Linguistics

(Introduction)

A. Hayrapetyan

(aram.hayrapetyan@gmail.com)

Table of Contents

Overview

Language and Speech

Linguistic communication can be viewed as a means of conveying the thoughts of one human being to another in the form of speech acts[1]. Currently, linguistics is viewed as a deductive science, very much similar to physics, chemistry, or molecular biology, that uses formal logic, mathematics, and computer science as descriptive tools.

In this textbook, natural language[2] is considered as a thought-encoding and -decoding system that creates messages for communicating mental states—rational thoughts and personal emotions—among individuals. It has two major components: 1) biological—evolutionarily developed "hardware" consisting of (a) the brain as a Central Language Processing Unit and (b) Peripheral Language Input/Output Units, that is, sound- and gesture-producing and -perceiving mechanisms such as the vocal cords, mouth, nasal cavities, hands, facial muscles, ears, and eyes; 2) cultural—evolutionarily developed "software" coding units, namely morphemes and combinatorial rules for combining morphemes into complex messages that represent thoughts, emotions, attitudes, and intentions and for analyzing and converting these complex messages back into thoughts and emotions. We will refer to all of these mental states or mental images as thoughts.

With these two subsystems, biological and cultural, human beings produce natural speech, making language the producer of speech and speech the product of language.

Currently, in addition to naturally developed and learned vocal and gestural speech, humans produce written speech – texts, based on technology: the writing. The most compact and widely used technology – the alphabet - maps phonetic representation of the morphemes into visual codes. Unlike the naturally acquired spoken mother tongue, which is learned without formal instruction, writing—even in a naturally acquired language—requires explicit learning.

In this textbook, we analyze language and speech at an abstract level, considering abstract, media-independent symbols—the morphemes—and combinatorial rules—the grammar—that allow the back-and-forth encoding and decoding of messages and thoughts.

Thinking

We will not be concerned with the emergence of thoughts, but will instead try to understand (or rather speculate) where thoughts come from? This exercise is geared towards positioning linguistic behavior within the overall picture of mental activities.

There are two primary sources (excluding linguistic communication) of thoughts formation: perception and imagination. Both are assumed to form thoughts (images) about reality, including other humans as part of that reality. However, imagination also allows us to form thoughts that have nothing to do with reality—fictional thoughts.

---

[1] Some workers consider non-verbal, non-linguistic communication, such as hunger strike, a speech act. In this context only linguistic communication is considered speech act. The latter is special cases of a more general communicative acts (Gmytrasiewicz, 2020): pantomime (dance), drawing, move on a chess board, etc.

[2] By language in this context we mean natural human language unless it is specified as artificial, formal, artistic: of music, of visual arts, of literature (poetry), etc. otherwise.

Every human being is able to form both types of thoughts. Perceptual thoughts are always about reality (disregarding pathological cases). Imaginary thoughts can be about reality or about fictional situations. Thoughts about reality are objective in the sense that they can be verified by others via their perceptions and our shared experience. In the end, as a community, we build up an objective image of reality that is mostly shared by the majority, even though some individuals report different experiences. We refer to this image as an objective worldview or objective reality.

Fictional thoughts are also objective in the sense that any human being can have any thought that another human being can have. However, we can know them only from someone's account, whose imagination is the subjective source of fictitious thoughts). While reality is *one*, there is an unlimited number of fictional worlds, and we can learn about them only via interhuman communication. This is how fictional worlds are different from reality.

The human worldview imposes a semantics on images of both real and fictional worlds by discerning things, changes in things, and relations between the things and the changes. Our perceptual image of reality consists of things in a three-dimensional space that change position and form over time. Humans perceive things and changes in the past and anticipate them in the future. We perceive things as farther or closer, moving faster or slower, being bigger or smaller, etc. On the basis of this semantics, a logic is built that imposes constraints on the semantic representation of the world: some things change, others do not; some relations are possible, others are not, and so on.

Humans "communicate" with reality via the interface of perception. We perceive colors, shapes, textures, etc. and their changes. We do not perceive space (distance or position) or time (duration or instance). We invent these abstract notions to describe the relevant locations of things (shapes) and the relevant speeds of changes in the locations (or shapes) of things. The interface of perception imposes the notions of things and changes upon us. These two opposite notions—things that represent stability and changes that represent instability—are fundamental for our understanding, description, and explanation of reality because they are inseparable. They exist because they negate each other. As a proverbial "chicken and egg" paradox[3] each of them is a precursor and a "killer" of the other.

We extract the "logic of reality" and the "semantics of reality" through our perception of the world. We apply this logic and semantics to the structure of our speech to make it comprehensible for our audience. We understand each other because we perceive the same things (nominals) and changes (verbs), extract the same relations of negation, equivalence, and causality (logic), and observe the relevant dispositions, interactions, and changes of things—the semantics of reality.

---

[3] The current author does not consider the primacy of chicken and egg paradoxical at all. It is a result of "false" perspective, a wrong question as philosophers would say. The right perspective is to look at evolution of the "chicken and egg" *loop* as a whole. At some point the "birth and death" loop evolved into "birth of egg death of egg birth of chicken death of chicken" loop. One "chicken and egg" *loop* creates another "chicken and egg" *loop*. We need to consider the primacy of loops, rather than the [arbitrarily discriminated] parts of them. The real mystery is how the loop – *something* - formed in the first place out of *nothing*. The speculations about necessary coexistence of opposites, discovered by Chinese philosophers in 14-th century BCE – the Yin (shade, dark, black) and Yan (sunny, light, white), can shed some light on the issues.

