

## ПРОДУКТИВНЫЕ СИНТАКСИЧЕСКИЕ МОДЕЛИ

Кедрова Г.Е., Потемкин С.Б.

Московский государственный университет им. М.В.Ломоносова, Филологический  
ф-т, лаб. Обеспеченного компьютером обучения,  
Россия, 119991, Москва, ГСП-1, Ленинские горы, д. 1, стр. 51,  
Тел.: (495)939-14-78  
E-mail: [prolexprim@gmail.com](mailto:prolexprim@gmail.com)

При оценке эффективности учебного материала, предназначенного для овладения языком иностранными студентами в рамках учебного курса «Русский как иностранный» важно понять, какие синтаксические структуры, предложенные для изучения, наиболее употребительны в ежедневном общении [1].

В качестве исходного материала используется русско-китайский разговорник [2], который представляет собой ряд слов и предложений на двух языках (в нашем случае, русский и китайский). Проведена токенизация и лемматизация русского текста с определением части речи (POS) и грамматических признаков (род, число, падеж) каждой словоформы. Затем выполнялся поверхностный синтаксический разбор каждого предложения на основании алогитма Нивра [3]. Полученные синтаксические связи сортируются по частоте встречаемости. Ожидаемо, наиболее частотной структурой оказывается форма субъект – глагол – объект (SVO). Именные группы представлены в основном формами сущ.им.падеж – сущ.род.падеж; притяжательное местоимение – сущ.; прилагательное – сущ., согласованные по роду, числу, падежу.

Каждая синтагма подставляется в запрос к НКРЯ [4]. В ответ получен список всех предложений, содержащих пару слов этой синтагмы на заданном расстоянии (от 1 до 3) и общее число таких предложений. Следующей задачей является определение продуктивности структуры предложения. В наиболее употребительных синтагмах будем заменять включенные в нее слова другими словами, принадлежащими к той же части речи, из словника разговорника. Полученные в результате синтагмы также проверяются на употребительность по НКРЯ. Сумма частот по каждой такой синтагме определяет ее продуктивность и, следовательно, целесообразность ее включения в учебный материал.

### Литература

1. Кедрова Г.Е., Потемкин С.Б. Лексикографическая работа с китайскими магистрантами-филологами в МГУ в сборнике: V Международная научно-практическая конференция: МАКС Пресс Москва, с. 231-233
2. Китайский за 1 час. Аудиокурс китайского языка (брошюра + CD) Издательство: Дельта Паблишинг. ISBN 5-94619-130-6; 2005 г.
3. Nivre, Joakim. 2003. An efficient algorithm for projective dependency parsing. In Proceedings of the 8th International Workshop on Parsing Technologies (IWPT), pages 149–160, Nancy.
4. Электронный ресурс: <http://www.ruscorpora.ru/> - дата посещения 23.10.2017.