Exploring the *N*-th Dimension of Language

Prakash Mondal

Cognitive Science Lab, International Institute of Information Technology
Gachibowli, Hyderabad 500032, India

Abstract. This paper is aimed at exploring the hidden fundamental computational property of natural language that has been so elusive that it has made all attempts to characterize its real computational property ultimately fail. Earlier natural language was thought to be context-free. However, it was gradually realized that this does not hold much water given that a range of natural language phenomena have been found as being of non-context-free character that they have almost scuttled plans to brand natural language context-free. So it has been suggested that natural language is mildly context-sensitive and to some extent context-free. In all, it seems that the issue over the exact computational property has not yet been solved. Against this background it will be proposed that this exact computational property of natural language is perhaps the *N*-th dimension of language, if what we mean by dimension is nothing but universal (computational) property of natural language.

Keywords: Hidden fundamental variable; natural language; context-free; computational property; *N*-th dimension of language.

1 Introduction

Let's start with the question "What exactly is the universal computational property of natural language?" The simple and perhaps a little mysterious answer is that nobody knows what it is. Then we may ask the other natural and subsequent question "Why?". Even here we are nowhere nearer to having any clearer grasp of the reason why the exact natural language computational property is beyond our reach. Perhaps it is not knowable at all. Or perhaps it is knowable, but the question on this issue is not a relevant one. Or even perhaps there is no single answer to this question. This is the reason why we have till now a plethora of grammar formalisms or models of natural language defining and characterizing different classes of language; some defining context-free languages, some context-sensitive, and some are of mixed nature with each type having weak or strong generative capacity. Here in this paper it will be argued that this seemingly unknowable computational fundamental having universal applicability and realization constitutes the *N*-th dimension of natural language. Hence, we may know a little bit about *N-1,...,N-m* (when *m* is an arbitrary

Special issue: Natural Language Processing and its Applications.
Research in Computing Science 46, 2010, pp. 55-66

Received 28/10/09
Accepted 16/01/10
Final version 08/03/10

number and $m < N$) computational properties of natural language, but we do not understand the N-th. This is the N-th dimension of natural language in that we do not yet know what N stands for. On the surface of it all, it appears that the exact computational property of natural language defies any attempt to uncover it.

However, it will also be argued that it may have, though not necessarily, something to do with the emergent nature of natural language since natural language is an emergent entity derived out of the interaction with several cognitive domains involving emotion, social cognition, vision, memory, attention, motor system, auditory system etc. And here at the linguistic level, language emerges through integrated and interwoven, but partially constrained, interactions between syntax, semantics, morphology, lexicon and phonology which form an overlapping network. This is a case of recursive emergence in that there are two layers of emergence: one at the level of language and another at the level of cognitive architecture as a whole.

Apart from all that, it will also be shown that as we do not know the N-th dimension of natural language, it hampers our progress in natural language processing in particular and artificial intelligence in general. That means we do not yet have a universally implementable grammar formalism or formal model of natural language that can characterize to the fullest possible extent *only and all* conceivable natural languages and all possible natural language expressions in some sense but metaphorically like Universal Turing machine. The lack of this hinders us in building a parser that will have in it such a grammar as being able to universally generate all possible natural languages and all possible natural language expressions. The same can be said of natural language generation system, natural language understanding system etc. However, it does not now exist, so it does not necessarily entail that we can never achieve this. Once achieved, it will perhaps serve as a key to towards solving most of our problems we are facing in building a universally implementable language model that may facilitate cost-effective solutions across languages. Let's now explore the possible nature, form and origin of the N-th dimension of language.

