

# When Language Survived, Music Resurrected and Computer Died: to the Problem of Covert Ontologies in Language

Kolmogorova <sup>[0000-0002-6425-2050]</sup> Anastasia

Siberian Federal University, Krasnoyarsk, Russia  
nastiakol@mail.ru

**Abstract.** Creating formal ontologies is one of the current information science trends. However, in the context of taxonomies existing in natural languages another type of **class** hierarchy seems to be more important – the so-called “covert ontologies” that categorize entities in terms of crypto classes or hidden classes.

The research aims to examine localization of the three entities in the Russian language natural ontology, which seem very different from **the** formal point of view. These entities are: “language” (i.e. belongs to the formal class of systems), “music” (i.e. represents the formal class of perception or activity), and, finally, “computer” (i.e. embodies the formal class of equipment).

According to our preliminary observations, all three entities under discussion are conceptualized in Russian language as living systems. Our further analysis of 500 occurrences in which the three entities’ names adjoin verbs designing different steps of vitality cycle showed that “music” enters the class of mythic heroes or Demiurges, “language” belongs to the covert class of Humans; at last, “computer” integrates the class of pets.

The revealed properties of natural categorization due to the effects of covert ontology also influence the eventual semantic roles of exploring entities’ names.

**Keywords:** Covert Ontologies, Formal and Natural Semantics, Crypto Classes.

## 1 Introduction

The present paper pursues the aim to highlight and then to discuss an assumption that a human mind is able to conceptualize the Universe not only by exemplifying all its entities in the strictly rational and formal way which we imitate building formal ontologies but also by using rudimentary classification principles which are less explicit and, at the same time, more natural. Such covert ontology manifests itself in natural languages world usage from which we can see arising so-called “crypto-classes”. Even if we compare the place of such disparate phenomena as language, music and computer in formal ontologies and the covert one, we will see that the mentioned entities are differently categorized. However, what categorization is more valid for such a living system as a Human? I argue that the response to this question does not

seem clear and [the hesitancy](#) invites discussion about the status of formal ontologies as a preponderant framework for the knowledge building systems.

## 2 Theoretical ground and methodology

### 2.1 Concept of Ontology

The concept of ontology is deeply rooted in Antique and European philosophy. Aristotle was the first to postulate the importance of the study of being in general that he called the “first philosophy” [1]. The Latin term of *ontologia* had [penetrated into](#) the European metaphysic discourse in 1606, [thanks](#) to the German thinker Jacob Lorhard [2]. It became a key term of the Husserlian phenomenology two centuries later.

For Husserl [3], logic is the theory of science because it is concerned with a closed collection of meanings and associated meaning-instantiating acts, so that the unity which is characteristic of the science must involve both: (1) an interconnection of truths (or of propositional meanings in general), and (2) an interconnection of the things to which these truths (and the associated cognitive acts) are directed. To grasp and to structure units of a science the latter uses formal ontology, its sister discipline, that relates to object categories such as object and property, relation and relatum, manifold, part, whole, state of affairs, existence etc. In this way, [modern science is due](#) to Husserl notions of formal logic and formal ontology: as formal logic deals with properties of inferences which are formal in the sense that they apply to inferences in virtue of their form alone, so formal ontology deals with properties of objects which are formal in the sense that they can be exemplified, in principle, by objects in all material spheres or domains of reality [4].

The ontological bias found its systematic treatment in the philosophical conception of Husserl’s disciple N. Hartmann [5], who, unlike his [predecessor](#), focuses on the objective and immanent aspect of world categorization postulating that categories deal with what is universal and necessary, they specify configurations, structures and contents, not forms of existence. For him, categories are immanent to the world: they do not form a second world. They form the network of internal, dynamic determinants and dependencies, which articulate the furniture of the world.

