regression and classification (fig. 7). Regression interpolates missing data based on given features. For example, taking as input a house's square footage, position and year of construction, the network can predict its price. Also, starting from pictures of dogs and cats, it can classify the pictures according to the represented animal. Everything that a neural network can do, must be done by means of these two techniques.

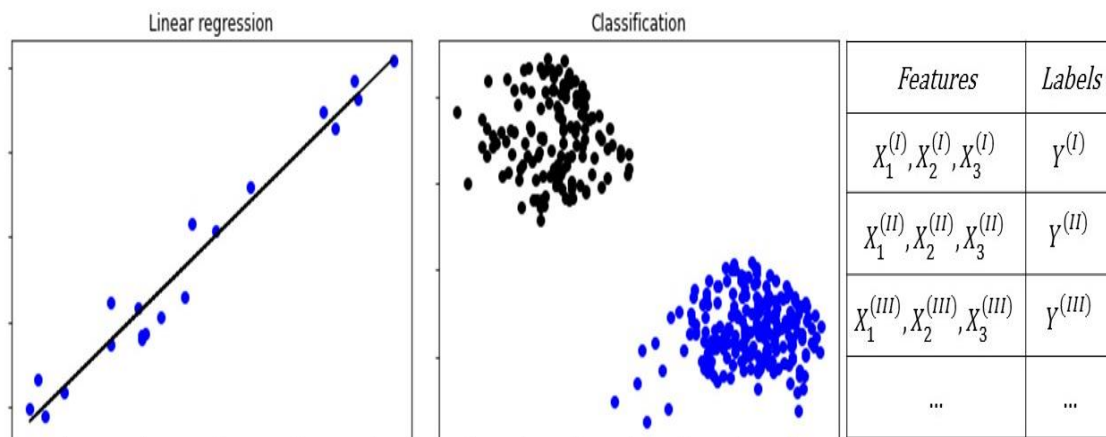


Figure 7. Linear regression and clustering of some random points

The network training process requires features (house features, images), and correct outputs (prices, image labels), all properly represented in numerical form. The network parameters are initially implemented with random values. The inputs are multiplied by weights in order to give each input feature a specific value for prediction purposes. The weighted features are then passed through the neurons layer, the process is repeated for each layer all the way to the output layer. The error of the predictions made in the training phase is calculated by comparison with the correct labels, the weights are modified accordingly (fig. 8). After several training cycles the weights of the network should be such that the operations between input, weights and neurons will result in the corresponding output.

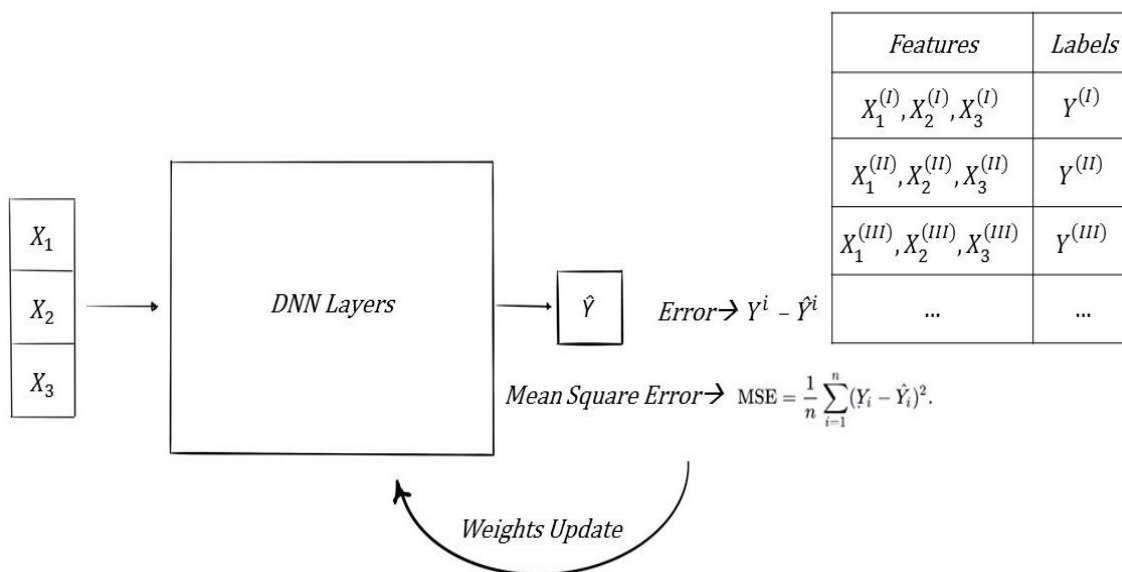


Figure 8. Example of a train cycle



If the weights are correctly calculated, the network will be able to predict the output of instances that were not in the train set. Briefly, during training, the inputs and the results of the function are available, but not the parameters (set randomly and corrected at each cycle). During prediction, the inputs and parameters are available, but not the output.

Clearly, such a tool has nothing to do with the mind, nor with supposed universal grammars. The first issue posed by the operationalization of verbal thinking is a problem of input representation; the second is how regression or classification might serve the purpose.

The Saussurian linguistics answers both these issues.

*Langage, Langue and Speech*⁶

Modeling verbal thought is mainly a problem of representation. The internalist paradigm has kept alive a very old tradition of language representation based on two categories of elements: lexicon and grammar rules. The modern internalist contribution has been to assume a universal grammar behind the grammar rules of each language (Chomsky, 1988, p. 61) and a set of semantic primitives from which the lexicon derives (Wierzbicka, 1996, p. 13; see also Osgood and Sebeok, 1954, p. 127).

Recent advances in NLP has definitively proven the impracticality of such a representation. Grammar is an *a posteriori* construction, the result of a reflexive relationship with speech, rather than a structure of rules that governs language. Similarly, the lexicon descends from a long work of differentiation within the speech (on the planes of expression and content) carried out by verbal thought. The externalist proposal aims to outline, within language, only those phenomena relevant to semiotic activity, and to represent them without relying on universal internal structures.

The term *langage* conceals an ambiguity: language is “a confused mass of heterogeneous and unrelated things” (Saussure, 2011, p. 9), the study of which involves many different disciplines, comprising various orders of problems. It is not possible to study semantics or verbal thought starting from language as a whole. Saussure proposes as the object of study the *langue* (the particular system of signs, i.e. Italian, English, Russian etc.)⁷ as distinct from language (*langage*), conceived as the faculty of articulating signs in general (the confused mass of things). The *langue* is defined as the essential part of language, its social product, “is a self-contained whole and a principle of classification” (Saussure, 2011, pp. 9–11). The *langue* is the system of reciprocal delimitations on the levels of content and expression, and as such it does not exist entirely in any individual, it is a social phenomenon resulting from the acts of speech of all the users. It is a treasure stored in the practice of the subjects of a community of speakers. The *speech* actualizes the signs of a language system by articulating them within syntagms (sentences). The relationships between signs within syntagms are what determine individuals’ learning of the structure of the language system, both expressive forms and contents. In turn, the mechanism of *langue* relies on these relationships in forming syntagms.

