

THE SHORT OVERVIEW OF THE AIMS, METHODS, AND RESULTS OF THE
LOGICAL GRAMMAR OF THE GEORGIAN LANGUAGE

Pkhakadze K., Chichua G., Chikvinidze M. Maskharashvili A.

Abstract. In the paper our aims, methods and results toward the construction Logical Grammar of the Georgian language are overviewed. Herewith: the complete construction of the logical grammar of the Georgian language is declared as our main scientific aim; the understanding of in us naturally and unconsciously existing knowledge of the Georgian language and its description by the help of the Sh.Pkhakadze's formally developable mathematical language is declared as our main methodology. In addition, in the paper, there are shortly considered some of our new views as the main results of our research. They are related to:

1. The Georgian speech alphabet and syllable structure;
2. The contracting morphemes which exist in the Georgian language and which are neither derivational nor inflectional ones;
3. The declination and conjunction in the Georgian language;
4. The mathematical structure of Georgian declarative sentences;
5. The isomorphic mathematical theory of the core part of the Georgian language.

Keywords and phrases: The logical grammar of the Georgian language; The Georgian speech alphabet and syllable structure; The existence of contracting morphemes which are neither derivational nor inflectional ones; The isomorphic mathematical theory of the core part of the Georgian language.

AMS subject classification: 03B65, 68T50, 68Q55, 91F20.

The short overview of the aims and methods of the logical grammar of the Georgian language. The logical grammar of the Georgian language aims at the construction of the mathematical theory of the Georgian language [4] and the Georgian intellectual computer system [5]. The method of the research is based on the view that the rules of the Georgian language exist independently from us. This means that the main method of our natural research is to understand the rules, which naturally exist in us, and to describe and formalize them as a mathematical theory [6], [7]. Herewith, it must be mentioned that our methods are based on Montagues approaches, however, our methods differ from his ones, which were created for the English language [13]. This means that instead of using λ -calculus, we are using Notation Theory [10]. Also, we can not use those methods of Chomsky, which are dictated to him by English [13]. Thus, instead of basing on the methods which are based on the principle of immediate constituent, our researches are based on the method of describing Georgian lingual datum as symbols of a formally developable mathematical language. We call this method the method of the direct formal-logical i.e. natural mathematical description of the language [8]. In addition, on the basis of our researches of the speech nature of the Georgian spoken language, we have defined the notion of the Georgian speech alphabet,

which gives new views on syllable structure and new methodological approaches of constructing text-to-speech and speech-to-text systems [3].

The short overview of the results of the logical grammar of the Georgian language. As we are studying Georgian phonetic, morphological, syntactic, semantic, logical and communicational systems, in the paper there are described new views on the Georgian speech alphabet and syllables, and, also, about morphemes, declination, conjugation and declarative sentences.

The speech alphabet according to the logical grammar of the Georgian language: We call the indivisible, i.e. non-separable, i.e. simple speech sound of a spoken language as speech unit of this language and we declare that the speech alphabet of any spoken language should be comprised from the speech units of this spoken language. Thus, the speech alphabet of the Georgian spoken language should be comprised from the Georgian speech units, the view about which will be presented below. Georgian written language is denoted by GWL, and its alphabet by GWLAlphabet. $\text{GWLAlphabet} = \text{GWLAlphabet}(V) \cup \text{GWLAlphabet}(C)$, where $\text{GWLAlphabet}(V) = \{a, e, i, o, u\}$ is the set of the Georgian vowels, and $\text{GWLAlphabet}(C) = \{b, g, d, v, z, t', k, l, m, n, p, \check{z}, r, s, t, p', k', \bar{g}, q, \check{s}, \check{c}', c', j, c, \check{c}, x, \check{j}, h\}$ is the set of Georgian consonants. By GSL we denote the Georgian spoken language and by GSLAlphabet, we denote its alphabet. Herewith, we are stating that by pronouncing any Georgian vowel, we are getting one speech unit, because none of the Georgian vowels can be decomposed into simpler speech sounds. So, in the GSL, there is 5 vowel speech units presented by the set $\text{GSLAlphabet}(V) = \{\underline{a}, \underline{e}, \underline{i}, \underline{o}, \underline{u}\}$ (here an underlined vowel denotes a speech unit attached to it). When dealing with Georgian consonant speech units, we are based on the fact that, there is no Georgian consonant which can be pronounced out alone, without a vowel. Thus, we state that a Georgian consonant cannot be understood as a notation of speech unit of the spoken language and that in the spoken language there are 10 speech units associated with it, from which 5 is received by pronouncing the syllables which are built up by adding to this consonant vowels from the right-hand side and 5 is received by the similar procedure in which vowels are added to this consonant from the left-hand side. We call them simple right and left syllables. In this way, in the Georgian spoken language there are $140 (= 28 \times 5)$ simple right syllables the set of which we denote by $\text{GSLAlphabet}(CV)$ and are $140 (= 5 \times 28)$ simple left syllables the set of which we denote by $\text{GSLAlphabet}(VC)$. For more clarity, we briefly present them here as follows: $\text{GSLAlphabet}(CV) = \{\underline{ba}, \underline{ga}, \underline{da}, \dots, \underline{xu}, \underline{j\check{u}}, \underline{hu}\}$ and $\text{GSLAlphabet}(VC) = \{\underline{ab}, \underline{ag}, \underline{ad}, \dots, \underline{ux}, \underline{u\check{j}}, \underline{uh}\}$ (here an underlined syllable denotes a speech unit attached to it).