### 2.2 Formal and Natural Ontologies

[As previously mentioned](#), Hartmann’s objectivist intention [was](#) to describe the “furniture of the world” [and](#) gave the recent emphasis on formal ontology building as the possibility of knowledge sharing and reuse across different applications of knowledge engineering. Today, formal ontologies are largely used in computer science, particularly, in the practice of knowledge-based systems building. Despite the enormous variety of definitions specifying the notion of ontology, there is a set of permanent features regularly ascribed to this phenomenon, such as the following: explicit description of the concepts and their hierarchical relationships that can exist for an agent

or a community of agents and can be designed for use in supporting information retrieval, analysis and integration in scientific and other domains [6-8]. An ontology defines a common vocabulary for researchers who need to share information in a domain. It includes machine-interpretable definitions of basic concepts in the domain and relations among them [9].

The most known ontologies are the upper level formalizations: SUMO (Suggested Upper Model Ontology), SENSUS, WorldNet. They all are based on cognitive agents' conceptual maps of the world governed by strictly logical rules and axioms.

However, [this formal approach is not](#) the only one used by Humans to categorize the world. Husserl was the first to distinguish natural or so-called "naive" ontology and formal one. As the name implies, natural "furniture of the world" does not demand any specific operation to present the observer's mind "work" more explicitly. On the contrary, it reflects a spontaneous human's world perception far from being accessible to any intentional analysis of the cognitive agent himself. Since the works of W. von Humboldt [10], linguistics has been aspiring to find out some of its traces or vestiges across natural language data. By introducing the notion of crypto-type, B. Whorf gave life to a new linguistic trend. Accordingly to his definition, "a crypto-type is a linguistic classification like English gender, which has no overt mark actualized along with the words of the class but which operates through an invisible "central exchange" of linkage bonds in such a way as to determine certain other words which mark the class, in contrast to the phenotype, such as gender in Latin" [11 : 78].

The idea of some covert classificatory map existence resulting from specific human cognitive strategies for interpreting their environmental niche, the world around them, was met by linguists with enthusiasm and found its development in the field. For many [researchers](#), language became a productive data to denude the hidden meshwork that is sometimes inherited from mythological conscience.

Thus, [the famous Russian scientist V.A. Vinogradov](#), used the notion of covert class to elaborate the typology of non-grammaticalized nominative classes in African languages [12]. By his works, he [proved](#), that despite the lack of formal morphological [marks](#), some names in African languages let the researchers see their categorical status through the tendency to meet in the context adjectives or verbs belonging to the same covert class, for example, class of "what is liquid" or "what is sharp". Such mechanism of semantic concordance was given the name of "echosemy". Its motivation is the universal Human's wondering to see the similarity in very different and heterogeneous things: river runs, time runs, thoughts run.

In similar vein, another Russian linguist A.A. Kretov, has developed Vinogradov's suggestion proposing the term of "crypto class". He defines the latter as "a lexico-grammatical category of noun that consists in [noun](#) distribution into classes according to their semantic features in the case if 1) the class corresponds leastwise to one morphemically expressed grammatical category in any language in the world; 2) the marker of class obligatory manifests in the utterance structure under the forms of grammatical construction or word combination, word form" [13 : 9].

While aided by the key concept mentioned above, O.O. Boriskina degages and describes 6 of such crypto classes in modern English language [14]. The researcher calls them by using Latin names Res Filiforms, Res Liquidae, Res Longa Penetrantes, Res

Acutae, Res Parvae, Res Rotundae. For example, class Res Liquidae embraces all the English nouns whose semantics manifests features “making allusion” to the properties of any liquid substance (water e.g.): information, knowledge, music, song, news, world, speech etc. The case illustrates the crucial principles of crypto classes building – first, it is based on the naive projection of the world formed in experience of everyday life and, secondly, it largely uses the principle of similarity inherited from a mythological world vision.

In the next section, I am going to compare in what classification case formal and natural or covert ontologies put such three different entities such as music, language and computer.

### 2.3 *Music, language and computer in ontologies and thesauri*

As it is shown below, searching for exemplifying the world formal ontologies build up a hierarchy of categories and their subcategories where each entity finds its localization.

