

WE HAVE EVOLVED FOR COORDINATED LANGUAGE BY DIALOGUE

Svetlana T. Davidova

<https://orcid.org/0000-0001-9190-5937>

Abstract

In current linguistic theorizing language is defined by most prominent theoretical perspectives as a code-like algorithm designed for providing internal structure of rational thought and exemplified by the linguistic production of literate humans in disseminating information mainly in monologues. That said, informing and socialization by linguistic engagements (language) via participation in dialogues is a universal of language, regrettably neglected in current linguistic thought. In the current article I argue that the dialogue and its discernible internal structure is the best representation of natural language in its primary role as advertising one's private thoughts by socialization. Moreover, language by dialogues comes naturally in humans of all ages, cultures, genders, ethnicities, suggesting innate predispositions. Consequently, knowledge of language must include ability for participation in dialogues and must be recognized as an essential component of the language faculty, a topic highly relevant for the study of the evolution of language. In this sense the evolution of language is better defined as evolution of capacities for language by dialogue.

Keywords: language system, language, dialogue, language evolution, multimodality

1. Introduction: defining language in contemporary linguistic thought

1.1. Language as algorithm

In contemporary linguistic thought mostly influenced by the generative/biolinguistic paradigm language is defined as a computation system based on algorithms, that is, artificial language-like symbolic codes, based on principles borrowed from Turing's theory of computation. The generative/biolinguistic model builds sentences out of binary choices between discrete elements, i.e. discreteness is pointed at as the defining feature of language. In this context the language algorithm builds symbolic structures of infinite length and complexity by recursive combinations under arbitrary principles of organization by recursive combinations (Hauser et al. 2002 and elsewhere). An algorithm is by definition mechanical manipulation of symbols and symbols are codified meanings, which makes language a code by definition. As such it has the following characteristics emblematic of a code:

- a. It is composed of linguistic primitives, understood as discrete, object-like abstract entities which stand in fixed relations with one another and have existence independent of their users.
 - b. These consist of equally discrete and finite component parts. So sentences are decomposed into clauses, phrases, words, morphemes, syllables, phonemes, phonological segments.
 - c. Members of a lexicon are one-to-one stable associations of a meaning and a form, i.e. synonymy and homonymy is non-existent. These are defined by their membership in discrete and well defined grammatical categories and organized into sentences according to predetermined and fixed principles of grammaticality.
 - d. The meaning of a sentence is the sum total of the meanings of the composing words and their place in the architecture of the sentence.
-

- e. A sentence is the encoding of a complete thought. Explicit and complete mapping between semantic structure and grammar is the norm. All thematic roles in the theta grid of a verb are expressed in grammatical categories. The agent consistently occupies the subject position in the sentence structure.
- f. The language system is self-contained, stands alone, independent of context of use. This facilitates the uniform decoding of the meaning by people with vastly different experiences and views at any place and time.
- g. The sentence structure is highly detailed, it contains multiple embedding of phrases and sentences and highly abstract grammatical forms.
- h. The message for the sender and the receiver are identical, thus, assuming that individuals have identical minds.
- i. The function of code systems is mainly to inform, ergo, sentences are mostly statements, usually articulated in monologues.

Moreover, the study of language has been done with theoretical tools which reflect the properties of written texts as these are determined to exhibit the main characteristics of language as defined by modern theories: compositionality, hierarchical organization and situation-independence. In biolinguistic context the language faculty is a code processor.

1.2. Language system through language use

Alternatively, linguists who subscribe to the usage-based /complexity perspective and who study languages as used in communication, define the language system as patterns of human communicative behaviour, a system of symbols formed by social conventions and periodically altered by language use. The linguistic symbols, from lexical words to the highly abstract grammatical rules are understood as grounded in the speakers' concrete experience with the world and with language. Utterances used in real communicative acts are regarded as potential sources for extracting the rules of the language system through generalization. In this way the gap between concrete and abstract, lexicon and grammar, grammar and use, is bridged and language is viewed as a continuum of lexicon and grammar and grammatical system and its use, as a smooth continuity of past and present forms.

In the usage-based view categories are gradient, not discrete entities. The idealized representation of a category is a prototype, the best example of a category or a representation of the most distinctive features of a class of entities. The degree of membership in a category is defined based on similarity to the prototype. For example, the distinction of regular and irregular verbs is a matter of degree (Bates, Goodman, 1997 and elsewhere). The continuity of linguistic forms is pervasive in all languages, suggesting that it is a universal property of language.

The most typical circumstances of language use is universally the dialogue. As such it is meaning-based and the intended meaning is recovered, or inferred with assistance from context. In this sense language is an inferential system with the following properties:

- a. The inferential system has information-based, not structure-based internal organization, that is, organized around information structure (topic vs. focus)
 - b. It exists mainly in spoken form, where intonation assumes some grammatical functions, e.g. the formation of questions without the use of question words.
 - c. The building blocks of the system are flexible associations of form and meaning as standard
-

meanings are interpreted with context-dependent flexibility.

d. These form utterances composed of the most frugal use of constructions, the choice of which varies with the circumstances.

e. Most verbs have incomplete argument structure with only a single argument and when an utterance is a full sentence the order of the elements is flexible to signal speaker's attitude.

f. Semantically vague words and phrases, e.g. 'that fellow', 'that thing', 'people' are often used, the intended meaning of which varies with the context.

g. The meaning of a sentence is different from the meaning of the utterance and the difference between the two cannot be stipulated in advance by a code-like rules. The meaning of an utterance is calculated as the meaning of a sentence and the speaker's communicative intentions. Utterances communicate the intended meaning in addition to the speaker's attitudes.

h. The meaning intended by the sender is most often different from the meaning understood by the receiver. An inferential system is based on the assumption that participants are individualities with different minds and different life experiences in different communicative circumstances, which creates the potential for different interpretations of the same linguistic forms.

i. Utterances form part of spontaneous spoken dialogues mainly conducted in speech.

Thus, language is understood here as a flexible system in constant flux as a process of constant adaptation to human communicative demands.

2. The dialogue as the best demonstration of the distinct features of natural language

2.1. Natural language: stability and flexibility, code and inference in the dialogue

The code system and the inferential system are abstract models and do not exist in their pure forms in linguistic communication. Any individual communicative act contains elements of both code and inference as all uses of language presuppose some context. Although the linguistic message codified in written monologues is meant to be received as intended, irrespective of social and cultural contexts, all messages are open to interpretation, e.g. laws are interpreted differently by different jurists, works of literature, philosophy and science are interpreted differently by different readers according to their cultural, educational, age, etc. backgrounds. It is an everyday occurrence that public statements in politics, media etc. are interpreted differently by different receivers and some deviate significantly from the message originally intended, often prompting additional clarifications. Thus, the use of code in language is different from the standard use of code, e.g. Morse code, mathematical symbols, artificial languages etc. where the message is identical for the sender and the receiver as the use of the language code in written monologues allows for some flexibility of inference as the intended message is almost never the same as the received. The inferential system has the code as a component as the standard meanings of constructions are creatively interpreted in spontaneous dialogues. Language cannot serve as a communication tool without some standardization both of meaning and structure. The difference then, is in the role of code and inference in the different communicative circumstances, e.g. professional communication is predominantly, although not exclusively, a code, the inferential aspect playing a contributing and clarifying role. Spontaneous dialogues, on the other hand, are predominantly inferential with elements of code playing a contributing role. Modality influences the balance of code and inference as

written texts are better suited for code and speech and gesture are better suited for dialogues.

In sum, any individual act of language use contains both fundamental elements of language although to various degrees. The language system is an abstract model, a general framework with the role of guiding, not determining, its use in real communicative acts. Its activation and communicative utility is determined by the context. Thus, language as a communication system is an integration of code and context, underdetermined by the code and enriched by inference from the context. This suggests that knowledge of language must include knowledge of the linguistic code, i.e. stable patterns of association of meaning and form, and knowledge of the communicative situation, cultural and social values and attitudes, likely to be reflected in the communicative attitudes of the individual participants, both senders and receivers.

2.2. The dialogue as the best illustration of natural language

All human communicators of any age, ethnicity, cultural traditions, education, profession, communicate by dialogues. Spoken dialogues are universally the most frequent form of language use. Casual conversations among individuals with close social ties occupy about 20% of all waking time of humans (Dessalles, 2007). So, the dialogue is the best representation of the most distinct features of natural language in its use.

