The Mechanism of the Form-Content Correlation Process in the Paradigm of Socio-Natural Sciences

Dorota Zielińska

Abstract In this chapter, I search for the mechanism correlating linguistic form with content in order to explain (in the sense of the word 'explain' used in empirical, i.e., natural and modern social sciences) how sentence meaning contributes to the utterance meaning. I do that against the background of two currently dominating positions on that issue: *minimalism* and *contextualism*. Minimalists regard language as a self-standing abstract system and claim that only weak pragmatic effects are involved in interpreting sentences. Contextualists believe that language can be described adequately only within a theory of language understanding and that strong pragmatic effects are also involved in interpreting sentences. The resultant controversy, presented in Sect. 1, has been pronounced by Michel Seymour the most important one in the 20th century. I begin Sect. 2 with Mario Bunge's argument that since abstract systems cannot change by themselves and only speakers of language do, an explanatory theory of language (one looking at language from the perspective of empirical (socionatural) sciences) must concern language understanding, i.e., view language as a bio-psycho-social phenomenon. However, language understanding needs to be incorporated in the theory of language in a more fundamental way than current contextualist models do.

D. Zielińska (✉)
Kraków, Poland
e-mail: dorotazielinska@yahoo.co.uk

These models assume the existence of language as self-standing, abstract structure with a list of symbol-reference pairings (Such assumption is legitimate as long as one regards such an abstraction as only a methodological device.) and model language understanding disregarding its psycho-social development process. Such assumptions, however, lead to a number of insurmountable problems. I conclude Sect. 2 by arguing that to solve these problems, as well as to be consistent with the evidence attesting to the fact that language self-organizes and self-regulates, (also reviewed in this section,), we need a model of language understanding and production to be coined within a developmental bio-psycho-social perspective. In Sect. 3, I propose a specific model of the form-meaning correlation process, based on a novel mechanism of a linguistic categorization, which is compatible with a bio-psycho-social developmental perspective advocated in Sect. 2. On this view, the utterance meaning is dependent both on the approximate conventional meaning of the construction components conveying it, and on the specific social function of the whole construction (a relevant *pragmeme*), which identifies feasible situation specific contents. The given construct selects one out of these options. I finish the chapter, Sect. 4, by preliminarily testing the mechanism of the form-content correlation process introduced in Sect. 3 both qualitatively and quantitatively to meet the methodological standards of empirical sciences.

> *Languages thrive in the hospitable environment of human brains and communities*
>
> Hurfort

The secret of natural language seems to be hidden in the way that conventional meaning contributes to utterance meaning. No wonder why Seymour (2010) said that the debate between contextualists and minimalists (who represent the two dominating opposing views on the issue) is the most important controversy to arise in the analytic philosophy of language in the past 30 years. This chapter is a contribution to the debate, offered from the perspective of the paradigm of empirical sciences. During this discussion, the author questions some of the philosophical assumptions taken for granted both by contextualists and minimalists and presents the issue at stake in a very new light.

The chapter starts by defining the controversy in its original form. Next, the author argues that, looking from the perspective of empirical (socio-natural) sciences (in which researchers aim at creating explanatory models implied by the functioning of material systems), language cannot be viewed as abstract, isolated structure as proposed by minimalists. Instead, a model of language must essentially involve accounting for understanding language in a social context, as postulated by contextualists. Current contextualist models, however, also fail to explain satisfactorily major linguistic phenomena such as the compositionality of meaning. The author goes on to claim that the inadequacy just mentioned is due to the fact that contextualist models currently proposed incorporate language understanding process in language interpretation models too late—as some sort of

adages to the view of language as an abstract system. Just like minimalists, contextualists start with the description of *langue*. They both treat language structure and content as independent from each other, and, most importantly, independent from the actual process of language creation, development, and understanding. This, I will argue, is the main reason of the limited adequacy of current contextualist approaches to modelling linguistic phenomena.

I claim that language, and in particular its structure, should not be viewed independently from its usage and people using it. It should be better viewed as a result of the simultaneous form and content creation and form-and-content correlation process taking place in human brains during social interaction between linguistic community members. This process is driven by (bio)psycho-social mechanisms. It is further argued that language creation, in the sense of there being statistically strongly dominant syntactic patterns and encodings in a linguistic corpus of *parole* produced by a given community, takes place through self-organization due to functional factors (like in biology) and not due to causal laws.[1] Consequently, language modelling in the empirical paradigm must imply a search for statistical patterns concerning the statistical characteristics of a corpus of *parole* and for mechanisms accounting for these patterns I conclude the discussion by proposing a specific developmental mechanism of linguistic form-and-content correlation process and indicate how to test it in a way that meets the standards of empirical (socio-natural) sciences.

1 The Contribution of the 'Minimalist Versus Contextualist' Debate to Understanding the Relation Between Conventional and Utterance Meanings

The discussion concerning the relation between conventional and utterance meanings carried on between minimalists and contexualists has brought us significantly closer to understanding that issue—the Holy Grail of linguistics. Therefore, I present my own views on the issue against the background of the debate just mentioned.

1.1 Defining the Controversy

The controversy between contextualists and minimalists requires taking a stand on what words in modern languages mean and what relation there is between word meaning and its contributed sense.[2] Once it was believed that the meaning of an

---

[1] after Altmann (1978).

[2] of which we know very little, as recently acknowledged by Recanati (2011).

utterance was essentially a token of the meaning of the uttered sentence. Nowadays, practically everybody agrees that even the conventional meaning of a sentence devoid of indexical expressions, does not determine the meaning of an utterance it helps to convey. As Seymour (2010: 2673) puts it "the fundamental question that remains is whether conventional meaning offers at least a kernel onto which the far richer intended meanings of the speaker are grafted, or if [strong—DZ] pragmatic features intrude in the very determination of what is literally expressed".

Pragmatic contribution to an utterance meaning has been divided into weak and strong pragmatic effects. We talk about weak pragmatic effects when the conventional meaning of a sentence, even one devoid of indexical expressions, contains an "algorithm" to extract additional information from the context. Such effects are acknowledged both by contextualists and minimalists. For instance, the minimalist Stanley (2007) proposes that the relevant algorithm is due to some hidden variables, which he terms 'indexes'. Stanley illustrates the relevant mechanism by offering the following examples. The sentence *It is raining* is taken to mean "it is raining here and now," because it contains hidden variables designating a place and time, which do not have specific values assigned to them before the sentence is uttered on a specific occasion. These variables (indexes) are assigned values defined by the time–space location of the relevant speech act. On the same account, the adjective *big* in the sentence *A four-year-old Jessica made a big snowman*, has a hidden index in it, which, when specified, turns 'big' into 'big for a snowman made by a four-year-old girl.'

While Stanley's proposal definitely brings us closer to describing the utterance meaning, the problem with an explanatory value of such an approach is that the number of indexes which need to be postulated cannot be limited if language is to express all possible situations ever to be encountered and therefore such parameters cannot be contained in a finite abstract system. On another occasion, for instance, the same adjective *big* may need to convey 'big for a snowman made by a four-year-old autistic girl with exceptional manual talent on a day when snow was not wet enough to make good snowmen and when she did not have a good day, either.'

Even the so-called descriptive adjectives,[3] as it is reminded below, often require similar filling in. Typically, when taking into account supposedly purely conventional context of a descriptive adjective, the number of parameters which would need to be postulated to interpret it, exceeds what the individual items intuitively encodes. Compare the meanings of the adjective *red* in the phrase *red eyes* used on the following two occasions: first by a photographer concerned with a red reflex in the photography—a portrait of a child, and second, as understood by a mother looking at her sick child with reddish whites of his eyes indicating the intensity of sickness. Thus, when interpreting the item *red* not only the shade of

---

[3] Consequently Recanati (2011) rightly questions the sense of the very division into descriptive adjectives and adjectives requiring "filling in".

the respective "rednesses" needs to be established, but so does the area it assesses. These values, in turn, both depend not only on the noun to which adjective *red* ascribes the value of colour, but also on the given situation, whose function defines possible meanings. And since the number of potential situations (defined by their functions), in which a given construct can be used depends on language use and not on the semiotic system itself, all potential meanings of a given linguistic item cannot be predefined by the given semiotic system, in particular by any set of indexes.

Recanati (2011),[4] a contextualist, proposes, in turn, that the change of conventional meaning such as the one just described is a psychologically instantiated change of conventional meaning due to context. Recanati introduces the terms 'modulation' to cover the concept just illustrated and 'syntactic flexibility' to refer to a parallel discrepancy in interpreting syntax. To illustrate that latter concept, I shall refer the reader to my analysis of the sentence "The visions of apocalypse have every right to scare us" in Zielinska (2007) and its logical interpretation stating that "we have the right to be scared by the visions of apocalypse",[5] whose syntax seems to have no conventional algorithmic relation with the original sentence. Syntactic flexibility introduced by Recanati (2011) also includes the instances of non-syntactic communication studied by Stainton (2005, 2006a, b). In any of the situations just mentioned, it is hard to think of any systematic way of grafting the syntactic structure of the sentence actually used onto the structure of the sentence expressing the content literally meant. In other words, occasionally, linguistic compositionality seems to break down without causing problems in communication.

Considering such examples as the ones mentioned above (exemplifying the issue of apparently regular compositionality), or well described cases of the novelty conveyed by compound constructions (c.f., deriving the concept of 'e-money' from the lexemes 'electronic' and 'money'), contextualists concluded that the influence of contextual information is not merely a matter of weak pragmatic effects such as saturation, (assigning contextual values to parameters determined by encoding). Instead, when words are used in expressions and sentences, their standing meanings are affected (modulated) in a fundamentally strongly context dependent way. (Therefore, as pinpointed by Recanati (2011), the controversy between contextualists and minimalists can be rephrased as providing the answer to the question whether (1) "both strong and weak", or (2) "only weak" pragmatic effects determine what is said by a sentence).

The concepts of modulation and syntactic flexibility, however, merely label, but do not explain the source of divergence between actual and encoded meanings. Such divergences have not been accounted for so far as explicitly admitted e.g., by

---

[4] Other researchers expressing similar views include: Capone (2005, 2006 and later), Carston (2002), Levinson (2000), Mey (2001, Recanati (2004, 2011), Searle (1983), Sperber and Wilson (1986), Stainton (2006) and Travis (2001).

[5] The sentence was overheard on a TV show.

Carston (2002). Carston (ibidem), explicitly gives up on accounting for the relation between what people know about word meanings and what they "literally" say, (which she illustrates analyzing selected meanings of the item *open*) the with the following words:

> The question which won't receive any answer here, is how the more general schema or indicator arises and how it comes to be the meaning of the lexical expression type. There must be some process of abstraction, or extraction, from the particular concepts associated with the phonological form /open/ to the more general 'meaning', which then functions as a gateway both to the existing concepts of opening and to the materials needed to make new 'open' concepts which may arise in the understanding of subsequent utterances. Carston (2002: 364).

1.2 The Source of the Controversy Between Minimalists and Contextualists

Saying that standing meaning is affected by context in the way unpredictable by the given standing meaning itself is equivalent to stating that something beyond the system of syntax and standing meanings influences utterance meaning (i.e., the input of the context goes beyond providing some values to encoded parameters of standing meaning). The only related entities that could be the source of unpredictability in the use of language system (itself) are language users functioning in a linguistic community. Therefore, the controversy whether only weak or both weak and strong pragmatic effects are relevant for establishing utterance meaning is really implied by the answer to the question whether we can explain language in an adequate way treating it as a set of qualitatively defined abstract sentence structures with independently defined sign-representation, lists thus disregarding language users and their lives. Minimalists assume 'we can', contextualists reply 'we cannot', and postulate that accounting for semantic non-combinatorial novelty requires modelling human psychological processing of language in context, i.e., it requires including the process of gaining knowledge in the very model of language.[6] As Dummett (1993: 12) states it "a theory of meaning must also be a theory of understanding" [cf. Searl (1983)].

The same conclusion must be reached when attempting to construct an explanatory model of language from the perspective of socio-natural sciences. Since abstract systems cannot change by themselves, to construct an explanatory model of language in this paradigm, language cannot be viewed as an abstract system isolated from its users and use. As Bunge explains:

---

[6] The debate between contextualists and minimalists covers a very similar terrain that the divide between Cartesian and non-Cartesian linguistics does as proposed in a research of Kopytko (1995, 1998, 2001a, b, 2004).

> Languages do not develop or evolve by themselves and there are no mechanisms of linguistic changes, in particular evolutionary forces. Only concrete things, such as people can develop and evolve. And, of course, as they develop or evolve, they modify, introduce, jettison linguistic expressions. The history of mathematics is parallel: mathematicians do come up with new mathematical ideas, which are adopted or rejected by the mathematical community, but mathematics does not evolve by itself Bunge (2003: 62).

From the perspective of empirical sciences, only the behaviour of material things can be explained (in the sense of providing its cause). Abstract systems alone are not capable of explaining anything in the sense of providing its cause, but can only offer summation rules—rules summarising typical experience (subsuming under generalization). Thus, from the perspective of empirical (socio-natural) sciences, the original controversy has an unambiguous answer. There cannot be an explanatory theory of a semiotic system.

1.3 Why Contextualists have Failed to Account Adequately for Strong Pragmatic Effects?

If we are convinced that contextualists are right in believing that modelling language requires modelling language understanding processes, why have they not found the mechanism of strong pragmatic effects yet? As already mentioned the reason is that they do not model language understanding deeply enough, i.e., they do not consider the understanding process involved in and leading to language creation and development. Instead, the dominant approach to language description, both among minimalists and contextualists, starts with describing *langue* in terms of patterns of symbols and lists of symbol-representation pairings, with disregard for epistemic concerns, i.e., answering the question how come we can get to understand the meaning of lexicon and of syntactically combined lexical items. Contextualists consider language understanding only at the stage of interpreting *langue as an abstract semiotic system* in a specific context, which, I shall argue below, is too late.

2 In Search of an Alternative Paradigm

In the last half of the century, independent research in linguistics, philosophy, neurology and systems theory provided us with a new way of perceiving language in comparison to the reductionist approaches explored in the 19th and 20th centuries. Below, I shall outline a new paradigm consistent with the state of the art in philosophy and sciences, and argue why it is more appropriate for modelling language.

2.1 Insufficiency of Defining Language as Sets of Patters of Symbols

A great initial success of approaches to describing language using qualitative mathematics came from the ingenuity of introducing recursive[7] rules of combining symbols. Presenting a model of language based on recursive rules operating on uninterpreted (mental) symbols allowed linguists to account for linguistic creativeness understood as combinatorial novelty, (i.e., new combinations of non-terminal, and eventually, of terminal symbols), which, at zero approximation, can be considered to be the essence of language. Consequently, first Artificial Intelligence (AI) specialists believed, as earlier Plato did, that formal logic and ideal form is the essence of thinking, and in particular—of language. This impressively fruitful step indeed allowed one to create a number of interesting linguistic models. Among others, it also guided early attempts of AI specialists to build a machine that would imitate brain functioning, including language. Yet, the results of such projects fell short of their expectations. In particular qualitative, classificatory formalisms that are used to flesh out such linguistic models cannot account for the change of a semantic category, or define the limits on category membership acceptance and thus account for novel categorization of a given exemplar; or account for non-combinatorial novelty, either syntactic or semantic. (What qualitative formalisms are good at, is writing down in an elegant fashion the regularities observed, thus describing neatly typical, common, core aspects of language).

The belief of formal linguists (shared by the main stream of AI specialists) that language can be adequately modelled as a set of patterns combining symbols and independent representation-symbol pairings is based on Putman's (1960) *Minds and Machine* article. In this article Putman concludes that since a Turing Machine can model any machine, it follows that reasoning is independent from the physical make-up of the machine that carried it. Thus, the same cognitive processes, such as thinking and reasoning, can be obtained by various physical set-ups. Therefore, mental states are functions of physical set-ups and the identification of the mental state is independent from the actual physical characteristics of the set-up used to carry them out.