Language is one of the tools – arguably the most powerful – for describing our thoughts about the reality and imaginary worlds.

Discourse analysis

The goal of conversation or discourse is to update the hearer's mental state by transmitting messages. The goal of discourse analysis is to take a given discourse, parse and analyze it, and, given the interlocutor's prior mental states, compute the new mental states that result from internalizing (comprehending) the discourse. Such analysis includes identification of symbolic structures (syntax) that are used to create a meaningful message (semantics) and identify the intent of the message (pragmatics).

In a very broad sense, theoretical linguistics is the science of discourse analysis.

Given the current state of the art, discourse can be defined as a sequence of connected linguistic units that together convey coherent messages. Following Halliday & Hasan (1976), cohesion links these units into a text, while coherence concerns how they form a meaningful whole in context. The contribution of inference and world knowledge to such coherence is discussed in Hobbs (1979) and Zufferey & Moeschler (2012: 143).

The units of speech are built according to syntactic rules and are interpreted through semantics, which maps linguistic form to meaning. The messages in discourse are realized as speech acts (Searle, 1965), where utterances perform actions such as describing, asserting, questioning, or requesting. Discourse analysis aims to assess the coherence and rational organization of such messages in actual language use.

The language-parsing and analysis procedures assume that discourse has an expression plane, corresponding roughly to surface structure, and a content plane, corresponding to deep structure[4]. The latter encodes most of the propositional content of the discourse, while additional contributions come from multilayered context (conversational, situational, and cultural). Syntactic analysis of the expression plane yields tree-like structures, whose interpretation produces a semantic representation. Combining this representation with contextual information yields a cognitive frame: a structured mental representation that organizes knowledge about a particular domain or scenario, determines what counts as relevant information within that scenario, and constrains the inferences that are treated as "common sense" in discourse (Lakoff, 2002, 2004, 2010). It integrates both the semantic content and the intended message of the discourse.

Terminological conventions

Precise and accurate definitions are very important, especially in deductive disciplines. However, they are not universal truths. They are simple, consistent statements that accurately, unambiguously define terms for deduction purposes. They play the role of axioms—the basic

[4] This view (Jahukyan, 1974) is different from Chomsky's view, who considers deep *syntactic* structure. In current context the deep structure caries *syntactic*, *semantic*, and *pragmatic* information – it is the structure of *content*.

"bricks" that we use to build the edifice of linguistic theories (see Defining in Hayrapetyan, 2026).

In general, there cannot be a common, universally applicable definition for more or less complex concepts. The role of definitions in this context is to make the content clear, rather than presenting an "everlasting truth". The goal is to make the presentation easily understandable. The scope of a definition in the main text or in the *Terms* section is intended to this presentation only. It does not mean that they cannot be "true" in other contexts. We are limiting the scope of a definition's applicability is to eliminate unnecessary arguments or criticism of definitions. The definitions in this textbook can be evaluated only for consistency and clarity in the context of this presentation.

This also does not mean that we redefine everything and do not use commonly accepted definitions. On the contrary: we redefine only those that might be inaccurate or inconsistent *in this context*.

Several terms mentioned above—for example, *speech*, *discourse*, and *discourse analysis*—are used in a fixed, technical sense. The "logic of reality" means the informal pattern of constraints and regularities that we extract from our perceptual experience of the world and project onto our mental models; it is not itself a formal calculus, but a cognitively grounded background structure. The same is true for the "semantics of reality": it reflects the relationships and behavior of objects and their changes in space and time as perceived in our experience of reality.

The book structure

In the *Introduction*, we review the logical representation of the linguistic "hardware"—the language faculty or "language organ," the "mental structure," and the communication system—to justify and clarify the definitions and approach. This review also clarifies the disposition of language and speech in mental behavior and in the communication process. We also review the basics of computer science, information theory, formal logic, and other mathematical tools and data structures that we use for the deductive presentation of theoretical linguistics. A very brief history of linguistics will also be presented to highlight the milestones on the path to our current understanding.

The main parts of the textbook are:

1. Book 1 – Introduction (current content). Disposition of language in the realm of mental activities. Brief history of linguistics
2. Book 2 - Philosophy of Language – We investigate what is language? How it came about? Human behavior and language.
3. Book 3 - *Use*: Pragmatics – What is the language for and how human use it? Analysis of human language usage.
4. Book 4 - *Content*: Semantics – How the meaning is attached to the units of speech – to lexemes, to phrases, to clauses, to sentences, to discourse?
5. Book 5 - *Form*: Syntax, Morphology, Orthography and Phonology - What are the smallest coding symbols – the morphemes ? How lexemes, phrases, clauses, sentences, are encoded and decoded?

6. Book 6 - Computational Linguistics – list of major algorithms that are used for language and speech generation and analysis; statistical linguistics: hidden Markov chains; basic statistics; OCR and ARS technologies.

Introduction

Overview

In this section, we will review hardware and software tools that underlie human linguistic behavior. We need to keep in mind that, in describing the "hardware," we use a logical representation very similar to the logical representation of computer units such as the CPU and I/O systems.

The goal of such representation is to separate biological and cultural components: the former is the "hardware", the latter – "software". Similar to computers, the culturally evolved "software" runs on biologically evolved "hardware". The "software" consists of natural language: a dictionary of morphemes and a grammar—combinatorial rules for combining existing sequences of morphemes into new ones. The "hardware," guided by language, sequences such semantically and logically related combinations into "executable code"—speech—and "executes" it, thereby changing mental states that manifest as human behavior.

The linguistic "hardware" and "software" produce, parse, and execute the code – change mental states or images.