A. The dialogue illustrates the nature of language as both a cognitive property and a social convention, i.e. a system for advertising private thoughts. Understanding and acknowledging this reality is crucial for rendering the cognition vs. communication debate superfluous.

B. It implies the active participation of at least two communicators, who alternate their activities and continuously adjust their input accordingly by turn-taking, therefore illustrating an important feature of language as communication by continuous interactions.

C. Communication implies sharing meaning through signs. Language is represented in dialogues as a system of symbols which as signs have an abstract dimension and a material dimension. It can only function as communication system through its externalization in material signs by human physiology. It is only through its material expression that we can glean knowledge about the language system. In this sense it demonstrates the interdependence of the human body and mind in sharing the private life of the individual.

D. It illustrates both the language system and its use in communication as linguistic meanings are interpreted as utterance meanings. Novel interpretations of standard meanings illustrate the infinite opportunity for creativity in language, and, crucially, illustrates the interdependence of the language system and its use.

E. The dialogue is the quintessential illustration of language proficiency of the typical human speaker.

F. Participation in dialogues is the behavioural frame for first and second language learning and language change.

G. Languages with long literary traditions as well as pre-literate languages which by definition exist only in spoken form, e.g. Piraha (Everett, 2005), Riau Indonesian (Gil 2009), Kalam, a language spoken in Papua New Guinea (Pawley in Givon, Shabatani, 2009) are represented by the universal features of dialogues. The common denominator among dialogues in all

languages is preference for grammatical categories representative of human experiences, e.g. classifications of material objects with significance to everyday life, persons and their interactions with objects and one another, processes and their development in time, properties (changing and permanent), etc., preference for simple and concise grammatical forms with clear contribution to meaning representation, significant reliance on the extralinguistic context.

H. The overwhelming majority of dialogues are conducted through the vocal-auditory channel in speech, thus, speech is the universal default channel for externalization of the language system in dialogues. The human organism has the ability to process spoken dialogues with remarkable speed and accuracy. A linguistic item is processed on average as follows: 65 milliseconds (msec) for the processing of a phonological form, 250 (msec) for processing the semantics, 1-2 sec for processing the grammatical properties of a sentence (Givon ,2002, p.74, see also Christiansen, Chater, 2016). The influence of the vocal-auditory channel on the organization of linguistic utterances in dialogues is as follows:

a. The demand for efficiency in articulation explains the fact that shorter forms are used with high frequency in all languages (Zipf 1949). Thus, the physiology of speech influences the choice of linguistic forms in an utterance.

b. The linguistic units are packaged in intonation contour. A sentence is organized to fit in a single prosodic contour. The boundary between a main clause and a compliment clause is marked by a pause.

c. Extensive use of intonation assumes some grammatical functions, e.g. the formation of questions without the use of question words and/or other forms of grammatical machinery, e.g. subject-auxiliary inversion in English.

d. Speakers regulate their communicative contribution to fit a prosodic contour.

“...spontaneous speech is typically not produced in a continuous stream. Speakers regulate the flow of information such that, in essence , they introduce just one new idea at a time per intonation unit or prosodic phrase. This new idea might be introduction of a new participant, action, time , place, or other new or significant item of information”(Mithun, in Givon, Shabatani, 2009. p. 67)

e. The high speed of processing of speech exerts influence on the cognitive aspects of spoken dialogues.

“an intonation unit can express no more than one new idea. In other words thought, or at least, language, proceeds in terms of one such activation at a time, and each activation applies to a single referent , event, state, but not to more than one” (Mithun, citing Chafe 1994, in Givon , Shibatani 2009, p.67).

Sign languages are alternative avenue for materialization of linguistic signs in a dialogue by utilizing the whole body as an instrument of linguistic communication. Gestures, like speech sounds, are ephemeral signs which makes them naturally suited for face-to-face communicative interactions, parallel to spoken dialogues.

f. Communication by spoken dialogues is complemented by non-linguistic body signals, e.g. pointing, body movements as a vital component of human multimodal communicative ecology.

To sum up, the dialogue illustrates the quintessential properties of language as a system of sharing experiences in a systematic and structured way and demonstrates the essence of

humans as cognitively advanced and inherently social and cooperative species. Importantly, engaging in dialogues is irresistible as humans engage in dialogues even with strangers, e.g. by using pidgins, and children spontaneously engage in dialogues with bare minimum of knowledge of language.

3. The dialogue in linguistic theorizing

A number of linguistic paradigms incorporate the dialogue in designing theories of language. The Relevance theory (Sperber, Wilson 2004 and elsewhere; Scott-Phillips, 2017) states that expectation of relevance is a property of human cognition to anticipate a potential contribution of any input to cognitive processes, be it perceptual, cognitive or communicative, applied intuitively in decision making by evaluation of available information in context to reach a positive cognitive effect, e.g. solving a problem, correcting a mistake, uncovering new knowledge, etc. Here relevance is understood as a matter of degree as some input may be more relevant than others and by zeroing on the most relevant input one reaches the best cognitive effect. The application of the principles of relevance to linguistic communication are in focus here. In the Relevance theory every utterance implies its own relevance which makes it worthy of processing effort as it has the highest degree of relevance among competing alternatives. Thus, the very fact of communicator's producing an utterance implies his/her desire to be understood. The information encoded in a sentence is the input which, under the automatic assumption of relevance, receives the interpretation determined to represent the sender's intentions most accurately. Thus, any instance of language use incorporates code and inference. The sender encodes the information in a construction of a sentence and conveys his/her intentions by suggesting its best interpretation in the given context. The receiver, under the automatic assumption of relevance, entertains hypotheses about the intended meaning given the current communicative context. The theory aims to identify the processes and principles by which the gap between the linguistic meaning, encoded in a sentence, and the utterance meaning intended by the sender and interpreted by the receiver, is bridged. In this sense the meaning encoded in a sentence is just a clue, guiding the receiver towards the correct interpretation of the message intended by the sender. The same sentence may have infinitely many utterance interpretations depending on the particular circumstances of the communicative act and its participants. The Relevance Theory applied to linguistic communication incorporates some of Grice's vision of intuitive expectation of relevance in utterance production and interpretation, while rejecting others, e.g. the need for Cooperative maxims and the notion of Maxims violations in the interpretation of metaphors and other figures of speech as deviations from literal interpretation.

Spoken dialogues are constructed by universal principles of cooperation in communication outlined in the theory of speech acts (Austin, 1975). It seeks to understand the universal principles of language use/ performance as the interface of code and context and the role of the human interpreter in the production and interpretation of the message. It distinguishes between sentence and utterance, i.e. the linguistic code and its use in individual acts of communication. The concept of "conversational implicature" is introduced in recognition that the message cannot be reduced to the code or what is explicitly said. Pragmatics is quite a heterogeneous branch of linguistics and includes a broad range of topics of research including

the formalization of referential aspects of grammatical forms .e.g. definiteness, deixis etc. as well as the use of language as verbalized action (Austin 1975.)

Conversation among at least two participants is the universal frame of linguistic communication as a chain of utterances , each organized around information structure based on the opposition new vs. old information or Topic and Focus. The internal organization of a conversation incorporates another layer of structure, organized around the rules and principles of conversation. Paul Grice (1989) has articulated the foundational principles of conversation as a joint activity and states that all participants voluntarily make cooperative contributions to the conversation by inferring each other's intentions and responding to these linguistically. These are :

- a.** Maxim of quantity, i.e. the information volunteered by the communicator is determined by the needs of the conversation, not less or more.
- b.** Maxim of quality, i.e. the assumption that the information given is truthful and not deceiving.
- c.** Maxim of relevance , i.e. the participation of all communicators must be relevant to the topic discussed.
- d.** Maxim of manner, i.e. communicators are bound by the demand of clarity to avoid ambiguity.

The theoretical perspective discussed here defines the general patterns of conversation and the cooperative contribution of participants, highlighting the fact that human speakers are inherently social beings.

And although language use is mainly the focus of the usage-based approaches, in generative context a less known version of generativism, the Parallel Architecture (Jackendoff, 2002 and Culicover and Jackendoff 2005) marks a significant departure from traditional generativism in acknowledging a role of performance as part of knowledge of language .The language system is hypothesized as a complex structure composed of various differently organized subsystems, each composed of interconnected and interacting tiers: phonological, semantic and syntactic, each independently organized in accordance with the characteristics of their basic units, connected to one another by interface rules, which map the components of each pair of structures and constrain the possible outcomes by licensing the well formed ones. The paradigm envisions that a strict one-to-one mapping between the three structures is the exception rather than the rule and shows that it is a correspondence of one-to-many, leading to homonymy, synonymy, etc. as explanation for the richness and diversity of linguistic communication. Significantly, the Parallel Architecture paradigm also incorporates formalization of performance.