2 Laying the Exploration out against the Backdrop

From the earliest days of Chomskyan hierarchy, the computational nature of natural language was assumed to be context-free. In fact, Chomsky's generative grammar started out with context-free grammar [1]. Even before that, language was characterized as context-free in the structuralist tradition. But later Chomsky [2] himself argued against the validity of context-free grammar in characterizing natural language. And with this the debate started about whether natural language is context-free or not. With this a cascade of papers came off arguing that natural language cannot be context-free [3], [4], [5], [6], [7], [8]. But Pullum and Gazdar [9] and Pullum [10] on the other hand threw skepticism over such arguments. In the middle, Joshi [11], [12] tried to strike a middle ground between mild context-sensitivity and context-freeness with his Tree Adjoining Grammar. Even Chomsky's updated versions of transformational grammar [13], [14], [15] had to constrain the power of his generative grammar by imposing filters in the forms of hard constraints. This is to enrich context-free phrase structure grammar. The issue did not, however, come to a halt. Head-Driven Phrase Structure Grammar [16] and Lexical Functional Grammar [17] are also of a sort of similar nature.

This shows that the question of what natural language belongs to as far as the class of formal languages is concerned is still a moot point. Is it because we classify languages (both artificial and natural) in terms of the Chomskyan hierarchy? Or is it because of some sort of natural order that has been imposed upon natural languages such that categorizing them into an exact class is not possible within the boundaries of classes of language that we know of. It seems that both questions are related to each other in that the available classes of languages in Chomskyan hierarchy do not take us much further in identifying and thereby understanding what computational property natural language universally possesses. The problem with such an approach is that it is not still clear what it would mean to revise and modify Chomskyan hierarchy in view of the fact that we cannot determine the exact computational dimension all natural languages have. In sum, what is clear enough is that this dimension which can be called the *N*-th dimension of language has not been broadly explored. People are debating whether natural language is context-free or context-sensitive; nobody seems to be concerned about whether natural language has some unknown computational dimension separate, segregated from and independent of context-freeness and context-sensitivity. Nobody wonders whether it is reasonable at all to get mired into the possibility of natural language being context-free or context-sensitive, because right wisdom among the computational linguistics community says language is not context-free at all. So one possibility may be to explore how much and to what extent natural language is context-free [18]. Nobody has yet done it, but its implications may be nontrivial in some respects. Let's now turn to some examples to get a flavor of the complication that the hidden dimension of language creates. If we have cases like the following in natural language,

(1) The girl the cat the dog chased bit saw ... the boy.
(2) I know he believes she knows John knows I believe .... we are here.

we have a class of languages like $a^n b^n$ which is exactly the class of context-free languages. In (1) we have a rule of the form $(NP)^n (VP)^n$ and $(NP\ VP)^n$ in (2). But it is quite clear that such cases are just a fragment of the set of *all* possible natural language expressions. Context-sensitivity also appears in a range of natural language expressions (especially in Dutch). Consider the following from English,

(3) What did I claim to know without seeing?
(4) What did you believe without seeing?

Cases like these show multiple dependencies between 'what' and the gaps after 'know' and 'seeing' in (3), and between 'what' and the gaps after 'believe' and 'seeing'. They may well fall into a pattern of context-sensitivity. Even such cases are pretty clear as to the way they reveal computational properties of context-sensitivity. What is not clear is that these examples neither explicitly demonstrate whether this can suffice for us to characterize natural language invariably as context-free or context-sensitive as long as one looks at it holistically, nor do they show the exact computational property within which both context-freeness and context-sensitivity can be accommodated. Overall, it seems that in showing different fragments of natural language expressions from different languages as context-free and context-sensitive, we are missing out some fundamental generalizations which might underlie the way natural language shows both context-sensitive and context-free properties. We still do not know what these generalizations may amount to, but they clearly point to the vacuum as it is unclear how to fit the property of context-freeness into that of context-sensitivity.

All this makes it quite evident that natural language is neither fully context-free nor fully context-sensitive. If this is the case, then it is of some unknown computational property which is intermediate between context-freeness and context-sensitivity; and it may well be possible it has some property of type 0 grammar as well. Nobody knows. This is what we may term as the *N*-th dimension of language. Let's now move over to its nature.