Mental Image Formation

We can talk about reality only through images formed in the human brain. "Our mental image of reality is the only one, since there is not yet (and it is difficult to imagine or reason if ever will be) an image of reality obtained by another transmitter different from our mind" (Hayrapetyan, 2022 : 17).

In mental images of reality, people distinguish two contradictory yet necessary, fundamental, and logically inseparable concepts: objects, which imply permanence and immutability, and actions, which imply movement, that is, transformation.

Our brain places things in a three-dimensional volume, where changes occur. Thus, we observe two main things: objects and changes. Objects are characterized by volume and distance (spatial relations (Hayrapetyan, 2022 : 20)), and changes are characterized by speed and, therefore, duration (temporal relations (Hayrapetyan, 2022 : 20-21)). To estimate and describe volume, we need three dimensions, three lengths, and for the speed of change, one dimension, duration.

"The transformations of objects are of two general types: becoming and being. Both types are closely related to our perception of duration: the process of transformation and the pause in transformation.

The duration during which we perceive sensory images or think about them, that is, what happens in parallel with our thinking activity, is perceived as the present.

That is, the transformations of objects, which occur in parallel with the transformations of our mental states, are placed in the present. Everything (thing or phenomenon) that has become or was before the present is perceived as the past, and things to become or phenomena to be as the future.

We perceive the beginning and end of pauses as events. They create a cause-and-effect relationship between parts of reality, states. The end of a transformation, an event, can be the beginning of a new transformation - a cause. Temporal sequence is the prerequisite of causality. The existence of a cause before the effect, earlier, is a necessary, but not sufficient, condition of causality. Such a dependence of cause and effect - causality, arranges the states of "reality" in temporal positions, creating our perception of time as a causal sequence of moments" (Hayrapetyan, 2022 : 20-21).

In images, in addition to objects (states), we distinguish their ontological, spatial, temporal, and logical (causal) relations (Hayrapetyan, 2022 : 16-28).

To characterize, distinguish and identify, and describe objects, relations, and changes, we attribute to them properties: large, green, soft or fast, slow, begun, oscillating, etc.

Images of changing objects, relationships, and properties are "registered" in the Concept Network (CN) (Hofstadter, 1979), which is formed in our mind (Hayrapetyan, 2022 : 190-192):

"Natural speech, together with the senses, conveys information about the surrounding. In this way, it participates in the construction of the CN.

Thus, the construction of the CN is carried out by 1) cognitive and 2) declarative processes.

1. Sense ⇒ Image ⇒ ((Consciousness)) ⇒ [[Thinking]] ⇒ Concept ⇒ CN.

2. Cognition. Concept ⇒ ((Language)) ⇒ [[Speech]] ⇒ ((Language)) ⇒ Concept ⇒ CN" (Hayrapetyan, 2022 : 224).

We will call these two channels of image formation, respectively, sensory and verbal. In addition to these, there is a third source of image formation - imagination. It allows us to construct images that may not actually exist, such as mythical images. Other examples include abstract mathematical objects. These are transmitted into the CN via a channel similar to the sensory channel (#1 above), but the source is imagination instead of Sense.

Both sensory and imaginary images can be transmitted via the verbal channel.

Classification of images

If we consider a mental state as an image, then our mental images can be classified according to their source—(a) sensory or (b) imaginary—and according to their type—(c) real, (d) mental, or (e) abstract.

Since our goal is not to study images neurologically, we will limit ourselves to assuming that they are analyzed into concepts: objects, relations, properties, changes, and stored in the CN. It can be argued that mental images do not turn into concepts. They are communicated in a "visual" way through dance, acting, recitation (artistic oral speech, rhetoric), literature (artistic [mostly] written speech), music, painting, sculpture, architecture, etc. Without disputing the validity of this statement, however, let us note that there are concepts that describe or even correspond to

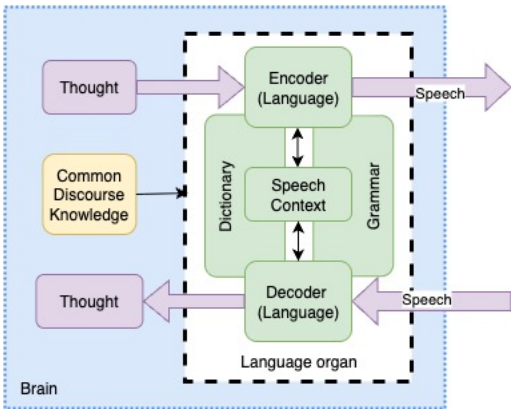
mental states: enthusiasm, joy, pain, despair, and so on. At least some mental states can be assigned concepts, and their classification can be considered a complete representation of mental images in the human brain.

"Language organ"

At the outset, it is important to explain why the term "language organ" appears in quotation marks. Unlike other organs, it has no clear temporal or spatial boundaries in the human body. In this essay, the "language organ" is used as an abstraction that represents human linguistic abilities. In this sense it is similar to the Turing "machine," which is not a physical machine but an abstract model of computation.

The hypothesis of a "language organ" is supported by a variety of biological and medical observations, most prominently by the ease with which children (a) acquire their first language with little explicit instruction and (b) develop grammatical creoles out of the relatively ungrammatical pidgins they hear from their parents. These and other observations motivate the assumption that all human languages share a common meta-grammar, already available in the infant's "language organ" before birth.