“ The competence grammar encodes the knowledge involved in the correspondences between phonology, syntax and semantics, and it is the establishment of these correspondences in real time that constitutes the computations that speakers and hearers perform in the course of using language”.
(Culicover, Jackendoff, 2005 p.163).

The Parallel Architecture incorporates fundamental theoretical innovations to the generative tradition as the model allows constant interaction between competence and performance at every level of the architecture of language. The sharp division of lexicon vs. grammar, as a distinction between learned and innate, or core/periphery and irregular vs. regular grammatical

forms is avoided here and is shown to be a matter of degree, as numerous grammatical forms and structures are viewed as derived through violations of grammatical rules and thus are mastered through learning.

4. The dialogue as a language universal

Ginsburg's pioneering work views the dialogue as a rule governed activity and a unit of the language system. He aims to outline a grammar of conversation in which the language system is only one of various other components. Importantly, at the centre of his search for a formal paradigm of the dialogue i.e. language in action, is meaning (Ginsburg, J. 2008 and elsewhere). The conversation as a chain of utterances united by a common topic and context, as a unit of language, is modelled. A conversation has a distinct organization, although it is a flexible unit of interaction with a distinct meaning and structure and its own “communication management” (Ginsburg, Poesio, 2016, p. 14). A common set of assumptions shared by the participants, i.e. initial common ground at the start undergoes alteration during the conversation as with each move a new repository of facts is added to the initial common ground. The initial topic of the conversation can change in the process as some communicative actions are clarifications of earlier statements and others are not directly related to immediately preceding moves.

ex. A. Several people showed up today.

B. Who? (Ginsburg, J. 2001)

Significantly, the extralinguistic context determines the range of topics and the semantic and grammatical organization of the dialogue, e.g. a conversation in a bakery prompts the topic of bread and a conversation in a courtroom prompts a discussion on witnesses, guilt pleas, etc., introducing the concept of *genre specificity*.

The meaning of an utterance in a dialogue is represented by a variety of linguistic forms, from complete sentences to non-sentential utterances. The majority of utterances are incomplete sentences, i.e. lacking overt predicate, which have meaning only in a specific context.

ex. That one.

In addition non-linguistic signals, e.g. a head shake and nod, laughter, eye rolling, etc. accompany and complete the linguistic message. Silence, i.e., the absence of explicit reaction, often carries a meaning.

A dialogue is described as a type of game, i.e. alternation of move and counter move, where every move triggers a countermove.

ex. A. We have a flat.

B. Uh. I see. (pause) Nice? (Pause). A flat. It's quite established, than, your..uh.... affair?

A. Yes.

B. How long?

A. Some time.

B. But how long exactly?

A. Five years.

B. Five years? (Ginsburg J. 2001)

A dialogue usually begins with a greeting word or a phrase.

e.x. Hi. Hello.

Some languages have also phrases used exclusively as a response to greeting.

ex. A. God give you health. B. God helthify you. (Ginsburg, Poesio, 2016). See also Enfield, Sidnell 2015 on the properties of the dialogue.

Thus, a conversation has a universal structure which is at the same time flexible and adaptable. In this context the theoretical distinction of competence and performance, on which current linguistic thought is based, appears unjustified and even detrimental for describing and theorizing of language, which is better understood in terms of grammar of interaction.

4.1. A human instinct for interaction: the interaction engine hypothesis

The argument for a language instinct as innate property of human mind to learn and process Universal Grammar is famously made by Pinker (1994). Alternatively the notion of instinct relevant to language is reinterpreted by Levinson (2019) who argues that humans are endowed not with an instinct for grammar, but with an instinct to communicate or instinct for interaction, a property of the human mind underlying all types of communicative interactions, including but not limited to linguistic interactions, a view outlined in the interaction engine hypothesis. It is demonstrated especially when language is non-existent or highly restricted, e.g. newborns communicate long before becoming proficient in a language, deaf children instinctively invent home signs, speakers of mutually unintelligible languages manage to coordinate goals without use of language.

The here postulated human interactive instinct is most clearly evident in the universal structure of dialogues, e.g. participants, often more than two face one another, and offer alternating contributions by turn-taking, concurrent with facial, gestural and body movements. Each participation is of short duration of ~2 sec and at each turn communicators try to guess and anticipate the current speaker's message and plan a response. Intervals between responses is estimated at ~200 ms placing significant pressure for fast and accurate comprehension and planning. Misunderstandings and clarifications are mediated by the repair system, assuring the maintenance of participants' intentions and priorities within the conversation, organized in pairs of specific sequence, i.e. initiating communicative action and a response, a basic frame which can become enriched with details and extensions. This universal structure is universally found in all languages, spoken and signed across cultural, ethnic, educational etc. environments, suggesting some innate support. In this context it is suggested that it is language which has adapted to fit the perimeters of the human instinct for interaction.

5. Dialogues as coordinated multimodal communication of language and non -linguistic body signals

Both spoken and signed dialogues use the human body for communication, although differently. Sign languages are multimodal by nature and spoken dialogues include elements of multimodality as the linguistic message is coordinated and complemented with non-linguistic signals, e.g. non-linguistics vocalizations, face expressions, body posture, eye movements etc. Thus, co-linguistic body signals are universal aspect of human communication, although there no agreement if non-linguistic body signals, accompanying linguistic messages are to be considered part of the language system or contributing extralinguistic layers of communication. For some authors co-speech body movements form integral part of conversations by aiding the

processing of the linguistic message, both in production and comprehension, a universal phenomenon formalized in the Information Packaging Hypothesis by Goldin-Meadow and Alibali (2013). Similarly, in Holler Levinson (2019) linguistic utterances are nestled within the “natural ecology” of multimodal communicative signals demonstrated in face-to-face interactions. It is made possible by “binding” of multiple articulators and modalities, which produce multimodal signals, perceived and processed by the brain as a package. Some scholars have proposed that the definition of constructions, i.e. indivisible associations of linguistic form and meaning, be extended and enriched by the recurrent kinetic multimodal contributions of non-linguistic signals, enriching experience with language. Thus, constructions as the building blocks of language be redefined as multimodal communicative units containing linguistic and non-linguistic elements. (Zima, Bergs, 2017)

More complex is the relation of co-linguistic body signals and linguistic signs in sign languages, currently a matter of debate. Kendon (2000 and elsewhere) argues for a continuum of linguistic sign and co-linguistic body signals accompanying signed utterances, known as “Kendon continuum”. Alternatively, McNeil finds clear discontinuity and focusses on the non-linguistic properties of body signals. For comparison see Muller (2018). The arguments for discontinuity are justified by the following major differences between body signals and sign language systems.

a. Most co-linguistic body signals are spontaneous, involuntary body reactions. Some are universal, e.g. a lie is accompanied by subconscious alterations in facial expression and voice modulation. Others are unique to each individual. This is why actors must learn the body reactions of the characters they portray as a signature of their individualities. It is reasonable to attribute a prominent role of nature, not nurture, as speakers perform gesticulations while speaking on the phone and blind speakers perform co-speech gesticulations while talking to blind speakers (Cartmill et al. 2012) which one could interpret as a confirmation for an instinct for interaction, referred to above. Emblems are culturally motivated, and thus, learned, but equally automatic body signals e.g. hand shake as greeting, head shake in some cultures is a sign of disagreement and head nod a sign of agreement. They are conventionalized and can be understood as equivalent to non-linguistic kinetic contributions to spoken and signed utterances as part of multimodal constructions (as in Zima, Bergs, 2017). In contrast, gestures in sign language are intentional, learned and conventionalized and demonstrate social and cultural diversity, suggesting a prominent role of nurture.

b. Body signals like non-linguistic vocalizations are mostly expressions of emotions, e.g. one subconsciously smiles when talking about pleasant things, an expression of disgust accompanies a conversation about unpleasant things, although emblems may convey propositional meaning in a holistic signal. In contrast sign languages express mostly propositional thought.

c. Given their instinctive nature, gesticulations are unchangeable vs. sign languages are susceptible to historical change.

d. Gesticulations are iconic and indexical signs, although as per Levinson, Holler, (2014) finger pointing has a symbolic component, while sign languages are predominantly symbolic systems.

e. Co-linguistic body signals are holistic signs, i.e pairs of holistic forms and holistic meanings while sign language is combinatorial, hierarchically organized system. Consequently the meaning of gesticulations is vague and ambiguous, which is why they are produced in coordination with linguistic signals, while the message in sign language is precise.

f. Co-linguistic body signals have a complementary role in communication, although emblems can be used on their own e.g. when speech is either inconvenient, as when interrupted by loud noise, or in silent films. Signs in sign languages are always components of a system.