Yet, after considering Goedel's theorem about the incompleteness of axiomatic systems, Putnam (1988) reconsidered his earlier views. Goedel's theorem shows that it is not possible to define every state there is with a limited number of operations specified by a given axiomatic system in a unique way. Therefore, it is impossible to prove that the same functional state of the brain was arrived at by the same procedure that the state of the machine was. Consequently Putman (1988) rejects the position that functional states (e.g., thinking) correspond to the states of

---

[7] Recursive rules were probably first proposed to capture linguistic compositionality in Ajdukiewicz's (1935) categorical grammar. The idea of recursiveness has been popularized in linguistics, however, through Chomsky's generative rules, starting with Chomsky (1957).

any physical machine that arrived at them, i.e., ontological functionalism, for the sake of epistemic interpretation of functionalism. Putnam (1988) states that computers working with different inputs and considering different analogies may have different descriptions of that data than those in a Turing machine formulae and may operate according to different rules therefore mental states cannot be equated with the states of the machine. Putnam concludes that mental states are plastic not only as far as physical machine processing them is concerned but also in terms of calculation. Consequently, the functional cognitive state of the human mind cannot be captured with a set of symbol manipulation rules.

Putnam's theses that mental states cannot be equated with the final states of the machine lets one conclude that hardware differences (differences in physical make-up) imply software differences of the physical set-up (differences in propositions arrived at by the physical set-up) and therefore mental states are constituted by the net of elements co-implying each other, not merely by isolated structures. Putnam (1988) notices, however, that equating a mental (brain) state with the state of a machine would be legitimate if that machine had the make-up of the brain.

2.2 What Type of Mechanism Can Support Language?

If we assume that mental (brain) state cannot be equated with the final state of the machine, unless the machine at stake has the make-up of the brain, then the knowledge of the development and functioning of the brain must inform successful attempts at language modelling.

As pointed out by Brook (1986) human brain with linguistic capabilities is the result of long evolution. Single cell forms of life arose about 3.5 billion years ago, photosynthetic plants a billion years later, fish about 550 million years ago. Next, at intervals of tens of millions years ago came about insects, reptiles and dinosaurs and mammals. Man appeared on Earth 2.5 million years ago and he invented agriculture less than 20,000 years ago. The oldest evidence that man could write comes from about 5,000 years ago and the expert knowledge, when symbolic thinking starts to be needed, has been accumulated for only a couple of hundreds years. If we assume that brain functioning depends on knowledge, how it evolved and how it is organized, it needs to be based on modelling skills once needed for survival such as hearing, moving around in space, interpreting signs indicating location in time. Brook (1986) concludes that "Such skills, in today's brain often unconscious, laid foundations for logical reasoning and understanding." In other words, the functioning of the brain depends on the history of its development, in particular, on specific type of survival challenges faced. As a result, as we well know, the brain is not merely reductionist in its structure and functioning. It consists of units of organization with considerable independence each, (i.e., each of them interacting with different environment) on the one hand, but on the other hand, with non-negligible interdependence between units of adjacent levels, which goes both ways up and down. Note, for instance, that if we cut out part of a tissue,

it will stop functioning—it will loose its vital characteristics. In other words, the brain is a system in the sense of empirical sciences (c.f., Bunge 2003)—its every level is built out of units of the lower level, each higher level is characterized by fulfilling a new function, which both depends on and conditions the characteristics of the units of the lower level.

Importantly, one part of the brain system is hardwired,[8] (as a result of the evolution that has been taking place over thousands of years), another, plastic part of the brain—that hosting language and many other cognitive skills—is a currently developing system. Such hypothesis has been corroborated by the results of recent neurological studies (cf. Karmiloff-Smith 1995; Karmiloff and Karmiloff-Smith 2010; Elman et al. 1997; Gopnik 2007, 2009) who have concluded that children are not merely unfinished adults, but are designed by evolution to change, create, learn and explore. We (and other altricial species) have much longer childhoods than precocial species. As a result of that long childhood involving a considerable amount of learning, we may differ significantly from members of the previous generation: improve our survival, reproduction and care-giving skills over our predecessors.

Karmiloff and Karmiloff-Smith (2010) observes that at the beginning, the brains of human children develop far more connections between neurons than needed, none of them particularly efficient, which specialize in time. She goes on to say that

> Infants are not born with pre-specified modules. Indeed, the infant cognitive system is less differentiated and thus less modular than the adult system, suggesting that modularity is an emergent property of the developmental process. So, domain specificity is not a built-in property of the brain but emerges over developmental time. And even if a modular organisation of the adult brain is the emergent outcome of development, even adult modules should not be viewed in terms of the rigid, static notion of a Fodorian module as outlined above. Thus, instead of the notion that a given brain module can only process proprietary inputs from a specialised domain, neuro-constructivism argues that the brain *becomes* very gradually more specialised over developmental time whereby it narrows its response to the types of inputs a given brain circuit may process, after initially processing many different types of inputs. This is also a relative rather than rigid concept. Indeed, brain circuits that have become relatively domain-specific may still attempt to process new inputs from other domains.

Karmiloff and Karmiloff-Smith (2010) concludes that "therefore a predominant amount of learning and language acquisition goes on in the plastic area of the brain. And the brain keeps developing—organizing in response to the challenges met."

Another strong argument against the thesis that language skills nested in the brain are fully deterministic, i.e., hardwired due to genes—is the evolution of Nicaraguan sign language, which took place within less than a decade. The location of linguistic skills in plastic areas of the brain is also supported by the fact that children with brain damage to the Wernica area of the brain learn to speak

---

[8] approximately speaking.

using the opposite hemisphere. The research into the results of brain damage also shows that among bilingual patients with brain damage who acquired a second language after adolescence, there are such ones who suffer from brain damage causing loss of only one of the languages. No such patients have been found among bilinguals who acquired a second language in early childhood. Importantly, recent developments in the theory of evolution place main emphasis on its developmental aspect. Griffiths and Gray (1994) stress that even genes to operate consistently need resources of the same type at the right moment of the organism development.

The evidence such as mentioned above indicates that the part of the brain supporting cognitive skills and, in particular, language is much less genetically pre-programmed than assumed by formal linguists, advocates of Universal Grammar.[9] It seems that linguistic skills are supported in an important way by the mechanisms located in the plastic area of the brain, as opposed to the hard-wired parts of the brain where the language mechanism instantiating a Turing machine manipulating meaningless symbols representing recursive rules of language would need to be exclusively located. The resultant structures of the brain constitute a system, and are part of a larger system, which, if not genetically determined, must emerge in a natural way, i.e., through self-organization.

2.3 Emergence in Material Systems

Can we explain the functioning of emergent systems? As already mentioned, changes, (e.g., emergence) in concrete things (systems), unlike in abstract systems, can be potentially modelled (thus explained in the sense it is done in natural and modern social sciences). At this time, I would also like to clear a common misconception that emergence is tantamount with no explanation for a given phenomena. This is not true. The phenomenon of emergence and our ability to account for it are two separate issues. It is true that sometimes we do not know how to explain an emergent phenomena, but sometimes we do. For instance the change of characteristics of liquids after their transition from liquids to solids when being frozen can be accounted for in terms of quantum chemistry [for more, see Bunge (2003)].

As pointed out by Bunge (2003), the source of non-combinatorial, emergent novelty in material systems—the change of characteristics of a group of elements which combine together to form a higher level unit requires input from the environment—outside the system itself, (c.f., the input of energy to defrost ice, sucking out heat to freeze water.) A new level will interact with a new subset of the total environment, and the emergent properties of that new level will let it perform a new function. By definition, a new level in a system is characterized by

---

[9] The theory of Universal Grammar requires the existence of an organ in the brain hosting UG that had evolved before language did. This, however, is evolutionary implausible.

properties absent from its parts fulfilling a new function. These new properties account for the interaction of the new level with different environments. Therefore, a compound unit of the system interacts with the environment with which its parts do not).

As explained by Bunge, the combination of parts leading to the creation of a new level can emerge in a system either through artificial assembly, or self-assembly. Car making illustrates artificial assembly. (A car has a novel function of transportation, which is missing in any of its parts). Examples of a self-assembly process resulting in novelty are vapour freezing or the coalescence of a street gang. A typical consequence of emergent novelty, especially in self-assembling systems, is mutual inter-level interaction. (More precisely, this is additional downward interaction of a higher level unit with new properties (and thus interacting with the new environment), onto lower level units and their old, "lower-level" properties. For instance, gases are characterized by temperature and pressure, none of which characterizes the constituting atoms. Yet, under sufficient pressure applied from outside, or if temperature is lowered sufficiently due to external energy sucking out heat, vapour will combine into ice, in the process changing the characteristics of the very atoms initially forming it—the lower level elements.[10])

Thus, understanding a system (in the sense of the term in natural sciences) means knowing its structure, composition (elements), the environment with which its elements interact, and the mechanism supporting its function in that environment. The interest in mechanisms explaining self-assembling (or self–organizing systems) has recently picked up. It has been studied, among others, by a newly developed field of synergetics, a branch of system theory, which models self-organization through HOT (highly optimized tolerance) mechanism (see Haken 2010). Another mechanism of self-organization, called self-organizing criticality (SOC), has been proposed by Bak et al. (1988). Both types of models allow one to form testable hypothesis and corroborate the self-organizing origin of the objects under study.

2.4 Has Language Got Characteristics of a (self-organizing) System in the Sense Just Defined?

Based on the assumption that functioning of software depends on hardware, as proved in Puntam (1988), and in view of undeniably systemic (and self-organized) nature of the brain, language can be expected to a have a system structure with considerably independent levels interacting with different environments, and with higher and lower levels influencing each other both ways, Has it really?

---

[10] The artificial and natural assemblies can also combine. The process of book production starting form farming trees can serve as an example.

Language is based on meaningless phonemes, some of which combine into morphemes and words with the emergent property of meaning. Meaningfulness and the frequency of use of specific words, in turn, along with biological capabilities of human beings, influence (interacting down) the shape of phonemes which need to be contrastive enough and easy to pronounce when one follows another. Next, some lexemes may combine together into phrases and instead of being related to a separate referent each, both may refer to the same referent and enter a new relation with each other, e.g., being a modifier and modified. A group of lexemes may also convey an emergent sense which is not a result of a simple addition of the composite lexemes' senses, i.e., the lexemes as a unit may interact within (refer to) a different environment (referent) than any of the lexemes composing a given unit. Consider, for instance, the relation between the meaning of the items *post*, the meaning of the item *card* and that of the compound *post card*. The meaning of the compound *post card* has a representation which cannot be calculated from the representations of the meanings of its components, but involves non-combinatorial semantic novelty based on information from beyond the system—the world of post offices. We may also say that the function of the phrase *post card* influences (modulates, interacts down onto) the meaning of the lexeme *post* as well as that of the lexeme *card* as used in the phrase *post card*.

Lexemes and phrases may combine into units capable of communicating something about something else, i.e., conveying propositions with the emergent property of truth value, i.e., being or not being true. Finally, a sentence used in (interacting with) a given situational or textual context acquires an emergent property of having an illocutionary force. For instance, the sentence *could you open the window*?, used in the situation when the speaker clearly would like to have the window opened, but for some reason does not want to do it himself, constitutes a request for opening the window. Because of the frequency of using similar questions in contexts with the differentiation frame indicating the desire that the addressee fulfils the action expressed by the respective verb, "could + (someone + do something)", the construction becomes correlated with a request to do the action expressed by the predicate directed to the listener, in addition to what this construction was correlated with so far, i.e., the question concerning the capabilities. Additionally, a specific sentence e.g., *Could you visit Jack at the hospital tomorrow?* uttered in a specific situation, adjust the referential content of the respective components. For instance this sentence assigns to the lexeme *visit* a specific meaning of a prototypical visit in the hospital, which differs from visiting healthy people at their own homes.

Next, sentences can be further organized into dialogs, or paragraphs, sections, chapters and texts, which again exhibit a collective purpose absent from individual sentences, which again may adjust their respective messages. Note, that this hierarchical organization is certainly an emergent phenomenon. Proto-languages must have had the form of single 'word-functional sentence'[11] correlations. With

---

[11] I use the term 'functional sentence' to refer to a sentence with a specific illocutionary force.

the growing complexity of technological and social life, the number of single 'word–functional sentences' grew to the point, that such a language became difficult to acquire and use. As pointed out by Kwapień (2010), simulations of the use of such proto-languages show that as the number of words reaches a certain threshold value, the number of mistakes in language communication grows to the point of considerably lowering its efficiency. The emergent grammatical organization allows a linguistic community to lower the number of words needed, while increasing the communicative efficiency of the language they speak.

In the above view, language is a system in the sense in which the term *system* is used by Bunge (2003). The examples just presented illustrate both relative independence of linguistic elements on each level, as well as the fact that they are influenced by higher levels. Importantly, the emergence of levels with emergent properties in language due to the interaction of higher levels with new environment and in a novel way (functional self-regulation) provides a space, in which to look for semantic novelty. There is no room for such novelty on a Turing machine view of language (i.e., language being a closed, and purely reductionist system). No wonder that although modulation and syntactic flexibility are widely observed, they remain unaccounted for in the latter framework. On the contrary, the emergentist view of language is, by definition, predisposed to model linguistic change over time (language development), as well as (in particular) the adjustment of linguistic meaning during its instantiation ('language efficiency', to use Berwise and Perry's (1983) term).

A strong argument for language being a self-organizing system comes from considering its quantitative characteristics. Recently the hypothesis that the source of self-organization in language comes from the mechanism resembling natural selection, during which process certain types of constructions and lexical meanings are selected for on economic bases has been considered again. (The hypothesis that language has been brought about by some sort of economy was earlier advocated by Zipf (1935), but later criticized on the grounds that the amount of calculation needed would exceed the capabilities of the brain[12]). Next, Zipf (1949) showed that there are a number of relations between certain quantitative characteristics of linguistic corpuses which follow power laws. Today we know [cf. Haken (2010), Bak (1988, 1996)] that Zipfian laws express[13] the characteristics of self-organizing complex structures, which cannot be derived from "first principles". The degree of conformity of Zipfian laws in a number of languages leaves no doubt that the assumption of language being a self- organizing structure makes a lot of sense. (By the way, since Zipfian laws concern self-organizing structures as such, such laws have also been attested in a number of other non-linguistic phenomena, such as the

---

[12] This argument is no longer valid if we transfer the burden of making economic choices from an individual to a natural selection process taking place in a community.

[13] Strictly speaking, Zipfian distribution could also be the result of a relatively simple, statistical processes. Yet, along with the information about the hierarchical structure of the object of study and the processes involved, establishing power laws relating some characteristics of that object is enough to indicate its self-organizing origin.

ranking of cities by size, income ranking, corporation sizes, the revenue of a company as a function of its rank, the distribution of the earthquakes as a function of the rank of their magnitude, and many more). Also, the association studies modelled by Bruza (2009, 2010) with Quantum Mechanical formalisms, which formalisms reflect some self-organizing principles, corroborate the hypothesis about the self-organizing nature of language. Interestingly, according to Bruza (ibid.), the shapes of many of distributions tested by synergetic linguists turned out to be describable even more adequately with the help of certain quantum mechanical formalisms.

2.5 Problems with Grounding Basic Encodings in the Reductionist (Saussurean) Systems

An important argument for the need to view language as a system (as it is understood in empirical sciences), comes also from considering basic encodings. Basic encodings are the atomic elements of meaning in the model of language based on the idea that language is a system in Saussure's (reductionist) understanding of the term. Saussure refined Franz Bopp's imprecise idea of a system, and to him a system is a context-free and downward-interaction-free (purely reductionist) non-developing abstract structure. To define language adequately in such terms (as it is done in formal linguistics), such systems require that two assumptions be met. The first assumption is that the human brain can be well modelled by a Turing machine. The second assumption is that we all have access to some common meanings of basic encodings. Yet, Putnam challenged both of these assumptions. In addition to challenging the former thesis, which he did in Putnam (1988) as already mentioned, Putnam (1975) proved there is no way of explaining the representation of basic encoding meaning from person to person, no way of conveying the meaning of basic encodings.

A positive solution to the latter issue, in line with the concept of a system in empirical sciences, comes from Bickhard and Campbell (1992a, b). Bickhard and Campbell (ibid.) propose that linguistic representation be expressed in terms of a control system, i.e., what can be shared among individuals, are functionally established categories. Therefore, at least some of the original linguistic categories must have been formed prior to, or simultaneously to the formation of individual linguistic representations.