⁶ The English edition of the Course in General Linguistics does not help to understand Saussure’s terminological distinctions. The paper presents the reading of De Mauro, editor of the French and Italian critical editions.

⁷ *Langue*, language system and system of signs will be used as synonyms.



It is these relations, or regularities, that the network must take into account when representing the system of signs, rather than ontologies or taxonomies of concepts.

The Signs System Mechanism

In the previous section, the oppositional relations between signs, at the foundation of the language system, have been mentioned. It is now necessary to understand how these relations can be described. This is crucial to understand how a language system can be implemented through a neural network.

In speech, signs are articulated according to two types of regularity, each of which produces a certain order of values (Saussure, 2011, p. 122–127). The relations between signs based on the linear structure of speech are called syntagmatic relations or *in presentia* relations. The position of signs within the syntagm implies a relation of similarity or dissimilarity between alternative signs, which are eligible to occupy the same position. Similarly, neighboring signs, in a dissimilarity relationship, still entertain a relevant relation for the purpose of ordering the language system (fig 9).

The	cat	lies	on	the	bed
The	dog	jumped	over	the	fence
The	car	drove	down	the	road
His	son	lives	in	the	UK

Figure 9. Syntagmatic relation within 3 sentences

Within language systems, similarity is not based on positive qualities of the signs, but only on their mutual opposition within the system. From this point of view, similarity and dissimilarity are two sides of the same coin. The signs *cat* and *dog*, though having different meanings, are in a sense similar (more similar compared to *car*, for example), occupying the same position within a syntagmatic chain, and relating to the same types of signs (similar signs). On the other hand, signs considered individually entertain associative relationships (or *in absentia*) with other signs of the system (fig. 10). These relationships can be morphological (*teacher, teaching, teachable*), or they can be consistent with the semantic aspect (*teach, class, degree*).

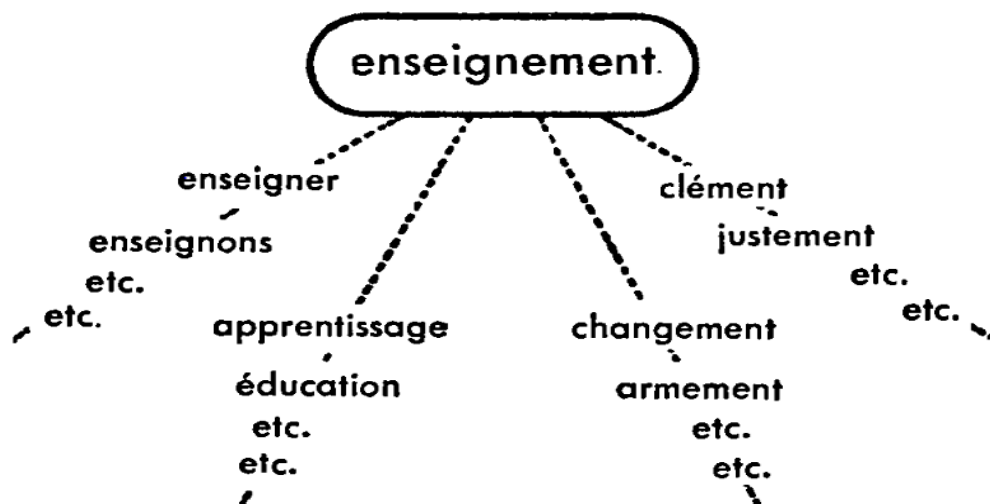


Figure 10. Associative relations (Saussure, 2011, p. 126)

Although in-presentia relations seem to prioritize syntactic features of the syntagm, this regularity may also reveal certain semantic properties. Within the sentence: “I am very sick, call a doctor”, the word *doctor* cannot be appropriately replaced by any given word, it can be replaced by *nurse*, maybe by *priest*, hardly by *architect*; even if the sentence would be grammatically correct.

The example is even clearer in commonly used stereotyped expressions such as: “break a leg”, “force the hand”, “it takes a village”, in which the relationships between signs depend on the precise meaning (one could say use) that these sentences have in the speech. In the language system, the distinction between semantic and syntax is not defined as it appears to those who rely on a grammatical (internalist) principle of language analysis, nor does it precede the formulation of full-sense propositions.

Sense and Meaning

To make the explanation of the functioning of the language system clearer, it is necessary to introduce the distinction between sense and meaning.

The meaning of a sign can be imagined as a differential zone in the plane (or space) of the content (De Mauro, 2019, p. 100). There cannot be an isolated meaning but only a system of meanings, of reciprocal delimitations (fig. 2 and 5). These delimitations are expressed by the syntagmatic and associative relations between signs. The meaning, consistently with its formal and systemic characters, is the form, or scheme, or rule of realization of a sense. De Mauro defines meaning as “the class to which a sense belongs” (De Mauro, 2019, p. 19, trans. by the author). The meaning of dog (a well-known animal, De Mauro, 2019, p. 186) can actualize into concrete utterances a potentially countless amount of unpredictable contingent senses through relations with other signs in a syntagm. Instead, sense is defined as “what in a particular moment, by a particular user, is indicated with a signal” (De Mauro, 2019, p. 7), the concrete actualization of a meaning in speech.



In order to manage the complexity of language systems, and not to limit itself to fixed representations of concepts, the language model must take into account this mobility of the sense within the general pattern of use of a word, the meaning (this problem is addressed in the section called Language Models).

DNNs AND LANGUE – WORD EMBEDDINGS

The main problem of language modeling with DNNs is the representation of inputs and the identification of an output that allows the network to learn parameters such that it can formulate correct and relevant propositions, categorize texts, summarize them, and complete tasks related to verbal thinking.

The first thing to do, in order to follow the externalist program, is to implement a model of the language system, i.e., a model of the differential space in which signs are opposed. Word2Vec was the first algorithm able to apply an externalist paradigm to NLP (Mikolov, Chen et al. 2013)

The network model is the same as illustrated in figure 6. The network's inputs consist of corpora (typically Wikipedia and BookCorpus). From these texts a vocabulary of words (tokens) known by the network is formed. Based on the vocabulary, the network can identify words as units within the text, but not their meanings yet.

Meaning, lacking positive properties, depends on syntagmatic and paradigmatic relations between signs within the text (meaning is a differential entity). These relationships establish the value of signs within the language system, determining their mutual similarities and dissimilarities, according to different orders of relationships (Mikolov, Chen et al. 2013). The network needs to model its content plane based on these relationships; to do this it relies on classification, trying to predict hidden words in the training corpora (in this case, the vocabulary tokens function as categories for the prediction).

It is possible to represent (with some approximation) the content plane as a homogeneous n -dimensional space. A n elements vector⁸ (or, as called by Mikolov, Chen et al. (2013), a *word embedding*), initialized with random numbers, is assigned to each vocabulary token. These n elements correspond to values in the various dimensions of the n -dimensional space for each token. Each token has its own place within this space (fig. 11). The mutual disposition of tokens in the space determines their meanings, providing a model of the language system.

