

**Костіков Микола Павлович**

**Костиков Николай Павлович**

**Mykola Kostikov**

ДЕКОМПОЗИЦІЯ МОРФОЛОГІЧНИХ ПРАВИЛ ПОЛЬСЬКОЇ МОВИ  
ДЛЯ АДАПТИВНОГО НАВЧАННЯ

**Ключові слова:** адаптивне навчання, граматики, електронні засоби навчання, інженерія знань, інтелектуальні навчальні системи, модель мови.

ДЕКОМПОЗИЦИЯ МОРФОЛОГИЧЕСКИХ ПРАВИЛ ПОЛЬСКОГО ЯЗЫКА  
ДЛЯ АДАПТИВНОГО ОБУЧЕНИЯ

**Ключевые слова:** адаптивное обучение, грамматика, инженерия знаний, интеллектуальные обучающие системы, модель языка, электронные средства обучения.

DECOMPOSING MORPHOLOGICAL RULES OF POLISH  
FOR ADAPTIVE LANGUAGE LEARNING

**Keywords:** adaptive learning, CALL, e-learning, grammar, intelligent tutoring systems, knowledge engineering, language model.

## **1. Introduction**

The modern informatization of the education calls for developing different approaches to the teaching process. To benefit from applying information technology, we should change the structure of learning material. Methods of artificial intelligence help make the learning process adaptive and personalized, so as to define the individual problems and demands of each user and react to them accordingly.

## **2. Need for the language model**

By now there is still a lack of language learning software for Slavic languages, and the existing ones have substantial drawbacks. The analysis of modern programs for Polish has shown that they often lack the feedback which is an indispensable condition of adaptive and personalized learning. Teaching Slavic languages in other Slavic-speaking countries also has its specifics, e.g. grammar becomes of much greater importance. Thus we should consider the peculiarities of the given grammatical system and develop a learning system based on a language model.

## **3. Decomposition of learning material**

Decomposition of learning material may be viewed in two ways. First of all, we can divide it into topics, units, lessons and so on, design separate learning objects so as to meet the SCORM technical standards or other formal requirements. Such decomposition may be called structural.

It is also possible to decompose the knowledge itself by developing a hierarchy of learning elements (ontology of terms and notions, semantic network etc.). At first we define the basic elements of knowledge and relationships between them. Thus previously acquired knowledge becomes the basis for solving more complex problems. Intelligent tutoring systems implement this step-by-step approach, being able both to solve problems and explain this process to users thanks to the domain model which lies underneath and is stored as a knowledge base.

## **4. Decomposition of Polish morphological rules**

The traditional approach to the morphology formalization implies using relational models. One of the co-developers of the Polish grammatical dictionary SGJP, M. Woliński notes that 'a relatively compact relational model can be used to describe Polish inflection in a uniform way' [1]. These models are useful for quick processing big amounts of data, e.g. in morphological analysis, but difficult to use in the language learning process due to its specifics. Relational models are based on inflectional patterns within which the inflectional rules are the same, thus the number of patterns may reach several hundred to take into account all the

subtleties, such as phonological alternations. Yet the conditions for choosing a specific pattern aren't described in these models.

As one of the closest approaches to the 'intelligent' grammar formalization we consider the 'smart paradigm', a notion introduced by the co-developers of MOLTO — multi-language online translation tool. Smart paradigms are 'heuristic functions that take just one or a few forms of a word as their arguments and infer the complete inflection table of the word from them' [2]. Such meta-paradigm 'inspects the base form and tries to infer which low-level paradigm applies' [3]. Unfortunately, the base form itself isn't enough to determine the inflectional pattern in many cases, so the smart paradigms use additional word forms provided by the users — linguists and speakers of the language.

Instead of this, we propose using additional parameters which are more likely to be known or found by the students, e.g. gender and animacy for nouns, which are presented in dictionaries or can be inferred from the context. They not only facilitate the morphological synthesis, but also make our problem solver algorithm more similar to human logic and thus easier to explain to the students. Furthermore, our main idea is to move away from processing a paradigm as a whole. We should analyze specific word forms and separate out the basic elements of morphological rules (adding certain endings etc.) and transformations (e.g.  $\acute{n} \rightarrow ni$ , interchange  $t \rightarrow \acute{c}$ ), so-called elementary rules. We should also investigate which of them are common and different between various patterns and define the factors which impact on their choice. Then we may analyze where and how often each of them occurs in a certain vocabulary, select certain words for lessons, build the hierarchy of learning elements and logical learning sequences.

By now we have decomposed elementary rules for Polish nouns and tested the model on the 350 most frequent (according to [4]) nouns. The experiment has proved that such decomposition and using additional parameters help to resolve ambiguities in inflection in most cases.

## **5. Adaptive grammar learning**

The presented model is being created as a part of the intelligent tutoring system for learning language in context. In theory, such model allows of using any word for generating learning examples and exercises, as it determines the inflectional rules based on the word itself and its general characteristics. The main drawback is the need for providing the additional

parameters by the knowledge engineer and/or the user. Besides, grammatical exceptions should still be taken into account.

The feedback based on the model should ensure the personalization of the learning process in two aspects. First of all, we can analyze the nature and reasons of the students' mistakes more precisely, as within the inflectional patterns model it's hard to detect and react to users' problems in applying specific components of a complex inflection process, e.g. interchange of consonants. Secondly, we can analyze the interaction between grammar and vocabulary and create individual learning plans based on the occurrence of certain rules in a given vocabulary (depending on user needs) and vice versa, since learning in context helps both to build grammar skills and learn words.

## 6. Conclusion

The described approach helps to model the inflection of Polish nouns. Applying it to other parts of speech and developing an intelligent tutoring system based on this model are the subjects to further research.

## References

1. Woliński M. *A Relational Model of Polish Inflection in Grammatical Dictionary of Polish* / Marcin Woliński // *Human Language Technology: Challenges of the Information Society*. – Berlin, Heidelberg : Springer-Verlag, 2009. – P. 96–106.
2. Ranta A. *How Predictable is Finnish Morphology? An Experiment on Lexicon Construction* / Aarne Ranta // *Resourceful Language Technology: Festschrift in Honor of Anna Sågvall Hein* / Edited by Joakim Nivre, Mats Dahllöf and Beata Megyesi. – Uppsala : University of Uppsala, 2008. – P. 130–148.
3. Détrez G. *Smart Paradigms and the Predictability and Complexity of Inflectional Morphology* : Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics, Avignon, France, April 23–27, 2012 / Grégoire Détrez and Aarne Ranta. – Avignon, 2012. – P. 645–653.
4. Kazojć J. *Słownik frekwencyjny leksemów V.06.2009*, 2009.  
<http://www.slowniki.org.pl/slownik-frekwencyjny-leksemow.pdf>