There is no doubt that such functional non-representational categorization indeed takes place during the evolution of organisms. Functional categorization, which would be quite sophisticated cognitively, is even possible in creatures without respective cognitive structures. Ants, for instance, have been long known to be capable of finding the shortest distance between their dwelling and a source of food, and obviously no one would credit them with having a representation of

the concept of the shortest route.[14] Next, such externally (functionally) defined categories may allow human community members to order their individual representations relative to each other to parallel the relations between the functionally established categories and thus ground the basic encodings. As a result, separate individuals will not necessarily share absolutely identical representations, but these somewhat idiosyncratic representations will let them refer to the same functional categories.

The hypothesis of functional substantiation of representation has been corroborated by neural evidence coming from Lin and Tsien's research. Lin et al. (2005, 2006) identified neural functional cliques underlying representational neuronal structures. Lin (2007) in turn, reports the discovery of neuronal cliques in mice responsible for mice's identifying a potential nesting object. These cliques are not activated when a real nest is covered with a transparent piece of plastic, but are activated when a mouse comes across a red, plastic cube with a sufficient indentation in its top to serve as a nest. Thus, depriving an object of its capacity to fulfil its typical function results in mice's inability to recognize its physical representation, while on the other hand, presenting a mouse with a highly a-typical object, yet one with the capacity to fulfil a given function, makes the mouse classify the given object as representing the given functional category despite its lack of appropriate physical characteristics (physical representation).

Not-surprisingly, functional interactions precede formation of basic representation both onto- and philo-genetically. In the last decade there has been a growth of interest in the way children acquire functional (pragmatic) linguistic skills along with, or even to some extent prior to acquiring semantic and syntactic ones. The point may be illustrated with the following example. A toddler has been reported to use the expression. 'This is …' as a request to have the object pointed to opened. He did it evidently without being aware of the semantic content of the words. His mother acknowledged that before opening something for the child, she used to explain what the object was using the structure *this is* "*x*". The child used the phrase *this is…* to functionally categorize the states of containers into "closed" and "open" ones, and next he correlated the phrase mentioned with the change between these states. The functional understanding of the phrase clearly preceded in this toddler the understanding of the sentential meaning and the meanings of components. (The same toddler used the phrase *Once upon a time* as a request to be read to, again apparently without being aware of the semantic content of the words).

Basic differentiation frames that initially subcategorize experience, using Campbell and Bickhard's terminology, allow one to ground basic encodings— attach labels. This position corresponds to Mey's (2001) role of *pragmemes*. Mey postulates the existence of *pragmemes*—situations motivated by social functions—

---

[14] Ants when walking, leave scent on the trail. The group of ants in search of food that has found the shortest route will cover the distance between their anthill and that source of food the largest number of times, making that trail most smelly. A new group of ants which has just left the anthill to search for food, will chose the most smelly path, which is the shortest one.

to argue that language interpretation takes place as if "outside- in" or downward, we might also say. For instance, one needs to know the relevant *pragmeme* (differentiation frame) to interpret e.g., *red eyes*, as already illustrated. In other words, knowing the functional subcategories of the relevant *pragmeme*, is prerequisite for assigning the semantic representational meaning to its parts. Capone (2005, 2006) shows us that the recognition of a *pragmeme* is necessary for the very concept of *explicature* to make sense, thus the concept of *pragmeme* cannot be excluded from forming a coherent model of linguistic encodings. [Capone (2010) further proves the utility of *pragmemes* when analyzing reported speech].

Initially, classes of pragmemic options which differed functionally (e.g., "chase it" vs. "stop chasing it") must have been differentiated by single signs. These categories, in turn, could have been used to subcategorize options defining other *pragmemes*, initially also marked with single signs, resulting in simultaneous development of a system and the temporary specification of individual meanings of elements in those *pragmemes*. As a result, the utterance subparts could assume approximate individual, temporarily encoded meanings, which are used next by these individuals to refer to functionally identifiably, thus commonly shared, *pragmemes*. These *pragmemes*, in turn, are used to specify the utterance meaning of the individually encoded subparts of the symbolic structure used on a specific occasion. At the current stage of language development, individually represented "conventional"[15] word meaning may serve to help identify *pragmemes*, which reflect back on the meaning of the constituents of the linguistic construct, which had identified the *pragmeme*. This is done by substituting the respective individualized "conventional" meanings of the components of the construct used to identify the *pragmeme* with the (utterance) reference meaning, which is determined with help from options identified by the function of the *pragmeme*. Thus, structure, representation, and function are inseparably interrelated to form language.

Note that the basic encodings identified through basic interactions will correspond not to semantic primitives assumed in formal semantics, i.e., the set of the most atomic meaning primitives needed to code all the lexicon, e.g., Katz componential analysis, but rather to the ground breaking idea of the universal lexicon posed by Wierzbicka (1972, 1985). Wierzbicka postulates the existence of "inborn *lingua mentalis*", an innate mini language of cognitive concepts, both lexicon and syntax, that allows a child to make a functional sense of a situation and speaker's intentions. The approach advocated, although denying the innateness part of Wierzbicka's hypothesis, shares with it its essence—the emphasis on functional origin of basic linguistic categories. The functional categorisation, along with general perceptual skills, and basic situations of usage results in the type of universal *ligua mentalis* identified by Wierzbicka. Importantly, however,

---

[15] I use the term "conventional" here in the sense of some sort of representation which is an average of the individual representations of the same functional category in a given linguistic society. .

Wierzbicka's empirical cross-linguistic investigations does not *per se* imply that the universal *lingua mentalis* is indeed inborn. It could just as well be derivative as postulated by the "outside-in" approaches, which is supported also by recent developments in the theory of evolution. The authors of the Developmental Systems and Evolutionary Explanation, Griffiths and Gray (1994) explain that there is no formal difference between inborn and developed.

> Developmental systems theory rejects the dichotomous approach to development: The genes are just one resource that is available to the developmental process. There is a fundamental symmetry between the role of the genes and that of the maternal cytoplasm, or of childhood exposure to language. The full range of developmental resources represents a complex system that is replicated in development. There is much to be said about the different roles of particular resources. But there is nothing that divides the resources into two fundamental kinds. The role of the genes is no more unique than the role of many other factors.

To sum up, fully subscribing to the resultant universal lexical units found by Wierzbicka, I see the proposition postulated in this chapter as expressing the gist of Wierzbicka's insight and intuition (initially formulated 40 years ago) in the language of the contemporary paradigm of empirical science. What also supports the functional focus of the proposition presented in this chapter is Wierzbicka's (2010) postulate of linguistic molecules, i.e., the lexemes containing more than one semantic prime combined together due to a specific *pragmeme* in which these primes originated. Wierzbicka explains the concept of a linguistic molecule with the following words:

> In addition to semantic primes ('atoms of meaning'), many NSM explications also rely (in a limited way) on 'semantic molecules', built from primes, especially in the area of concrete vocabulary. In particular, body part concepts often function as 'semantic molecules' in the meaning of verbs of physical activity, such as *walk* ('legs', 'feet'), *lick* ('tongue'), *bite* ('teeth'), and *eat* and *drink* ('mouth') ....[while -DZ] color words rely, to a considerable extent, on environmental and bodily molecules such as 'sky', 'sun', 'day', and 'blood', as well as on the molecule 'color'.

Grounding language in a material system categorized by its function (supported by some mechanism) and environment solves yet another cornerstone problem in linguistics—that of linguistic categorization. Aristotelian definition of the linguistic category, in addition to requiring everyone to share a common representation of a category, imposes limits on future applications of the symbol of a given category and precludes metaphoric usage. Introducing a prototype as the pattern defining a given category does not solve that problem. Therefore, Roch [in Lakoff (1987)] renounced her earlier claim that prototypes define category membership and stated that a model of a linguistic category must reflect the phenomenon of prototypes. The proposition that individual category members fit the category well, or not so well, as proposed by Lakoff (1987) and Langacker (1987), results in every item being at least a bad exemplar of any category, which is not good either. A proposition to model linguistic category with fuzzy sets must be rejected on the same grounds as the previous one. Therefore, these models of a linguistic category cannot model linguistic compositionality adequately, either. In Sect. 3, I propose a

category model within the empirical sciences paradigm advocated here, one coined as a developmental (emergentist, self-organizing) system, which does not suffer from the above difficulties.

2.6 How to Model Language in the Empirical (socio-natural) Paradigm?

So far we have concluded that natural language as a semiotic system is closely (integrally) related to evolving material systems (part of the plastic areas of the brain), and therefore language needs to be treated as a result of a process in which the evolution of form and representation are inseparable and determined by a mechanism dependent on the social function of language. In other words, we argued that language as a semiotic system is inseparable from its function mainly as a tool for communication. Therefore, we may repeat after Bunge (2003) that the form and content of language are the integral[16] parts of the history of specific purposeful interactions between linguistic community members involved in speech acts. During that process certain aspects of participant's brain state which was evoked as a response to bio-socio-environmental conditions, become correlated with the symbolic elements of language through changes to some plastic areas of the speaker's brain. So natural language as a semiotic system is a reflection of a dynamic system of individual speech acts produced by human agents located in and interacting with social situations. Grzybek (2006: 12) expressed that idea by saying: "Genesis and evolution of these systems must be attributed to repercussions of communication upon structure." In other words, an explanation of the existence, properties, and changes of linguistic, (more generally speaking, of a semiotic) system is not possible without treating it as an aspect of the (dynamic) interdependence between structure and function, or in Bunge's (2003) language, without understanding the mechanism supporting that function of the given semiotic system in a specific environment.

This is so because, as Bunge (ibid.) stresses, the changes in material system components involved in their combining to become a higher level unit **are always the result of some input from their environment—outside of the system.** Consequently, in an empirical paradigm, the material system (neural connections) supporting linguistic behaviour of an agent is determined not only by bio-psychological principles (as implied by Chomsky (1986) style view of language), but also by external, socially established principles. In other words, the perspective on language just advocated makes linguistics necessarily an inter-science that

---

[16] Bunge (2003) opposes an integral structure, such as the one present in a cardiovascular system, to a combinatory structure, such as that present in a car. While the latter one was put together from parts, the former one evolved by subsequent evolutionary steps and cannot be substituted fully by plastic elements. It is integrally related to all other elements in a human body and specific history of evolution. .

straddles biology, psychology, and equally importantly, social sciences (Bunge 2003: 63).

In view of the above, linguistic explanation is not likely to be possible by means of casual relations. Instead, Altmann (1978) proposes that a likely strategy is looking for a functional explanation similar to ones offered in biology. Therefore, as Altmann continues, it is plausible to assume that language is an aspect of a self-organizing and self-regulating system of members of a linguistic community engaged in communication promoting their co-existence—a special kind of a dynamic system with particular properties brought in line as a result of some sort of economy related to the communicative behaviour of its participants. Or as Grzybek (2006) puts it—"the economic result of communicative processes".[17]

In the empirical paradigm, the assumption of the self-organizing nature of language implies the fact that the laws modelling language must not have reference to specific objects but be statistical. Another reason for statistical nature of linguistic laws in the empirical paradigm is that since the formation of linguistic objects depends strongly on the history of contingencies and exact data is not available, only statistical hypothesis can be formed. Similarly, Bak (1996: 10) talks about life, (which is also characterized by the variability of its exemplars resulting from its self-organizational character): "A theory of life is likely to be a theory of a process, not a detailed account of utterly accidental details of that process such as the emergence of humans."

2.7 Epistemic Concerns

Note also that the perspective outlined so far is the result of considering the manner of gaining knowledge about the phenomenon studied (language), i.e., in establishing (linguistic) facts—(recall grounding basic encodings and inadequacy of interpreting the brain states as the final states of a Turing machine). We initially learn about the environment, which consists of physical and social components, via physical interaction. Such knowledge can be shared whenever functional interactions with the environment can be shared.

The previous approaches with explanatory ambitions, (such as Chomsky's initial proposal that the biological make-up of man generates sets of syntactically restricted uninterpreted strings of symbols along with a separate list of symbol-representation pairings) required taking God's eye view into linguistic meaning on the one hand, and disregarding the make-up of the machine using language (when equating its final states with mental states), on the other hand. Both of these assumptions have been proven to be false, as reviewed earlier.

---

[17] This view may remind of 19th century concepts in linguistics, but here, language is not viewed as an independent organism. The "organism" considered here is not the semiotic system *per se*, but a linguistic community with the semiotic system (language) being an aspect of its behavior.

The epistemic problem with models based on God's view of basic encodings in linguistics resembles the situation faced by physicists who proposed a cosmological model of the atom. It turned out that the parameters needed to test the model could not be in fact measured. (The measurement of the position of an electron with an already measured momentum on the orbit around an atom, which will typically be established by shooting a photon with a known velocity into the electron, will affect the momentum of that electron to a degree which cannot be considered negligible. Thus, the measurement of the position of an electron will deprive us of the knowledge of its momentum). The epistemic impasse in physics just mentioned resulted in a search for a totally new paradigm of the description of the micro-world—quantum mechanics—an approach guided by epistemic concerns of gaining knowledge, a theory built on new measurable concepts. The same seems to be taking place when proposing a developmental psycho-social model of language. Thus, it would be hard not to agree with the following words:

> Most reasonably, language lends itself to being viewed as a specific cultural sign system. Culture, in turn, offers itself to be interpreted in the framework of the evolutionary theory of cognition or of evolutionary cultural semiotics, respectively. Culture, thus, is defined as a cognitive and semiotic device for the adaptation of human being to nature. In this sense, culture is a continuation of nature on the one hand and simultaneously a reflection of nature, on the other—consequently, culture stands in isologic relation to nature and can be studied as such.
>
> Therefore, *langue* viewed as a cultural sign system cannot be seen as being ontologically different from nature because the nature we know can be observed only through culturally biased theories and perspectives. ….Thus, both culture and nature are cultural constructs co-determining each other Grzybek (2006:8).

3 Modelling Language in the Empirical Paradigm: A Dual Model of Linguistic Form and Content Correlation

Within the empirical (socio-natural) paradigm outlined in Bunge (2003), a model of language, in addition to specifying linguistic composition and structure, must describe the mechanism that (creates and) supports it. Relevant mechanisms must account for both the process of correlating symbols with representations (lexicon) (and constructs with representations) and for ordering symbols (account for syntax). In this section, I propose, first, a general mechanism of correlating form with representation, which in fact must reflect modelling the process of linguistic categorization. Second, I shall illustrate the essence of the mechanism proposed in this chapter by illustrating briefly "the soft way" of approaching a wide range of selected linguistic problems from the perspective advocated.

3.1 A Qualitative, Developmental Theory of the Form-
Meaning Correlation Process

To account for both the stability and flexibility of meaning-form correlation in natural language, I postulate that the form and content of linguistic items are correlated via two largely independent, but co-dependent mechanisms situated in the central nervous system. The first mechanism, which I call the 'encoding mode of language use', ensures the stability of the linguistic system. *Via* this mechanism, with every use of a given linguistic item, the cases of attested similarity between the brain state considered and the brain states correlated with the given sign so far are added up (wired up) in the plastic brain area. The brain states reflect the socio-environmental stimuli filtered through the human perception and cognitive systems—represent both the object referred to and the relevant aspects of the respective situation. This encoding mechanism of categorization can be modelled in a sort of Aristotelian way.

With time features encoded more often, statistically speaking, correlate with the given linguistic sign/pattern strongly enough to be recalled by the next prompt of the given sign/pattern. The encoded representations are idiosyncratic, resembling Aristotelian representations of conventional meaning (sense). They differ somewhat between individuals, depending on their individual history of language acquisition defined by functionally established categories. Although no identical meaning corresponding to the same sign can be represented in the brain of another individual, yet the second mechanism, (which I call a selective mode of language use, and which I introduce below), ensures that the individually encoded meanings correlated to the same sign can be functionally equivalent for different individuals, i.e., different individuals using their idiosyncratic representations correlated with the same sign will make predominantly the same choices between functionally provided categories. The set of properties occurring in the sum of the representations correlated with a given sign for all members of a given community frequently will correspond best to social, conventional meanings assumed as the encoded value of that sign both by minimalists and contextualists.