By analogy with the Central Processing Unit (CPU) in computer systems, the "language organ" can be described as an abstract Linguistic Processing Unit (LPU) to outline conceptual architecture and interactions between major functional blocks:



In the LPU, the following components are distinguished:

1. *Dictionary* – a list of morphemes together with their meanings and functional roles (patterns of use).
2. *Grammar* – rules for combining and analyzing lexemes via morphemes, and for composing lexemes and phrases into larger phrases and sentences. The resulting structures are tree-like and are linked by connectives (conjunctions and relational lexemes) to form sentences, extended speech, and discourse.
3. *Decoder* – a device that takes speech as input and converts it, using grammatical rules, into a mental image or thought.
4. *Encoder* – a device that takes thought as input and converts it into speech using grammatical rules.
5. *Speech Context* – a temporary store that records information relevant for speech production and comprehension, supplied by both Encoder and Decoder.

The main processes associated with the LPU are:

1. Language acquisition, during which the Dictionary of morphemes and the grammatical rules of the ambient language are formed in the child's brain.
2. Speech production (synthesis).
3. Speech comprehension (analysis).

The LPU is not intended as a neuroscientific theory of brain anatomy. It is a *conceptual architecture* that organizes the functional roles we need in order to talk about discourse production and comprehension: lexicon and grammar, encoding and decoding procedures, and a dynamically updated context store.

Each of these components underpins phenomena examined in the sections below. The Dictionary and Grammar are where vagueness, polysemy, and category mismatches in definitions arise. The Encoder and Decoder, together with Speech Context and background knowledge, are responsible for resolving ambiguities of relational verbs like *is* and *have*, and for "filling in" unstated information. When System 1 runs most of this pipeline automatically, the resulting discourse logic can diverge from the stricter patterns enforced when System 2 deliberately inspects and regiments these same representations.

The generic meta-language of natural languages or meta-grammar[5] of natural language grammar is assumed to be "wired" into the LPU. If CPU is able to run "machine code", then LPU can run "linguistic code": perform grammatical operations on morphemes, lexemes, phrases, clauses, and sentences. "Language organ" is *hardware,* implemented in neural tissue, while linguistic code is *software,* built or implemented as set of symbols (morphemes) combinatorial rules (grammar).

The grammar of the particular language dominant in the child's environment is "uploaded" into the LPU, somewhat like loading an operating system—or more precisely, a compiler—into computer memory. In a computer, the compiler is pre-built and loaded at once, whereas in humans the grammar is actively constructed, using the pre-wired meta-grammar, during roughly the first four to five years of life and then refined more slowly throughout the individual's lifetime. During acquisition, the Dictionary of morphemes and the community-specific patterns of their use are configured; these patterns specify the type of morpheme as a pairing of meaning and grammatical role in lexemes, phrases, and sentences.

In speech production, the Encoder receives a Thought (mental image) and, using the Dictionary and Grammar, turns it into speech. In addition, it relies on *common knowledge* (Lewis, 1975), which is crucial because, without shared background knowledge, verbal communication would be not merely inefficient but impossible. A Generating Device (within this architecture) makes the relevant entries in the Speech Context store and keeps them active over the span of a conversation or reading episode; this contextual information, among other things, licenses the use of pronouns and other anaphoric expressions.

---

[5] Not accurately named *universal grammar* and sometimes incorrectly identified with *deep structure*. See *Terms*.

In speech perception, the LPU's Decoder analyzes the incoming signal, updates the Speech Context, and, using this context together with General Knowledge and Grammar, converts the received speech back into Thought. Other linguistic activities, such as learning and using a non-native language or using written language (reading and writing), are not strictly biological processes; rather, written language is a cultural technology that exploits the "language organ" to extend a person's verbal memory and communication range.

Mental structure: System 1 and System 2

Overview

There are two distinct systems of thinking (Kahneman, 2011): System 1 and System 2. System 1 operates automatically and quickly, with little or no effort; it is fast, intuitive, and emotional. These are the major characteristics of System 1:

1. Functions quickly, effortlessly, and often unconsciously.
2. Handles everyday judgments, perceptions, and responses based on patterns, intuition, and emotions.
3. Responsible for immediate impressions and gut reactions.

Examples: Detecting that one object is farther away, understanding simple sentences, or reacting to a loud noise.

System 2 allocates attention to effortful mental activities, requires conscious reasoning, and is deliberate, analytical, and logical. It

1. Operates slowly and with effort, requiring concentration and self-control.
2. Engaged in complex computations, critical thinking, and problem-solving tasks.
3. Needed for activities that require logic, planning, and reasoned decisions.

Examples: Parking in a narrow space, calculating a complicated math problem, or comparing two products for quality.

In the context of this paper, System 1 (intuitive) and System 2 (deliberate) are used as abstractions[6] that represent the groups of mental functions listed above; they are abstract conceptual representations of some cognitive abilities, analogous to the LPU.

In that sense, System 1 largely "runs" the LPU in ordinary discourse, while System 2 selectively re-analyzes its outputs in deductive contexts.

Relationship and Division of Labor

Both systems are active, but System 1 mainly runs the show, providing suggestions to System 2.

System 2 is usually in a low-effort mode and becomes fully engaged when System 1 encounters a novel situation or a problem it cannot solve.

[6] The author is aware of dual-system model criticism: e.g. De Houwer, 2019; Hochman, 2024.

This division optimizes mental processes. In the majority of everyday situations, people rely on System 1, making quick decisions based on fast but coarse measurements performed by the sensory apparatus. Without System 1, human beings would not be able to function and, more importantly, could not speak their native language. Without System 2, it would be impossible to perform creative activities such as writing, painting, imagining, and calculating—simply put, thinking.

This optimization balances the advantage of System 1 in solving simple problems quickly but imperfectly, often with errors, against the advantage of System 2 in solving complex problems accurately but slowly, at the cost of substantial time and energy.