3 The Nature of the *N*-th Dimension

A little bit of precision is now necessary. Hence, for the sake of precision, we will model this *N*-th dimension as a 4-tuple

$$\mathcal{D} = < D_n, E_p, \mathcal{P}_w, \infty >. \tag{1}$$

D_n is the set of hidden sub-dimensions characterizing and representing the components of the *N*-th dimension. E_p is the unknown expressive power of the *N*-th dimensional computational property of natural language. $\mathcal{P}_w$ is the Cartesian product of $D^i_n \otimes, \dots, \otimes D^m_n$ where $m \neq n$. And ∞ is the distribution of the other three in a linear (or real or complex) space $\acute{S}$.

What is required at this stage is the stipulation that D_n determines how the unknown computational property appears whenever natural language is put under analysis, since this set contains the hidden components constituting the *N*-th dimension. So $D_n \models (\forall x)\ \lambda y\ [x \triangleq y]$ when $x \in D_n$ and $y \in P(D_n)^C$, the complement of the power set of D_n. This makes sure that it may well be possible that some components of the hidden *N*-th dimension of natural language correspond to some other components in the other class of computational properties in the Chomskyan hierarchy.

This is what leads us to make such claims that natural language is context-free or context-sensitive. No moving off, let's now say something about E_p. It can describe and generate all possible natural languages and all possible natural language expressions with its absolute generative capacity. Let's first denote $\mathcal{L}$ the set of all possible languages (both artificial and natural). And then let $\mathcal{L}^C$ denote the set of only and all possible natural languages. And $\boldsymbol{E}$ is the set of all possible expressions in natural language. So we can now say that

$$E_p \succcurlyeq \boldsymbol{E}(\mathcal{L}^C) \mid \acute{\boldsymbol{E}} \subseteq \boldsymbol{E} \wedge \acute{\boldsymbol{E}} = \phi \text{ when } E_p \not\succcurlyeq \acute{\boldsymbol{E}}. \tag{2}$$

Here read the sign $\succcurlyeq$ as 'generable from'. This guarantees that there cannot exist any subset of the set of all possible expressions of natural languages which is not generable from E_p. At last ∞ orders the 4-tuple in configurations we are still unaware of. Roughly for the purpose of modeling, let's say there can possibly be $\mathbf{C} = \{C_1, \ldots, C_k\}$ such configurations where k can be any arbitrary number. So we have

$$O_1(D_n), O_2(E_p), O_3(\mathcal{P}_w) \vdash O_i(\infty). \tag{3}$$

This is what expresses the ordering O_i of all the other three in the 4-tuple with respect to ∞. Now it becomes clear that the *N*-th dimension of language is quite complex in nature. Even if it can be mathematically modeled, it is at best patchy. We can never be sure that we have grasped it. For the sake of simplification, it was vital to model it. And now a bit of focus will be shed upon how it might work.

3.1 How the *N*-th Dimension Possibly Works

The description of how the *N*-th Dimension can possibly work is provided to facilitate an understanding of the way it works in natural language in general. In addition, this will also help us in proceeding one step ahead in realizing why it is so elusive as well that it has cudgeled so many brains of computer scientists and linguists both. Given that D_n is the set of hidden sub-dimensions characterizing and representing the components of the *N*-th dimension, we are now able to say that

$$\int f_1, \ldots, f_m(x_1, \ldots, x_m)\, d\,x_1, \ldots, d\,x_m.\ \Delta\infty. \tag{4}$$

where $x_1, \ldots, x_m \in D_n$ when $m \leq n$ such that $|D_n| = n$. This shows that $x_1, \ldots, x_m \in D_n$ may appear in different distributions which determine differential probabilities associated with each distribution. The nature of such very differential probabilities themselves space out the way the *N*-th dimension appears in our analysis. Hence it is not simply a product of a linear function mapping a class of computational property into something, for example, from context-freeness into the *N*-th dimension. Using type predicate calculus [19], we can say that the following