The proposition outlined so far (the encoding mode of language use) presents a generally accepted rough picture of conventional linguistic meaning. The representation of the sense of a linguistic item reflects, roughly, an Aristotelian model of a category and the mechanism of its installation in the brain is compatible with a received view that learning a fact is equivalent to the emergence of a specialised system of neurons held together by excitatory plastic synaptic junctions, which arises after a sufficient number of respective excitations. That rather uncontroversial mechanism is responsible for the creation of a core, prototypical part of language. It is hypothesised to be grounded in the plasticity of the brain, which is the key to behavioural and social plasticity.

What makes my proposal different from such classical propositions is that the encoding mechanism of form-content correlation described above is integrally combined with, (i.e., results from) and influences, the second mechanism of form-

and-content creation, which I call a selective mode of language use. This mechanism is presented below. The two mechanisms jointly allow one to model the form-meaning correlation process in a more satisfactory way than previous models did. They let us avoid the major pitfalls of conventional approaches based exclusively on the former, encoding mechanism of sense installation.

The major problems faced by mechanisms of categorisation having solely Aristotelian roots are the following ones. First, the Aristotelian type of definition requires precisely delimited and ideally shared by linguistic community members core meanings, as well as requires establishing the limits on the allowed departure of the item being classified as a given category member from the set of properties defining that category. None of these requirements can be met in the case of typical linguistic categories. Second, Aristotelian definitions of a category do not model the development of meaning and cannot account for modulation and flexibility involved in linguistic compositionality[18] or instantiation. (However, liberal we decide to be in posing the limits on the allowable departure from the prototype, we may eliminate a future use of the given sign that will require an even more considerable departure from the standard. On the other hand, if we allow any degree of departure from the set of properties defining the given category as proposed by a fuzzy set approach, or Langacker's' 'schematicity' or Lakoff's 'motivation', then virtually anything can be considered to be a bad member of any other category).

The essence of the second mechanism—the selective mode of language use, (the second component of the mechanism supporting the meaning-form correlation process)—builds on a commonly accepted observation that the human brain encodes **relational** (co-occurring, associative) meaning in addition to encoding core meaning, (i.e., meaning that enumerates the properties and functions of the named object). The co-occurring, (relational) meaning that I talk about, (similar to that studied by Leibniz), is close to what psychologists refer to by the term "association", or more recently "cueing", or "priming", (which can be said to reflect the co-occurrence of elements in *pragmemes*).

Unlike the case of encoding mechanism, which operates as if "in a vacuum", (i.e., is context free), and as a result provides a representation of the concept that is correlated statistically most frequently with a given linguistic item, the selective mode of language use starts from what we already know about a given situation ["a situated speech act" to use Mey's (2001) terminology]. The relevant situational information, associations formed from cues in the verbal text, let one make predictions about the content, (interpretation) including the function, of the item to come next in the linguistic construction being formed, (or interpreted) before even considering the encoded value of the form actually used). For instance, consider interpreting the item *shrimps* in the utterance *I love shrimps* uttered when sitting with friends around a table during a party. The sentence fragment as | I love… |

---

[18] For instance, the Polish for *guinea pig* is "*swinka morska*" (literally: a sea piglet), which animal, of course, is neither a pig, nor has anything to do with marine life.

pronounced in the given situation lets us guess that the next item in the linguistic construct being formed will be a name of a food item on the table. It could also be, however, a food name qualified with *fresh*, *cooked by the host*, etc. Consequently, we end up with a set of orthogonal possibilities defined by the function of a given element within a given situated speech act, (which could potentially fill up the slot of the item under interpretation), each with a specific probability of its occurrence. A set of such options, along with the probabilities of their occurrences, will be further referred to as a communicative field. Now, in its selective mode, the function of a linguistic construct whose contribution we are assessing will not be to add its encoded information to the interpretation of the sentence being revealed (act encodingly), but **to use its encoded information to select among the options in the communicative field.**

It is postulated that one selects with a given form among the available options (i.e., out of the communicative field), the option whose encoded content is more like the encoded content of the form used for selection than the encoded content of any other option. Since the selective mechanism selects options out of any set of data, including sets of novel data, it overcomes the major problems of the Aristotelian model of linguistic categorisation mentioned earlier. The selective mode of language use can select an item, which is not a given category member according to the Aristotelian definition. This way it does not impose any limits on category memberships, and yet gives an unambiguous result in any specific situation. Consequently, it can model non-combinatorial novelty, instantiation, modulation and syntactic flexibility, as well as allow change of an encoded meaning of a given linguistic category Besides, while Langacker's (1987) or Lakoff's (1987) propositions, in which category members may resemble a category pattern only partially, leads to the problem that any item may be a bad member of any category and categorization breaks down, the selective mode of language use poses no such problem. This is so because the communicative field enumerates all possible meanings, and the selection can be carried out among these meanings in a rigorous manner, i.e., with help of a Supervised Learning (SL) technique. (These latter algorithms are used for instance, for classifying medical images). Finally, a selective mode of language use concerns the organization of individual representational data, thus avoiding the problem of deriving shared basic encodings—it does not require different individuals to share identical category patterns for them to make the same selection out of a set of predefined options. For instance, a person for whom a prototypical dog is a German Shepard Dog and a prototypical cat is a Siamese will classify 'a Poodle', or 'a mongrel' as a [dog], and not as a [cat], just as a person for whom a prototypical dog is 'a dachshund' and a prototypical cat is 'a tabby cat'.

As already mentioned, the two mechanisms just proposed are integrally related because the content that was identified selectively with a given symbolic item, (which can be quite novel and not-fitting (in the Aristotelian sense) the current encoding representation of the symbolic item considered), is added in the encoded (statistical) meaning of that item, thus affecting it. Therefore, after every occurrence the selected (utterance) content affects encoded content of the linguistic

items used. The current encoded content, in turn, influences the result of the selection among the option of the forthcoming communicative field.

Importantly, there need not be a direct, surface, correlation between syntactic and semantic structure of a given sentence uttered and the options of the communicative field on which the selection process operates, because the options are generated for a specific *pragmeme*. As elucidated in *the Extended Functional Analysis*, (Zielinska 1997), the sentence *She is a ski instructor* written in a letter of application for a job requiring a reasonable, but not outstanding, physical strength (such as a summer camp councillor), i.e., when a physical fitness of the candidate needs to be assessed, will mean roughly { [the degree of physical fitness] = that of a ski instructor}. In other words, the communicative field stipulates that the predicate *is a ski instructor* selects among [possible degrees of fitness relevant for the job at stake]. Note also that selecting and assessing may take place simultaneously. (see Zielinska 2007b for examples).

To recap, one might say that the model of categorization proposed combines what Stainton's (2010) terms System Perspective with Use Perspective, (thus in a way combines early with late Wittgenstein (1961, 1963) views). While, as Stainton (2010) points out, there is an ontological gap between the two Perspectives derived within language as an abstract system view, the model proposed derived within an empirical paradigm seals that gap.

3.2 A Preliminary Qualitative Illustration of the Mechanism Postulated

To illustrate qualitatively (in "a soft way") the mechanism just postulated, let's consider the contribution of the item *red* as used in the phrase *red car,* assuming that we have never seen a red car before. In this case, we recall what we identified as 'blue cars', 'black cars', and 'green cars' in the past, to form expectations, as to which parts of the car can be of different colours. Only now can we interpret the contribution of the adjective *red* to the meaning of the phrase *red car*. Note, that a Martian who speaks English, but who has never been to Earth, (he has observed cars from such a distance from his space ship that everything looked grey to him), will not be able to understand, what 'a red car' is even if we show him a red paint in a jar and point to a car on Earth. Neither will he draw a 'red rose' properly, not having known what flowers on the Earth look like colour-wise.

Next, let us also have a look at the contribution of the item *red* to the phrase *red barszcz,* which designates a type of Polish soup made of beetroots (a soup of certain taste, which has a crimson, or sometimes brownish colour). On the selective mode of language use 'red barszcz' is pointed out due to the fact that it is more red than the other Polish 'barszcz' is. The other type of Polish 'barszcz', 'white barszcz', is yellowish/grey in colour. Thus, the encoded value of the item *red* allows one to easily differentiate between these two options defined

functionally in the Polish culinary world. Redness, on this occasion, has a referential content of crimson, but even more importantly, indicates a certain taste provided by the respective culinary *pragmeme*, in which the parameter of taste is crucial.

Note, that this time the taste has not been encoded for the future uses of the item *red* with other nouns for statistical reasons. Since a similar taste is not shared by other red food items, it is statistically insignificant in other contexts. If some selected characteristics were more common, it could have gotten encoded, as it happened with the item *green* to encode 'not ripe' in relation to fruit and vegetables via the selective mechanism postulated above.

Obviously, on some occasions both selective and encoding modes of language use may offer a separate relevant interpretation each. For instance, when my son was recovering after his appendix had been removed, a new doctor approached me asking "Are you the mother of the boy with appendix?". "No"—I answered. "I am the mother of the boy without an appendix"—I joked. Selectively, in the situation described, the phrase *a boy with appendix* obviously chooses the patient being treated for appendicitis, which singles him out from other patients in the ward. Encodingly, however, a person after surgery no longer has an appendix, and he is not a person with an appendix.

To finish the illustration of the functioning of the mechanism postulated, let me apply it to motivating the thesis that epistemic modality originated from deontic modality and compare its effectiveness with some alternative account. The alternative account that supports the above claim, and which I am going to refer to, comes from Sweetser (1990). Sweetser argues for that claim by stating that the concept of "forcing someone to do something", present in deontic modals, is metaphorically extended onto the content "some evidence forces the subject to reach a conclusion, to have a thought". A problem with such an explanation is that it does not say a word about the mechanism that causes the metaphorical extension postulated. In other words, Sweetser describes what happened without explaining why, i.e., by what means it happened. Thus, Sweetser's account is not an explanation in the sense of the word "explanation" used in sciences, but simply a description, or to use Bunge's (2003) term: "subsumption of particulars under generalization" at best. I propose that that the claim needs to be argued in a different manner.

On the model proposed, the mechanism explaining how deontic modality gave rise to epistemic modality can be illustrated in the following way. Telling someone to do something happens most often in situations in which one has power to enforce the order. Therefore, most of the time, the action at stake will be actually performed and thus what will also became pragmatically correlated with the modal construction "A must do y", is "a big likelihood that A will do y".

Now, let us look at the following illustration of such a *pragmeme*. Parents are worried about their daughter's being late when coming back home from a university for a week-end. Mom says: *Jane hasn't arrived yet*. Now, the *pragmeme* under consideration is defined by a discussion concerning parental worries. The major issue, the function of the *pragmeme*, is to decide whether Jane is all right, or

whether something happened to her. Her parents want to consider all likely situations, which could have retained her and which do not involve any mishap before calling the police and hospitals.

In such a context as described, when we mention 'trains' and 'lateness' in a modal construction "The train must be late", since the option of forcing the train to be late is unavailable as a plausible interpretation, given the correlated issue of the likelihood that the subject will perform the action expressed by the predicate in the construction considered, there are two feasible options having to do with "trains, lateness and likelihood" and thus building up the communicative field: (1) it is likely that the train was late. (2) It is not likely that the train was late. Since, as just recalled the construction "A must do y" expresses the likelihood of action y taking place, that likelihood-feature of the construction considered will select the former one between the options 1 and 2 specified externally, i.e., it will select the option "It is likely that the train is late." This unconventional usage must have happened for the first time purely selectively.

The following situation may serve as another example of a purely selective usage of language. A four year old boy was reported to use the sentence "Open the light" in a garage without any windows. He used that sentence, which encodingly lacks any logical meaning, to select the message "open the garage door to let the light in." Unlike that latter construction that does not stand a chance of being commonly needed, the novelty involved in the pattern 'The train must be late' to indicate "the likelihood of the action being performed" was frequently useful and spread by being repeated (imitated), thus eventually inscribed in the neural system of the brain.

The above reasoning shows qualitatively, what could have happened. It shows that according to the mechanisms postulated, it is possible for epistemic modality to have arisen from deontic modality, but not precisely how it happened. In other words, the above example shows that it is logically possible for the transition to have happened via a selective mode of language use—but of course not that it did take place precisely in the circumstances presented. This is, however considerably more than labelling that process as a metaphorical extension.

3.3 Analogy

The essence of the mechanisms postulated can be also presented in the following way. While in the encoding mode of language use speakers use pre-established meanings, as if Lego Blocks, and place them together one next to another on an empty table, in the selective mode of language use speakers use encoded content to select out of options. The second mechanism assumes that the table is not empty, but that the non-verbal situation accompanying the utterance, as well as the building blocks placed so far, fill the table top up with shapes expressing viable situations, which form a specific discrete space of options. Now, instead of gluing the new building blocks to be interpreted to the ones already placed on the table,

on the selective mode of language use, the new blocks serve to identify elements of the space just described, very much the way two points will identify a specific line out of all possible lines. Consequently, the content of the lexeme used selects a part of that table top space and as a result of selecting it, the value of that linguistic construct on that occasion gets adjusted to the relevant part of the table top space.

The "building block" employed in the selective mode of language use, in addition to carrying "an encoded" "core"[19] content, carries contextual information from past uses—sort of the memory of *pragmemes*, in which it was used previously. This contextual information fills up (modifies, enriches) the space, in which the succeeding "block" will be placed. Thus, "building blocks" used selectively behave more like electrons in an electric field rather than pieces of plastic. Every electron generates an electric field around itself and therefore, if we place an additional electron in its vicinity, the behaviour of that new electron will be affected by the field (as well as the second electron will influence the former one via the field it generates).

The difference between the two modes of language use introduced in this chapter also resembles the difference between Newtonian dynamics and that proposed by the general theory of relativity. Newtonian dynamics assumes the existence of an abstract endless space, in which material bodies are placed and interact with each other. The existential hypothesis of the existence of an empty space, however, is not testable, thus philosophically cumbersome. In the general theory of relativity material objects generate time–space with testable parameters. Therefore, an object placed next to another one will interact with the field of that other object. (The reverse effect will take place, too).

Mind you that the last two analogies concern the respective sets of relations only, and not the mechanisms.

3.4 Supporting the Hypothesis of the Existence of the Communicative Field and a Selective Mode of Language Use

What is crucial when arguing for the feasibility of the mechanism proposed is motivating the existence of biological mechanisms generating a communicative field and accounting for the existence of a selective mode of language use. Below I shall introduce briefly some of arguments corroborating the hypothesis of the existence of the communicative field and a selective mode of language use.

First, the evidence from eye tracing experiments shows that the scope of our attention shifts when we proceed with decoding utterances—i.e, we might say—along with the change of the respective communicative field (c-field). Second, as

---

[19] The core meaning can be defined as the part which is statistically "significant". This is done in elementary particle physics, when identifying short living particles (it is possible to give statistically an approximate cut off point and state what is 'core').

reported in Bunge (2003), an important feature of the nervous system, which if assumed to support a communicative field will account for its potential usefulness for communication is lateral inhibition. Lateral inhibition means that neuronal excitations remain confined and do not propagate far as is the case with, e.g., electromagnetic field propagating extremely far at the speed of light. And if a communicative field is to be of any use, it needs to generate only a restricted number of options.

Third, Bruza (2010), models association patterns evoked in response to groups of linguistic items with the help of Quantum Mechanical formalism, which additionally testifies to the hypothesis that self-organization processes structure our associative knowledge. The research done by Horst et al. 2006), in turn, shows that fast mapping (identification from context), which seems to be an example of the mechanism of selection proposed here, does not result in memorization. This fact corroborates the hypothesis that the selective mode of language use performing on single occasions only does not influence the encoded meaning.