Overall co-linguistic body signals and sign language differ in : **a.** bio-cognitive aspects, **b.** in semiotic aspect, **c.** structurally, **d.** functionally. In sum, the dialogue is a composite form of communicative interaction where linguistic signs are used within a communicative ecology in combination with non-linguistic signs.

6. The evolution of language as evolution of languaging by dialogue

For most linguists, especially the proponents of the generative formalism, still highly influential in modern linguistics, language is viewed as a code with the primary function of informing, mainly exemplified in monologues, while the interactive function is viewed as distant second.

That said, the linguistic properties of the linguistic code and its application in monologues differ significantly from these of dialogues both in meaning and structure, as illustrated above. Crucially, all humans of every race, ethnicity, profession, have been engaging in communicative interactions by participating in spontaneous dialogues as part of daily life since the beginning of humanity, while public speaking by monologues with the explicit intention to inform, e.g. speeches, lectures, weather reports, etc. are reserved for rare occasions. And given that participating in dialogues is presumed by some pre-human species, e.g. Neanderthals, and by pre-linguistic infants, it is safe to argue that languaging by dialogue is the original function of language, very likely associated with hunter-gatherer lifestyle, while acts of public speaking by monologues are derivative from the first and signalling a later stage of evolution of humanity with the development of settled life and complex culture, and proliferated by the emergence of writing. This furnishes the argument that defining natural language in terms of dialogues and its evolution from pre-human pre-linguistic communication must be the goal of inquiries into its evolution.

To note, although dissemination of information by code in monologues is deliberate and explicit, through participating in dialogue information is shared inadvertently by divulging one's state of mind, state of knowledge and worldviews, age, ethnicity, gender, education, first language, etc.

6.1. The evolutionary roots of multimodal communication

The communication of various primate species resembles to various degrees human dialogues. Some turn-taking abilities in vocalizing , i.e. ability to participate in dialogues are detected in monkeys and apes. In great apes, a comparison most relevant to the current topic, turn-taking is evidenced in the gestural modality, similar in speed and flexibility to human vocal negotiations during verbal dialogues(Levinson 2019). Bats have dialects as markers of group identity and

their communicative interactions resemble human dialogues in turn-taking. (Vernes, 2017). Great apes and even monkeys anticipate calls to have a communicative function, suggesting a primitive form of Gracian principles of signal interpretation. They demonstrate awareness of social structure as they tailor their calls to reflect the social rankings of the receivers. Apes demonstrate ability to interpret calls in context-dependent way.

Multimodality is at the heart of primate signalling. The communication of primates, both in natural habitats and in captivity, is a composite signal, i.e., a combination of various components which act in concert to deliver the complete message (Slocombe, Waller and Liebal 2011). Chimps in captivity communicate with signals emitted simultaneously from a variety of modalities. They recognize and produce visual, tactile, vocal signs in their communication with caregivers. Chimps are also known to combine alarm calls with gaze and are shown to have voluntary control over manual gesticulations as well as control over individual vocal organs, e.g. control of breathing, larynx, supra-laryngeal articulators and capacities for vocal learning (Perlman 2017). Importantly, although chimps have displayed a large repertoire of innate gesticulations, they have demonstrated ability to communicate by learned gestures even in the wild, which suggests, among other things, a capacity for social learning(Pika, 2014). Macaque's communication is cross-modal as well, suggesting that “communicative complexity may be less about how each single modality is used, and more about signal integration.” (Slocombe Waller, Liebal, 2011, p. 540) .

Given these findings the unimodal approach to primate communication is inadequate for understanding a complex and integrated system as studying each component in isolation ignores the influence of the rest. For example, “ When a slap is paired with a playface it leads to play. Isolating the slap from the playface will not, therefore, help us understand the signal better.”(Slocombe, Waller, Liebal 2011, p. 540). See also Boe, et all. 2017)

Communication in general is multimodal and a multimodal approach is to be preferred to isolationist one as the function of a single modality may not be clear without considering its use as a component of a multimodal communicative ecology. And given that language as part of the human multimodal communicative complex, human linguistic communication can be better understood as emergent in the context of a multimodal primate communication. On the other hand, inadequate understanding of primate communication is clearly an obstacle for the study of language evolution. Moreover, by understanding the communicative functions of the multimodal complex as a whole, we can understand the phylogenesis of the mechanisms which make this multifaceted behaviour possible.

6.2. Evolution of languaging by spoken dialogues as restructuring of pre-human multimodal communication

From the early starting point of multimodality some form of speech is likely to have gradually evolved as the major component in the multimodal complex , while relegating to body posture, non-linguistic vocalizations, a contributing role. The leading role of speech is explicable from both physiological and cognitive perspective, as the most energy efficient, as well as from communicative perspective as allowing for human thoughts, encased in a symbolic and combinatorial system, to be materialized effectively with speed and precision. In this context one could anticipate pre-human multimodal communication systems as a mixture of holistic and combinatorial signs. The fact that modern languages preserve some holistic structural

elements suggests that there never was a complete displacement of one system by another, but a gradual transformation where some elements assumed a leading role while others remained as a contributor. The fact that modern languages preserve some indexical, iconic and symbolic signs, inherited from an older, more primitive and less effective multimodal stage, suggests a gradual shift giving priority to symbols.

That said, although speech is the predominant modality, multimodal elements of linguistic communication have been preserved in the language systems of today, which are universally made explicit in spontaneous spoken dialogues, embedded as a component of the multimodal communicative complex humans use to communicate. In fact, today hunter-gatherers of Central African Congo Basin use a mimetic language where spoken message is mixed with singing, whistles, imitation of animal vocalizations, forest sounds, reenactments of events, dancing, etc. thus, multimodality occupies a role equal, if not larger than spoken language (Lewis, 2014). Thus, spoken language can only be understood as part, although very significant part, of multimodal human communication. In this context the evolution of spoken language is better understood in the broader context of human multimodal communication.

And given the adaptive advantages of effective communication, it is reasonable to suspect that these structural alterations in the language system by prioritizing discreteness and combinatoriality, have triggered internal restructuring in the multimodal participation of the human body by coordinated behavioural adjustments among the relevant aspects of physiology ending up in the dominance of speech. Most importantly, restructuring in human communicative behaviour would have triggered adaptations in the human organism to the new kind of cognitive and physiological demands of language processing and learning, culminating in the evolution of a Language Faculty. This perspective suggests a new angle to the continuity debate in language evolution.

On a different but related topic, a contribution to understanding the bio-cognitive roots of the advanced manual dexterity and language skills offered by Stout et al. (2008) in a study of the neuroanatomy of Early Stone Age tool making and language, suggests an evolutionary path towards increased coordination of perception and action towards *immediate goals* by interaction of both hemispheres. These findings underscore the significance of coordinated evolution of perception and action in “ongoing hierarchically structured action sequence” (ibid. p.1946) and challenge the established view of language processing in terms of abstract thought involved in conceptualization and planning performed by left hemisphere. In my interpretation the study contributes to our understanding of neuroanatomical aspects of languaging via participation in dialogues underscores the combined participation of both hemispheres in coordination of perceptual and structured motor activity in attaining short term communicative goals, e.g. of speakers during a dialogue.