$$\frac{\Gamma \vdash t_1 : T_1, \ldots, \Gamma \vdash t_n : T_n}{\Gamma \vdash F(t_1, \ldots, t_n) : O_i(T_1, \ldots, T_n)}. \tag{5}$$

does not hold for the *N*-th dimension of natural language when Γ is a context, $t1, \ldots, tn$ are computational properties of formal languages in the Chomskyan hierarchy like finite-state*ness*, context-freeness, context-sensitivity etc. which can be labeled as being of different types $T1, \ldots, Tn$. Let's us suppose here that the operator O_i fixes how the *N*-th dimension arises out of such a mapping from those computational properties $t1, \ldots, tn$ into the types $T1, \ldots, Tn$. But even this does not capture the *N*-th dimension as analyzing natural language as being a complex of different computational types does not help us much in that we do not know how O_i really works. We have just a vague idea of how it may function given that natural language has the properties of the types *T*1, ..., *Tn*.

Other dimensions of natural language also do not take us any farther in such a situation. We now know that a language has a grammar along with a lexicon and it generates expressions of that language. This is one of the (computational) properties of (natural) language. Given a grammar and a lexicon, the generated expressions of the grammar in that language are string sets. Of course language users do not think of their languages as string sets. However, this is another dimension. So is infinite recursion of string sets in natural language. Constraints on natural language constructions are determined by a set of universal principles. Pendar [20] has emphasized that natural language is a modularized computational system of interaction and satisfaction of such soft constraints. Optimality theory [21] a theory of natural language has built into it just such a system. These are some of the dimensions of language. Now consider the question: what if we add them together and then derive the *N*-th dimension since above all cumulative summing over all dimensions is equivalent to a higher dimension? A little bit of common sense shows that this may well be wrong. Let's see how. Let's sum all the existing dimensions of natural language

$$S^{(d)} = \sum_{i=1}^{n} d_i \qquad (6)$$

Here $S^{(d)}$ is never equivalent to the *N*-th dimension. Had it been the case, then by principle we would obtain the following

$$S^{(d)} = < D_n, E_p, \mathcal{P}_w, \infty >. \qquad (7)$$

But a careful consideration of the matter suggests that we do not get E_p out of $S^{(d)}$ even if the equivalence shows that we should. We do not get E_p out of $S^{(d)}$ mainly because $S^{(d)}$ may well generate some artificial language which does not have E_p. Analogically, let's suppose that entities in the world have only two dimensions, namely, width and length, but not height. In such a case combining the known will perhaps never give us the third one, height.

All this indicates that the *N*-the dimension is much more complex than has been recognized. However, it does not give us any reason for rejoicing in such a nature of natural language. Rather, it challenges us to decode its complexity which is so overwhelming that it is at present out of the bounds of our rational hold. This brings us to the next issue, that is, the relation of complexity to the *N-th* dimension.

3.2 Complexity and *N*-th Dimension

It may now be necessary to draw up connections between the *N*-th dimension and complexity. Is there any connection between them at all? How are they related? Of course, it is true that answers to such questions are not easy to get. We shall not try to answer those questions over here. Instead, an attempt will be made to map out the space over which the *N*-th dimension can be said to be complex. Now let's proceed. The notion of Kolmogorov complexity will be used here as it is handy enough in that it has been proved that it is machine independent [22]. We know that Kolmogorov complexity is the shortest possible program for a Universal Turing machine to compute the description of an object. Now let's assume that the object under consideration is the *N*-th dimension. What would its measure of complexity be with respect to Kolmogorov complexity? Mathematically speaking, we have