Music, language and computer are limited within rather different categories (except for the Worldnet ontology where music and language are perceived as constituents of the same communication class): language is categorized within the classes of Property or Process, Communication or System; music – the classes of Activity, Communication and Art; computer enters those of Machine, Internet or Electronic device (see Table 1).

**Table 1.** Music, language and computer in meshwork of formal ontologies and thesauri

Entity / Type of ontology	SUMO	Worldnet	RuThez	Roger’s Thesaurus
Language	<b>Property or Process</b>	Communication	<b>System</b>	Communication
Music	<b>Activity</b>	Communication	<b>Art</b>	Art
Computer	<b>Machine</b>	Internet	<b>Electronic device</b>	Machine

Even from the view of formal logic, the examined phenomena are far from being classified always in the same way. This example is meant to highlight the assumption that the formal categories are neither absolutely objective nor immanent to the real world configuration. They are more explicit due to the deeply rooted metaphysical tradition, but they do not represent the only way to perceive and understand the world around us. Their alternative, being less explicit, but, as it seems, having the same relevance and importance for the human mind, is covert “natural” categorization that manifests in natural language functioning.

### 2.4 *Music, language and computer through natural language collocations (on the material of the Russian language)*

The observations of linguistic data of Russian language prompted me to the preliminary conclusion that from a naïve cognitive agent’s point of view, language, music

and *computer* are conceptualized as living organisms. In this way, we can speak about upper level conceptual metaphor LANGUAGE, MUSIC and COMPUTER ARE LIVING ORGANISMS in Lakoff's and Johnson's sense of the term.

To be able to retrieve the metaphorical occurrences relevant to this study from our corpora we adopted the Metaphorical Pattern Analysis method (MPA) proposed by Stefanowitsch. The researcher defines "metaphorical pattern" as 'a multi-word expression from a given source domain (SD) into which one or more specific lexical items from a given target domain (TD) have been inserted' [15 : 66].

In this way, I was preoccupied in the search for metaphorical patterns for music, language and computer, as I discovered a large range of collocations of these words (in Russian) containing classificatory marks of living organism class. As such classificatory markers, I consider verbs denoting the main processes and properties of living systems: 1) main properties – nutrition, breathing, motion, reactivity, reproduction; 2) main processes – birth, growing up, reproduction, aging and death. As a result, the corpus of 97 Russian verbs that contain semantic components related to the above-mentioned properties or processes was built: *родиться, рождаться, есть, кушать, вкушать, питаться, завтракать, обедать, ужинать, дышать, вздохнуть, выдохнуть, дыхнуть, расти, взрослеть, возмужать, вымахать, стареть, стариться, дряхлеть, бежать, прыгать, падать, идти, двигаться, съежиться, отпрянуть, вздрогнуть, родить, заболеть, болеть, стареть, стариться, умереть, сдохнуть, подохнуть, загнуться* etc..

Using a corpus-based approach to metaphor, I retrieved more than 500 instances with collocations containing lexemes *язык* 'language', *музыка* 'music', *компьютер* 'computer' and verbs as their metaphorical predicates from the list mentioned above.

As sources for my data corpus, the contexts from National Corpus of Russian Language, Russian Corpus of Leeds University, the database of Modern Russian Papers "East View.com", services of "Google ngramm viewer" and "Google Trends" were used.

Having checked the list of 97 verbs (i.e. eventual markers of the class of living systems) by applying the MI score (a measure of how strongly two words seem to associate in a corpus, based on the independent relative frequency of two words [16]), I selected 14 verbs from my data collection that regularly occur in the same contexts with three words under discussion. All retrieved collocations were with the MI score more than 3 that is for Russian language the sufficient score for considering such collocates as statistically relevant word combination [17]. The 14 verbs list includes the following lexemes: (1) *родиться* 'to be born', (2) *жить* 'to live', (3) *расти* 'to grow up', (4) *выжить* 'to survive', (5) *пережить* 'to face, to experience', (6) *страдать* 'to suffer', (7) *болеть* 'to be ill', (8) *(по)родить* 'to give birth', (9) *стареть* 'to get old', (10) *умереть* 'to die', (11) *сдохнуть* / (12) *помереть* / (13) *кинуть кони* / (14) *околеть* – everyday speech synonyms of *to die*.