Fourth, the final group of evidence concerns the assumption that the communicative field is arranged around functionally differentiable *pragmemes*, which implies that in addition to representational meaning, the brain records functional meaning. It turns out that indeed a relevant property of the nervous tissues [in Bunge (2003)] has been found. Mountcastle (1998) discovered that neurons group into systems acting as wholes with emergent properties (functions) and with relative independence. The same hypothesis is supported by Lin et al. (2005, 2006, 2007) research concerning functional cliques mentioned earlier. Priming research, in turn, shows beyond doubt that much of the information related to a given linguistic item is hidden in the context, thus, the other way round, can be retrieved from that context, is "encoded" by it. A copious amount of psychological data, started with Heart's frames, supports the claim that people have knowledge of whole functional structures generated by prior knowledge and expectations and that people utilize these functional structures in creating representations of an incoming individual perception. It has been demonstrated, for instance, that people "remember" false facts which fit common frames.

Fifth, the everyday observations presented below also support the hypothesis of the existence of the selective mode of language use in communicative space (thus indirectly the existence of a communicative field itself). Note, for instance, how much more difficult it is to read nonce words, or simply words new to us, (cf. deciphering doctor's prescriptions) than words composing meaningful texts, e.g., letters even when written in the same handwriting. I propose that on the latter occasion, we are helped by a task of distinguishing from few feasible options with considerable differing forms, which is faster than recognizing all individual letters and decoding their respective pronunciations. In turn, subjects reading texts including words whose spelling has been slightly altered (some letters transposed) often do not even notice the mistakes. Another common phenomenon that can serve well to illustrate the operation of the mechanism of selection out of limited options is the metaphoric use of language. A given construct can refer to an item it does not encode, i.e., metaphorically, because the relevant item has already been

partially identified by the remaining contextual information and the item used metaphorically merely points out one of these options

Sixth, considerably different functioning of the left and right hemispheres gives support to the very possibility of there being two modes of language use as postulated here. While the left hemisphere tends to focus on details, i.e., decoding and logic; the right one looks for the global picture, is action oriented and, we might say, it looks for the best available fit in potential *pragmemes*. (Note, that having two ways of organizing information, e.g., as is the case with separate types of operation of the left and right hemispheres, opens the possibility of modelling the change of the correlation between the form and representation, while at the same time maintaining its short-time stability.) Additionally, the right hemisphere selects from options, which are often functionally motivated. It relies on emotions, feelings, and intuition and it is non-verbal. Note also that, historically speaking, the hardwired processes concerning, e.g., emotions, took place before those leading to the formation of plastic areas in the brain, which found more advanced cognitive processes. This all gives additional support to the claim that pre-representational early categorization was functionally defined and thus allowed one next to represent and label the categories which had already been singled out.

Finally, the model proposed is compatible with a widely accepted model concerning our memory operation (Grzybek 2006:160), which says:

> After having extracted the meaning of an actual clause, its verbatim form (words and syntax) is rapidly lost from memory, while the meaning is preserved and affects the interpretation of the following clause.[20]

3.5 Relating the Proposal to the Contemporary Mainstream Linguistic Scene

How does the proposal advocated in this chapter differ from and/or resemble major received views? Let us look at several best known proposals.

Let us begin with Chomsky's grammars. Chomsky's models, starting with Chomsky (1965), were all purely reductionist and nested exclusively in biology. The truly explanatory aspects of these propositions were to be found in psychology, which lied beyond the interests of the author. Chomsky and his followers also looked for some descriptive generalizations among data. Yet, such practices is not what is meant by explanation in empirical sciences.

Capone and Mey, in turn, have recently reintroduced the foundational role of social dimension into the main stream language modelling, when they introduced and developed, respectively, the concept of a *pragmeme*. Mey (2010: 2884) says:

---

[20] This hypothesis is motivated by Sachs (1967) in Grzybek (2006), while Luther and Fenk (1984) in Grzybek (2006) further showed that this strategy operates under "normal condition" i.e. when there is no motivation to concentrate on the form.

> Nomenclature aside, it is clear that the final question is to determine what things mean in a situation. But this meaning can be conceptualized in different ways. For a pragmaticist, this is not first and foremost a matter of determining the generalized (or even particularized meaning of an utterance), or of its individual segments; what these segments mean is always a function of their use in the situation, and in how they merge towards constituting the situation's pragmatic relevance.

In other words Mey (2010) (cf. also Mey 2001) proposes, as he terms it, an "outside-in" approach, i.e., that semantic (representational) meaning depends on functionally (originally interactively) defined *pragmemes* in which it is used. The functional (interactive) aspect of pragmatic meaning is crucial for epistemic reasons, which were already explained in the section on basic encodings. Interestingly, skin (the most primitive sensory organ) and nervous system hosting cognitive skills, originate from the same part of an embryo, which at least suggests a close correlation between the two. The crucial difference between Mey's proposal and mine is that the qualitative theory outlined here is evolutionary/developmental and both the pragmatic and semantic content continually co-develop and co-define each other—make each other more and more precise. Thus, I propose that the first imprecise representation was possible due to purely interactive categorization, and current representations and *pragmemes* modify each other with every use. Mey's proposal will coincide with mine under the assumption that the time span considered is short enough to allow one to regard *pragmemes* as being stable. This assumption holds true e.g., during the interpretation of a given text at a given moment. And this is precisely the assumption I make in this chapter when talking about qualitative (plausible) explanation and hypothetical interpretations of sentences meant to illustrate the mechanism of interpretation. Yet, neither Capone, not May is interested in quantitative tests of their proposals, i.e., tests in the empirical paradigm, which I illustrate in Sect. 4.

Mey3 (2010) points out Jaszczolt[21] (2005) as the approach to pragmatics reflecting his "outside in" perspective on language best. Jaszczolt (ibid.) proposes that encoded, default and pragmatic/cultural information all contribute simultaneously and merge into a final representation. She coins her hypothesis in the paradigm of formal linguistics based on qualitative tools, which she handles with great dexterity and sophistication. As a result, her proposition reflects the spirit that both *nature* and *nurture* matter equally, as stressed by Bunge's *systemism cum emergetism* introduced earlier. Yet, despite the same conclusion, the underling philosophical assumptions made by Jaszczolt are completely different from those of Bunge and mine. So are Jaszczolt's goals and testing methodology. For instance, as the author admits herself, the concept of default meaning lacks any proposal of a mechanisms supporting it. (On the contrary, a functionally similar concept of a communicative field has a clear psychological and philosophical justification, and a clear mechanism relating it both to pragmatic and to semantic

---

[21] Another interesting proposal crossing the boundary of semantics in the process of interpreting language in a novel way is Distributive Grammar proposed by Andre and Helene Włodarczyk, cf. www.celta.paris-sorbonne.fr/anasem/indexASMIC.html.

meanings). Therefore, unlike in Jaszczolt's, in my proposal cultural knowledge and pragmatic knowledge are not simply postulated as independent sources of information known to a language user, (which makes sense on a descriptive level and is appropriate given the goals set by Jaszczolt), but constitute an integral part of language creation process. For epistemic reasons, pragmatic and semantic knowledge in my proposal co-determine and co-create each other. Besides, while Jaszczolt's tools are suitable for descriptive analysis only, the model proposed is capable of posing some explanatory hypothesis concerning the relationship between encoded and functional (situated) meanings (see Sect. 4). These differences, however, are the differences not between the proposal advocated here and specifically that of Jaszczolt's (2005), but between the proposals coined within the developmental, empirical paradigm and all the approaches based on tackling language as a self-standing abstract structure describable with qualitative formalisms.

Another proposal with an established reputation, which covers similar ground to that covered by my proposal, is presented in Recanati (2011) and is based on the concepts of flexibility and modulation, The mechanism of selection[22] propounded in this chapter seems to meet the goals put forward by these concepts. It also seems to meet the goals set out by Carston (2002) quoted earlier. Similarly, I believe that when Stainton (2005, 2006) argues for cases of non-syntactic, or elliptical usage of words and phrases, these, too, can be seen as an instance of selective use of language, where a word or phrase serves to identify an option form parameters of potentially relevant instantiated speech acts. As already mentioned, non syntactic communication resembles, in turn, very much an early stage of language acquisition in children—when they develop a proto language. It is not unlikely then that such non elliptical communication described by Stainton (2005, 2006b) as mentioned above is the essence of children's speech, preceding their acquisition of syntactic language and eventually gives rise to linguistic structure through self-organization.

The concept of selection seems also to flesh out the concept of replacement postulated by Ariel (2008: 308). Ariel (ibid.) proposes that one of the mechanisms of explication of linguistic constructions on many occasions must be that of replacement. This is needed to cover situations of partial replacement of meaning, such as that which took place in the case of meaning change in the lexeme *gourment* when used in *gourment garage*. While Ariel mentions the concept of replacement only in passing (she is not interested in its mechanism), selection is clearly able to model it. Since we select the most similar items out of the given set, the item selected does not need to share all of the content of the item used for selection. In this way the item selected replaces the content of the item used for selection, rather than adds to it.

Ariel (ibid) concludes her book by posing the hypothesis that the same processes which are responsible for synchronic meaning creation have been *en force*

---

[22] An early version of the present model appeared in Zielinska (1997, 1999).

during diachronic change., i.e., that a diachronic change is the summation of pragmatic history of the same processes and does not involve any additional laws.

> … the very same grammar pragmatics interface representations functional in the ephemeral discourse time are also the input for the diachronic transfer of the pragmatic into grammatical. Ariel (2008: 308)

And this is precisely the position assumed here. Only one type of law—statistical laws driving the behaviour of linguistic community members during their interaction with each other are postulated. All diachronic linguistic changes are the result of self-organization processes taking place in the material system made up of members of a given linguistic community, as explained in the section discussing the concept of emergence.

As to the research concerning specific applications of the encoding and selective modes of the use of language, the position advocated can explicate the difference between descriptive and classificatory uses of adjectives, respectively. This difference has been known at least since Bolinger (1967) and recently was nicely summarized in a formal language by Kennedy (2007). The selective mode of language use can account for the semantic novelty appearing in categorizing (classificatory) usage of adjectives, while the encoding mode of language use can account for modelling the descriptive use of adjectives. The difference between the previous propositions and the account advocated here is that the selective and descriptive uses of adjectives do not exclude each other. Both can and often do function simultaneously. The evidence from Polish, where categorizing use of adjectives is marked by merely statistical preference for their postpositional usage in noun phrases, supports my view and undermines Kennedy's "either, or" stand.

Finally, the proposal presented in this chapter fits Mey's observation concerning the transition in perspective on the concept of class in pragmatic research over the first 25 years of Mey and Haberland (2002) observe that while earlier "class was essentially thought of as a product of historical developments (which it certainly is, too) the shift of perspective has to do with a more developmental view of class, class as 'work in progress'." While encoding reflects the product of historical development, selective mode of language use reflects that 'work in progress' as Mey put it—creating classes *ad hoc* for a given pragmeme. And selective mode of language use postulated in this chapter models the creation of such an *ad hoc* class.

In its general spirit, I find the proposal introduced to be somewhat similar to RT. The major similarity is that both rely heavily on context believing, as Wilson and Sperber state, that "semantics of natural language might be too weak to encode human thought" and both propose the mechanism of reaching that thought despite the shortcomings of semantics. RT does it with the help of the theory of relevance, the proposition advocated in this article does it by postulating the selection mode of language use to be modelled formally by an Supervised Learning technique. RT covers both the area of traditional Gricean pragmatics and of recovering the semantics of sentences. The proposal presented here, on the other hand, focuses primarily on accounting for the enrichment stage of RT, but the

same mechanism may operate on already explicated propositions, too. While RT worked out a very detailed mechanism of inferencing and finding relevant implicatures considering interpreted sentences, it is much less specific when applied to modelling enrichment (interpreting sentences). As already stated, the proposition presented in this chapter focuses on finding the enriched content of the utterance, to use RT terminology.

To illustrate the difference in the methodology between the two approaches, let me recall the analysis of the sentence: *It will take some time to repair your watch* uttered by a watch-maker responding to a customer who has brought a watch out of order for repair, offered by Sperber and Wilson (Sperber and Wilson 1986: 178). Sperber and Wilson (1986) say "It goes without saying that watch-repairing is a process with a temporal duration and a speaker aiming at optimal relevance must have intended to express something more then what goes without saying. In general the utterance of the form in 23 [the sentence just mentioned–DZ] should be interpreted as conveying not the truism that the job in question will take some time but that it will take an amount of time it would be relevant to remark on. i.e., longer than it would otherwise be expected."

But the authors offer not a word on why the amount of time it would be relevant to remark on is "longer than it would otherwise be expected"? RT leaves this unaccounted for—proposing no mechanism for drawing that conclusion. On the approach advocated, a qualitative explanation indicating a relevant mechanism goes as follows. In the situated speech act of a customer talking to a watchmaker when depositing his watch, the relevant options concerning the amount of time taken by the repair could have the following rough, functionally distinct (established by the *pragmeme*) values: {Done on the spot, done within a shorter time than usually, done within a usual amount of time, done within a longer than usual time, needs a lot of time to be completed}. Now, the adjective 'some' will select the "longer then usual option" by virtue of its meaning taken in comparison to the meanings of other common terms denoting values, let's assume: {no, little, some, a lot} resembling the one to the last of the five options assumed above most. (We proceed by taking the scale of conventional terms encoding values (the latter scale), next by creating the *pragmeme* relevant scale of the relevant values (the five possibilities described earlier) and next we correlate items from both scales. This way, we will temporarily assign the terms correlated with typically encoded values to actual values).

Importantly for the proposition presented, however, such an explanation such as just suggested is given only as a pre-theoretical one, offered at a descriptive stage of the research, which indicates that the proposal seems to be sensible enough to merit scientific investigations. The final goal for the proposal is to describe linguistic phenomena in an objective, quantitative fashion. Therefore, we shall be looking for theories allowing us to construct models having quantitative implications, which could be squarely born out, or rejected, by measuring relevant statistical characteristics of available corpuses. This step will be preliminarily illustrated in Sect. 4.

4 Qualitative and Quantitative Substantiation of the Hypothesis of the Mechanism of Form and Content Correlation in Natural Language

Any proposition to be evaluated within an empirical framework needs to be tested (verified.). In the preceding section, I have described the mechanism of a form-content correlation process in most general terms. Yet, (it will be explained more thoroughly in 4.1 below), following Bunge (1972), even a detailed theory of a form-content correlation mechanism will be too general to be tested *per se*. A general theory like the one postulated can only be confirmed by testing its application to models of specific phenomena. These models are described with the help of that general theory, as well as of some additional assumptions and laws concerning the specific phenomenon at stake. Therefore, to finish the presentation of the mechanism of the form—content correlation process, I shall choose a specific linguistic phenomenon to investigate, which has characteristics of its own; propose a law which could account for such specific characteristics, formulate a hypothesis this law along with the theory of form-content correlation mechanism implies, and finally test that hypothesis. The linguistic phenomenon I have chosen for this purpose is the ordering of selected classes of adjectives in noun phrases.

In other words, I shall propose a specific hypothesis that will allow me to account for some aspects of the ordering of adjectives found in corpuses. This time, however, to meet the "hard" standards of the methodology of empirical sciences, I shall account for the observations not only in a qualitative fashion to make the hypothesis proposed plausible, but I shall also propose and check the validity of some of its quantitative, statistical implications expressible in terms of objectively measurable parameters. Before doing that, however, let me recap briefly the essence of empirical linguistics advocated here and the essence of the methodology of empirical sciences.

4.1 The Foundations of Empirical Linguistics

The formal branches of linguistics treat language as a set of sentences with structures assigned to them, and, accordingly, use the formalisms of qualitative mathematical (algebra, set theory) and logics to model structural and semantic properties of language. Therefore, on the one hand, such formal approaches to language cannot hope to account for quantitative observations, on the other hand, structure and meaning being totally independent, none of them can serve to investigate the other one. Consequently meaning is left out from rigorous inspection. Since, for formal linguistics language is a closed system, neither can it hope to account for non-combinatorial novelty in language.