Apparently, native speakers produce speech using the 'hardware' of the 'language organ,' which is a conceptual component of System 1. We speak without explicit ‘thinking’. When conscious reasoning is necessary, System 1 transfers the discourse context to System 2 for contemplation, logical analysis, and rational decision making.

Communication

Three layers of message

Language is a tool for communicating thoughts—more generally, for communicating mental states from one individual to another.

Natural language is predominantly used for communicating rational thoughts, but it can also be used for artistic purposes. Literary prose and poetry use linguistic constructions to convey emotional states as well. Frequently, especially in poetry, the emotional content is linked to the “melody” of speech, abstracting almost completely from the literal meaning. Such use of speech is akin to use of sounds, tones, and rhythms in music or shapes and color combinations in painting.

Artistic communication is not in the scope of this content.

We concentrate on the epistemological and intentional nature of speech: its ability to describe our mental images and to influence the course of the world[7]. That is, in addition to meaning, speech contains a message—it has a communicative plane (where the intent lies), which is parallel to the content plane (where the meaning lies). The message is encoded by morphemes and by grammatical rules for combining and sequencing them; this encoding constitutes the expression plane. Language is a toolkit for building the expression plane of speech, which expresses the meaning and intention of the content plane.

“When examining the relationships between the members of a sentence, it is necessary to distinguish at least two layers: objective and subjective or communicative and conversational. In the first layer, the verb-predicate acts as a manifestation of the relations of objects, and since the

[7] We do not subscribe to the opinion that speech [act] changes the world directly. The speech, even in Declarative mode or use, changes the states in mind of the members of community and that, in its turn, creates or becomes a new reality. Here we follow the accepted wording – one might argue that mental states *are* part of the world - but at the same time keep this interpretation in mind.

objects in a given sentence act in precisely this relationship, the verb-predicate becomes a kind of center. This center, naturally, exists in and through the objects. In the second layer, the so-called sentence subject acts as the center of the sentence - the mental object that is in the center of the speaker's attention, and to which some characteristic is predicated (in a broad sense). The subjective plain of the sentence is the direct manifestation of the communicative situation" (Jahukyan, 1974 : 333).

If the content encodes information about the world and, one might say, answers the question of *what* the speaker is saying, then the message encodes the speaker's intention, that is, *why* the speaker is saying it, and the expression plane shows *how* the speaker is saying it.

Thus, speech has an expression plane, which is a code built from symbols according to combinatorial laws, and a content plane that can be divided into objective and communicative layers (levels).The community of speakers of a language is able to attribute meaning to linguistic messages.

Speech act theory

Speech act theory assumes that speech does not solely describe the world, but it also changes it. In either case, speech conveys information.

Within the framework of speech act theory, the following communicative effects – acts or forces - are attributed to what is said: a) the expression of literal meaning - Locutionary act, b) the expression of the speaker's intention - Illocutionary act or force, and c) the effect on the listener - Perlocutionary act or force (Grice, 1957; Austin, 1962). Communicative acts are:

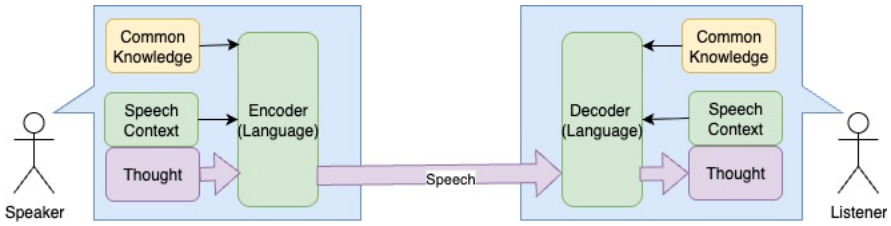
1. Descriptive: Recording a fact: *It is hot outside*
2. Imperative: Expressing a request, an exhortation: *Please pass the salt*
3. Messaging: Devotion: *I promise to keep the flames burning*
4. Expressive: Expressing an emotion: *I am happy for you*[8]
5. Declarative. Changing the state of the affairs through expression: *The meeting is adjourned*

All these actions correspond to the linguistic (grammatical) means, mainly modality. However, it is not always possible to determine the correspondence of speech to behavior by linguistic means.

Communication system

Since language is often described as a communication system, it is necessary to clarify what the technical term communication system means. Shannon described the communication system and analyzed it in depth (Shannon, 1948). The diagram below shows the adaptation of Shannon's communication system to linguistic communication.

[8] In the case of this particular sentence it is hard to discern this category from the Descriptive (the naming of categories is also a hint why this is happening). In this exposition we will develop symbolic, deductive, formal theory of speech acts – rather intentionality.



Shannon's goal was to study communication over a noisy channel. In the case of natural-language communication, the types of noise include mispronunciation, unfamiliar accents, and lack of differentiation of letters in a particular font or handwriting, and so on. However, this is beyond the scope of our consideration.

The Transmitter and the Receiver of Shannon's communication system have been replaced by the "Language organ" or language, which is the Transmitter in the case of speech production (speaker), and the Receiver in the case of perception (listener).

The diagram is a representation of the thought-communication channels described in the *Mental Image Formation* section. The “Language organ” receives as its input the Thought that the Speaker wants to communicate to the Listener. The “Language organ” converts the Thought into a tree-like structures and then converts them into speech. Common Knowledge is used to remove pieces of speech that the Listener can unambiguously restore, and Speech Context allows the removal of additional pieces of speech or their replacement with shorter ones, for example, pronouns. Pronouns can stand for sentence members and complex phrases and clauses.