7. The evolution of the human body and mind in adaptation for languaging by dialogue

7.1. Speculations on a language faculty and its evolutionary history

The nature of the language faculty is a matter of debates with two major alternative views. In biolinguistic circles the standard view of the representation of linguistic knowledge and processing advocates for spatial and functional segregation of language in the brain in Broca's

area, identified as the grammar organ, or the location of syntactic computations (Chomsky, 1968 and elsewhere ; Pinker 1994; Bickerton 2014). The segregationist argument was based on the argument for modularity of human cognition (J. Fodor 1983). Alternatively the brain is understood as a flexible multipurpose processor, where linguistic functions are highly distributed and coordinated by experience (MacWinney, 1998, Deacon 1997, Ph.Liebermann 2000)

At the same time recent empirical studies conclude that the localization of language in the brain is difficult to pinpoint given that a large portion of the brain is involved in language-relevant functions, including subcortical regions such as striatum, cerebellum, thalamus, among various others (Fisher, Marcus, 2006). Moreover, the young developing brain is flexible and able to compensate for damaged abilities, linguistic and other wise, as demonstrated by recoveries from injuries, which makes attempts to isolate language areas even more challenging. The finding that there is no one-to-one correspondence between syntactic phenomena and brain functions adds to the difficulty, e.g. the brain does not differentiate between pronouns and reflexives (Kaan, 2009). The language faculty is said to be “ the most invasive”,(Gulyas, 2009, p.59) i.e.the most widely distributed cognitive faculty in the human brain (Ph. Liebermann, 2000; Bickerton, Szathmari, 2009). Moreover, it is suggested that the widely extended and diversified interconnectivity of the human brain, compared to that of chimpanzees and macaques, makes possible the integration of various types of information, i.e. phonological, semantic, lexical, grammatical, stylistic, in the verbalization of thought (Sherwood et all. 2008). That said the established view that linguistic functions in most normal individuals are asymmetrically concentrated in the left hemisphere is still valid (Fedor, et all. 2009). In addition, it has been argued that ,as a general tendency, domain specific mechanisms arise during ontogeny under the influence of experience, forming domain-specific configurations of neuronal connectivity(Sherwood et all.2008). And experience with language during language learning happens by participation in dialogues. At the same time, although scholars do not question the prominence of the Broca's and Weernicke's regions in linguistic functions, as per Fedor et all. (2009) “It seems more correct to say that some areas of the normally developing human brain are more prone (in quantitative sense) to host and process different components of language than others “ (Fedor et all 2009. p. 300). In the same vain Friedericcci (2017) made abundantly clear that the human brain is a unified whole of functionally integrated anatomically diverse neuronal assemblages, where perception, emotion, reasoning, socialization, imagination etc. are all interconnected. In this sense the representation of language in the brain is well beyond the physical location and cognitive specialization of a LF and integrates multiple processing mechanisms in both hemispheres. Thus, the brain's processing of language engages the brain as an integrated whole to which the author refers as “ language network”.

Its properties in the adult brain begin their formation early in childhood by initially engaging primarily the right hemisphere, most actively involved in perception of prosody and learning basic vocabulary, to gradually integrate the right hemisphere and Broca's region in processing of complex grammar, only after a decade of experience with language as it reaches adult proficiency. The early development of language is made possible by a number of developmental properties, uniquely tailored to facilitate language learning, e.g. developmental instinct to babble as training of the articulatory apparatus for speech by mimicking the speech

production of adults , a critical period for language learning, where proficiency in a first language is gradually achieved by witnessing and/or participating in dialogues (Hurford 1991; Hurford, 2012; Hurford and Kirby 1999) and developmental instinct for word learning, triggered by witnessing the referential use of words in communicative interactions (Bloom, 2000). So, experience with language begins by experimenting with participation in conversations, illustrating the crucial formative role of engaging in dialogue in the formation of the LF.

Further, Liebermann(2008) demonstrates that the human natural potential for language is a web of interrelated and spatially distributed groups of neurons where the same biological entities participate in various neural circuits and support multiple related behaviours.

“ ...a particular neural structure may play a part in different neural circuits that regulate different aspects of behaviour...(and at the same time) distinct neural populations may occur that project to neurons in different brain structures contributing to circuits that each regulate a different aspect of behaviour. “ For example, the basal ganglia was found to “ switch from one sequence of motor pattern generators to another” (Lieberman 2008, p.218- 219)

Liebermann also argues that the basal ganglia participates in regulating coordinated movements in speech, dancing, walking and other behaviours, unique to humans, an indication that these have evolved for a purpose broader than language alone. Moreover, the basal ganglia, which regulates the vocal tract's ability to form complex vocal gestures, is possibly the same ability which regulates also aspects of human cognition and syntax, which involves the recombination of primitives and formation of structures. Thus the basal ganglia is “ the sequencing machine “ of the brain, which participates in different neural circuits , regulating different behaviours(Liebermann 2008 and elsewhere).

In short, the natural potential for language, i.e. the language faculty, is a web of neuronal connections with shared functions which have co-evolved in tandem.

To remind, the linguistic computations in UG are designed to rely on the existence of the lexicon as a storage of detailed descriptions of real linguistic experiences, which makes it an essential component of the computation procedure. If this is so, the argument for spatial and structural isolation of the language faculty cannot stand as the memory storage does incorporate sensory representations reflecting material reality.

Moreover, experimental studies confirm words are stored in memory as rich descriptions of individual examples of perceptual experiences , including phonetic properties of words, where linguistic properties are combined with extralinguistic details(Port 2007). The ingredients of such representations are very concrete, e.g. the idiosyncrasies of the speaker's voice and other sensory representations. The vocal representation of a word in memory is articulation - based , not abstract feature -based. That is, words are stored in memory in the form of specific events with idiosyncratic details, not as structured combinations of abstract prototypical categories. This process is subconscious and universal. Similarly, Pulvermuhler (2005, 2013, 2018) demonstrates that linguistic forms, with both concrete and abstract meanings are stored in conjunction with perceptual experiences by witnessing and/or participating in conversations. One would expect that this innate feature would be used for grounding linguistic symbols in extralinguistic activities as potential referents.

Consistent with these empirical findings is Deacon's argument (1998) for an innate human

capacity for symbol formation and symbolic reference, a complex combination of various types of referential relationships, e.g., among symbols as members of a symbolic system, between a symbol and its referent, and among objects in reality as perceived by the human mind. And given that the very existence of symbols implies communication, i.e. interactive exchange of information between at least two individuals, interaction with reality by engaging in conversations is inherent to the very functioning of the language algorithm.

This integration of abstractions and concrete experiences in the representation of language is reflected in the organization of the brain at the neuronal level. Neuronal networks processing linguistic symbols are associated and interact with networks processing symbolic referents, i.e. the symbol grounding in aspects of extralinguistic reality. In this context the semantic features of linguistic forms as prototypical examples of linguistic symbols are “only suggestions in need of further elaboration and modification” (Pulvermüller, 2018, segment 6).

The integration of language system and language function in the brain is also found at genetic level where an integrated capacity for speech and structure is a phenotypic expression of various genes (Dediu, Levinson, 2018) most prominently FOXP2, involved in the formation of the human basal ganglia and responsible for rule-governed complex structured behaviours in cognition and praxis, e.g. ability to dance, produce and manipulate tools, use language, demonstrating the innate link between language use and extralinguistic activities. (Liebermann, Ph. 2016 and elsewhere)

But perhaps most relevant to the current topic is the argument for human hyper-socialization (Donald, 1993) and a capacity for anticipation of communicative intention and communicative relevance as ability to participate in dialogue as part of a Theory of Mind (Sperber, Wilson 2004 and elsewhere ; Scott-Philips, 2017) In sum, the interaction of code and context during participation in dialogues is reflected in the mind of the language user. To sum up, the concept of a language faculty must include capacities for coordinated languaging by participation in dialogues.

And although not explicitly underscored, the conceptualization of language as a communicative technology implies language use in conversations. In this sense linguistic theorizing would be enriched by universal principles of pragmatics as an integral part of linguistic theory in reflection of the nature of natural language as integration of code and inference, a topic currently underestimated by most linguists.

From evolutionary perspective it is argued that during human speciation various aspects of the human body and mind have been altered to various degrees as adaptation to language learning and processing within the frame of a dialogue. In fact, some aspects of the human organism are especially well attuned for processing of interactive multimodal behaviours.

7.2. Broca's region, a processing hub for multimodal behaviours

Broca's area is standardly stipulated as the locus for processing of abstract structures and the left hemisphere is stipulated to be the locus for integrating semantic, grammatical information, instrumental in the formation of individual words and their further integration into larger structures, phrases and sentences. That said, Broca's area as part of the Inferior Prefrontal Gyrus (IPG) in humans is also known to integrate processing of vocal and gestural signs. Moreover, speech, sign language and co-speech body signals are processed by the human brain

in a coordinated way (Frohlich, M. et al. 2019). Thus, the brain processes multimodal human communication, linguistic and extralinguistic, as a unified complex, demonstrated by engagement in dialogues. In spontaneous spoken dialogues linguistic signals are co-articulated with multimodal body signals.