The perspective advocated here, (shared, among others,[23] with synergetic linguists working in the paradigm developed by Bunge (2003); with Bak (1996), who models self-organizing phenomena; and 'The Fife Graces Group'), is that language is a semiotic system,[24] which is an aspect (the result) of the communicative processes in linguistic communities. Therefore, language is subject to evolutionary processes in analogy to biological organisms. As the 'Fife Graces Group' puts it:

> Language change is a cultural evolutionary process, (Christiansen and Chapter in press, Croft 2000). According to the general Analysis of Selection (Hull 1988, 2001), evolutionary processes take place on two linked levels: replication and selection. Replicators are replicated, but with culminating errors resultant from mutation and recombination, and this way variation is generated. Selection is a process by which interactors in interaction with their environment cause replication to be differential: that is some replicators are replicated more than others, which in extreme case leads to fixation of the former and extinction of the latter. In language, linguistic structures—sounds, words, and constructions are replicated in utterances every time we open our mouths. That is, replication and variation occurs when we use language in the service of joined actions between human beings in a community. Due in part to the indeterminacy of communication described above, the replication process produces variation. Speakers differentially replicate certain structures through interaction with the environment, namely the situation being communicated and their interlocutors. In the former case, changes in life styles (e.g. the rise of *cell* (phone) and the fall of *harquebus*). In the latter case, the social identity and the social contexts of the interaction lead to the rise and fall of linguistic forms that are associated with various social values by speakers.

On the above assumption, what is especially relevant for modeling language—the development and the functioning of a linguistic system—is considering the multitude of its quantitative properties very much the way that statistical characteristics of biological groups allowed one to approach biology within the paradigm of empirical sciences. As mentioned, many quantitative characteristics of language testify to a self-organizational character of language.

> Moreover, it can be shown that these properties of linguistic elements and their interrelations abide by universal *LAWS OF LANGUAGE*, which can be formulated in a strict mathematical way—in analogy to the laws of the natural sciences. Emphasis has to be put on the fact that these laws are stochastic; they do not capture single cases (this would neither be expected nor possible), they rather predict the probabilities of certain events or certain conditions in a whole. It is easy to find counter-examples with respect to any of the examples cited above. However, this does not mean that they contradict the corresponding laws. Divergences from a statistical average are not only admissible but even lawful—they are themselves determined with quantitative exactness. This situation is, in principle, not different from that in the natural sciences, where the old deterministic ideas have been replaced by modern statistical/probabilistic models. The role of QL is to unveil corresponding phenomena, to systematically describe them, and to find and formulate the laws which explain the observed and described facts.

---

[23] The members of the Fife Graces Group, the proponents of the thesis that language is a Complex Adaptive System view, are, among others, Clay Beckner of the University of New Mexico, Richard Blythe, Edinburgh University, Joan Bybee, University of New Mexico, Morten, H Christensen, University of Cornell, William Croft, Universityu of New Mexico, John Holland Santa Fe Institute, Nick N. Ellis, University of Michigan, and others.

[24] in Bunge's (2003) understanding of the concept 'system'.

> […] the development and the application of quantitative models and methods is indispensable in all cases where purely formal (algebraic, set-theoretical, and logical) methods fail, i.e. where the variability and vagueness of natural languages cannot be neglected, where gradual changes debar the application of static/structural models. Briefly, quantitative approaches must be applied whenever the dramatic simplification caused by the qualitative yes/no scale is inappropriate for a given investigation. (Koehler, Lectures in Quantitative Linguistics, Trier University)

And such is a general philosophy of modeling language in the empirical paradigm as advocated here. Yet, before proceeding to test the form-meaning correlation law postulated in this chapter, I would like to clarify the understanding of the concepts of a theory, (law), and of a model in empirical sciences. Empirical sciences are built around theories, i.e., sets of compatible laws. The laws concern general fundamental characteristics of an aspect of a type of phenomena, and as already said, are too general to be tested *per se*. A theory is tested by its application to models of specific phenomena, which have additional characteristics and properties implied solely by the respective models. (Therefore, if the test does not support the theoretical predictions, we can never be sure whether it was the theory, or a model that failed.)

Let me illustrate what is meant by applying a law to a model. For instance, one of the Newton laws states that two material points attract each other with the force that is proportional to the product of their respective masses, to some constant G, and inversely proportional to the distance between the two points. This law, however, is not testable *per se*. It can be tested only indirectly by applying it to some model. This Newton's law can be applied, for instance, to model the movement of the Earth around the Sun. To this end, however, we will need to construct a model of the Earth orbiting the Sun, which could be written down in terms of Newton's law, i.e., we need to approximate the Sun and the Earth as two material points with specific masses. This assumption, however, is part of the cosmological planetary model, not of the Newton's law. To measure the distance between these planets, in turn, we assume that the planets are placed in empty Euclidian space, which is another assumption independent from Newton's Third Law.

Thus, if we want to test some law, we need to identify and test some of its implications for modelling a specific phenomenon—i.e., a hypothesis concerning a specific model, which model is characterized by some additional constrains absent from those defining the law. To meet the criteria of empirical sciences, such hypotheses need to be expressible in terms of data gathered with objective, received measuring techniques.

To describe a phenomenon under consideration in terms of measurable characteristics, we may need to resort also to theories and hypothesis other than the ones being tested, i.e., to the hypothesis which have already been well tested and which can assess objectively the aspects of the situation described by the new theory. We shall illustrate what that last statement means by considering a test of Hook's law (the law stating that the force exerted by a squeezed spring (coil) is proportional to the deformation caused by that force.) The relevant force can be

measured only in such a situation which describable not only with the law pos-
tulated, but also independently from the law being tested. In the case of a Hook's
law, this will take place, if we model a situation in which some weight is placed on
top of a vertically positioned spring placed on a table. In this case Hook's force is
balanced by the well tested force of gravity, the measuring of which involves
conventional and objective rules i.e., is repeatable by others following the same
instructions.

To repeat, the above example illustrates that we test theories (laws) indirectly
by testing some of their predictions in relation to a specific situation, here a spring
squeezed by a weight placed on top of it (or more exactly, to the model of that
situation.) A given situation to be of use must have characteristics measurable in
objective fashion—both with the help of the hypothesis being tested and of some
other well tested theory.

Before proceeding to propose a model of a specific linguistic phenomena that
would validate the mechanism of the form-meaning correlation process hypothe-
sized in this chapter, one more comment is in place. A given model is not expected
to cover the phenomenon at stake perfectly. "The best theory of a theoretical
model is not a copy but a theoretical model or conceptual reconstruction con-
taining concepts without a concrete counterpart (such as logical concepts), as well
as hypothesis that at best, are approximately true." (Bunge 1972: 171). Therefore,
when proposing some linguistic laws in next subsection, I do not mean to claim
that no other laws are in operation. Yet, any model creation can start only from
postulating some characteristics, which will be confirmed or rejected by experi-
ment. Only later the model can be fine-tuned and tested again, and so on.[25] (This is
a stand taken by critical realism.) Bunge (1972:171) writes about such a continual
process in the following words

> If neither experience nor reason were necessary to conduct scientific research, we could
> resort to wild intuition or to mystic communion. If theory were sufficient we would waste
> no time with empirical tests, and would give the triumph to idealism. If scientific theories
> were of no need of theory, empirism would win. In other words, critical realism assumes
> that the thing in itself is not knowable as such without any distortion. It necessarily
> involves proposing some traits and next keeping correcting them to the point of giving
> some of them up completely. As it is, factual knowledge consists of a set of theories and a
> set of data, such that the former must be compatible with at least some of the latter, while
> the data must be sought and processed with the help of some theory. Moreover, data are in
> principle as corrigible as theories, in the light of both further data and other theories.

An interesting category of object models are ones based on assumptions sim-
plifying reality to the point of significantly distorting it. This is done in situations
in which the application of the received theory to a more realistic model results in
equations which cannot be solved. Such models can be exemplified by those used
in solid state physics, in which specific phenomena are modelled in 1D (one
dimension). The resultant models cannot answer all problems, but allow one to

---

[25] Those conversant in Polish are advised to see Grabińska (1993, 1994, 1998) for especially
illuminating presentation of the relation between models, theories, and reality.

solve a certain class of problems. The usefulness of such conclusions is testified by having a special journal devoted solely to such 1D models of solid states. Apparently, the failure of a well put, good idea is often more telling than the success of an imprecise one full of *ad hoc* hypothesis mending it.

Note, also that the status of a law as concerning the most basic properties of the reality studied is relative. It is not uncommon that two laws seen at firs as independent ones later on turn out to be instances of a more fundamental one. For instance, the theory of magnetism and the theory of electricity have turned out to be special cases of the theory of electromagnetism.

4.2 Coining a Model to Verify Quantitatively the Mechanism of Form and Meaning Correlation Process

The mechanism of the form-content correlation process introduced in Sect. 3 was outlined in most general terms there. In Sect. 4, we have been arguing that form and content correlation in natural language is the result of self-organization and self-regulation and therefore the likely processes of form-meaning correlation could be constrained by laws optimizing linguistic effort and effect.

Linguistic laws proposing the optimalization of speaker and listener's effort have been present at least since Zipf (1935, 1949). Zipf, for instance, accounted for diachronically observed phonological reduction in the following way. He said that since frequently occurring items are generally more predictable than the ones rarely occurring, listeners can decode the message coded by a frequently occurring item, even if it is not very carefully articulated. Consequently, speakers may afford to articulate these items less carefully, and as a result these items will undergo faster phonological change (e.g., reduction) than the ones used less frequently. This will lead to a more economical semiotic system.

These days, many linguists agree, [cf. Keller (1994), Kirby (1999), Haspelmath (2006, 2008)], in addition to those already mentioned earlier, that "a diachronic change is the necessary link between patterns of language use and grammatical structures" (Haspelmath 2006). Many researchers further agree that the diachronic change leads to the optimization of language, although such an optimization need not be a conscious goal of language users. Haspelmath (ibid: 18), for instance, elaborates on that saying

> Speakers do not intend to create well-designed grammars, but they behave purposefully and rationally in selecting from available variants and in creating new variants—they mostly opt for the most useful variants for their particular purposes. Through an invisible-hand process in language change, the cumulative effect of many individuals' behavior leads to useful language structures (cf. Keller 1994) …. So how do economical patterns arise in language change? There are two rather different routes by which this can happen: differential phonological reduction (§6.1) and differential expansion of a new construction (§6.2). Moreover, a minor route, morphological analogy, must also be recognized (§6.3).

When processes such as differential phonological reduction, differential expansion of a new construction, or morphological analogy concern purely local phenomena, (i.e., do not influence other subsystems), the hypothesis that a language user chooses an optimal variants is plausible. (Never mind whether that decision how to choose the optimal variant is rational, as Haspelmath has it, or there is an unconscious[26] mechanism ensuring such a choice.) Yet, first, the development and optimalization of language is not limited to the reduction of existing forms and their propagation but, importantly, involves the creation of novel forms and meanings. Second, the optimalization in language in addition to being opportunistic and in some respects optimized in local niches, also has a global character. On the other hand, plausibility of a purely cognitive mechanism being capable of making global optimal decisions relevant for specific situations, which would optimize the whole system, is unlikely, because such a task would require unrealistically complex calculations involving an extensive consideration of the whole system. Such global optimizing would mean, for instance, deciding to keep the English item *her* in the situation, where it does not convey any new content because of the benefits of the resultant pattern in other constructions. And so, in the following example: *Mary arrived late. She parked her car right in front of the main entrance and entered the school building,* the item *her* is highly predictable. Note, that in a Polish translation, the lexeme *her* would be skipped. Yet, Poles are not better at economizing natural language than the English are, one would think. The reason for the difference in the decision weather to keep the pronoun *her* or not in English and Polish, respectively, is that keeping the pronoun in English in the situation illustrated allows a given linguistic community to save more effort "across the board", among others, by not needing to add verbal suffices informing of the gender of the speaker, which is the case in Polish. However, this is not the end of the story, because languages, which inflect verbs tend to add suffixes to nouns, marking the case as well as the gender, and this all allows these languages to have a relatively free word order. That, in turn, allows them to mark 'new' and 'old' by the positioning of a given phrase in the sentence, thus do without 'a/the' type of articles, and so on. But such optimizing decisions concerning complex, interrelated parameters cannot be calculated exclusively locally. Therefore, likely some global optimizing mechanisms not requiring carrying out complex cognitive calculations concerning the whole system on the part of the speaker are also at stake during a language formation process.

I shall argue, as brought up already when quoting Altmann (1978), that these global mechanisms resemble natural selection. Note, that such global laws (based on mutation and selection of the fittest) are likely not only to account for global optimizing effects, but also for novelty in content and form (both combinatorial and non-combinatorial). And for any reduction or analogical differential spreading to take place, first some variants, e.g., morpho-syntactic patterns, need to be

---

[26] Linguistic research employing analogical modeling introduced by Skousen, puts me in that second camp.

established some way or other, thus a relevant mechanism of their creation needs to be additionally explicated.

Thus, for instance in reference to the emergence of the order of some existing elements, at the stage when there was no preferred order for that set of items, the order of such items could be selected in the following "evolutionary" way. Some sub-groups of speakers, on purely statistical grounds, start selecting one of the possible orders chosen, which results in the development of at least two sub-varieties of the current linguistic system, possibly one more efficient, the other less efficient. Next, speakers of the more instead of most efficient sub-variety being better communicators, became, statistically speaking, more successful in life. As a result, the order present in that most effective sub-variety of that system, will be repeated relatively most often due to the fact that their speech will reach a wider population, more often, and possibly will also be more esteemed thus additionally easier to memorize. Next, the decision which pattern to follow, made by an individual speaker can be purely local, i.e., determined exclusively by the relative frequencies of the options heard by that speaker. Eventually, the community will adopt the ordering of the most efficient sub-variety along with the remaining elements of that most efficient sub-variety. Establishing a new ordering of some items may lead to further unconscious reorganization because of the change in the distribution of data. Such mechanism of syntax formation could have been first implemented already on a proto-language stage, when the first combinations of previously single referent items were being tried out. Later on, this mechanism could be still taking place in relation to the ordering of items still at free variation. Among others, such too, must have once been the situation when two adjectives were used to modify the same noun for the first time.

Now, let us go back to accounting for the form-content correlation mechanism outlined in Sect. 3. As already mentioned, I have chosen to apply the proposed mechanism of form-content correlation process to explaining the order of adjective categories in 'adjective adjective noun' phrases. The phenomenon of the adjectives order in noun phrases is what I have chosen to apply the proposed evolutionary mechanism of form-content correlation to—in order to allow for its quantitative tests. More precisely, I intend to account for a statistical preference in the adjectives order between adjectives expressing semantic categories of (1) "size or shape" and (2) "age or colour" and (3) "origin or material" which is observed in English noun phrases of the A_1A_2N type. A similar statistically observable preferred order of adjectives in noun phrases has been attested, among others, in such diverse languages as Chinese, Hungarian, German, Polish, and, in a mirror reflection, in Italian and French, which suggests a universal mechanism for the phenomenon considered.

The main empirical hypothesis

In line with what has been said so far, I propose that, other factors being equal, statistically speaking, a more efficient language variant is one, which allows its speakers to express messages more precisely[27] both on specific occasions, and as

---

[27] Until the precision arrived at is sufficiently good and further increase in the precision of encoded lexemes (or their intended content) does not increase the functionality of language.

far as the resultant encoded content of lexemes is concerned. In particular, the ordering of adjectives is such that, statistically speaking, it optimizes locally the precision of the messages expressed with A_1A_2N on specific occasions, and the precision of the resultant encoded[28] values of the adjectives used. For reasons explained below, this will take place if, statistically speaking, the head noun in A_1A_2N phrases is modified with the adjective, which is more categorizing[29] first (A_2 is supposed to be more categorizing) and, by the one, which is most relative—last (A_1 is supposed to be more relative).

Qualitative justification of the hypothesis

The reason why the above order increases the precision of the encoded content of adjectives is the following. The second adjective applied (A_1) is more likely to apply to an atypical situation defined by A_2N, as a result selecting an atypical value for A_1 and thus skewing the average value of an encoded meaning of A_1. For instance consider the reference of the phrase *a red big bird* used by a visitor to a Krakow zoo. The lexeme *red* when selecting only among 'big birds', will select a flamingo, which is pink. As a result, the referred colour will add the value 'pink' to the colours encoded with the item *red* so far. In the reverse order, if we looked for a *big red bird*, i.e., for a big one among red birds, we will end up selecting a 'red bird' out of all birds thus ending up with the redness of an Ara Parrot, whose red colour represents a focal red. Therefore if we let non-relational adjectives act first, this will result in their statistical average being less dispersed, without affecting negatively the relative adjectives, which have operatorial character thus, with every use, their values depend on the category to which they are applied.