$$\mathcal{K}_u(\mathcal{D}) = \min_{p\,:\,u(p) = \mathcal{D}} l(p) \tag{8}$$

when $\mathcal{K}_u(\mathcal{D})$ is the Kolmogorov complexity computed by a Universal Turing machine u and p is the shortest possible program that can give a description of $\mathcal{D}$, the *N*-th dimension. If according to Kolmogorov complexity, the complexity of $\mathcal{D}$ is the length of the shortest possible program for computing $\mathcal{D}$, then such a program p does not independently exist. The term 'shortest possible program' is relative. Unless we find out a set of such programs P= $\{p_i, \ldots, p_k\}$, we can never measure the complexity of $\mathcal{D}$. But obviously we are sure that the $\mathcal{K}_u(\mathcal{D})$ is lower than l(this paper itself)! Now suppose the 4-tuple model that has been provided above to model $\mathcal{D}$ is actually equal to $\mathcal{K}_u(\mathcal{D})$. What is the implication then? So the length of any program that can model the 4-tuple is the complexity of ($\mathcal{D}$). Are we done with it? Possibly no. We still do not know whether this is *the* shortest program in all possible worlds. But if the program above is by the present criteria the complexity measure for $\mathcal{D}$, we have to accept that. However, the problem is still open ended, because we have not yet fully understood $\mathcal{D}$ at all despite the progress that has been made in formal language theory. This leads us to give the following definition

Definition (Unboundedness of the Complexity of $\mathcal{K}_u(\mathcal{D})$): Given that there may well exist a set of programs P= $\{p_i, \ldots, p_k\}$ whose actual length we are yet to figure out, $\mathcal{K}_u(\mathcal{D})$ is not bound from above by any constant *C*.

This definition ensures that we are still in uncertainty as to what the actual complexity of $\mathcal{D}$ can be. Science consists in simplification of complex phenomena. Here the *N*-th dimension of natural language is such a case. Even if we aim at simplifying it, it defies such attempts at simplification. That is what this section teaches us. This may be because of another important property of natural language, it is an emergent entity. Let's see what it has for the *N*-the dimension of natural language.

3.3 Emergence and the *N*-Dimension of Language

Now we have come near to the last candidate that has any claim that it can be related to the *N*-th dimension. It is not, of course, necessary that a link has to be forged between the *N*-th dimension and the case of emergence according to which the whole is not a sum of its parts. The case of natural language being emergent has been made quite forcefully by Andreewsky [23]. It appears that language is recursively emergent in that language emerges out of the interaction with several cognitive domains involving emotion, social cognition, vision, memory, attention, motor system, auditory system etc. at the higher level. And then at the linguistic level spanning out at the lower level, language emerges through integrated and interwoven, but partally constrained, interactions between syntax, semantics, morphology, lexicon and phonology which form an overlapping network. One type of emergence is thus embedded into another. The cognitive architecture of language at the higher level has been shown below

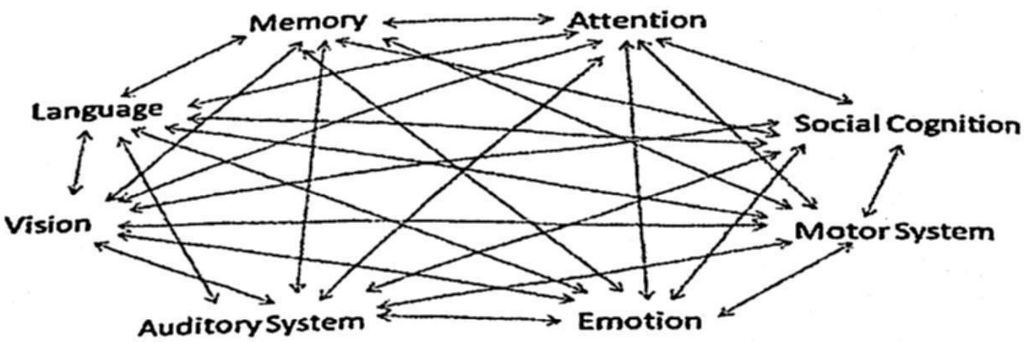

Fig. 1. Interconnection of language to other cognitive domains

Fig.1. above is symmetric with certain constraints on mutual interaction of the cognitive domains/systems such that not all information passes from one domain/system to another. Such constraints- computational, algorithmic and implementational in Marr's [24] sense- are specified by our genome. Here all these domains have varying interactions with respect to each other; what this means is that all these cognitive domains have differential mutual interactions with respect to each other. And by following Mondal [25] this can be represented through a distribution of joint probability in the following equation The interaction potential I(φ) can be characterized as

$$I(\varphi) = \sum_{i=1}^{N} P_i \int d_{c1,\dots} d_{cN}\, \delta(_{c1 \dots cN}).\ \Delta\psi. \tag{9}$$