From a slightly different but related perspective, it is commonly stated (mistakenly) that all and any thought can be communicated through language. On the other hand, Dor and Jablonka, (2000) point out that only selected fractions of the human semiosis is communicable through language, i.e., language is a system designed to communicate a broad although limited range of concepts. Certain scenes, events and their participants are universally easier to express in linguistic forms than others. As per Dor and Jablonka “ Natural language is a communication tool, which is structurally designed for the communication of a constrained set of meanings.” (ibid. p.41). The universal patterns of categorization are universally reflected in the grammatical systems of all languages, e.g. the distinction of animate vs. inanimate, female vs. male, near vs. distant, human vs. nonhuman, etc.

“ ...semantic categories which are reflected by grammatical complexities in natural languages belong to a very constrained subset of all the categories we use to think, feel and conceptualize the world ” and “...some semantic categories turn out to be grammatically marked in language after language, whereas some others consistently do not participate in the grammatical game ” (Dor, Jablonka, ibid. p.39).

In this context the participation of non-linguistic signs in face-to-face conversations is explicable as a compensation in the externalization of meanings for which language is not designed. Moreover, in modern societies where long distance communication among individuals separated by space, time and communicative context is prevalent, pockets of the human conceptual universe are found to be difficult or impossible to express in linguistic forms, e.g. emotions, spatial orientation, some scientific concepts, etc. As compensation for these deficiencies civilization has invented additional, non-linguistic technologies e.g. maps, graphs, formulas, drawings, etc..

8. Evolutionary roots of human multimodal behaviour

The evolutionary roots of human multimodal communicative behaviour are found by comparing brain architecture and functions in primate and human brains. Multimodality in non-human communication is made possible by brain functions at neuronal level as various neuronal populations at various brain regions are engaged in processing signals from multiple modalities simultaneously (Slocombe et al. 2011). Thus non-human communication is multimodal both in terms of behaviour and brain functions. Significantly, the area F5, Broca's homologue in the primate brain, has very similar functions as integrating centre for processing of various communicative signals. A spacial and functional overlap is found between neurons processing vocal, visual and gestural signals (Frohlich, M. Et al. 2019). It also has similar functions of integration of perception and motor functions essential in observation, imitation, planning, in macaques and humans (C. Sherwood et al. 2008) Consequently, deficits and/or damages would impair a number of functions. Thus, Broca's region is engaged in processing multimodal communication in both human and pre-human species and continuity in brain functions of Broca's region and its pre-human homologue suggest evolutionary continuity in processing multimodal communication. But most importantly for the purposes of the present article, the evolution of language has exploited Broca's diverse functionality in the processing

of dialogues.

8.1. The mirror neurons and multimodality of language origins

To remind, the mirror neurons, first discovered by Rizzolatti and Arbib(1998), become excited both when an individual performs a manual action and when he/she observes the same action, thus, providing link between cognition and communication. The discovery of mirror neurons in monkeys has been interpreted as suggesting that mirror neurons as an evolutionary innovation first evolved in monkey species, where the behavioural representation of mirror neurons is argued to be demonstrated by copying of manual gesticulations. That said, a fairly recent and less popular study finds behaviour suggestive of mirror neurons in avian species, e.g. neurons in the forebrain of sparrows, become activated both by performance and perception of bird songs, involved in vocal learning(R.Garcia et al. 2014), suggesting a much older evolutionary history.

In humans mirror neurons occupy a small area within Broca's region in the left hemisphere. The mirror neurons are attributed a crucial role in the phylogenesis of the Language Faculty in terms of transformation from automatic and subconscious copying of observed praxis routines into compositional expression of meaning and structure as “ action grammar”, intentionally emitted and directed to a receiver (Arbib 2004 and elsewhere). In addition, the fact that the mirror neurons in non-human and human species share similarity of function as a link between the external reality and its representation in cognition and communication and the fact that language is embedded in multimodal communicative behaviour suggests that the evolution of speech and its current dominance in externalization of language by dialogues is better explicable with continuity with multimodal communication rather than an event of switch from gesture to sound, which is currently the dominant understanding, espoused by Corballis,(2009 and elsewhere), Tomasello (2008), Arbib(2004, 2007 and elsewhere).

8. 2. The evolution of human theory of mind and its use in dialogues

The ability to correctly detect and anticipate the thoughts, desires, aspirations, of fellow humans, is known as social intelligence, or “ theory of mind”, an innate cognitive ability. The term *theory of mind* was coined by M Donald (1993, 1999) to mean a specialized cognitive capacity of every normal human being which allows him/her to instinctively detect the mental states (beliefs, intentions, etc.) of fellow conspecifics and use this knowledge in communicative interactions . This capacity is demonstrated behaviourally in : **1.** intentional communication, **2.** ability to repair failed communication, **3.** ability to teach , **4.**to persuade , **5.** ability for intentional deception , **6.** work on shared plans and goals, **7.** share focus of attention, **8.** pretend.

Having a theory of mind is vital to the ability to communicate, including linguistically. Communication usually means exchange of information between a speaker and a listener. Language can be used for communication only if each participant has understanding of the state of mind of the others. The very act of communication is the attempt of the speaker to intentionally influence the state of mind of the listener. An impaired theory of mind results in impaired language learning and use. Autism is a developmental disorder characterized by lack theory of mind. Autistic individuals often have knowledge of semantics, syntax and phonology,

but lack understanding of speaker's intentions, which prevents them from understanding non-literal meaning. This makes them less than fully competent speakers.

Some form of theory of mind is a necessary precondition for language learning in ontogeny. And indeed human infants have demonstrated some elements of a theory of mind, e.g. ability for imitation, joint attention, some form of ability to guess the intentions of the other. Given that, one is prompted to suspect that a primitive version of the modern theory of mind must have preceded the evolution of language. The mind-reading abilities of modern primates may give a clue of earlier stages of modern human theory of mind which include ability to follow gaze, ability to guess intentions of others, etc.

As the development of theory of mind is so crucial to the communicative function of language, the question of the evolution of theory of mind as a cognitive capacity may shed light on the evolution of language. It is clear that the two are interconnected, both developmentally and phylogenetically, although there is no agreement among experts on the matter which one of the two was the cause and which the consequence. A plausible argument for a co-evolutionary scenario of a language faculty and a theory of mind envisions a co-evolutionary scenario (Givon and Malle, 2002) where components of each mutually encouraged each other's further evolution.

“A plausible model of co-evolution will have to include the specification of elementary components of both language and theory of mind, because it is likely that those components influenced one another in a gradual escalation.” (Givon, Malle, *ibid.* p.272).

In a multistage scenario “one of the faculties emerged first in a primitive form and facilitated a primitive form of the other, and soon the two plunged into a race in which advances in one faculty repeatedly enabled or demanded advances of the other. What we have here is a version of the arms-race argument.” (in Givon, Malle, *ibid.*, p. 274) In more recent publications the evolution of language is envisioned as evolution from asymmetric persuasion by pantomime, to conversation as “reciprocal persuasion” by evolution of grammar and speech demonstrated in dialogues, identified as paramount in the evolution of complex language (Ferreti, Adornetti (2021).

In sum, language is a highly complex behaviour as it is as much a biological as a social phenomenon. The human individual is equipped with cognitive abilities to learn a code and interpret it in context. Its formation in the adult individual is a complex integration of multiple contributing factors in the formation of the human language user as a biological body and mind and a socially integrated member from conception to adulthood which incorporates innate predispositions, spontaneous learning and social interaction. In this sense linguistic theorizing would be enriched by universal principles of pragmatics, an integral part of linguistic theory in reflection of the nature of natural language as integration of code and inference, stability and flexibility.

From evolutionary perspective, given that the human body and mind have evolved as an integrated unity and given that language-relevant functions are highly integrated in the human brain and mind, it is natural to suspect that the evolution of language is better understood with an integrative approach. It is also natural to suggest that the evolution of language is only explicable as a consequence of participation in dialogues.

