

Computer model of the Tibetan language morphology

Morphological changes in the Tibetan language are inherent in a number of enclitic and function morphemes: case markers showing the genitive, ergative and terminative meanings, markers of converbial and nominal verb forms, the quotation marker, the indefinite particle, the finite particles of narrative, interrogative and imperative sentences, the derivational affixes of the adjective and noun (in total 14 markers). The selection of an allomorph depends on the final of the preceding syllable. The number of allomorphs (degree of variability) is from 2 to 11. A rarer phenomenon in the Tibetan morphology is the presence of positional allomorphs of verbal and nominal roots. If the root ends in a vowel, then when the derivational affix of a noun or adjective is added, it can append the final *n*. The verb in the Tibetan language has categories of modality (grammemes of imperative and indicative mood) and time (grammemes of the past, present and future time, only for the indicative mood). These grammatical meanings are expressed using consonant prefixation and suffixation, as well as vowel and consonant gradation in verbal roots. However, not every verb has 4 distinct forms, because some of them may coincide. The limiting case is the verbal root having a single allomorph.

To explain the described phenomena of positional morpheme interchange in the Tibetan language, a computer model of the Tibetan morphology was created by solving the following tasks: 1. Development of the faceted classification of observed interchanges according to the type of variation (progressive and regressive, complete or partial assimilation or dissimilation, etc.), types of initials and finals of morphemes and other possible reasons. 2. Development of an object-oriented model reflecting the created classification, and allowing to automate the work of observed rules of gradation. 3. Development of the system of automatic regression testing of the model, which makes it possible to guarantee its compliance with linguistic material. 4. Identifying the dependent grounds for classification, reduction of their number to a minimum set of independent grounds, reworking the model in accordance with the obtained set of grounds.

The created computer model of the Tibetan morphology was evaluated using a regression testing system, which ensures that the model conforms to the observed morphological phenomena. The automatic test system ensured compliance of the model with the existing descriptions. In addition, the accuracy and completeness of the created model was evaluated on the material of the electronic corpus of Tibetan grammatical texts, developed at St. Petersburg State University.