Thus, we state, that $\text{GSLAlphabet} = \text{GSLAlphabet}(V) \cup \text{GSLAlphabet}(CV) \cup \text{GSLAlphabet}(VC)$ and it is comprised only from the above presented $285 (= 28 \times 5 + 5 \times 28 + 5)$ speech units, from which 5 are vowel and 280 are syllable speech units.

The text-to-speech systems for Georgian, which is constructed based on the Georgian speech alphabet, is one of the best systems of such type for Georgian [1]. This fact together with the fact that only 5-10% of this methods is implemented in this system, definitely points out the high perspective value of this method. It is also supported by the fact that based on this method, there is already elaborated ways of improvement

of the speech recognition system for the Georgian language [3]. Herewith, this view on speech alphabet gives a new view on the structure of syllables. Namely, according to nowadays dominating view [14], a syllable with the form $CR_1 \dots CR_m VLC_1 \dots LC_m$ is comprised by nucleus, which is denoted by V vowel. Differently from this, in a syllable $CR_1 \dots CR_m VLC_1 \dots LC_m$ the nucleus is either RC_1V , which is simple right syllable, or VLC_1 , which is simple left syllable. For example, in Georgian, in syllable $k'u$, which, at the same time, is a word, vowel u is not a nucleus and, accordingly, consonant k is not an onset, but $k'u$ is one speech unit and its nucleus is itself it¹.

About Georgian morphemes, declination, conjugation, and declarative sentences: According to the logical grammar of the Georgian language, it is considered to be the result of extension of primary mathematical language. That is why, we always take into account the mathematical language of lower level. In this way, our view differs from the ones, which consider Georgian language without this mathematical language [9, 11, 12]. For example:

1. According to existing views, a morpheme is either derivational or inflectional. An inflectional morpheme operates on the word (root) and gives the word with the same lexical and different grammatical meaning. Derivational morpheme operates on the root (word) and as a result gives the word, which has different lexical meaning, whose part of speech differs or does not differ from the initial one. According to our researches this classification of morphemes is not complete. For instance, the word (root) *citel* (red) is A type word, i.e. belong to the adjectives, while the word *citel-s* (red in Ns case (i.e. in DAT case)) contracts noun phrase *citel Ns* (the/a red Ns), which, obviously, is not a grammatical form of adjective *citel* (red), and, also, it cannot belong to any class of part of speech. Thus, we prove that the right 1-place morpheme $(_{-1})s$ in the word $(citel)s$ is neither derivational nor inflectional. We call this type of morpheme a contracting morpheme. Moreover, we call *citels* as an intra-lingual contracting word, because it contracts an expression which is fully expressible by means of the language. Such words are in great deal in the Georgian language. Namely, most of the verbs are like this. For instance, a verb *citelia* contracts expression *aris citeli*. However, in Georgian, if we do not take into account the words denoting basic, i.e. none definable sets (as *citel* (red), *vašli* (apple), etc.), most of the remaining words are contracting words. For instance, according to us, $citeli_{(-1)}$ is right 1-place operation and it contracts $citeli \cap_{-1}$ form, which is built up with the set $\{citeli\}$ ² and with the set theoretic operation of intersection \cap . In this way, *citeli* as an adjective, is a contracting word and we call it extra-lingual contracting word, because it contracts the expression constructing with the help of low level mathematical language.

2. According to us, a noun has as many declinational forms as it can be presented in the sentence as a necessary member of it. A simple declarative sentence is an

¹This fact is additionally proven by the fact that in Georgian syllable $k'u$ takes less time to pronounce than to pronounce alone vowel u . Such syllables are not rarely encountered in the Georgian [1].