⁸ A vector is a data structure, consisting of an array of numbers. In the case of Word2Vec it usually consists of 300 values.

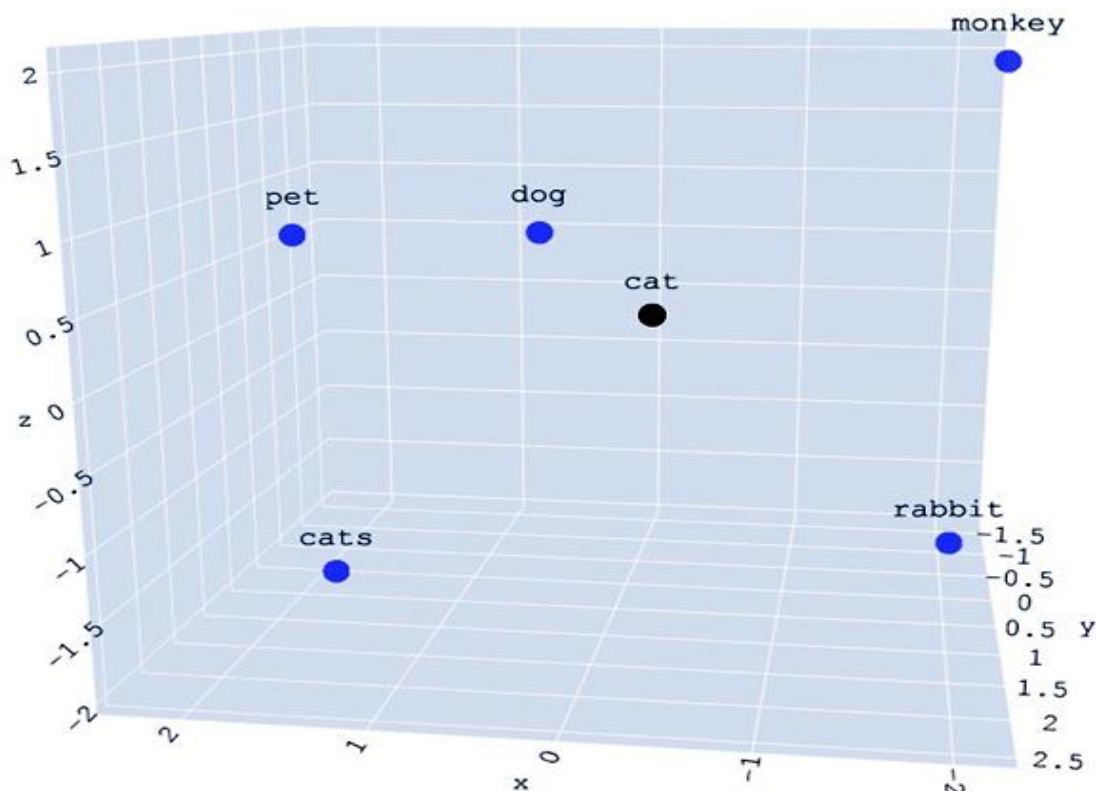


Figure 11. The five “cat’s” nearest neighbor tokens. The 300 value vectors corresponding to the tokens have been reduced to 3 dimensions by PCA (Principal Component Analysis).

The training can be described as follows. The network processes the text one batch of words at a time. At each step the network tries to predict the target word (the batch’s central token) based on the tokens surrounding it (the context). In the example (fig. 12) the network takes in a 3-token batch (usually wider), the embeddings of the first and third tokens constitute the input, the vector of the second token is the output to be predicted. The output of the network is a n dimensional vector (a word embedding), this embedding is confronted with the vocabulary tokens’ embeddings to make a prediction (basically the network classifies the hidden token as a specimen of a vocabulary token). Afterward, the predicted token is compared to the target token and the error is computed. Thus, the word embeddings are modified accordingly. At the end of training, the token disposition should reflect the word distribution in the training corpora⁹.

⁹ This training method is called CBOV (continuous bag of words). It is not the only possible method, usually several methods are used together (in the case of Word2Vec, SkipGram is also used). What is important is that all these methods are based on the analysis of the distribution of words in the corpora, representing syntagmatic relations between



Tokenized sentence: [sos] [l] [am] [sick] [,] [please] [call] [a] [doctor] [eos]

3 token batch: [please] [call] [a]

$Y = [call]$

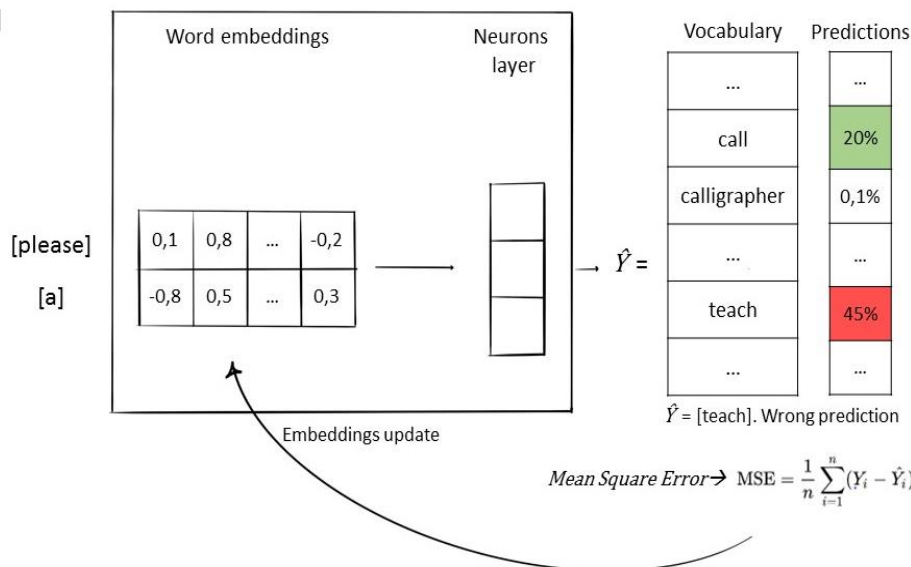


Figure 12. Word2Vec training ([sos] and [eos] stand for start and end of sequence).

In the example the network must predict the embedding for *call*, based on the context tokens. However, other embeddings with similar meaning are likely to occupy a neighboring space to that of the target token in the n -dimension space (fig. 13). The output embedding should be as close as possible to that of the target word.

contiguous signs and paradigmatic relations between signs that are likely to occupy the same position within similar contexts.

