

Рафаева А.В.

НИВЦ МГУ

Система РУСЛО (РУсское СЛОвообразование) может быть использована для словообразовательного анализа и синтеза производных слов русского языка, в том числе неологизмов, построенных по редким словообразовательным моделям (см. [Перцова 1998]).

Лингвистическое обеспечение системы РУСЛО разработано Н.Н. Перцовой. Первый вариант программного обеспечения системы, написанный А.В. Черемхиным, работал в среде DOS. В настоящих тезисах описывается второй вариант системы, работающий в среде Windows 98 и использующий некоторые функции и алгоритмы среды STARLING, разработанной С.А. Старостиным. В этом варианте грамматика отделена от алгоритмической реализации, что позволяет менять лингвистическую информацию, не затрагивая программной части системы, а также задавать различные режимы работы системы, в частности:

- словообразовательный анализ слов, входящих в состав словаря А.А. Зализняка;
- анализ слов, не входящих в состав словаря Зализняка, но построенных по регулярным словообразовательным моделям;
- анализ слов, построенных по редким словообразовательным моделям (может использоваться, в частности, для анализа поэтических неологизмов В. Хлебникова).

Система РУСЛО состоит из трех главных блоков, а именно:

- блок данных, который содержит базы данных в формате DBF, описывающие различные словообразовательные модели, а также компьютерную версию словаря Зализняка;
- блок алгоритмов, содержащий функции словообразовательного анализа и функции обработки запросов к словарю;
- интерфейс пользователя, позволяющий настраивать систему, задавать режимы ее работы, а также формы ввода анализируемых слов (ручной ввод и ввод из текстового файла) и вывода результатов анализа.

В процессе работы с системой (при ручном вводе анализируемых слов) можно просматривать БД, содержащие лингвистическую информацию и запрашивать правила, примененные на каждом шаге вывода. В дальнейшем предполагается добавить возможность внесения изменений в указанные БД непосредственно в процессе работы с системой.

Система РУСЛО написана на языке программирования C++. В программном обеспечении системы используются алгоритмы С.А. Старостина и А.В. Черемхина.

Литература

[Перцова 1998] – Перцова Н.Н. Анализ окказиональных слов в системе РУСЛО // Труды Международного семинара Диалог'98 по компьютерной лингвистике и ее приложениям. Т. 2. Казань, 1998. С.846 – 847.

[Перцова, Черемхин 1992] – Перцова Н.Н., Черемхин А.В. Эксперименты по построению формальной модели русского словообразования // Труды машинного фонда русского языка. Т.2. М.: ИРЯ РАН, 1992. С. 86 – 103.