Here *c1* ... *cN* are differential probability functions coming out of the interaction dynamics of language with other cognitive domains and N in *c1* ... *cN* must be a finite arbitrary number as required for the totality of cognitive domains; *P* is a probability function; $\Delta\psi$ is a temporal distribution. The cognitive domains are functionally coherent units of emergent self-organizing representational resources realized in diffuse, often shared and overlapping, networks of brain regions as mediated by the bidirectional interactions in Fig. 1.

Now it may be asked what connection there exists between such cognitive emergence and the *N*-th dimension. Note that what is emergent is so complex that its complexity can be neither represented nor predicted. The above discussion on complexity and the *N*-th dimension just shows this much more clearly. In addition, emergent entities are chaotic and random at most times. Natural language is no exception. The way the lack of grasp of *N*-th dimension has been the centerpiece of debate on the computational property of natural language evidences the fact that it has appeared randomly as different to different researchers. That is why it has been once claimed to be of finite-state nature or sometimes context-free; sometimes mildly context-sensitive and so on. This is perhaps because of the emergent randomness and an apparent lack of order in the dimensions of natural language. One may also argue that here two separate issues- the computational property of linguistic structures and the evolution of language – are being conflated. But one should also note that whatever the (computational) properties of natural language are, they have come only through the mechanisms of evolution. Even many distinguishing properties of natural language like ambiguity, vagueness, redundancy, arbitrariness have appeared in natural language only as products or bi-products of evolution which is generally blind and so not goal-driven [26]. And it may well be plausible that it is such properties that confound and mystify the exact computational property of natural language. However, whatever the *N*-th dimension turns out to be, it will certainly be one of the greatest challenges for natural language processing, mathematical linguistics and theoretical computer science in general. This brings us nearer to the issue of what *N*-th dimension means for natural language processing.

4 What the *N*-th Dimension Means for Natural Language Processing

Grammar formalism is at the heart of natural language processing. Whether we are doing parsing, natural language generation, natural language understanding or machine translation etc., grammar formalisms are inevitable. Even if in recent years grammar induction through statistical techniques has been possible, the combination of robustness and minimum computational costs in terms of time and storage, still a goal aimed at, demands that a grammar formalism be available at hand. Even in cases where grammar induction is possible through machine learning, having a grammar formalism is always indispensable in that before the training period a standard model must be available for modeling [27]. In fact, there can be, in a more general sense, no parsing without a grammar. And parsing is at the heart of a range of natural language processing tasks ranging from natural language generation, natural language understanding or machine translation to dialog systems etc. All present types of parsing use one or another grammar formalism based on different types of formal grammar [28], [29], [30], [31], [32], [33], [34].

So let's us confront the question of the *N*-th dimension in a rather straightforward manner. If we become successful some day in future in finding out this dimension, however chaotic and random it may well be, it will be one of the greatest achievements in the entire field of natural language processing and AI in general; for this will place us on a pedestal in building grammar models that can be implemented across languages universally. There will be no variation, no fractionation in the way grammar formalisms and models are built and implemented computationally. Once we coast to the *N*-th dimension of natural language processing, it will enable us in finessing and enhancing the effectiveness of our natural language processing tools which are still far behind compared to humans [35].