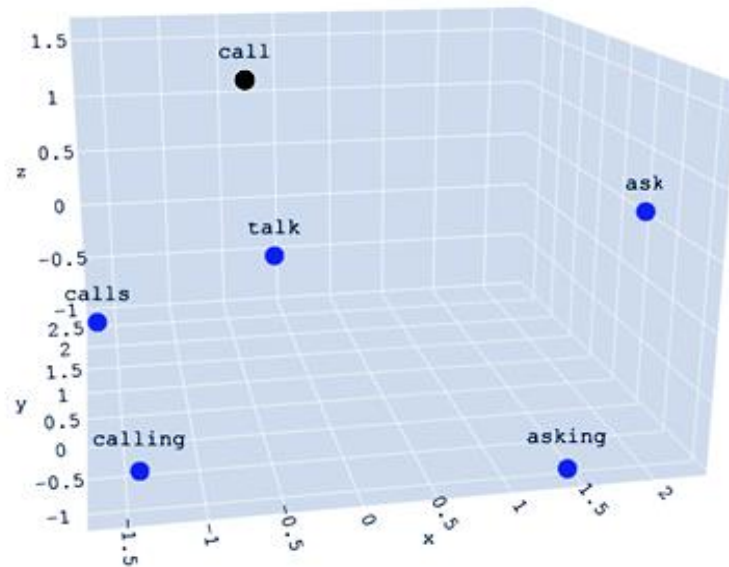


Figure 13. Five nearest neighbors for the token “call”

The most impressive result is that, once trained, word vectors show to represent general relationships between concepts. Mikolov, Yin et al. (2013) report a couple of examples. Pairs of tokens which stand with each other in an analogous semantic relation are found grouped in the same way within the vector space (fig. 14).

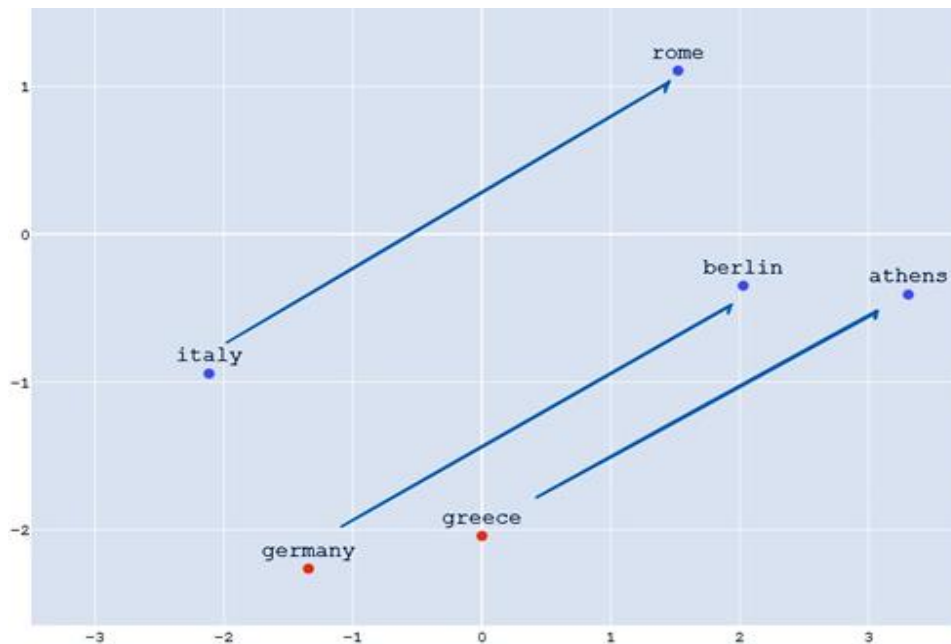


Figure 14. Grouping of states and capitals tokens



It is also possible to obtain meaningful results through operations between vectors. In the example of Mikolov, Yih et al. (2013), the vector *man* is subtracted from the vector *king*, then the vector *woman* is added. The result is an embedding very close to the vector corresponding to the *queen* token (fig. 15).

```
result = model.most_similar(positive=['woman', 'king'],
                           negative=['man'], topn=1)
print(result)

[('queen', 0.7698541283607483)]
```

Figure 15. A line of code reproducing the “king/queen experiment”

Another interesting application is sentiment analysis. Through the word embedding algorithm it is possible to train a network to recognize the emotional polarity of a text by simply giving in input texts and labels that indicate the polarity. Once the embeddings are computed, the network can estimate the emotional polarity of a text. Similar training is used for text classification, using topics as labels.

Despite its interesting applications, Word2Vec is simply a DNN that, based on a specific task, calculates a series of embeddings whose application is limited to the task on which they were computed. Furthermore, the embeddings are static, in prediction phase each word will always correspond to the same embedding regardless of the context of occurrence. The network merely computes a weighted average of the distribution of words within the training text, something very similar to an average of the usages, or better, of the relations between signs in the corpus.

In a nutshell, Word2Vec computes an approximation of the content plane, leaving out entirely the problem of *sense*, i.e. the concrete actualization of word meanings in syntagms.

Two problems to be solved:

- The embeddings contextuality, an essential factor to determine the sense of a token (and of a syntagm),
- Model generality, i.e., the implementation of a LM (Language Model) capable of generating contextual embeddings that can be used for any task.

The Transformer model addresses these problems from an externalist perspective.

Language Models

The current state of the art in NLP is dominated by LMs based on transformer, the DNN architecture that exploits the algorithm called Attention Mechanism (Vaswani et al., 2017), for the representation of words contextual meaning¹⁰. This section focuses only

¹⁰ Chomsky has recently been critical in discussing NLP applications of deep neural networks, <https://www.youtube.com/watch?v=ndwIzPBs8Y4>.



on Bert (Bidirectional Encoder Representations from Transformers), a particular model that exploits the transformer architecture (Devlin et al., 2019). Bert is a reduction of the standard transformer model, which is specialized in natural language encoding.

In order to provide contextual meaning (sense) representations, it is necessary for the network to take into account the wide variety of contexts in which a token may occur and a very large context within the syntagmatic chain (not just neighboring signs as with Word2Vec). The embeddings' contextuality also solves the problem of the model's generality, as its embeddings will no longer be task-oriented but re-computed at each prediction according to the task at hand. In addition, Bert provides a better integration between content and expression planes. Instead of using word tokens, it uses word-piece tokens (similar to morphemes, extracted from corpora thanks to a special algorithm). This makes the vocabulary less redundant, provides better handling for compound words and allows prediction of unknown words meaning based on context (fig. 16).

```
tokenizer = BertTokenizer.from_pretrained("bert-base-uncased")
print(tokenizer.tokenize("transubstantiation is a difficult word"))

['trans', '##ub', '##stan', '##tia', '##tion', 'is', 'a', 'difficult', 'word']
```

Figure 16. Bert model tokenizer

The transformer architecture is at the heart of many downloadable pre-trained models. Usually they consist of a word-pieces vocabulary with its static embeddings, plus a set of heads and layers. These are sets of parameters, trained by the model's developer, through which the network computes the contextual embeddings (fig. 17). Their training is similar to that of Word2Vec.

