

Functional-Semantic Analysis of Grammaticalized Verbal Forms in Georgian and Homonymy Caused by Them in Corpus Linguistics

Mariam Rukhadze

Ivane Javakhishvili Tbilisi State University, Georgia

Abstract

The language is a system of signs and represents a complex phenomenon itself. The full analysis of the language means, on the one hand, the study of the linguistic system, and on the other hand - the study of the mechanism of its functioning, the analysis of the interaction between the deep and surface structures of the language. At the same time, the language is a constantly changing system. Changes are made at each level of the language, changes take place in phonological composition of the language, in its morphological system and syntactic structures, changes apply to the field of semantics and pragmatics as well.

The primary function of language - its use in communication - is carried out through linguistic inventory. Linguistic inventory consists of **lexical** units and **functional** elements. The inventory in the language is divided into different paradigmatic classes. During the changes in natural languages a sign can transposed from one paradigmatic class to another. This type of change in the current language system is called **grammaticalization**.

Grammaticalization is a well-known phenomenon in typology. Hopper explains that: "A term referring to the change whereby lexical items and constructions come in certain linguistic contexts to serve grammatical functions" [Hopper, 2005:17]. During the process of **grammaticalization** lexical unit losses its historically developed semantics and after desemantization acquires some specific grammatical function, consequently, we obtain the functional element. Generally, formation of lexical units to grammatical/functional elements, is quite common in languages and it is characteristic to both names and verbs; their research methodology is different. In typological terms, it is particularly interesting and specific to form verb forms as the functional elements. See the examples: *There is a flower in the garden that is not seen from here* – vs - *It seems Philippe's business will go wrong*; *Come tomorrow, I will be waiting* – VS - *Come on, everyone dreams about the power. He got up in the morning very early* – vs- *He Got up and spoiled everything.* (the idea is: he did not hesitate and destroyed everything).

The process of **grammaticalization** involves a change of paradigmatic class for linguistic sign. Change of this type does not occur at once, it is gradual, which causes grammaticalization process's. As Hopper says: "a change is only incompletely achieved at any given stage of language" (Hopper; 2005)

The paper refers to the process of **grammaticalization**, functional and semantic analysis of the **grammaticalized** item in Georgian, and the ways to tackle the problems arisen due to the **grammaticalized** element of verbal origin in computer linguistics. The goal of the paper is to demonstrate **grammaticalized** verbal form *čan-s (it_is_visible-PRS.S.3.SG) – It is visible vs. It seems*; semantic and functional aspects of this item. The research is based on the corpus method, both the corpus-based and the corpus-oriented methods and is carried out on the basis of the Georgian National corpus (GNC) (www.gnc.gov.ge). Empirical data is collected from old, middle and modern Georgian language corpuses, as well as the law and political texts corpuses.

The aim of the paper is to search **functional and semantic** characteristics of the **grammaticalized** item *čan-s (it_is_visible-PRS.S.3.SG) – It is visible vs. It seems* in Georgian language; For this purpose there are used methods of lexical substitution, and elimination, to describe the functions of the item we used a language competency test. The conclusion of the test is illustrated by the charts.

Functional elements create special difficulties in processing language by computer methods, particularly in computer linguistics. Today no one argues about the establishment of digital bases of language and the importance of its processing with digital methods: in the 21st century scientist's efforts are aimed at creating and improving artificial intelligence. Creation of automatic recognition - analyzer is one of the main tasks today. This requires a thorough analysis of the linguistic system on phonological, morphological, syntactic, semantic and pragmatic levels. Since the Georgian language material is not functionally properly processed, Georgian National Corpus (GNC) does not have an algorithm to subtract the functional elements from the elements belonging to other paradigmatic class, which makes the homonymy and makes the problem of disambiguate unsettled.

The goal of the paper is to describe the problem of the homonymy caused by **grammaticalized** item a-dg-a (PR-stand-AOR.S.3.SG) in corpus linguistics and find the ways of removing the ambiguities. The paper suggests the rules created by analyses of collected data, rules are illustrated by the charts. We hope that suggested rules will be the base of improvement of the working mechanism of Georgian National Corpus morphological algorithm.

Keywords: corpus linguistics; functional element; grammaticalization; homonymy; semantic analysis