### 3 Results

The analysis of collocations shows that not all 14 mentioned above Russian verbs do occur with *music* (*музыка*), *language* (*язык*) and *computer* (*компьютер*). It means, not all the three entities possess the same properties of living systems (see Table 2).

**Table 2.** Verbs' occurrences with music, language and computer in research data corpus

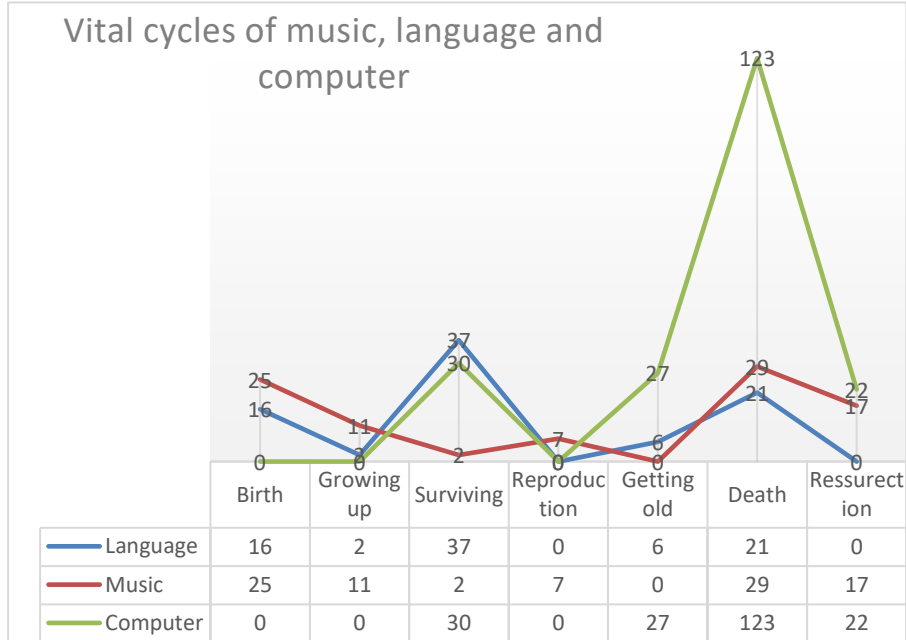
Entity / Verb	Be born	Live	Grow up	Survive	Experience	Suf- fer	Be ill	Give birth	Get old	Die/ its everyday synonyms
Language	+	+	+	+	+	+	+	+	+	+/-
Music	+	+	+	+	+	-	-	+	-	+/-
Computer	-	+	-	+	+	-	+	-	+	+/+

The further metaphorical patterns analysis also revealed some particularities in using *music*, *language* and *computer* in collocations:

- In collocations where *music* is ascribed to be born the verb **is predominantly used in the Present Tense** (*музыка рождается* 'music is born'). In this way, music is conceptualized as someone who is appearing in front of our faces so that we can observe his/her birth. It is similar to the scenes of god's or goddess's birth frequent to Greek mythological narrations: the birth of Zeus, the birth of Apollonius, the birth of Athena etc..
- Only *music* (*музыка*) and *computer* (*компьютер*) can occur with the verb *воскреснуть* 'to rise from the death' as if they were divine beings or if they had experienced some divine influence:
  - Музыка воскресла. Музыка встала на дыбы и рассыпалась* (*Music has risen from death. Music has reared up and has broken*) (V. Nabokov "King, queen, knave").
  - Компьютер воскрес! Ура, товарищи* (*Computer has risen from death! Hurrah comrades!*) (from the technical support blog).
- Only *computer* (*компьютер*) combines with some rather rude and low-colloquial synonyms of the verb *умереть* ('to die'), such as *сдохнуть*, *око-леть*, *помереть*, *кинуть кони*. The Russians largely use all of them when speaking about pets, but, in contrast, they employ such words to talk about people only in pejorative way.
- Only two of the entities under discussion could be perceived as being ill or becoming sick (*болеть*, *заболеть*):
  - Русский язык заболел суффиксами* (*The Russian language became sick of suffixes*) ("Nevskoe vremja", 15.04.2017);
  - Заболел компьютер. Что делать?* (*My computer is ill. What I can do?*) (from the technical support blog).