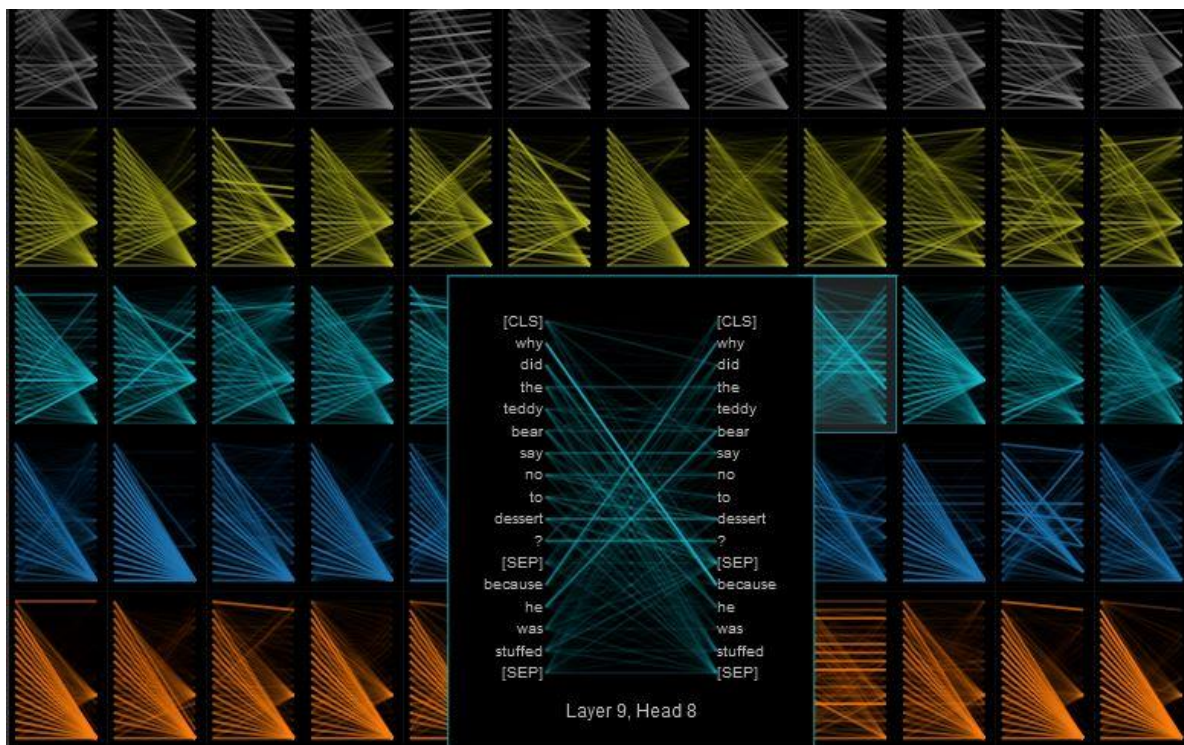


Figure 17. The last five layer of Bert model (12 head x 12 layer) processing a sequence of two sentences. In the box, the detail of head 8 processing at layer 9. At this point the network highlights the links between “why” and “because”, as well as “he” and “bear”. Picture obtained with BertViz (Vig, 2019).

The process of contextual embedding computation (encoding) works as follows (fig. 18). A long sequence of text is processed. The Bert model (Devlin et al., 2019) for instance, can process up to 512 tokens at time, representing each token with a 768 elements embedding. At first, a positional and a sequential embedding are added to the static embedding of each token, to incorporate information about its position in the sequence. Afterward, the embeddings sequence is processed on each head in parallel starting from layer zero. In the following layers, along each head, the embeddings of the sequence undergo several calculations and are multiplied by each other¹¹, so that the vector of each token represents information about the context in which it occurs. This generates a set of partial embeddings per head. Partial embeddings pass through all the layers of their respective head, multiplying by each other at each layer. At the end of the

¹¹ This is not a simple multiplication. In each layer, based on the input embedding of each token, three indices (vectors) are calculated: Query (Q), Key (K) and Value (V). The network calculates the “score” of each token with respect to each other in the examined chain (in the batch of tokens). This score is calculated through a complex matrix operation in which the dot product between all vectors K and each vector Q is divided by the square roots of the dimensionality of K. A softmax function is applied to the resulting vectors, they are multiplied by the vectors V and the resulting vectors are added together. In this way a score is obtained for each token in the batch. However, it is not necessary to go into the details of these operations for the purpose of this paper.



layers, all the partial embeddings of all the heads are merged to form the actual contextual embeddings.

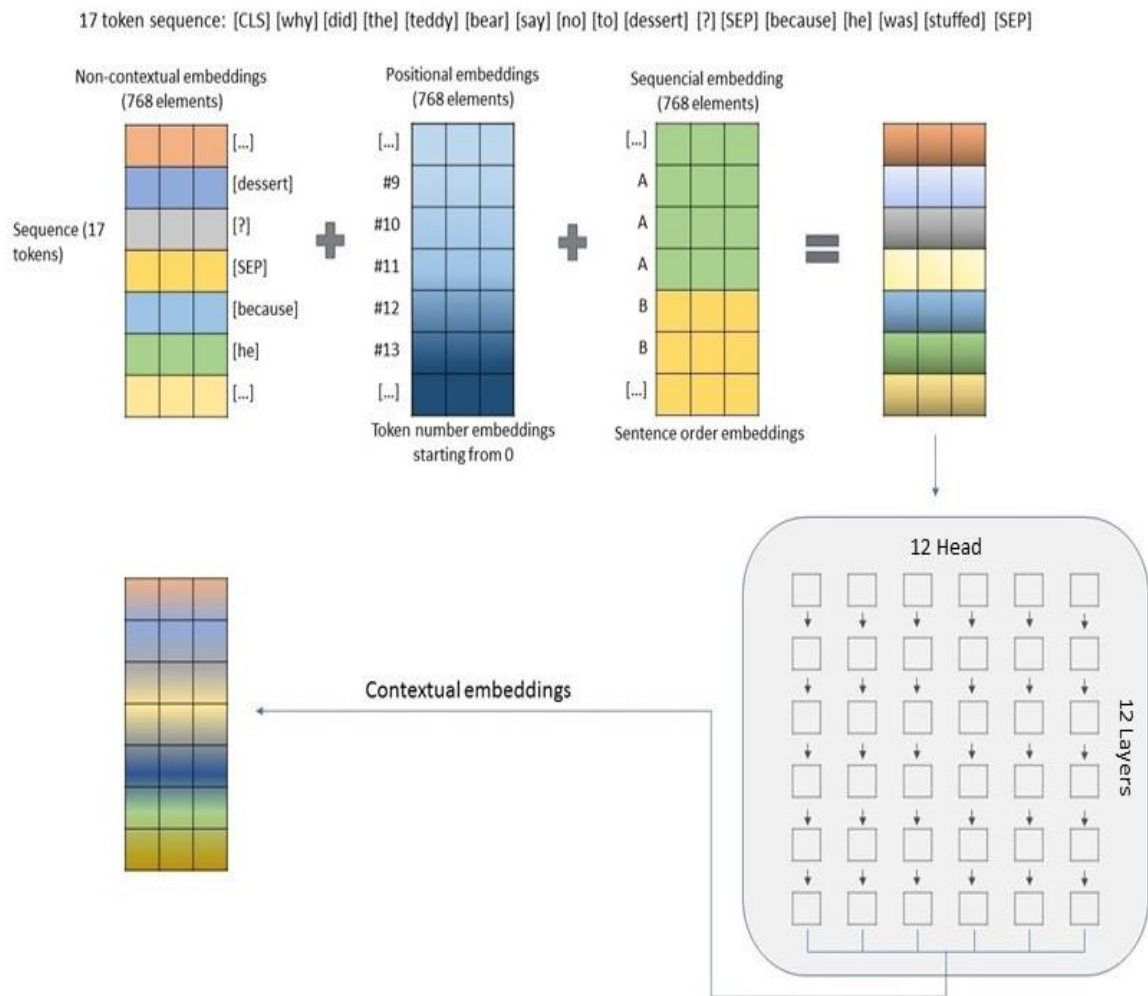


Figure 18. Contextual embeddings computation

Having different parameters, each head processes the input differently, highlighting certain relationships between signs and ignoring others (fig. 19).