The speech thus formed is communicated via communication channel to the Listener, whose “Language organ”, using Common Knowledge and Speech Context, assigns meaning to the received message, the speech, and turns it into Thought.

The above description of the communication system has several goals:

1 To show that language is not itself a communication system but a component which, at one end of the system, plays the role of transmitter or encoder of the message, and at the other, the role of receiver or decoder of the transmitted code—the speech.

2 To confirm the principle of the autonomy of grammar (Hayrapetyan, 2022 : 31): it shows that speech has no meaning during transmission. Speech, which in Shannon’s system is the message, has no meaning outside of thought—meaning is “attached” to speech in the Mind.

3 To emphasize that Common Knowledge guarantees that the meaning and interpretation of a concise (“discounted”) speech are identical for the Listener and the Speaker. Common knowledge is important for language mastery or awareness (competence), which the "Language organ" uses "without thinking" - "instinctively".

4 In speech (the speech flow), grammatical anaphora is provided by Speech Context, which the “Language organ” also constructs and uses “instinctively.”

“The ultimate task of communication is to record (or update) the subnet of the CN of one person - the speaker - in the isomorphic CN of another person (the listener). In this process, the speaker successively a) converts the CN subnet into a semantic structure - Context Tree (CT), b)

converts the CT into a syntactic structure, and c) linearizes the syntactic structure into speech (sentences), while the listener d) parses the speech into a syntactic structure, e) converts it into a CT, and f) records the CT in the listener's CN" (Hayrapetyan, 2022 : 225-226).

Linguistic structures

Overview

Structure of Language

Language is a system of linguistic units and combinatorial rules for creating new linguistic structures from existing ones. There are two major types of linguistic structures: lexical structures—lexemes—and syntactic structures—phrases, clauses, and sentences.

The set of combinatorial rules for constructing lexemes is know as linguistic morphology. The set of combinatorial rules for constructing lexemes is known as linguistic morphology, and the set of rules for building syntactic structures is syntax. The combination of these two sets constitutes the grammar of a language.

The smallest units of language are morphemes. Lexemes are combinations of morphemes.

These are the types of morphemes:

1. *Stem, typed by independence):*
    a. *Dependent stem (combining form),*
    b. *Independent stem (simple lexeme)*
2. Suffix, typed by relative location to the stem:
    a. *Prefix,*
    b. *Infix,*
    c. *Interfix*,
    d. *Postfix,*
    e. *Circumfix*
3. *Particles:*
    a. Adpositions
    b. Negations (oppositions)

Morphological rules combine morphemes into lexemes—minimal syntactic units that can be assigned a grammatical role in the sentence or can bear ontological meaning.

There are two morphological types of lexemes: (a) simple—words that are continuous sequences of morphemes—and (b) complex—word combinations separated by short pauses in oral speech, which in writing are represented by spaces or dashes.

Simple lexemes are built paradigmatically by two systems: declension and conjugation. In these systems, new lexemes are created by attaching different suffixes to stems. The former is used for nominal forms, while the latter is used for verbal forms.

Complex lexemes are usually combinations of words and adpositions. They can be applied to nominals (for example, in Armenian) as well as to verbs (for example, in English).

In the case of nominals, word–adposition combinations are extensions (or replacements) of the declension system. The declension system produces nouns, adjectives, and adverbs to be used as sentence members, or, in other words, as heads of phrases.

The typology of the lexemes includes:

1. *Adjective* – predominantly denoting attributes or predicates; they can act as nouns and attributes of attributes (adverbs)
2. *Adverb* – can be used as an attribute to attributes or as a description of circumstances of actions.
3. *Adposition* – morphemes that are used in word groups to produce new lexemes.
4. *Article* - predominantly act as quantifier (?).
5. *Classifier (Counter)* - morphemes (?) that categorize and count entities within a nominal domain,
6. *Conjunction* - predominantly denoting logical operations between sentence members (heads of phrases),
7. *Determiner* - morphemes (?) that specify reference or quantity of nouns (for example, demonstratives and possessives)
8. *Imitation (Interjection)* – copies of sounds in the natural or cultural environment, such as animal (including human) calls, sounds of moving air and water, and collisions of solid bodies, including machinery.
9. *Noun* - predominantly denoting "things", but also qualities (attributes), abstractions, actions, processes, etc.
10. *Numerals* – have no unique syntactic identity; they are either nouns (cardinal or absolute numbers, fractions) or adjectives (ordinal numbers). In cultures with developed notions of quantity, there are special morphological rules (grammars) for building lexemes that represent unlimited quantities from basic numeric morphemes.
11. *Participle* – verbal forms denoting tense and aspect of action. They typically denote "thing" (gerund) or attribute (verbal adjectives).
12. *Pronoun* – lexemes that stand for other lexemes: adjectives, adverbs, and nouns in phrases, or phrases., (clarity: punctuation; number agreement), Pronoun – lexemes that stand for other lexemes—adjectives, adverbs, and nouns in phrases—or for entire phrases.
13. *Relational* - lexemes that encode relations between entities, events, or properties.
14. *Verb* – lexemes that characterize change of state, appearance (shape, color), location of "things" (nouns and adjectives) in a time duration relevant to present frame of speech production.

Structure of Speech

Speech is a hierarchical (tree-like) structure built from four types of phrases: adjectival, adverbial, nominal, and verbal. Each type of phrase contains a head—the main lexeme in the phrase—which is combined with complement phrases.

The head of the Noun Phrase (NP) is a noun, with an optional complement Adjective Phrase (AjP). The adjective head of an AjP can have a complement NP and an Adverbial Phrase (AdP), and an AdP can itself have another complement AdP.