Note also, that the above conclusion is also consistent with the predictions of the theory of complexity stating that lowering of the complexity (such as that caused by limiting the number of allowable syntactic patterns) is accompanied by the increase of the informational content of components.

The hypothesised order will also ensure that the situation specific values of adjectives convey a more precise message, because the adjective that is more categorizing establishes the scale for gradable adjectives to operate on. Consequently, e.g., the gradable adjective *long* used in *long wooden bridge* will have different meaning from *long* used *a long steel bridge*, because steel bridges can be much longer than stone bridges. For the same reason, *old stone bridge* can be much older than a *wooden stone bridge* thus the value provided by the adjective *old* is

---

[28] Encoded value is understood as a statistical average of past values.

[29] What needs to be made clear here is the definition of the terms "categorizing" and "relative". By a categorizing adjective I mean one which when applied to the noun results in the selection of a distinct subcategory. For instance 'a blue crayon' differences from 'a red crayons' in colour only thus the adjectives *red* and *blue* as used in the examples above are not categorizing. On the contrary, the adjective *high* when modifying the noun *chair* selects 'a high chair', which item has a number of characteristics (including its novel function) singling out the subcategory of 'high chairs' from among all chairs, therefore the adjective *high* in the phrase *high chair* can be termed 'categorizing'. By 'a relative adjective', in turn I mean one whose actual value depends on the range of the given property in the items modified with it. For instance, the value of an adjective *big* changes depending whether it modifies a star or a mouse.

different in both cases. If we reversed the order of adjectives application and said, e.g., *a wooden old bridge*,* the lexeme *old* would select from the scale of ages of all bridges thus would not be as precise as when selecting from the scale of ages of wooden bridges only. To sum up, according to the postulated mechanism of form-content correlation presented in this chapter, if the categorizing adjective is applied to the noun first, the resultant A_cN generates a communicative field of options defined with parameters of certain range each. The relative adjectives A_r used second serves to establish an appropriate point on the scale established by the options generated with A_cN. To recap, the order of adjectives A $_{relative\ (gradable)}$ A $_{caterorizing}$ optimize the system, at least locally.

Quantitative (hard) confirmation of the hypothesis

The model of the adjectives order in noun phrases proposed above implies that the preferred order of categories (1) "size and shape", (2) "colour and age", (3) "material and origin" categories in English AAN phrases reflects the fact that the relativity (gradability) of the successive categories decreases with the number of that category, while their respective categoriability increases.

Tapping into semantic intuition and having manually tagged the adjectives, Zielinska (2007) found that when considering AAN phrases consisting of two adjectives, each belonging to one of the above defined categories, the dominance of the phrases in which the adjective further from the noun belongs to the category designated by a lower number than the adjective closer to the noun was highly statistically significant both in Polish and English corpuses. Yet, this test relied partially on human semantic intuition when categorizing adjectives, which causes some discomfort. Could we avoid semantic classification by a human completely?

If we were able to formalize the concepts of relativity and categoriability, the hypothesis we want to confirm would imply that there is a positive correlation between the degree of the difference in relativity between the first and second adjectives and the initial position of the first adjective in AAN phrases, and a negative correlation between the degree of the difference in categorizability between those adjectives and the initial position of the first of them. Or there should be a correlation between the degree of relativity and being the first adjective, and a negative correlation between the degree of categoriability and being the first of the adjectives in a AAN phrase.

Wulf (2003) set out to confirm practically the same hypothesis by providing corpus-linguistic operationalizations of concepts close to that of "relativity" and to that of the degree of being "non-categorizing". More precisely, she operationalized a concept complementary to relativity, i.e., the independence from comparison index (IndComp), which she defined as the ration of the number of occurrences of a given adjective in non-comparative degree to the number of all occurrences of the given adjective in a given corpus. By analogy, we can define the degree of relativity as the ration of the number of occurrences of a given adjective in the comparative (and superlative) degrees to the total number of its occurrences in the given corpus.

Wulf (2003) finds out that the mean IndComp values for adjective$_1$ and adjective$_2$ in her study differ highly significantly ($p < 0.001$). Adjectives standing further from their head noun occur with more forms of degree than adjectives

directly preceding the head noun, which supports the results from previous works. Wulf (ibidem) also reports that knowing the adjectives' IndComp values improves the prediction accuracy of the order of a given pair of adjectives by 35.78 % (her total prediction accuracy is 67.89 %).

Now, moving on to the concept of categoriability, the idea that adjectives expressing concepts with a high degree of categoriability are placed closer to the noun is similar to that expressed by one of Behaghel's Laws which states that things belonging close together in mind are also put closely together in communication. The concept of the degree of being categorizing can also be related to Ziff's (1960) concept of the adjectives' different "privilege of occurrence:—the degree to which adjectives may occur in different contexts", as well as to Wulf's (2003) concept of semantic closeness. Wulf gives the following corpus-linguistic operationalizations to semantic closeness:

> Accordingly, the semantic closeness of the adjectives in the present data sample was measured via the number of different head nouns that the adjective in question collocates with. For all 1,154 adjectives in the present data sample, it was checked in the whole BNC how often they occurred with any noun. More precisely, three concordances and corresponding frequency lists were produced, as not only the adjective in its positive form, but also in its comparative and superlative forms had to be included to achieve a representative picture of the span of nouns the adjective collocates with. The resulting frequency lists had to be checked manually for potential double counts of nouns, i.e. cases where a noun collocates with an adjective in its positive as well as with any/both of its compared forms. The resulting number of different noun collocates was relativized against the corpus frequency of the adjective in question because adjectives which are generally more frequent than others will automatically have a greater number of different noun collocates.

Surprisingly, Wulf (2003) finds practically no influence of the above corpus linguistic operationalization (CLO) of the degree of semantic closeness onto the adjective's being positioned closer to the noun. Yet, in the same study, she finds a strong correlation between the closeness to the noun and a membership in Dixon semantic category of 'origin, and composition (material)', which are intuitively category-forming thus, semantically close. This latter correlation, tapping also into one's intuition when categorizing adjectives, in addition to many technical arguments brought up by Wulf herself[30] against the operationalization proposed, makes one suspect that this particular CLO is not adequate. It also shows that capturing the semantic concepts purely in numerical terms is not easy and thus tapping somewhat to semantic intuition operationalized through psychological tests' results might sometimes be a better solution.

Nonetheless, Wulf (2003) comes up with statistically significant results showing there being inter-dependences between frequency information and symbolic data. Such a correlation certainly cannot be captured by qualitative formalisms.

---

[30] One more, intuitively better way to try when CLO-ing 'semantic closeness' would be to consider the actual frequency of the occurrence of specific collocates, and not the number of types, as Wulf did.

Wulf has shown that AO, although superficially a phenomenon with purely syntactic, i.e., generatively definable patterns, in fact depends on quantitative characteristics of a variety of variables from different levels of linguistic analysis. (Wulf considered also the influence of other factors such as the length of words.) These quantitative dependencies clearly point to the insufficiency of purely qualitative descriptions. Importantly, the more influential, by far, of these variables is the degree of relativity. Relativity is also the variable with the least controversial operationization. The second part of the hypothesis concerning the influence of the degree of categoriability of a given adjective on its position in an AAN string has not been clearly supported by Wulf's (2003) data, but given a questionable operatorization of what might come closest to the concept of the degree of categoriability presented in the hypothesis posed, Wulf's (2003) data does not rule out a more significant influence of that latter variable on the AO order in AANs, either.

The problems with the operationalization of the categoriability concept reported above, prompted Zielinska (2007b) to choose an approach to the issue of adjectives order based partially on tapping into semantic intuition, (as Wulf (2003) did that too, when considering the Dixon's semantic categories). Zielinska (2007b) set out to confirm the hypothesis concerning the dependence of the position of adjectives on the degree of their categoriability and relativity in the following way. She demonstrated the role of the degree of relativity (= gradability) and or categoriability of an adjective in an AAN phrase by subcategorizing the Dixon's categories into more and lees gradable (relative) subcategories and by checking the influence on the relative order between such subcategories. For instance, Zielinska (ibid.) shows both for English and for Polish that the division of the colour category into a category of intuitively highly relative colour terms, such as *light, pale, vivid, dark* and intuitively less relative ones such as *red, blue, yellow*, results in the category 'relative colour' being statistically more likely to precede other semantic categories than the category 'non-relative colour', (or let's call these 'descriptive colour' terms) does. Similarly, the subdivision of a given category of adjectives containing the information about age into a subcategory of the adjectives more and less categorizing, respectively, e.g., into {*pre-war, renaissance, baroque,*etc.} and such ones as {*one-year old, 20 year old,* etc.}, results in more categorizing subcategory following other selected categories statistically more frequently than the other subset of the category 'Age' does. As a matter of fact, in the BNC, the subcategory "descriptive colour" follows statistically the category of the "less categorizing colour" in spite of the category 'Age' as a whole preceding the category 'Colour' as a whole. (By saying that the category 'Age' precedes the category 'Colour', it means there is statistically highly significant difference in the number of AAN phrases in which 'Age' precedes 'Colour' and the number of AAN phrases in which 'Colour' precedes 'Age'.)

As reported by Zielinska (ibid.), similar correspondences have been noted also for Polish—based on the IPN corpus, (although in Polish the effect was weaker for Dixon's "middle categories" than in English). The results presented by Zielinska concerning Polish are especially interesting because so far, since Polish is a

language with a considerable free word order, there had been no[31] prior sugges-
tions as to the corresponding ordering among Polish adjectives in noun phrases.

The biggest difference between the studies of Wulf (2003) and of Zielinska
(2007), is that while Wulf (2003) was merely interested in establishing quantitative
correlations in her data, Zielinska is also interested in looking for laws that could
imply these correlations. While Wulf (2003) collected previous qualitative anal-
yses of AO, the descriptions of dependencies on various linguistic properties and
"CLO-ed" them, Zielinska searched for an explanation to these quantitative
observations. She posed a hypothesis how (via what sort of mechanism) they could
have arisen. Thus, looking at it from the perspective of an empirical paradigm, if
we note that the qualitative studies of AO may resemble data collection by Tycho
Brache, then Wulf's studies correspond to finding empirical principles (finding
patterns) such as those proposed by Kepler in relation to the data gathered by
Tycho Brache, and the current study [as well as Zielinska (2007b)] corresponds to
looking for laws explaining (implying) the numerical relations observed (the way
Newton's laws are in relation to Kepler's laws).

Zielinska (2007b) proposed that the mechanism resulting in AAN constructions
used in the right order conveying more precise messages, and in increasing the
precision of encoded adjectives values, is grounded exclusively in cognitive
human capabilities, i.e., these are speakers who calculate, consciously or not,
optimal solutions. In this study, in view of the discussion in Sect. 4, I see that the
specific mechanism leading to the optimalization could also be different than
suggested in Zielinska (2007b), i.e., as presented below.

The order of lexemes A and B, which has not been fixed yet (grammaticalized),
self-organizes through the following mechanism resembling natural selection,
which increases the efficiency of language. At first, the majority of speakers order
these items at random, on a purely statistical bases.[32] Next, on purely statistical
grounds, there may form two subgroups of people, whose idiolects show a strong
preference for one of the possible two orders, AB or BA, respectively, in addition
to those who still place these items at random. If so, the speakers of a more
efficient dialect (let us say AB vs. BA vs. (AB vs. BA)), i.e., better communicators,
will be statistically selected in the sense that they will become more influential in
life and thus their speech will receive wider reception, statistically speaking. In
other words, their speech will become an input to the corpora of a larger than

---

[31] At the same time when I published my results, Tabakowska (2007) also published a study
concerning the ordering the adjectives in Polish. Tabakowska, however does not consider any
hard frequency based evidence, but carries out a purely intuition based (cognitive) analysis of
meaning of selected AAO phrases. She concludes "that there is a preferred order of adjectives in
Polish AAN phrases, which, however, can be overridden by stylistic reasons" thus, her study is
unrelated to the empirical paradigm.

[32] In many statistical models, cf. Skousen's (1989) analogical modeling, after the system has
reached a certain level of preference of a given type, self-organization takes place. Speakers are
predicted not to chose between options at random any more, but to select one of them. Such
regularizations have been long described in language –e.g. the regularization of past tense in
Finnish modeled by Skousen (ibid.).

average number of speakers—(e.g., through mass media, education, also because the forms they use are associated with various positive social values, etc.). Consequently, the more efficient dialect (let us say that with AB order) will propagate due to the mechanism of the form content correlation process introduced in this chapter.

The explanation just proposed looks appealing because it is global in character. The self-organization of language will take place this way only if it makes the whole system more efficient, yet it is not an individual speaker who needs to carry out such complex global calculations. The speaker does not need to be aware of all benefits involved in a given choice, i.e., what let him respond more efficiently, because, in addition to making purely local judgements (based on his individual expertise), he is influenced by the frequency of the same or similar forms that he hears. Note, that the more efficient sub-variety of speech brings in "the wisdom of the crowd[33]"—its collective experience derived from independent individual experience and independent[34] individual expertise, yet evaluated statistically.

Finally, the mechanism postulated allows one also to introduce novelty—convey functions, which did not exist at an earlier stage of language development. Importantly, the quantitative data presented clearly corroborated the hypothesis that the ordering of adjective in noun phrases is not random, but characteristic of a more self-organized system. The lack of a fully deterministic ordering of adjectives in AAN phrases found by Wulf likely results simply from the dispersion in the relevant values calculated for adjectives: we resort to global parameters to form hypothesis concerning individual items, (as if we resorted to relating the volume, temperature, and pressure of gasses to form hypothesis concerning an individual particle characterized by mass and velocity). Additionally, there is also a possibility that not fully deterministic ordering is beneficial for the system as a whole—comes from the interaction with other elements of the semiotic system.

5 Conclusions

The purpose of this chapter was to position linguistics in the empirical paradigm (as a socio-natural science) by drawing conclusions from the relevant research done by others, as well as to contribute to the research in that paradigm by proposing some specific solutions. These two types of contribution can be summed up in the following way.

1. **Linguistics in the paradigm of empirical sciences**

Linguistics in the empirical paradigm is characterized by the search for explanatory laws concerning linguistic data (observed pattern) which are implied

---

[33] to use the title of Surowiecki's book.

[34] This independence is crucial for the wisdom of the crowd to be efficient.

by relevant material systems. The trend to look for explanatory laws in linguistics took place for the first time on a big scale, when Chomsky (1957, 1965) proposed that generative grammar reflects the genetic makeup of a man. It has been argued in this chapter that the reason this approach failed to account for linguistic facts well enough and faces a number of insurmountable philosophical problems is that Chomsky's approach is purely reductionist, limited to considering the biological/ psychological make up of man with disregard of the crucial role of social influences (emergent phenomena). It was also argued that to arrive at better, explanatory laws concerning linguistic data from the perspective of empirical paradigm a cognitive-social approach to language is needed, just as advocated early on by Bronisław Malinowski and contemporarily by some main stream linguists, notably Capone (2005, 2006, 2009), Jaszczolt (2005), Kecskes (2010), Kopytko (1995, 2001a, b, 2004), Mey (2001, 2010), Włodarczyk (2011). However, I argue that for constructing better linguistic models, pragmatics cannot be treated merely as some independent, separate source of data from that constituted by language as an abstract semiotic system, as propounded by main stream linguistics resorting exclusively to formal qualitative tools, cf. Jaszczolt's default semantics. What is indispensable is integrally "connecting individual features with societal features", to use Kecskes' words. This can be done by assuming that language is a self-organizing semiotic system, an integral result of the history of communication processes in a given linguistic community, (cf. Altmann (1978), Grzybek (2006), Koehler (2005), Zielinska (2007), the Fife Graces Group)—a result of some sort of economy between the speaker and listener. In this vein, Altmann (1978) suggested that a possible self-organizational mechanism could resemble of natural selection. Any way, the validity of the assumption of self-organizational nature of language was corroborated both by considering the epistemic concerns presented in this chapter, as well as by pointing out the existence of quantitative data attesting to the fact that power laws characterize language.[35] The wealth of additional quantitative data copiously gathered, e.g., in the *Journal of Quantitative Linguistics*, or, for instance, in Koeler (2005), Grzybek (2006), Wulf (2003), Zielinska (2007), etc., indicates the existence of undisputable connexions of many kinds between symbolic data and frequency information. In an empirical paradigm, such connections can also be used to operationalize linguistic concepts through statistical characteristics of corpora, leading to an objective measurement also of selected semantic characteristics of language.