8. 3. Early child development, a demonstration of innate abilities for multimodal dialogues

Innate abilities are often demonstrated in early stages of human growth soon after birth(Dehaene-Lambertz, 2017). In the first weeks and months after birth the infant displays affinity for socializing and participating in communicative behaviour by visible body actions, e.g. smiling in response to a smile, imitate facial expressions, turn-taking in non-verbal sound making and playful interactions , facial expression, gesticulations, body posture, directing attention by sound making, looking, pointing. Newborns' brains are argued to register speech and grammar(Dehaene-Lambertz, 2017). Moreover, some propensities for participation in proto-conversation by turn-taking are demonstrated first by 3 month olds, suggesting some innate support for a human instinct for interaction, hypothesized by Levinson(2019).

The youngster's first experimentation with language begins with babbling, followed by one-word utterances where children 's communicative interactions are predominantly by body movements with language having a limited and supporting role which broadens with the next stages of language development with two-word utterances and beyond. As development progresses, although the child's linguistic abilities become more sophisticated, his/her linguistic output continues to be complemented by non-linguistic body signals. During development of syntactic competence, as children often omit arguments or elements of constructions which compromises the overall understanding of the message, the lack of full linguistic competence is mediated by body actions, e.g. pointing, etc. Linguistic behaviour is constantly synchronized with communicative use of other modalities as a component of a multimodal unified system (Morgenstern A.,2014). And although the use of body gesticulations as compensation for linguistic immaturity gradually decreases as the child's linguistic skills develop in full, to paraphrase Morgenstern A.(2014), body signals continue to illustrate, specify, reinforce, modify the meanings of linguistic production (ibid. p. 18). In sum, as early language development is viewed as providing insights into the evolution of language, the multimodal participation in early linguistic behaviour in young children suggests that early language also emerged as a component of a multimodal communicative ecology. Importantly, the argument that ontogeny represents a condensed version of phylogeny is well known in the literature and , although it has its supporters and critics, it provides a possibility to infer information on evolution, permanently lost and non recoverable, from observation of development . In this sense the multimodal nature of the first communicative engagements at early age could shed light at initial stages in language evolution.

9. The nature of communicative interactions by dialogue influence the language systems over space and time

9 .1. Social organization influences the language systems

A language system is predicated on a community where speakers interact by dialogues. This is why the idea of a genius-mutant, a fabulously fertile barer of a grammar gene, who has become the first speaker of language, proposed by Pinker (1994)is not a plausible argument for the evolution of language. The availability of at least two potential communicators, capable of engaging in a dialogue is essential to the origination of language. This is what, among other things, we learn from the case of Genie, who during the time of her isolation, lacking the

opportunity to engage in dialogues, did not invent a language of any kind.

Moreover, the social organization of a community determines the nature of interpersonal contacts of communicators engaging in dialogues and with that influence the language system. This is because language is a system based on shared conventions, i.e., a social agreement, formed through interpersonal contacts formed by participating in dialogues. Scholars have detected correlations between the communal types and language types. Hurford (2012) reviews studies which show that there is a correlation and causal connection between the social structure and linguistic structure. For example, Trudgill, Wray, Grace, Marsh (1967), Keenan and Comrie, (1977) and other renowned experts, quoted by Hurford (2012) show that languages spoken by small isolated communities tend to develop complex morphologies and simple syntax. They also develop less complex deictic systems, given that communicative engagements are usually direct, i.e. face-to-face and communicators rely to a great extent on shared assumptions about the details of a communicative situation, which eliminates the need for deictic markers. The lexicons and grammars of small communities are also less transparent, i.e. they contain more irregular forms, which, despite creating difficulties for young learners, are a tendency, explicable with the fact that social norms, one type of which are linguistic norms, are easily enforceable in a small closed group. In contrast, languages spoken in large communities open to migration, develop simpler morphology and complex syntax, transparency of word formation. Thus these languages must become simplified as they must adapt to the needs of newcomers who are second-language speakers.

Trudgill (2009) argues that modern languages vary in complexity and the explanation for this is language contact or lack of it, which depends on social factors like size of the community (small vs. large), the nature of the social interactions (dense or intense vs. loose) and the degree of contact (extensive vs. isolation). For example pidgins are formed in communities with loose ties, who only occasionally engage in conversations, given that they are speakers of mutually unintelligible languages. On the other hand, small, close-knit communities with dense social ties who interact frequently have languages with excessive use of elaborate linguistic forms.

9. 2. Language between instinct and reason: the “invisible hand” hypothesis

Keller (1994) argues that a stable language system is emergent from the interactions of communicators which engage in dialogues with a specific goal in mind, a goal relevant to the specific participants in a specific situation. In this way its initially local purpose acquires global dimensions not foreseen or intended at the local level. They exist only in a society of rational beings, but rational beings follow social rules blindly, unconsciously, automatically, almost by instinct. So language combines characteristics of both natural and artificial entities, thus, it is located “between instinct and reason” (in Keller, *ibid*, p. 40). Thus, linguistic conventions are unintended global results of intended local interactions by engaging in dialogues. And although sociolects differ as communities converge on different linguistic conventions, the role of the dialogue is universal and indispensable. Similarly, Christiansen and Chater's vision (2016 a.b.) for the evolution of language in the interaction of three interrelated contexts, i.e. on-line processing, learning and historical change, explicable with the limitations of human memory which triggers a “now-or-never bottleneck”, resulting in the formation of the multilevel structure of language, implicitly assume the face-to-face dialogue during which

the bottleneck is created. In this context the dialogue is at the centre in formation of the defining properties of language at every level.

9.3. Simulating the beginning of language in interacting artificial networks

Because direct evidence for the onset of the first sociolects is unavailable scholars resort to indirect sources of information as a substitute. One alternative is the use of artificial intelligence where artificial agents are designed to interact and communicate as they mimic the behaviour of human individuals who to a degree sufficient enough to justify inferences on the very beginnings of language. Dowman (2007, 2008) simulates initial stages of human language where robots are programmed to behave and communicate closely imitating the attitudes of early humans and shows that the interacting agents, in the process of attempting to communicate with each other, came up with a communication system consisting of a mixture of holophrastic forms, single concept forms and various mixtures of the two, forming a continuum. The so produced communicative units also showed different degrees of generality and specificity of meanings, resembling the continuum of lexical words and grammatical forms in a lexicon of constructions. Given that languages of today are also a mixture of linguistic meanings of various degrees of generality and forms of various degrees of combinatoriality along a continuum, I find the results of the experiment sufficiently realistic and informative.

A similar experiment shows, through computer simulation using the game theory, how individual agents through coordinated interactions converge onto a unified symbolic system. Here “individual agents select an action from a set of possible actions, seeking to align their selections with those of other agents”(Barr, 2004). The experiment demonstrates how through repeated coordinated interactions among individual agents common patterns of language use emerge as cultural conventions. Other experiments by Luc Steels (1995) labeled as language games, where artificial participants with characteristics designed to closely resemble some of the characteristics of human beings, interact and communicate with one another, demonstrating that during “conversations” as a result of self-organization processes a common vocabulary emerges. Yet other experiments demonstrate the emergence of speech (Oudeyer,2006).

And given the assumption of close similarity between the simulated behaviour of artificial interacting units and that of real humans, the experiments have implications in understanding language evolution, confirming that language is likely to have originated in small communities where the members are constantly in contact and interact with one another repeatedly and frequently, comparable to human dialogues. In short, in absence of direct evidence human behaviour and, most relevantly, communicative behaviour, can be imitated through the use of modern technological innovations, to a degree significant enough, to infer information about the initial stages of the evolution of language.

Summary and conclusions

In modern linguistics the attention to the language system as a code has taken central stage, while the contribution of the ostensive-inferential component instrumental in language use and the role of context in dialogues is marginalized. In this way the language system is divorced

from the language user and context, ignoring the obvious fact that the language code has no independent existence without its use in communication and that all communication happens in context.

In attempt to minimize this gap I have argued that natural language is best identified in terms of a language system as implemented in dialogues as both stable and flexible system, integrating code and context. In this sense it seems natural that the dialogue should take central place as a unit of language in theoretical approaches.

Moreover, the dialogue is the best representation of language as a system for sharing private thoughts and illustrating the human individual's two most essential traits as a rational and social being, therefore eliminating the controversy over the primacy of cognition vs. communication in defining language and the human mind.

Following the same line of thought I have argued that the evolution of language must be understood as evolution of the dialogue., i.e language was designed from its inception for use in dialogues. Moreover, humans have evolved for participation in dialogues, i.e. coordinated languaging, which shapes and reshapes the language system over time.