The figure below (see Fig.1) displays the frequency of different vital cycles verbal markers occurring with each of the three words. We can see that the blue line of language has little pics on the three relevant events and properties: birth, surviving and

death. In contrast, this lexeme is never associated with the ideas of growing up, reproduction nor resurrection. The red line of music is rather smooth – its focal points on the vital cycle vector are birth, death and resurrection. Finally, the green line of computer rose twice – on surviving and death, but also reveals the importance of proper-



ties to get old and to rise from death.

**Fig. 1.** The frequency of verbal markers of the category “the living system” for music, language and computer

## 4 Discussion

As can be seen from the linguistic data, the three entities under discussion are conceptualized in the naïve projection of the world within a large source domain of living organism. However, they “represent” different subspecies of such organisms. To continue with the idea of crypto classes, we argue, that music, language and computer belong to the different crypto classes of living being in frame of this natural covert ontology. The corpus of linguistic collocations shows that they have specific sets of classificatory markers.

Focusing on such processes or events in the living system vital cycle as BIRTH, SURVIVING, GETTING OLD and DEATH, language’s existence, as naïve cognitive agents perceive it, seems to be very close to human life. In other terms, Russians speak about *язык* (*‘language’*) as if it was a human being. They use verbal markers of this crypto class to predicate properties to the language. Indeed, Russian speakers conceptualize language as a member of covert category of “human being”, which is

less explicit than, for instance, categories of “systems” or “communication” in formal ontologies, but, instead, more natural and cognitively relevant for them.

In contrast, the Russian collocations containing the lexeme *музыка* (*‘music’*) lead us to the conclusion that in naïve world categorization of Russians music enters such crypto class as “divine beings, gods or goddesses”. In our corpus, *музыка* (*‘music’*) meets predicates that highlight three key-points of its being: BIRTH, DEATH and RESURRECTION. As such, music escapes from any sort of suffering, it *does not* face any problems or difficulties of surviving. Instead of it, in narrations about music, its birth represents a focal point designed, in the most cases, by the Present Tense, which makes clear that this scene is observable for the narrator: *и вот теперь музыка рождается во мне* (and just now I feel the music coming into the world).

Finally, the functioning of the lexeme *компьютер* (*‘computer’*) shows that its referent belongs to the class of domestic animals: it faces only the processes of SURVIVING, GETTING OLD and DYING. The last event seems to be the most relevant for its conceptualization by naïve cognitive agents. It is to stress, that thinking about computers people adopt the same strategy that they use for thinking about domestic animals: as long as an animal works well helping its master in his labor, its existence is taken for granted, but when it becomes old or dies, its master pays attention to its existence or, already, absence. In most contexts, *компьютер* (*‘computer’*) is perceived as being already dead but there are some collocations that suggest the idea that sometimes the Divine Grace can influence it and can give it a second birth.

The three crypto classes discussed above (“human beings”, “divine beings” and “domestic animals”) that embrace, respectively, the entities of language, music and computer on the level of covert categories displayed in Russian, influence the semantic properties of the lexemes themselves. Indeed, the ideas of non-autonomy of the computer, of autonomy of the language and of the upper or supernatural position of the music determine the mentioned lexemes’ central semantic roles, according to Fillmore’s theory of semantic cases, respectively, of an Experiencer, Patient and Agent or Stimulus e.g. *уход за компьютером, забота о языке, восхищение музыкой*.