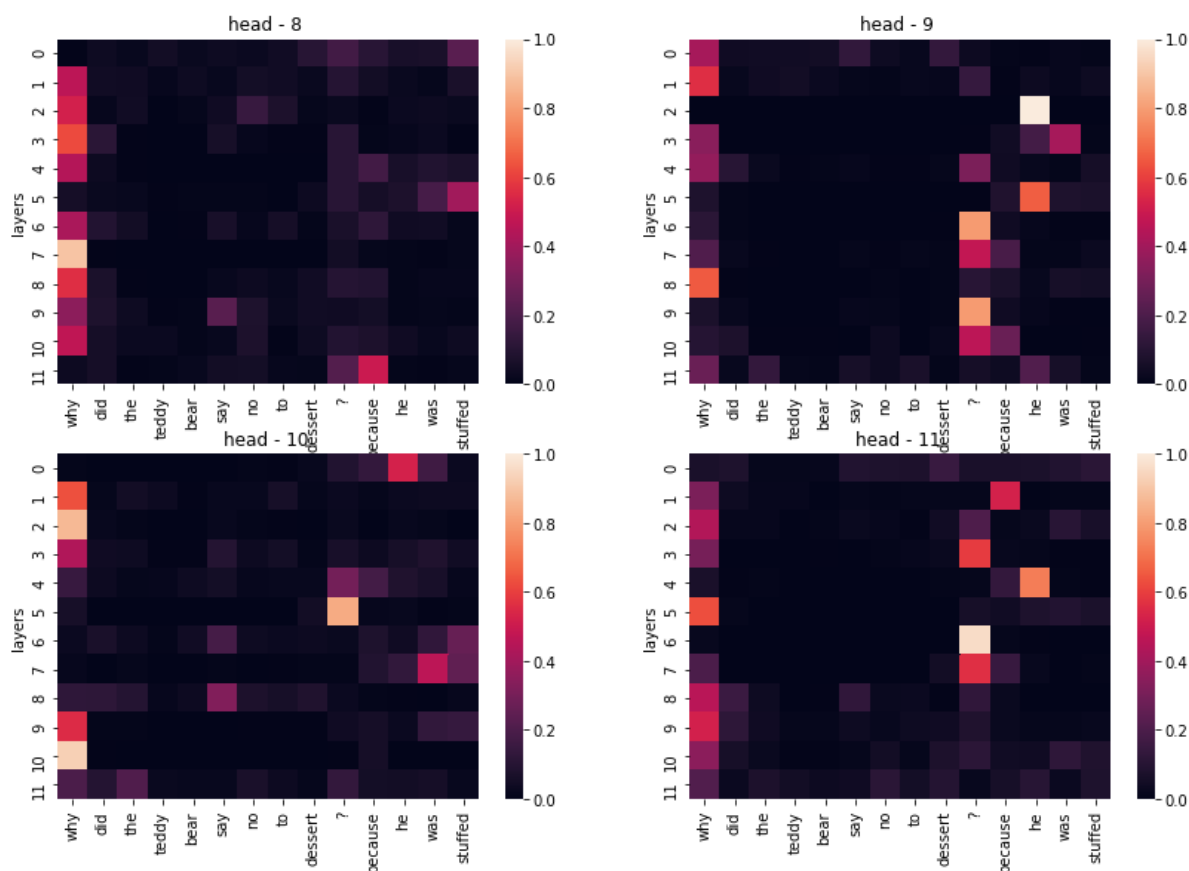


Figure 19. Heatmap of the attention of heads 8,9,10,11 related to the token “because”

Afterward, the contextual embeddings can be processed by a simple multilayer perceptron (a simple neural network) for specific tasks, but there is more. The ability to predict unknown tokens by relying on contextual embeddings makes these models’ performances in Cloze test questions and zero-shot classification tasks (text classification without training) worthy of notice. These tasks give an idea of what these models *know*. Two examples follow.

Masked Language Modeling Question Answering

It is possible to ask questions to a non-fine-tuned LM as long as they are formulated in the form of a text to be completed (Cloze Test) (Schick and Schütze, 2021). The network will simply estimate the missing word following the syntagmatic and associative relations learnt during training (fig. 20).

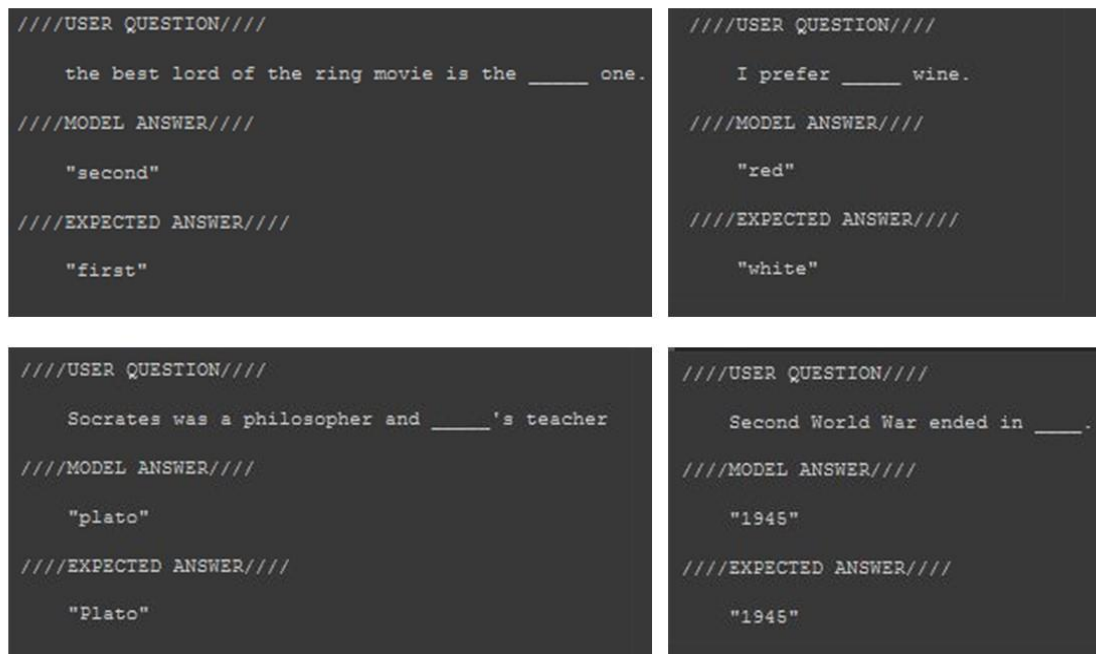


Figure 20. “Personal taste” and general culture answers from Bert model

Since it is a non-fine-tuned model, it will not be possible to get overly precise answers. Furthermore, in this example the network has a limited context (few words) to predict the desired answer. However, the answers obtainable through general knowledge questions are of interest, considering that the model does not make use of the Internet or any database from which to extract answers. The only resources available to the network are the vocabulary, and the parameters.