The most complex type of phrase is the Verbal Phrase (VP). The head verb of a VP can be complemented with an NP and an AdP. A VP with a finite verb becomes a clause. The combination of a subject NP, morpho-syntactically alligned with the clause head verb, constitutes a sentence.

All types of phrases can be connected to each other by connectives—special lexemes that include conjunctions and relational lexemes. These typically considered meta-sentential lexemes – they are not members of sentence.

Another type of meta-sentential unit is address, which typically consists of NPs (and can also be an interjection).

Sentences are combined into discourse, which is a semantic structure characterized by close relationships among its constituent sentences.

Types of combinatorial rules

Natural speech is a subordinating linguistic structure that builds up from bottom up according to this diagram:

morpheme → *lexeme* → *phrase* (*clause*) → *sentence* → *text* (speech, discourse)

|- -^

Lexemes are typically built according to paradigmatic rules: declension paradigms for nominal forms and conjugation paradigms for verbal forms. In both cases they produce lexemes that are combinations of stems with attached suffixes and detached adpositions or other particle morphemes. A stem itself can be a combination of two or more (typically three) stems.

Syntactic rules determine how phrases are constructed and linked to each other. Typical speech is a sequence of tree-like structures coordinated by the alignment of phrase heads. We can always compose a sentence using an atomic process of merge: concatenation of two lexemes. Typically, the two merged lexemes are diagrammed as an upside-down letter V = Λ. These are some examples of sentence structures:

```
   car /           Λ          Λ          Λ
 the Λ red        Λ         Λ  Λ        Λ  \
                 /          Λ          Λ Λ Λ
```

The first structure can represent, for example, the NP "the red car"; the next can be "the bright red car stopped".

This is the list of different types of grammatical rules:

1. Morphemes are combined into lexemes by combinatoric morphological rules.
2. Lexemes are combined into phrases by the rules of positioning (*առդրություն*), complementarity (*խնդրառություն*), or connectivity (? *շաղկապում*).
3. The pairing of the subject NP head noun and the VP head verb obeys the rules of morphosyntactic alignment.

The structure of the sentence is the result of applying the above rules. Sentence is a VP containing a finite verb aligned with the subject NP, and it can be divided into three layers: 1) the verb at the top, 2) the subject and object NPs, and 3) the adverbial and adjectival complements in the VP and in the subject and object NPs (Hayrapetyan, 2022 : 149-158).

Terms

References

Aumann, R.J. (1976). Agreeing to Disagree. The Annals of Statistics. 4 (6): 1236–1239.

Belnap, N. D. Jr. (1977). A Useful Four-Valued Logic. In Modern Uses of Multiple-Valued Logic, edited by J. Michael Dunn and George Epstein, 5–37. Dordrecht: Reidel.

Blackburn, P., de Rijke, M., & Venema, Y. (2001). Modal Logic. Cambridge: Cambridge University Press.

Bobzien, S. Ancient Logic, The Stanford Encyclopedia of Philosophy (Summer 2020 Edition), Edward N. Zalta (ed.)

Crupi, V., Fitelson, B., & Tentori, K. (2008). Probability, Confirmation, and the Conjunction Fallacy. Thinking & Reasoning 14(2): 182–199.

De Houwer, J. (2019). Moving Beyond System 1 and System 2. Experimental Psychology. 66. 257-265.

Devlin, K. (1997) Goodbye, Descartes. Wiley, John & Sons. NY

Doležel, L. (1998). Heterocosmica: Fiction and possible worlds. Johns Hopkins University Press.

Dunn, J. M. (2019). Two, Three, Four, Infinity: The Path to the Four-Valued Logic and Beyond. In Hitoshi Omori and Heinrich Wansing (eds.), New Essays on Belnap-Dunn Logic. Cham: Springer.

Frege, G. (1879). Begriffsschrift: Eine der arithmetischen nachgebildete Formelsprache des reinen Denkens (Conceptual notation and related articles). Halle: Nebert.

Gigerenzer, G. & Hertwig, R. (1999). The ‘conjunction fallacy’ revisited: How intelligent inferences look like reasoning errors. Journal of Behavioral Decision Making, 12(4), 275–305.

Gmytrasiewicz, P. (2020). "How to Do Things with Words: A Bayesian Approach". Journal of Artificial Intelligence Research. 68: 753–776.

Gupta A., Mackereth S. Definitions, The Stanford Encyclopedia of Philosophy (Fall 2023 Edition), Edward N. Zalta & Uri Nodelman (eds.)

Hayrapetyan, A. (2022) On Syntactic Structure Representation (in Eastern Armenian: *Քնական խոսքի շնորհանրական ներկայացման մի տարբերակի մասին*). Agoulis, Concord. Autoreferat.

Hayrapetyan, A. (2025). Conjunctions in Eastern Armenian, www.academia.edu/129638433

Hayrapetyan, A. (2026). The_Logic_of_Discourse, www.academia.edu/165955739

Hochman, G. (2024). Beyond the Surface: A New Perspective on Dual-System Theories in Decision-Making. Behavioral Sciences, 14(11), 1028.

D.R. Hofstadter. (1979) Gödel, Escher, Bach: An Eternal Golden Braid. (20th anniversary edition) Basic Books, 1999.

Jahukyan. G. (1974). The Foundations of the Theory of the Contemporary Armenian. Ջահուկյան. Գ. Ժամանակակից հայերենի տեսության հիմունքները. ՀՍՍՀ գիտությունների ակադեմիայի հրատարակչություն, Ե.