This will also help us in having a uniformity in our grammar models thereby eliminating the barrier of cross-linguistic transfer. This will propagate into multiple channels of progress in natural language technology in general. Just as the discovery of formal language hierarchy opened up new vistas for computation in general which ultimately gave birth to the field of natural language processing or NLP, the discovery of the *N*-th dimension will take us one revolutionary step farther toward a 'golden age' of natural language technology. One may complain that too much optimism is misplaced; but consider the sort of achievement in formal language theory this will amount to, as formal language theory has long been embroiled in the debates on the exact computational nature of natural language. After a lot of years, the efforts have been futile in that nobody has yet discovered the *N*-th dimension. What if we discover it some day? Would not it constitute a bigger achievement than even the discovery of Chomskyan hierarchy? Well, only the future will tell us. We are then nearing the conclusion with this ray of hope.

5 Concluding Remarks

This paper has sketched a brief sketch of what the *N*-th dimension is, its nature, its possible functioning and its links to complexity, emergence and NLP. After this brief tour it seems that we have achieved very little about the computational nature of natural language, even if we may know a lot about formal or artificial languages. The reason behind this is perhaps the *N*-th dimension itself. It is a major bottleneck in our path towards being able to compute all natural languages on earth. It may sound too ambitious, but any NLP researcher may be asked whether he/she does not want to see how his/her algorithms fare in other languages. The answer must at better be yes.

Moreover, it is also a dream for formal language theorists and perhaps mathematical linguists. In such a situation there is still enough scope for the claim that we may never be able to discover the *N*-th dimension of natural language because it is emergent, random and chaotic. We think such claims are as plausible as the claim that there does exist the *N*-th dimension of natural language to be discovered. Nobody can predict the path of the trajectory that research takes on in any field. The same applies to NLP as well. Even if ultimately it transpires that the entire research program following on from the search for the computational class of natural language is flawed and wrong, this will still be fruitful in showing us the right path and a gateway towards a greater understanding of natural language.

References

1. Chomsky, N.: Syntactic Structures. Mouton, The Hague (1957).
2. Chomsky, N.: Formal Properties of Grammars. In: Luce, R. D., Bush, R. R., Galanter, E. (eds.) Handbook of Mathematical Psychology. Vol. II. Wiley, New York (1963).
3. Postal, P. M.: Limitations of Phrase-Structure Grammars. In: Fodor, Jerry, Katz, Jerald. (eds.) The Structure of Language: Readings in the Philosophy of Language. Prentice Hall, New Jersey (1964).
4. Langendoen, D. T.: On the Inadequacy of Type-2 and Type-3 Grammars for Human Languages. In: Hopper, P. J. (ed.) Studies in Descriptive and Historical Linguistics. John Benjamins, Amsterdam (1977).
5. Higginbotham, J.: English is not a Context-Free Language. Linguistic Inquiry. Vol. 15, 225-234 (1984).
6. Langendoen, D. T., Postal, P. M.: The Vastness of Natural Language. Blackwell, Oxford (1984).
7. Langendoen, D. T., Postal, P. M.: English and the Class of Context-Free Languages. Computational Linguistics. Vol. 10, 177-181 (1985).
8. Shieber, S. M.: Evidence against Context-Freeness of Natural Language. Linguistics and Philosophy. Vol. 8, 333-343 (1985).
9. Pullum, G. K., Gazdar, G.: Natural Languages and Context-Free Languages. Linguistics and Philosophy. Vol 4, 471-504 (1982).
10. Pullum, G. K.: On Two Recent Attempts to Show that English is not a CFL. Computational Linguistics. Vol. 10, 182-186 (1985).
11. Joshi, A.: Factoring Recursion and Dependencies: An Aspect of Tree-Adjoining Grammars and a Formal Comparison of some Formal Properties of TAGs, GPSGs, PLGs, and LFGs. In: Proceedings of the 21st Annual Meeting of the Association for Computational Linguistics, pp. 7-15. Association for Computational Linguistics, Cambridge, MA (1983).
12. Joshi, A.: Tree Adjoining Grammars: How much Context-Sensitivity is Required to Provide Reasonable Structural Descriptions? In: Dowty, D., Karttunen, L., Zwicky, A. M. (eds.) Natural Language Processing: Psycholinguistic, Computational and Theoretical Perspectives. Cambridge University Press, New York (1985).
13. Chomsky, N.: Aspects of the Theory of Syntax. MIT Press, Cambridge, Mass. (1965).
14. Chomsky, N.: Rules and Representations. Columbia University Press, New York (1980).
15. Chomsky, N.: The Minimalist Program. MIT Press, Cambridge, Mass. (1995).
16. Pollard, C., Sag, I.: Head-Driven Phrase Structure Grammar. Chicago University Press, Chicago (1994).
17. Bresnan, J.: The Mental Representation of Grammatical Relations. MIT Press, Cambridge, Mass. (1982).
18. Pullum, G. K.: Systematicity and Natural Language Syntax. Croatian Journal of Philosophy. Vol. 7. 375-402 (2007).
19. Turner, R.: Computable Models. Springer, London (2009).
20. Pendar, N.: . Linguistic Constraint Systems as General Soft Constraint Satisfaction. Research on Language and Computation. Vol. 6:163–203 (2008).
21. Prince, A., Smolensky, P.: Optimality Theory: Constraint Interaction in Generative Grammar. Rutgers University Centre for Cognitive Science, New Jersey (1993).