² $\{citeli\}$ denotes the basic, i.e. none definable set of all red entities that means that this set cannot be given neither by naming its elements, nor by using recursion, nor by pointing out a property which defines it.— here we have the similar situation as it is in Euclid's Geometry with the notion, i.e. word point.

expression, which is built up with the verb of declarative mood and with those forms of the nouns, which from the logical point of view are necessarily attached to this verb in the thinking level, and these forms of the nouns are called declinational forms, or cases. A necessary member of a declarative sentence is a noun which is (from the logical point of view) necessarily attached to the verb by which this sentence is built up. The placeness i.e. arity of a verb of a declarative mood is defined by the number of those nouns which have to be necessarily attached to this verb; and a place type of the verb is defined by the type of filler noun of the place. Based on these notions, we came up with 14 declinational forms, i.e. cases for nouns (N). We denote them as: Ni, Nis, Nisḡen, Nisgan, Nistvis, Nit, Nidan, Ns. Nḡi, Nze, Nad, Namde, Ntan, Nma cases. This system of declination is called the logical declination. Moreover, it was also established that in the core part of Georgian, i.e. in the “singular fragment”, there are m different V_phrases. They are as follows: $[V_1(N_1)=t], \dots, [V_m(N_1, \dots, N_m)=t]$ ($N_k \in \{Ni, Nis, Nisḡen, Nisgan, Nistvis, Nit, Nidan, Ns, Nḡi, Nze, Nad, Namde, Ntan, Nma, me \text{ (I and me)}, ḡen \text{ (you}_{sing})\}$, $k=1, \dots, m$). Hereby, we consider V_phrases as the set of under operational propositions. Namely, it was established that operators of substitutions are realized by declinational morphemes of nouns, which are also attached to pronouns *me* (I and me), *ḡen* (*you_{sing}*) without marking them morphologically. This guarantees the fact that pronouns and N_phrases operate on V_phrases and give in result V_phrases back, which means that under this operation V_phrases are operationally closed. According to us, in Georgian, adjectives (resp. quantifiers) i.e. A (resp. Q) words have 3 declinational forms, i.e. cases. They are as follows: Ai (resp. Qi), A- (resp. Q-), Ama (resp. Qma) cases. Non-declinational A (resp. Q) words are denoted as A* (resp. Q*). Herewith, it was established that by declinational forms of the A (resp. Q) words there are given their area of definitions, i.e. domains. For instance, the domain of Ai and Qi operations is Ni, Nis, Nisḡen, Nisgan, Nistvis, Nit, Nidan, and the domain of A- and Q- operations are Ns. Nḡi, Nze, Nad, Namde, Ntan, and the domain of Ama and Qma operations is Nma. It was also established that Ai, A-, Ama, A* (resp. Qi, Q-, Qma, Q*) operates on the domain elements or on AN α _phrases of N α _type, which are on the right side of these operators, gives back an AN α _phrase (resp. QN α _phrase) of N α _type, which shows that the set Ni, Nis, Nisḡen, Nisgan, Nistvis, Nit, Nidan, Ns. Nḡi, Nze, Nad, Namde, Ntan, Nma is closed with respect to these operations. Thus, one of our main results is that there is an already constructed mathematical theory, which is isomorphic to the core part of the Georgian language, and whose alphabet is comprised of:

1. N type words: Ni, Nis, Nisḡen, Nisgan, Nistvis, Nit, Nidan, Ns, Nḡi, Nze, Nad, Namde, Ntan, Nma;
2. A type words: A-, Ai, Ama, A*;
3. Q type words: Q-, Qi, Qma, Q*;
4. Personal pronouns: *me* (I and me), *ḡen* (*you_{sing}*),
5. V type words: $[V_1(N_1)=t], \dots, [V_m(N_1, \dots, N_m)=t]$ ($N_k \in \{Ni, Nis, Nisḡen, Nisgan, Nistvis, Nit, Nidan, Ns, Nḡi, Nze, Nad, Namde, Ntan, Nma, me \text{ (I and me)}, ḡen$

$(\text{you}_{sing})\}, k=1, \dots, m).$

We consider this mathematical theory, which is isomorphic to the core part of the Georgian language, as a foundations of the logical grammar of the Georgian language and we call it as natural logic of the core part of natural Georgian language. The importance of this result is highlighted by the fact that the approach that is based on this theory completely solves the problem of the automatic translation of the core part of the Georgian language into the mathematical language, which is the one of the main tasks in the areas of modern linguistics, the logic of natural languages and natural language processing. Namely, by this approach it is already built [2]:

1. An experimental version of the two way automatic translator system of the simple declarative sentences from Georgian into the mathematical language;
2. An experimental version of the Georgian problem solver (verificator) system, which processes a problem by preview translating it into the mathematical language;
3. An experimental version of the Georgian-English-German two way automatic translator system, which, with the use of mathematical language as mediator language, translates from one language into others;
4. An experimental version of the checker and analyzer of the complex and simple declarative sentences of the core part of the Georgian language, which is the first and only such type system, in spite of the fact that the task of building up a syntactic spell checker for Georgian is in the active phase for more than 40 years (an experimental version of the checker and analyzer only of simple declarative sentences of the core part of the Georgian language was made by us in 2008-2009).