Kahneman, D. (2011). Thinking, Fast and Slow. Farrar, Straus and Giroux. NY

Karimjonova, Sh., & Usmonova, D. 2021. The Causative Verb to Have in Modern English. International Journal of Humanities and Applied Sciences 7(4): 41–45

Krifka, M. & Teichert, H., 2019. Interpretation of Approximate Numerical Expressions: Computational Model and Empirical Study. Journal of Pragmatics 148: 1–21.

Kripke, S. A. (1963). Semantical analysis of modal logic I: Normal modal propositional calculi. Zeitschrift für mathematische Logik und Grundlagen der Mathematik, 9(5–6), 67–96.

Lakoff, G. 2002. Moral Politics: How Liberals and Conservatives Think. 2nd ed. Chicago: University of Chicago Press.

Lakoff, G. 2004. Don't Think of an Elephant! Know Your Values and Frame the Debate. White River Junction, VT: Chelsea Green.

Lakoff, G. 2010. Why it Matters How We Frame the Environment. Environmental Communication 4(1): 70–81.

Langer, S.K. (1953). An Introduction to Symbolic Logic. 2-nd Edition. Dover Publications. Inc. NY.

Lee, C. J. 2006. Gricean Charity: The Linda Problem Reconsidered. Philosophy of the Social Sciences 36(2): 193–218.

Lewis, D. (1975). Languages and language. Minnesota Studies in the Philosophy of Science.

Lewis, D. (1978). Truth in fiction. American Philosophical Quarterly, 15(1), 37–46

Maienborn, C. 1999. Stage Levels, States, and the Semantics of the Copula, ZAS Papers in Linguistics 14: 65–91.

Mares, Edwin, Relevance Logic, The Stanford Encyclopedia of Philosophy (Summer 2024 Edition), Edward N. Zalta & Uri Nodelman (eds.)

Mikkelsen, L. (2011). Copular Clauses. In Semantics: An International Handbook of Natural Language Meaning, vol. 2, edited by Klaus von Heusinger, Claudia Maienborn, and Paul Portner, 1805–1829. Berlin: De Gruyter Mouton.

Miyamoto, J. M., Gonzalez, R., & Tu, S. (1995). Compositional Anomalies in the Semantics of Evidence. In J. Busemeyer, R. Hastie & D. Medin (eds.), The Psychology of Learning and Motivation, Vol. 32, pp. 217–247. San Diego, CA: Academic Press.

McCarthy, J. (1980). Circumscription—A form of non-monotonic reasoning. Artificial Intelligence, 13(1–2), 27–39.

McDermott, D. (1981), Artificial intelligence meets natural stupidity. In J. Haugel, Mind Design. MIT Press. pp. 5-18.

Poundstone, W. (1988). Labyrinths of Reason. Paradox, Puzzles, and the Frailty of Knowledge. Anchor Press, Doubleday, NY.

Priest, G. (1979). The Logic of Paradox. Journal of Philosophical Logic, 8

Reiter, R. (1980). A logic for default reasoning. Artificial Intelligence, 13(1–2), 81–132.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. Cognitive Psychology, 7(4), 573–605.

Rosch, E. (1978). Principles of Categorization. In Cognition and categorization, E. Rosch and B.B. Lloyd (eds), 27-48. Hillsdale, NJ: Lawrence Erlbaum.

Routley, R., R.K. Meyer, V. Plumwood and R. Brady, (1982). Relevant Logics and its Rivals (Volume I), Atascardero, CA: Ridgeview.

Russell, B. (1903). The Principles of Mathematics. Second Edition, 1938. Cambridge University Press.

Russell, B. (1905). On Denoting. Mind 14(56): 479–493.

Searle, J. (1965). What is a speech act?. M. Black, ed. In Max Black, Philosophy in America. Routledge. pp. 221--239. 2004

Searle J. R., Vanderveken D. (1985). Foundations of illocutionary logic. Cambridge University Press, Cambridge.

Sides, A., Osherson, D., Bonini, N., & Viale, R. (2002). On the Reality of the Conjunction Fallacy. Memory & Cognition 30(2): 191–198.

Solt, S. (2014). Vagueness in Quantity: Two Case Studies from a Linguistic Perspective. In The Cambridge Handbook of Formal Semantics, edited by Maria Aloni and Paul Dekker, 85–114. Cambridge: Cambridge University Press.

Tarski, A. (1944). The semantic conception of truth and the foundations of semantics. Philosophy and Phenomenological Research, 4(3), 341–376.

Tarski, A. (1941). Introduction to Logic and to the Methodology of the Deductive Sciences. Oxford University Press, Martino Publishing, NY. 2013

Thomasson, A. L. (1999). Fiction and metaphysics. Cambridge University Press.

Tversky, A. & Kahneman, D. (1982). Judgments of and by Representativeness. In Kahneman, D., Slovic, P., & Tversky, A. (Eds). Judgment under Uncertainty: Heuristics and Biases. Cambridge University Press, (pp. 84-98). https://apps.dtic.mil/sti/tr/pdf/ADA099502.pdf

Tversky, A. & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. Psychological Review, 90(4), 293–315.

Zadeh, L. A. (1965). Fuzzy sets. Information and Control, 8(3), 338–353.

Zadeh, L. A. (1975). Fuzzy logic and approximate reasoning. Synthese, 30(3–4), 407–428.

Похлебкин, В.В. (1996). Кулинарный словарь. Центрополиграф. М.

History of Revisions

Name	Date	Reason for Changes	Version
Aram Hayrapetyan	07/04/26	First draft	0.1