

РАЗРАБОТКА ЭЛЕКТРОННОГО ПЕРЕВОДЧИКА НА ПЛАТФОРМЕ МАШИННОГО ПЕРЕВОДА APERTIUM

Васильева Ю.В.

Харьковский гуманитарный университет

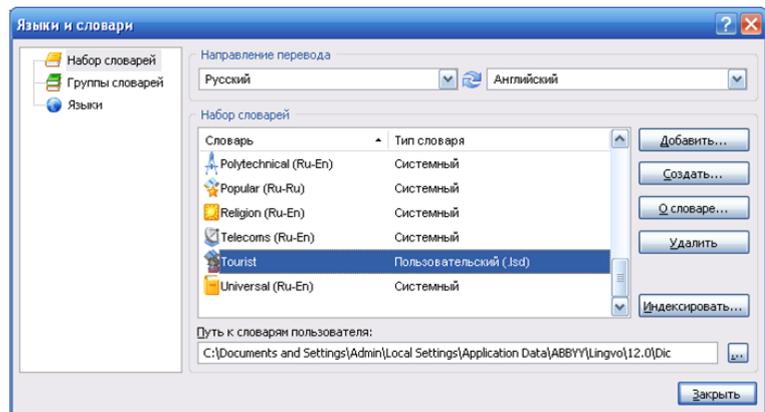
«Народная украинская академия»,

г. Харьков, ул. Лермонтовская, 27, тел. 704-10-37,

e-mail: yuljav90@mail.ru

В условиях современной информационной среды все популярней становится участие пользователей в работе над различными интернет-проектами. Этот процесс получил свое развитие благодаря переходу на новую концепцию Сети, получившую название Web 2.0. Суть этой концепции состоит в возможности и эффективности работы над проектами путём совместной деятельности пользователей. Ярким примером таких проектов являются системы перевода и системы словарей, которые открыты для разработки и пополнения.

Разработанный нами в предыдущие годы русско-английский электронный словарь «Tourist», интегрированный в оболочку программы Lingvo в ХГУ «НУА», может быть с успехом применен в словарных системах, доступных широкому кругу



пользователей, с целью использования для перевода как отдельных слов и фраз, так и для применения в системе машинного перевода. Для дальнейшего использования разработанного нами словаря мы выбрали платформу машинного перевода Apertium (<http://www.apertium.org>), в которой есть возможность размещения словарей в процессе создания системы машинного перевода для различных языковых пар. На данной платформе создано уже около 40 автоматических переводчиков и ещё многие находятся в процессе создания. Команда Apertium проявляет большой интерес к региональным языкам и активно поддерживает работы по созданию новых систем автоматического перевода, для которых разрабатывается лингвистическая база из словарей и правил в чётко указанных форматах XML. С этой целью используется простой XML-стандарт на основе форматов для кодирования необходимых языковых данных [1].

Система Apertium представляет собой платформу машинного перевода поверхностно-трансферного типа. Данная платформа имеет дело со словарями и правилами поверхностного трансфера. Правила представляют собой операции с группами лексических единиц, входящих в три словаря: два морфологических



словаря для языков рассматриваемой пары, содержащих информацию о словоизменении (склонении или спряжении) на каждом из языков (apertium-rus-en.rus.dix и apertium-rus-en.en.dix), и двуязычный словарь, который содержит переводные соответствия слов и символов двух языков (apertium-rus-en.rus - en.dix) [2].

```
<section id="main" type="standard">
<e><p><l>автобус<s n="n"/></l><r>bus<s n="n"/></r></p></e>
<e><p><l>троллейбус<s n="n"/></l><r>trolleybus<s n="n"/></r></p></e>
<e><p><l>стол<s n="n"/></l><r>table<s n="n"/></r></p></e>
<e><p><l>шкаф<s n="n"/></l><r>cupboard<s n="n"/></r></p></e>
<e><p><l>аромат<s n="n"/></l><r>aroma<s n="n"/></r></p></e>
<e><p><l>телефон<s n="n"/></l><r>phone<s n="n"/></r></p></e>
<e><p><l>балкон<s n="n"/></l><r>balcony<s n="n"/></r></p></e>
```

Кроме того, составляющими языковую пару являются также два файла с правилами трансфера. Правила трансфера русского языка на английский язык описывают, каким изменениям подвергнутся предложения русского языка при переводе на английский язык (apertium-rus-en.rus -en.t1x). Правила трансфера английского языка на русский язык описывают преобразования, которые должны быть осуществлены при переводе с английского языка на русский язык (apertium-rus-en.en- rus.t1x) [2]. Следует отметить, что часто такие языковые ресурсы, как корпуса, словари, грамматики, морфологические анализаторы, списки лемм, находятся в свободном доступе либо с лицензией на возможность свободного использования, и могут быть использованы повторно, что значительно сокращает время разработки нового переводчика.

Таким образом, учитывая имеющийся опыт работы в данном направлении, можно утверждать, что идея создания русско-английского переводчика на базе платформы машинного перевода Apertium является достаточно актуальной и перспективной сферой деятельности в области развития современных языковых компьютерных технологий.

Список литературы

1. Apertium [Электронный ресурс]: – Режим доступа: <http://ru.wikipedia.org/wiki/Apertium>.
2. Виртуальная лаборатория Apertium [Электронный ресурс]: – Режим доступа: <http://ru.vlab.wikia.com/wiki/Apertium>.