Zero-Shot Classification

A very useful task in NLP is text classification. Traditionally, a classifier was trained through texts labelled according to relevant categories to be predicted. These models were not general but task-oriented (and topic-oriented). In contrast, a zero-shot model is a model that, without any prior topic-oriented training, must be able to predict to which of the user-provided categories a text belongs.

A hypothetical user may need to know what has been tweeted about in the last month. By providing tweets and desired categories as input, the model will categorize all posts automatically (fig. 21).

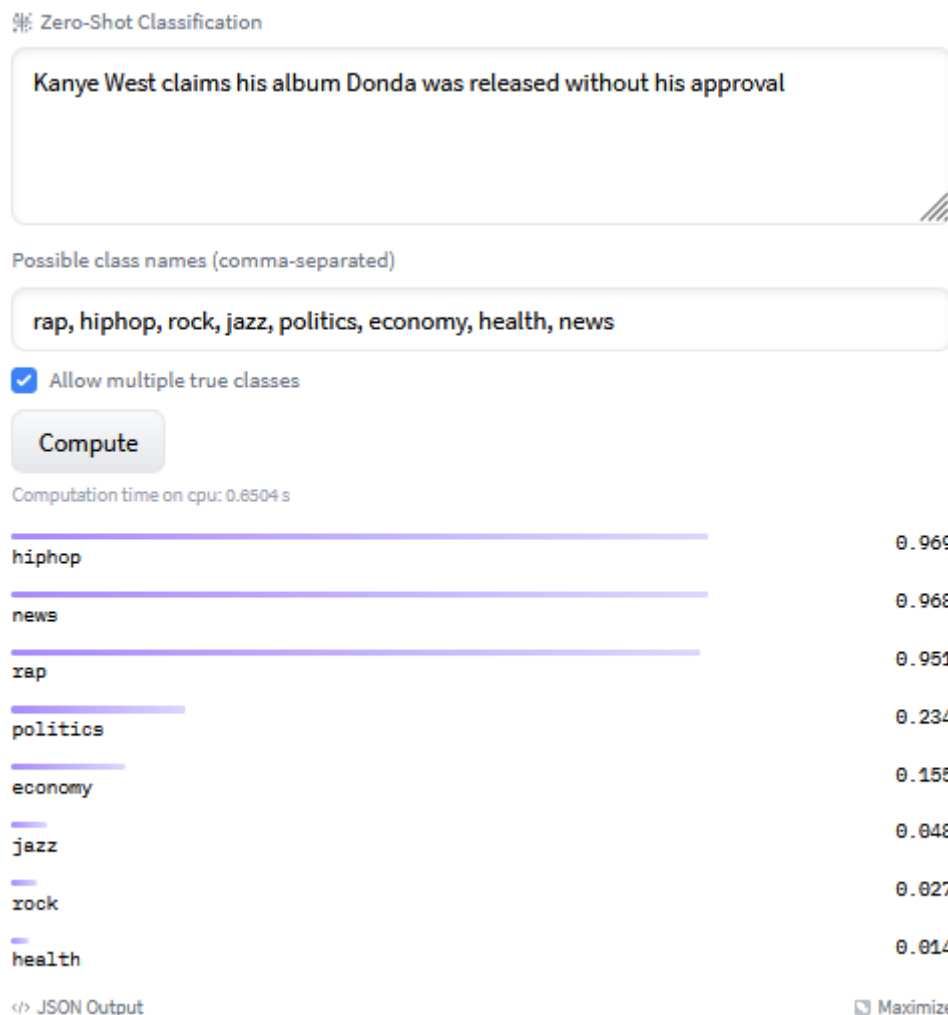


Figure 21. Screenshot of the huggingface API¹² running the Yin et al (2019) LM for zero-shot classification. The model is processing a 30th August 2021 BBC tweet¹³. The classes are chosen by the author. It appears that Bert knows Kanye West.

The training of this specific model is more complex. The model is trained on pairs of propositions labelled as mutually entailed or non-entailed. This training method is called Entailment Model Training (Yin et al., 2019). However, the relationships between tokens are still at the basis of the model’s cognitive capabilities in this training.

The most interesting applications remain those involving a dialogue between user and machine. There are several platforms where people without coding skills can interact with these models. Apart from the *huggingface API* mentioned above, *AI Dungeon*¹⁴ is a

¹² <https://huggingface.co/facebook/bart-large-mnli>

¹³ <https://twitter.com/BBCWorld/status/1432341614962356224>

¹⁴ <https://play.aidungeon.io/main/about>



role-playing platform managed by a transformer model, and *Eleuther AI*¹⁵ is a *text2text* LM that produces sentences based on a prompt provided by the user.

CONCLUSION

The success of the externalist approach to NLP proves the validity of the structural conception of language as a differential system of signs. On the other hand, from a functional point of view, the use of signs for the accomplishment of tasks by LM demonstrates that sign systems are not merely a communicative tool, but a mediating interface necessary for verbal thought and its replication.

In this chapter, the language system representation proposed by Saussure was compared with the LM provided by the transformer architecture. The following points were established:

- Word embeddings are able to differentiate meanings within a continuous n -dimensional space, providing a good structural representation of *langue*.
- The transformer architecture can model a system of meanings and take into account (to a certain extent) the difference between sense and meaning in the execution of tasks.
- Contextual meaning representations of tokens can be used by the network to generalize certain semantic relations and eventually reproduce verbal thinking features.

The direction taken by the research seems promising, but the full replication of verbal thinking is still a long way off. It is clear, however, that the progress of AI is closely linked to language and semiosis, understood as the technical activities of operating with signs.