22. Cover, T. M., Thomas, J. A.: Elements of Information Theory. Wiley, New York (1991).
23. Andreewsky, E.: Complexity of the Basic Unit of Language: Some Parallels in Physics and Biology. In: Mugur-Schachetr, M., van der Merwe, A. (eds.) Quantum Mechanics, Mathematics, Cognition and Action. Springer, Amsterdam (2002).
24. Marr, D.: Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. W H Freeman, San Francisco (1982).
25. Mondal, P.: How Limited is the Limit? In: Proceedings of the International Conference on Recent Advances in Natural Languages Processing, pp. 270-274. RANLP, Borovets (2009).
26. Kinsella, A. R., Marcus, G. F.: Evolution, Perfection and Theories of Language. Biolinguistics. Vol 3.2–3: 186–212 (2009).
27. Ristad, E. S.: The Language Complexity Game. MIT Press, Cambridge, Mass. (1993).
28. Goodman, J.: Semiring Parsing. Computational Linguistics. Vol. 25, 573.605 (1999).
29. Charnaik, E.: A Maximum-Entropy-Inspired Parser. In: Proceedings of North American Chapter of the Association for Computational Linguistics-Human Language Technology, pp. 132-139. Association for Computational Linguistics, Washington (2000).
30. Eisner, J.: Bilexical Grammars and their Cubic-Time Parsing Algorithms. In: Bunt, H., Nijholt, A. (eds.) Advances in Probabilistic and Other Parsing Technologies, pp. 29.62. Kluwer Academic Publishers, Amsterdam (2000).
31. Nederhof, M. J.: Weighted Deductive Parsing and Knuth's Algorithm. Computational Linguistics. Vol. 6. 135.143 (2003).
32. Pradhan, S., Ward, W., Hacioglu, K., Martin, J. H., Jurafsky, D.: Shallow Semantic Parsing Using Support Vector Machines. In: Proceedings of North American Chapter of the Association for Computational Linguistics-Human Language Technology, pp. 233-240. Association for Computational Linguistics, Boston, MA (2004).
33. Charnaik, E., Johnson, M.: Coarse-to-Fine n-Best Parsing and MaxEnt Discriminative Reranking. In: Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics, 173-180. Association for Computational Linguistics, Michigan (2005).
34. Huang, L., Chiang, D.: Better k-Best Parsing. In: Proceedings of International Workshop on Parsing Technologies (IWPT), pp. 53-64. Vancouver, Omnipress (2005).
35. Moore, R. K.: Towards a Unified Theory of Spoken Language Processing. In: Proceedings of 4th IEEE International Conference on Cognitive Informatics, pp. 167-172. IEEE Press, Irvine, CA (2005).