This view could sound controversial to some die-hard defenders of the classical view of language as a system for codifying rational thought and its direct mapping by monologues. That said, I am open to the challenge as by challenging theoretical establishments progress is made and knowledge is gained.

References

- Austin J. (1975)** How to do things with words, Harvard University Press
- Barr D. (2004)** Establishing conventional communication systems. Is common innate knowledge necessary?, *Cognitive Science*, 28, p.937-962
- Bates E., Goodman J.(1997)** On the inseparability of grammar and lexicon, evidence from acquisition, aphasia and real time processing, *Language and Cognitive Processes* , 12(5/6) , p. 507-58
- Boe J.L, Fagot, J. Perrier P., Schartz J.L.(2017)** Speech production and perception, vol.4 Peter Lang Edition
- Cartmill E. , Beilock S., Goldin-Meadow S.(2012)** A word in the hand: gesture and mental representation in humans and non-human primates, *Philosophical Transactions of the Royal Society*, 367, p.129-143
- Cullicover P. Jackendoff R. (2005)** Simpler Syntax, Oxford University Press
- Dehaene-Lambertz G.(2017)** The human infant brain: a neural architecture able to learn language, *Psychonomic Bulletin Review*, vol.24, n. p. 48-55
- Dessalles J. L.(2007)**Why we talk, evolutionary origins of language, Oxford University Press
- Donald M.(1993)** Origins of the Modern Mind, Harvard University Press
- Donald M. (1999)** Preconditions for the evolution of protolanguages , in M. Corballis, I. Lea, *The descent of mind*, Oxford University Press, p. 355-365
- Dor D., Jablonka E.(2000)** From cultural selection to genetic selection, a framework for the evolution of language, *Selection*, 1, 1-3, 33-55
-

- Dowman M. (2007)** Protolanguages that are semi-holophrastic, European Conference on Artificial Life: Advances in artificial life, p.435-444
- Dowman M. (2008)** The nature of words in human protolanguages: It's not a holophrastic-atomic meanings dichotomy, Artificial life, 14(4), p.445-465, MIT Press
- Enfield N., Sidnell J. (2015)** Language structure and social agency: confirming polar questions in conversation, Linguistic Vanguard 1(1) 131-143
- Ferreti F., Adornetti I. (2021)** Persuasive conversation as a new form of communication in Homo sapiens, Phil. Trans. R. Soc., B 376, doi: 10.1098/rstb.2020.0196
- Frohlich M.et all. (2019)** Multimodal communication and language origins: integrating gestures and vocalizations, Biological Reviews, Cambridge Philosophical Society, doi: 10.1111/brv.12535
- Garcia R., Zamorano F., Aboitiz F.(2014)** From imitation to meaning: circuit plasticity and the acquisition of conventionalized semantics, Frontiers in human neuroscience, vol 8 art.605, doi: 10.3389/fnhum.2014.00605
- Gil D.(2009)** How much grammar does it take to sail a boat? In G. Sampson, P. Trudgill, eds. Language Complexity as an Evolving Variable, Oxford University Press
- Ginsburg J. (2001)** Dynamics and semantics of the dialogue, Semantic scholar, Corpus ID 119010931
- Ginsburg J. Poesio M. (2016)** Grammar is a System that Characterizes Talk in Interaction, Frontiers in psychology, vol.7 art. 1938
- Givon T. (2002)** Bio-linguistics, the Santa Barbara lectures, John Benjamins
- Goldin-Meadow S., Alibali M. (2013)** Gesture's role in speaking, learning and creating language, Anu. Rev.Psyciol. 64, p.257-283
- Grice P. (1989)** Studies in the Way of Words, Cambridge MA, Harvard University Press
- Hauser M., Chomsky N., Fitch T.(2002)** The Faculty of Language, What is it, Who Has it, How did it Evolve, Science 298, p. 1569-1579 ,
- Hurford J. (1991)** The evolution of critical period for language acquisition, Cognition, 40 p. 159-201
- Hurford J., Kirby S.(1999)** The co-evolution of language size and the critical period , in Birdsong, Second language acquisition and the critical period hypothesis, Lawrence Erlbaum Assoc.
- Hurford J. (2012)** The origins of grammar, Oxford University Press
- Jackendoff R. (2002)** Foundations of Language, Oxford University Press
- Keller R.(1994)** On language change, the invisible hand in language, Rutledge
- Kendon A. (2000)** Language and gesture, unity or duality? , in McNeil, D. ed. Language and Gesture, Cambridge University Press, p. 47-63
- Levinson S., Holler J.(2014)** The origin of human multimodal communication , Philosophical transactions of the royal society, B,vol. 369, issue 1651
- Levinson S. (2019)** Interactional foundations of language: the interaction engine hypothesis, in Hagoort P. Et all. eds. Human language, from genes and brains to behaviour, chap.14, p.189-200, MIT press
-

- Lewis J. (2014)** BaYaka pygmy multi-modal and mimetic communication traditions, in Dor, D., Knight, C. Lewis, J. eds. The social origin of language, Oxford University Press, chap. 7
- Liebermann Ph. (2008)** Old time linguistic theories, Science direct, Cortex 44, p. 218-226
www.cog.brown.edu/.../Lieberman.%20P.%20.2008.Old%20time%20linguistic%20the.
- Mithun M.(2009)** Re(e)volving complexity, Adding intonation, in Givon,T., Shibatani, M. Eds. Syntactic Complexity, Diachrony, Acquisition, Evolution, Typological Studies in Language, John Benjamins, p. 53-80
- Morgenstern A.(2014)** Children's multimodal language development, in Facke,C. ed. Manual of language acquisition, De Gruyter, p. 124-142
- Muller C. (2018)** Gesture and sign, cataclysmic break or dynamic relations? Frontiers in Psychology, vol.9 art.1651
- Oudeyer P.Y. (2006)** Self-organization in the evolution of speech , Oxford University Press
- Pinker S. (1994)** The Language Instinct, how the mind creates language, Harper, Collins
- Port R. (2007)** How are words stored in memory? Beyond phones and phonemes, in New Ideas in Psychology, 25 ,p. 143-170
- Pulvermuller F. (2018)** Neurobiological mechanisms for semantic feature extraction and conceptual flexibility, Topics in cognitive science, vol 10 , issue 3, p. 590-620
doi: 10.1111/tops.12367
- Rizzolatti G., Arbib M.(1998)**“ Language within our grasp “, Trends in neuroscience,vol. 21Issue 5, p. 188-194
- Scott-Philipps T. (2017)** Pragmatics and the aims of language evolution, Psychonomic Bulletin and Review, 24(1), p.186-189
- Sherwood C., Subiaul F., Zawidzki T.(2008)** A natural history of the human mind, Tracing evolutionary changes in brain and cognition, Journal of Anatomy, Apr. 212 (4) p. 426-454, doi: 10.1111/j.1469-7580.2008.00868.x.
- Slocombe K., Waller B., Liebal K. (2011)** The language void: the need for multimodality in primate communication research, Animal behaviour, 81 (2011) 919-924
- Sperber D.,Wilson D. (2004)** Relevance Theory, in L.Horn, G. Ward, eds.The Handbook of pragmatics, Oxford, Blackwell, p. 607-632
- Steels L.(1995)**A self-organizing spatial vocabulary, Artificial life, 2(3), p.319-332, MITPress
- Stout D., Toth N., Schick K., Chaminade T. (2008)** Neural correlates of Early Stone Age toolmaking: technology, language and cognition in human evolution, Phil. Trans. R. Soc. B, 369, 1939-1949
- Tomasello M.(2008)** Origins of Human Communication, MIT Press
- Trudgill P. (2009)** Sociolinguistic typology and complexification, in Sampson, Gil, Trudgill, eds., Language complexity as an evolving variable, Oxford Univ. Press, p. 98-109
- Vernes S. (2017)** What bats have to say about speech and language, Psychonomic Bulletin and Review, vol. 24, issue 1, p. 111-117
- Zima E., Bergs A. (2017)** Multimodality and Construction Grammar, Linguistic Vanguard, doi:10.1515/lingvan-2016-1006
- Zipf G. (1949)** Human behaviour and the principe of least effort, Addison-Wesley
- Zuberbuhler K.(2015)** Linguistic capacities of non-human animals, Wire's Cognitive Science,
-

vol.6.issue 3, Wiley Online, p.313-321

*** address for correspondence: svetlana.t.davidova@gmail.com