REFERENCES

- Aristotle. (2014). *Complete Works of Aristotle: The Revised Oxford translation, one-volume digital edition* (J. Barnes, Ed.). Princeton University Press.
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I., & Amodei, D. (2020). *Language Models are Few-shot Learners*. arXiv. <https://arxiv.org/abs/2005.14165>
- Cambria, E., & White, B. (2014). Jumping NLP Curves: A Review of Natural Language Processing Research. *IEEE Computational Intelligence Magazine*, 9(2), 48–57. <https://doi.org/10.1109/MCI.2014.2307227>
- Capone, L. (2020). Tecnica e Rappresentazione. L'essere Umano Come Estensione del Media Digitale [Technique and Representation. The Human Being as an Extension of Digital Media]. *Pólemos. Materiali di filosofia e critica sociale*. <https://doi.org/10.48247/P2020-2-018>

¹⁵ <https://6b.eleuther.ai/>



- Capone, L. & Bertolaso, M. (2020). A Philosophical Approach for a Human-centered Explainable AI. *Proceedings of the Italian Workshop on Explainable Artificial Intelligence 2020*. <http://ceur-ws.org/Vol-2742/short1.pdf>
- Ceccato, S. (1961). Operational Linguistics and Translations. In *Linguistic Analysis and Programming for Mechanical Translation* (pp. 11-27). Feltrinelli.
- Ceccato, S. (2003). Correlational Analysis and Mechanical Translation. In S. Nirenburg, H.L Somers, & Y.A. Wilks (Eds.), *Readings in Machine Translations* (pp. 137–156). The MIT Press. <https://doi.org/10.7551/mitpress/5779.003.0015>
- Chomsky, N. (1959). A Review of B.F. Skinner’s Verbal Behavior. *Language*, 35(1), 26–85.
- Chomsky, N. (1988). *Language and Problems of Knowledge: The Managua Lectures*. MIT Press.
- Chomsky, N. (2015). *Aspects of the Theory of Syntax: 50th Anniversary Edition*. The MIT Press.
- De Mauro, T. (1967). *Ludwig Wittgenstein: His Place in the Development of Semantics*. Springer Science + Business Media B.V.
- De Mauro, T. (1999). *Introduzione Alla Semantica* [Introduction to Semantics]. Laterza.
- De Mauro, T. (2019). *Minisemantica dei linguaggi non verbali e delle lingue*. Laterza.
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). Bert: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv:1810.04805 [cs]*.
- Dreyfus, H. L. (1992). *What computers still can’t do: A critique of artificial reason*. MIT Press.
- Dreyfus, H. L. (2007). Why Heideggerian AI Failed and How Fixing it Would Require Making it More Heideggerian. *Philosophical Psychology*, 20(2), 247–268. <https://doi.org/10.1080/09515080701239510>
- Fellbaum, C. (2006). The Princeton Wordnet. *Encyclopedia of Language & Linguistics*, 13, 665–670.
- Havelock, E. A. (2010). *Preface to Plato*. Belknap Press, Harvard University Press.
- Heidegger, M. (1996). The Question Concerning Technology. In *The question Concerning Technology and Other Essays* (W. Lovitt, Trans.) (pp. 3–36). Harper and Row.
- Hjelmslev, L. (1969). *Prolegomena to a Theory of Language* (F. J. Whitfield, Trans.). University of Wisconsin Pr.
- Ihde, D., & Malafouris, L. (2019). Homo faber revisited: Postphenomenology and Material Engagement Theory. *Philosophy & Technology*, 32(2), 195–214. <https://doi.org/10.1007/s13347-018-0321-7>
- Jurafsky, D., & Martin, J. H. (2009). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition* (2nd ed). Pearson Prentice Hall.
- Leroi-Gourhan, A. (1993). *Gesture and Speech*. MIT Press.
- Liu, H., & Singh, P. (2004). Conceptnet—A Practical Commonsense Reasoning tool-kit. *BT Technology Journal*, 22(4), 211–226. <https://doi.org/10.1023/B:BTTJ.0000047600.45421.6d>



- Malafouris, L. (2007), Before and beyond Representation: towards an Enactive Conception of the Palaeolithic Image. In *Image and Imagination: A global Prehistory of Figurative Representation* (Renfrew C., Morley, I. Eds.) (pp. 287–301). Cambridge: McDonald Institute.
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). *Efficient Estimation of word Representations in Vector Space*. arXiv. <https://arxiv.org/abs/1301.3781>
- Mikolov, T., Yih, W., & Zweig, G. (2013). Linguistic Regularities in Continuous Space Word Representations. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 746–751). Association for Computational Linguistics. <https://aclanthology.org/N13-1090.pdf>
- Montani, P. (2020). The imagination and its technological destiny. *Open Philosophy* 3(1). <https://doi.org/10.1515/opphil-2020-0107>
- Ogden, C. K., & Richards, I. A. (1946). *The Meaning of Meaning: A Study of the Influence of Language upon Thought and of the Science of Symbolism*. Harcourt Brace & World Inc.
- Ong, W. J. (2004). *Orality and Literacy: The Technologizing of the Word*. Taylor & Francis e-Library.
- Osgood, C. E., & Sebeok, T. A. (Eds.). (1954). Psycholinguistics: A Survey of theory and Research problems. *The Journal of Abnormal and Social Psychology*, 49(4, Pt.2), i–203. <https://doi.org/10.1037/h0063655>
- Russell, S. J., & Norvig, P. (2021). *Artificial Intelligence: A Modern Approach* (4th ed.). Pearson.
- Saussure, F. de. (2011). *Course in General Linguistics* (P. Meisel & H. Saussy, eds.; W. Baskin, Trans.). Columbia University Press.
- Schick, T., & Schütze, H. (2021). Exploiting Cloze Questions for Few Shot Text Classification and Natural Language Inference. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics* (pp. 255–269) Association for Computational Linguistics. <https://doi.org/10.18653/v1/2021.eacl-main.20>
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, LIX(236), 433–460. <https://doi.org/10.1093/mind/LIX.236.433>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). *Attention is All You Need*. arXiv. <https://arxiv.org/abs/1706.03762>
- Vig, J. (2019). A Multiscale Visualization of Attention in the Transformer Model. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: System Demonstrations* (pp. 37–42). Association for Computational Linguistics. <https://doi.org/10.18653/v1/P19-3007>
- Vygotsky, L. S. (1987). *The Collected Works of L. S. Vygotsky* (R. W. Rieber & A. S. Carton, Eds.). Plenum Press.
- Vygotskij, L. S., & Luria, A. R. (1993). *Studies on the History of Behavior. Ape, Primitive, and Child*. Lawrence Erlbaum Associates, Inc., Publishers.

Special Topic:

Technology as Language – Understanding Action in a Technical Condition

Спецвыпуск

“Техника как язык: понимание и действие в техническом мировоззрении”



- Whorf, B. L. (2012). *Language, Thought, and Reality. Selected Writings of Benjamin Lee Whorf*. The MIT Press.
- Wierzbicka, A. (1996). *Semantics. Primes and Universals*. OUP Oxford.
- Wittgenstein, L. (1968). *Philosophical Investigations*. Basil Blackwell.
- Wittgenstein, L. (2002). *Tractatus Logico-philosophicus*. Routledge & Kegan Paul.
- Yin, W., Hay, J., & Roth, D. (2019). Benchmarking Zero-shot Text Classification: Datasets, Evaluation and Entailment Approach. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)* (pp. 3914–3923). Association for Computational Linguistics. <https://doi.org/10.18653/v1/D19-1404>

СВЕДЕНИЯ ОБ АВТОРЕ / THE AUTHOR

Лука Капоне, l.capone@unicampus.it

Luca Capone, l.capone@unicampus.it

Статья поступила 12 июня 2021
одобрена после рецензирования 12 ноября 2021
принята к публикации 28 ноября 2021

Received: 12 June 2021
Revised: 12 November 2021
Accepted: 28 November 2021