To sum up, we may quote Heylighen (2008) and state that according to the current philosophical understanding of the world, the cutting-edge results

---

[35] Originally, Zip inferred the power laws in language from the principle of minimal effort. Later on this principle was expressed as the optimalization of the effort involved in information transfer between the speaker and the addressee. It must be acknowledged, however, that the possibility of there being also some other sources of power laws, which would not imply the self-organizational character of language, has also been considered. Yet, these other models, cf, an overview in Kwapień (2010), require making a number of assumptions contradictory to what we know about language.

concerning language modelling should be expected in the research of language as a self-organizing system.

2. Modelling a form—meaning correlation mechanism within the paradigm of empirical sciences

My contribution to modelling language within an empirical paradigm consists primarily of two hypotheses. First, I have proposed a general mechanism of the form-meaning correlation process (categorization) that would fit a self-organizational perspective onto the nature of language. The mechanism postulated relies on two modes of language use: encoding and selective ones. (Selection takes place from options generated by expectations, goals and associations related to a specific situated speech act.) This model, as demonstrated, overcomes a number of major problems faced by an Aristotelian category relating form with content, (and its extensions, such as a prototype model of a category, or a fuzzy set model, etc.). The crucial problems present in the classical models of a linguistic category, which are avoided in the approach proposed, can be exemplified by the impossibility of crossing with basic encodings between separate individuals, and by the impossibility of defining the limits on allowable departure from the definition of a given linguistic category to verify the membership of a given exemplar in that category. Importantly, the mechanism proposed has been corroborated both in a qualitative terms and in a quantitative way. The quantitative verification took place by showing that the mechanism, (along with some additional law governing the word order in AAN phrases), has some statistically significant implications for the statistical order of adjectives in the AAN phrases.

Second, I proposed an additional mechanism of the form-content correlation formation, which is effected not only by purely cognitive processes, but also by social ones. Postulating such a mechanism is especially important in order to account realistically for the possibility of arriving at globally optimal solutions when creating language, i.e., ones valid for the whole system of interrelated phenomena. This latter mechanism shifts the burden of carrying complicated estimations resulting in creating a globally optimal semiotic system, (identical for the majority of community members faced with individual input data each) from the brain alone to the brain along with "the wisdom of the crowd" (a social process similar to natural selection, where gene transmission is substituted by linguistic replication and where the idiolects of socially successful individuals affect the decisions of individual speakers due to the frequency of the respective input data they face.) This can happen in the following way.

The idiolects of successful individuals reach a proportionally wide audience,[36] thus become highly represented in the corpuses of a larger than average number of members of a given linguistic community. This translates into a relatively high probability of the choice of a given variant by these individual speakers. (Cf.

---

[36] Additionally, the community members are motivated to remember the successful idiolect better due to prestige involved.

Skousen's (1989) modelling of past tense in Finnish). As a result, complex calculations concerning the influence of a given variant onto a global economy of a given semiotic system to be carried out by an individual are substituted by his assessment of the relative frequency of a given option. Importantly, by observing a new statistically significant correlation between some, possibly modified, linguistic construction and a new type of situated speech acts, functional novelty can be identified and established.

References

Ajdukiewicz, K. (1935). Die syntaktische Konnexit. In S. McCall (Ed.) *Polish logic 1920–1939* (pp. 207–231). Oxford University Press. (Translated from *Studia Philosophica, 1*, 1–27).

Altmann, G. (1978). Towards a theory of language. *Glottometrica, 1*, 1–25.

Ariel, M. (2008). *Pragmatics and grammar. In the series Cambridge textbooks in linguistics*. Cambridge: Cambridge University Press.

Bak, P. (1996). *How nature works. The science of self-organized criticality*. New York: Springer.

Bak, P., Tang, C., & Weisenfeld, K. (1988). Self-organized criticality. *Physical Review E, 38*, 364–374.

Berwise, J., & Perry, J. (1983). *Situations and attitudes*. In J. Barwise, & J. Perry. Cambridge, Mass: Brad-ford/M.I.T. Press.

Bolinger, D. (1967). Adjectives in English. Attribution and predication. *Lingua, 18*, 1–34.

Brickhard, M. H., & Campbell, R. L. (1992a). Some foundational studies concerning language Studies. *Journal of Pragmatics, 17*, 404–458.

Brickhard, M. H., & Campbell, R. L. (1992b). Clearing the grounds. *Journal of Pragmatics, 17*, 472–486.

Brook, R. (1986). Intelligence without representation. Unpublished research report, Cambridge Mass MIT Artificial Intelligence laboratory, 3.

Bruza, P. D., Kitto, K., Nelson, D., Cathy, L., & McEvoy, C. (2009). Is there something quantum-like about the human mental lexicon? *Journal of Mathematical Psychology, 53*, 362–377.

Bruza, P. D., & Peter, D. (2010). Quantum memory. *Australasian Science, 31*(1), 34–35.

Bunge, M. (1972). *Method, model and matter*. Heidelberg: Springer.

Bunge, M. (2003). Emergence and convergence. Qualitative Novelty and the Unity of Knowledge (Toronto Studies in Philosophy) Toronto: UTP inc.

Capone, A. (2005). Pragmemes (a study with reference to English and Italian) *Journal of Pragmatics 37*, 1355–1371.

Capone, A. (2006). On Grice's circle (a theory-internal problem in linguistic theories of the Gricean type.) *Journal of Pragmatics 38*, 645–669.

Capone, A. (2010). On the social practice of indirect reports (further advances in the theory of pragmemes). *Journal of Pragmatics, 42*(2), 377–391.

Carston, R. (2002). *Thoughts and utterances. The pragmatics of explicit communication*. Oxford: Blackwell Publishers Ltd.

Chomsky, N. (1957). *Syntactic structure*. The Hague: Mouton.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge: M.I.T. Press.

Chomsky, N. (1986). *Knowledge of language: its nature, origins, and use.* New York: Praeger.

Dummett, M. (1993). *The seas of language*: Oxford: Oxford University Press.

Elman, J. L., Bates, E. A., Johnson, M. H., & Karmiloff-Smith, A. (1997). *Rethinking innateness: A connectionist perspective on development (neural networks and connectionist modelling)*. Cambridge: The MIT Press.

Gopnik, A. (2009). *The philosophical baby: What children's minds tell us about truth, love, and the meaning of life*. New York: Farrar, Straus and Giroux.

Gopnik, A. (2007). Bayesian network, Bayesian learning and cognitive development. *Development Science, 10*(3), 281–364.

Grabińska, T. (1993). Teoria, Model, Rzeczywistośc. Wrocław: Wrocławska Oficyna Wydawnicza Politechniki Wrocławskiej.

Grabińska, T. (1994). Poznanie i modelowanie. Wrocław: Wrocławska Oficyna Wydawnicza Politechniki Wrocławskiej.

Grabińska, T. (1998). Od Nauki do metafizyki. Warszawa: PWN.

Griffiths, P. E, & Gray, R. D. (1994). Developmental systems and evolutionary explanation. *The Journal of Philosophy 91*(6), 277–304.

Grzybek, P. (2006). *Contributions to the science of text and language. Word length studies and related issues*. Netherlands: Springer.

Haken, H. (2010). *Synergetics: Introduction and advanced topics*. Berlin: Springer.

Haspelmath, M. (2008). Frequency vs. iconicity in explaining grammatical asymmetries. *Cognitive Linguistics 19*(1), 1–33.

Haspelmath, M. (2006). Creating economical morphosyntactic patterns in language change. In J. Good (Ed.) *Language universals and language change.* Oxford: Oxford University Press.

Heylighen, F. (2008). Complexity and self-organization. In M. J. Bates, & M. N. Mack (Eds.) *Encyclopedia of library and information sciences* (Taylor & Francis).

Horst, JS., McMurray, B., & Samuelson, L. K. (2006). Online processing is essential for learning: Understanding fast mapping and word learning in a dynamic connectionist architecture. *The Poceedings of the 28th Meeting of the Cognitive Science Society.*

Hull, D. L. (1988). Replication and interactors in his Science and Selection. Cambridge: Cambridge University Press.

Jaszczolt, K. (2005). *Default semantics: Foundations of a compositional theory of acts of communication*. Oxford: OUP.

Karmiloff-Smith, A. (1995). *Beyond modularity: A developmental perspective on cognitive science*. Cambrigde: Bradford Books.

Karmiloff, K., & Karmiloff-Smith, A. (2010). *Getting to know your baby*. London: Carroll & Brown.

Kennedy, C. (2007). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy, 30*, 1–45.

Keller, R. (1994). *On language change: The invisible hand in language.* London and New York: Routledge.

Kecskes, I. (2010). The paradox of communication: socio-cognitive approach to Pragmatics. *Pragmatics and Society, 1*(1), 55–73.

Kirby, S. (1999) . Syntax out of Learning: The cultural evolution of structured communication in a population of induction algorithms. In D. Floreano, J. D. Nicond, & F. Mondado (Eds.), ECAL99, 694–703.

Koehler, R. Synergetic Linguistics. In R. Koehler, G. Altmann, Rajmund G. Piotrowski (Eds.), *Quantitative Linguistik- Quantitative Linquistics*. Ein internationales Handbuch, Berlin/New York: de Gruyter, 760–774.

Kopytko, R. (2001a). From Cartesian toward non-Cartesian pragmatics. *Journal of Pragmatics, 6*, 783–804.

Kopytko, R. (2001b). The great dissolution a la Michel Foucault and the non-Cartesian pragmatics. *Journal of Pragmatics, 10*, 1637–1642.

Kopytko, R. (2004). The affective context in non-Cartesian pragmatics: A theoretical grounding. *Journal of Pragmatics, 3*, 521–548.

Kopytko, R. (1998). Relational pragmatics: Towards a holistic view of pragmatic Phenomena. *Studia Anglica Posnaniensia, 33*, 195–211.

Kopytko, R. (1995). Against rationalistic pragmatics. *Journal of Pragmatics, 23*, 475–491.

Kwapień, J. (2010). Fizyczne charakterystyki złożoności. The Institute of Nuclear Physics of Polish Academy of Science in Kraków. Raport 2036/AP.

Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge: MIT Press.

Lin, L., Osan, R., Shoham, S., Jin, W., Zuo, W., & Tsien, J. Z. (2005). Identification of network level coding units for real time representation of episodic experiences in the hippocampus. In *Proceedings of the National Academy of Science* USA, 102 no 17, pp. 6125–6130.

Lin, L., Osan, R., & Tsien, J. Z. (2006). Organizing principles of real-time memory encoding neural cliques assemblies and universal neural codes. *Trends in Neurosciences, 29*(1), 48–57.

Lin, L., Chen, G., Kuang, H., Wang, D., & Tsien, J. Z. (2007). Neural encoding and the concept of nest in the mouse brain. In *Proceedings of the National Academy of Science*, USA, 104 no 14, pp. 6066–6071.

Lakoff, G. (1987). *Woman fire and dangerous things*. Chicago: CUP.

Langacker, R. W. (1987). *Foundations of cognitive grammar*. Stanford: SUP.

Luther, P., & Fenk, A. (1984). Wird der Wortlaut von Saetzen zwangslaeufig schneller vergessen als ihr Inhalt? Zeitschrift fuer Experimentelle und Angewandte Psychologie, *31*(1), 101–123

Mey, J. L. (2001). *Pragmatics*. Oxford: Blackwell.

Mey, J. L., & Haberland, H. (2002). Linguistics and pragmatics, 25 years after. *Journal of Pragmatics, 34*(12), 1671–1682.

Mey, J. L. (2010). Reference and pragmeme. *Journal of Pragmatics, 42*, 2882–2888.

Mey, J. L. Forthcoming. Modular, cellular, integral: A pragmatic elephant? In A. Capone, F. Lo Piparo, & M. Carapezza (Eds.) *Perspectives on pragmatics and philosophy*. Dordrecht: Springer.

Mountcastle, V. (1998). *Perceptual neuroscience; the cerebral cortex*. Cambridge: Harvard UP.

Putnam, H. (1960). Minds and machines. In S. Hook (Ed.) *Dimensions of mind* (pp. 148–180). New York: New York University Press.

Putnam, H. (1975). The meaning of meaning. In K. Gunderson (Ed.) *Language, mind, and knowledge* Minneapolis: University of Minnesota Press.

Putnam, H. (1988). *Representation and reality*. Cambridge: MIT Press.

Recanati, F. (2011). *Truth-conditional pragmatics*. Oxford: OUP.

Recanati, F. (2004). *Literal meaning*. Cambridge: Cambridge University Press.

Searl, J. (1983). *Intentionality: An essay in the philosophy of mind*. Cambridge: Cambridge University Press.

Sachs, J. S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Percept Psychophys. 2:*437–442.

Seymour, M. (2010). Speech act pluralism, minimal content, and pragmemes. *Journal of pragmatics, 42*(11), 2870–2881.

Skausan, R. (1989). *Analogical modelling of language*. Dordrecht: Kluwer Academic Publishers.

Stainton, R. J. (2010). Philosophy of language. In L. Cummings (Ed.), *The pragmatics encyclopedia* (pp. 348–353). London: Routledge.

Stainton, R. J. (2005). In defense of non-sentential assertion. In Z. Szabó (Ed.) *Semantics vs. pragmatics* (pp. 383–457). Oxford: Oxford University Press.

Stainton, R. J. (2006a). *Words and thoughts: Subsentences, ellipsis, and the philosophy of language*. Oxford: Oxford University Press.

Stainton, R. J. (2006b). Neither fragments nor ellipsis. In L. Progovac et al. (Eds.) *The syntax of nonsententials.* (pp. 93–116). Philadelphia: John Benjamins.

Stanley, J. (2007). *Language in context*. Oxford: Clarendon Press.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Oxford: Blackwell.

Sweetser, E. E. (1990). *From etymology to pragmatics. Metaphorical and cultural aspects of semantic structure*. Cambridge: CUP.

Travis, C. (2001). *Unshadowed thought: Representation in thought and language*. Cambridge: Harvard University Press.

Wierzbicka, A. (1972). *Semantic primitives. Linguistische Forschungen no 23*. Frankfurt: Athenaeum.

Wierzbicka, A. (1985). *Lexicography and conceptual analysis*. Ann Arbor: Karoma.

Wierzbicka, A. (2010). Bilingualism and cognition: The perspective from semantics. In V. Cook, & B. Bassetti (Eds.) *Language and bilingual cognition* New York, NY: Psychology Press.

Włodarczyk, A. (2011). Distributed grammar. www.celta.paris-sorbonne.fr/anasem/intexASMIC.html.

Wulf, S. (2003). A multifactorial corpus analysis of adjective order in English. *International Journal of Corpus Linguistics, 8*(2), 245–282.

Wittgenstein, L. (1963). Philosophical investigations. Oxford: Basil Blackwell. Originally 1953.

Wittgenstein, L. (1961). Tractatus Logico-Philosophicus, London: Routledge and Kegan Paul. Originally 1921.

Zielinska, D. (1997). A note on the extended functional analysis. *Journal of Pragmatics, 27*, 841–843.

Zielinska, D. (1999). The selective mode of language use; the way nature language adopted itself to describing the word around us. *Zeszyty Naukowe UJ, 119*, 173–176.

Zielinska, D. (2007). The selective mode of language use and the quantized communicative field. *Journal of Pragmatics 39*(5), 813–830.

Zielinska, D, (2007b), Proceduralny model języka. Językoznawstwo z pozycji teorii modeli nauk empirycznych. WUJ. Kraków.

Ziff, P. (1960). *Semantic Analysis.* Ithaca, NY: Cornell University Press.

Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Boston: Addison-Wesley.

Zipf, G. K. (1935). *The psycho-biology of language*. London: Routledge & Sons Ltd.