## 5 Conclusion

Increased attention to the formal ontologies building is due to the interest of modern computer science to the universal inventory of the “furniture of the world”, a sort of window on a portion of reality that specialists in very different domains could share. However, searching for a ready way to complete and to unify human knowledge specialists in computer science seem to be excessively attracted by the metaphysical idea of the immanent nature of world classes and categories. All concur, therefore, that cognitive agents such as humans implement not only formal principles to build up strongly hierarchical systems of classes to structure their environment. They have some deeply rooted cognitive strategies enabling them to conceptualize the world in terms of crypto classes. We believe that for naïve cognitive agents, such covert natural ontologies are very important. Seen through natural language collocations, crypto



classes provide the most comprehensive framework for perceiving and understanding the world in our everyday life.

## References

1. Wyatt Aiken, D.: *Essence and Existence, Transcendentalism and Phenomenalism: Aristotle's Answers to the Questions of Ontology*. *Review of Metaphysics* 45 (1), 29–55 (1991).
2. Øhrstrøm, P., Andersen, J., Schärfe H.: *What Has Happened to Ontology?* In: Dau, F., Mugnier, M.-L., Stumme, G. (eds.) *CONCEPTUAL STRUCTURES 2005, ICCS*, pp. 135–147. Springer, Heidelberg (2005).
3. Husserl, E.: *Formal and Transcendental Logic*. Martinus Nijhoff, Netherlands (1969).
4. Smith, B. *Logic and Formal Ontology*. In J. N. Mohanty and W. McKenna (eds.), *Husserl's Phenomenology: A Textbook*, Lanham, pp.29–67. University Press of America (1989).
5. Hartmann, N.: *New ways of ontology*. Henry Regnery Company, Chicago (1953).
6. Hovy, E., Knight, K., Junk, M. *Large Resources. Ontologies (SENSUS) and Lexicons*, <http://www.isi.edu/natural-language/projects/ONTOLOGIES.html>, last accessed 2017/05/24.
7. Gruber, Th. *What is an Ontology*, <http://www-ksl.stanford.edu/kst/what-is-an-ontology.html>, last accessed 2017/05/31.
8. Guarino, N. *Understanding, Building, and Using Ontologies*, <http://ksi.cpsc.ucalgary.ca/KAW/KAW96/guarino/guarino.html>, last accessed 2017/06/05.
9. Protégé, [http://protege.stanford.edu/publications/ontology\\_development/ontology101.pdf](http://protege.stanford.edu/publications/ontology_development/ontology101.pdf), last accessed 2017/06/01.
10. Humboldt, W.v.: *On Language. On the Diversity of Human Language Construction and its Influence on the Mental Development of the Human Species*. Cambridge University Press, Cambridge (1999).
11. Whorf, B.L.: *Language: Thought and Reality*. Cambridge University Press, Cambridge (1956).
12. Vinogradov, V.A.: *Nominative categories in songai*. In: Vinogradov, V.A. (ed.) *Foundations of African linguistics. Nominative categories, vol.1*. Aspect Press, Moscow (1997).
13. Kretov, A.A., Titov, V.T.: *The role of covert categories in typological description of Romance languages*. *Bulletin of Voronej State University. Linguistics and Cross-cultural communication* 1 (6), 7–12 (2010).
14. Boriskina, O.O.: *Crypto classes in English*. Publishing House “Istoky”, Voronej (2011).
15. Stefanowitsch, A.: *Words and their metaphors: A corpus-based approach*. In: Stefanowitsch, A., Gries, Th. (eds.) *Corpus-based Approaches to Metaphor and Metonymy*, pp.63–105. Mouton de Gruyter (2006).
16. Church, K., Hanks, P.: *Word association norms, mutual information, and lexicography*. *Computational Linguistics* 1(16), 22–29 (1996).
17. Zakharov, V., Khokhlova M.: *Study of effectiveness of statistical measures for collocation extraction on Russian texts*. *Computational Linguistics and Intellectual Technologies* 9 (16), 137–143.