R E F E R E N C E S

1. Pkhakadze K., Chichua G. The notion of Georgian spoken unit and the overview of methods the construction and aims of Georgian multi voice reader system. *The Works of Georgian Language Technology Center*, (2011), 21-40.
2. Pkhakadze K., Chikvinidze M. The overview of the methods and aims of the construction of the logical syntax and spell checker of the Georgian simple and complex sentences. *The Works of Georgian Language Technology Center*, (2011), 41-70.
3. Pkhakadze K., Chichua G., Vashalomidze A., Abzianidze L., Maskharashvili A., The Georgian Reader-Listener System with User's Possibility to Build in an Own Synthetic Voice, *Rep. Enlarged Sess. Semin. I. Vekua Appl. Math.*, **23** (2009), 100-104.
4. Pkhakadze K., Abzianidze L., Maskharashvili A. Georgian Language's Theses, *Seminar of I. Vekua Institute of Applied Mathematics REPORTS*, **34** (2008), 108-121.
5. Pkhakadze K., Chichua G., Abzianidze L., Maskharashvili A., About 1-Stage Voice Managed Georgian Intellectual Computer System. *I. Vekua Institute of Applied Mathematics*, **34**, (2008), 96-107.
6. Pkhakadze K. About Logical Declination and Lingual Relations in Georgian. (Georgian) *Journal "Georgian language and logic" N1, "Universali"*, (2005), 19-77.

7. Pkhakadze K. Pre-verbal Semantic Unit, Problem of Personal Signs, Integral and Non-Integral Verbal Semantics and Incomplete or First Semantic Classification of Georgian Verbs. (Georgian) *Main and Additional text-books in Contemporary Mathematical linguistics, N1, "Universali"*, (2004), 72-152.

8. Pkhakadze K., Chankvetadze G., Tibua L., Ivanishvili M., Lekiasvili L., Soselia E. About the Main Ideas of the Direct Formal-Logical Description of the Georgian Natural Language System Through the Examples. *Proceedings of the V Tbilisi Symposium on Language, Logic and Computation, ILLC*, (2003), 129-137.

9. Chikobava Arn. The Problem of Simple Sentences. *The academy of Sciences*, 1961.

10. Pkhakadze Sh. Some Problems of the Notation Theory. (Russian) *Tbilisi University Press, Tbilisi*, 1977.

11. Shanidze A. The Fundamentals of the Georgian Grammar. *TSU*, 1980.

12. Harris A.C. Georgian Syntax - A Study in Relational Grammar. *Cambridge University Press*, 1981.

13. Partee B., Meulen A., Wall R. Mathematical Methods in Linguistics. *Kluwer Academic Publishers*, 1990.

14. Taylor P. Text-to-Speech Synthesis. *Cambridge University Press*, 2009.

Received 30.05.2012; revised 15.09.2012; accepted 28.11.2012.

Authors' addresses:

K. Pkhakadze

I. Vekua Institute of Applied Mathematics

Iv. Javakhishvili Tbilisi State University

2, University St., Tbilisi 0186

Georgia

Open Institute of Georgian Language, Logic and Computer

Web-site: www.gllc.ge

Georgia

Scientific- Educational Center for Georgian Language Technology

Georgian Technical University

Web-site: <http://gtu.ge/gluc/index.htm>

Georgia

E-mail: gllc.ge@gmail.com

G. Chichua

Open Institute of Georgian Language, Logic and Computer

Web-site: www.gllc.ge

Georgia

Scientific- Educational Center for Georgian Language Technology

Georgian Technical University

Web-site: <http://gtu.ge/gluc/index.htm>

Georgia

E-mail: gllc.ge@gmail.com

M. Chikvinidze

Iv. Javakhishvili Tbilisi State University

2, University St., Tbilisi 0186

Georgia

Open Institute of Georgian Language, Logic and Computer

Web-site: www.gllc.ge

Georgia

Scientific- Educational Center for Georgian Language Technology
Georgian Technical University
Web-site: <http://gtu.ge/gluc/index.htm>
Georgia
E-mail: gllc.ge@gmail.com

A. Maskharashvili
Loria, Universite de Lorraine, INRIA Nancy Grand-Est
Nancy, France
Open Institute of Georgian Language, Logic and Computer
Web-site: www.gllc.ge
Georgia
E-mail: aleksandre.maskharashvili@inria.fr