

ПРЕДИСЛОВИЕ

В настоящем номере АРГУВ печатаются материалы семинара «Армянская корпусная лингвистика», организованного в Горисе (Армения) с 19 по 21 августа 2007 г. проектом «Востоочноармянский национальный корпус» (Москва), Национальным институтом восточных языков и цивилизаций (Institut National des Langues et Civilisations Orientales / INALCO) (Париж) и Ереванским государственным лингвистическим университетом им. В. Брюсова (Ереван).

В настоящее время корпусная лингвистика является одним из наиболее интенсивно развивающихся направлений современной науки о языке. Как правило, под корпусом понимается собрание текстов на данном языке, представленное в электронной форме и снабженное специальным научным и поисковым аппаратами, а под национальным корпусом – самый большой и представительный корпус, характеризующий язык данного национального коллектива в целом, во всей его полноте и на всех этапах его существования.

Необходимость создания крупных (в том числе национальных) корпусов для языков мира уже давно не оспаривается. Сферы, в которых корпус может быть использован и быть полезен, самые разнообразные – литературоведение, журналистика, история и т.д., но, в первую очередь, это, конечно, лингвистика, как теоретическая, так и прикладная (статистика речи, автоматическая обработка текстов и т.п.). Корпус является мощным инструментом для проведения самых разнообразных исследований лексики и грамматики языка, его истории, социальной и диалектной вариативности и т.п., он дает возможность сравнительно легко осуществлять такие операции с большими массивами текстов, которые в «докорпусную» эпоху были или вовсе невозможны, или требовали очень существенных затрат времени и сил.

На сегодняшний день имеются десятки национальных корпусов разных языков: например, Британский национальный корпус (www.natcorp.ox.ac.uk), Чешский национальный корпус (www.ucnk.ff.cuni.cz), Национальный корпус русского языка (www.ruscorgo.ru), и т.д.; каждый из этих корпусов имеет свою инфраструктуру и свои поисковые возможности.

На нынешнем этапе развития научная арменистика остро нуждается в новых современных методах и инструментах изучения лексической и грамматической структур языка, которые позволили бы исследователям с достаточной степенью надежности верифицировать те или иные гипотезы и модели, а также достичь принципиально новых теоретических и описательных результатов. Корпусно-ориентированные исследования могут значительно способствовать разработке слабоизученных и постановке новых исследовательских проблем.

Востоочноармянский национальный корпус (ВАНК), открытый для свободного доступа в июле 2007 г., был призван восполнить дефицит армянских корпусных лингвистических ресурсов и предоставить в распоряжение лингвистов справочно-информационную систему по литературному варианту востоочноармянского языка.

Поскольку корпусная лингвистика является относительно молодым лингвистическим направлением, то существует ряд проблем, связанные с

отсутствием четкого представления у арменистов того, как устроен корпус и как надо им пользоваться, с определением целей, для которых корпус является незаменимым помощником, и, наконец, с необходимостью осознания значимости привлечения корпусного материала для проведения ряда лингвистических исследований.

Проведенный семинар был первой попыткой, с одной стороны, собрать людей, которые к тому времени уже провели корпусные исследования, и могли поделиться своим опытом, с другой стороны, он был призван предоставить специалистам возможность обсудить перспективы развития армянской корпусной лингвистики. К сожалению, по ряду причин не все участники семинара смогли представить свои статьи для публикации, поэтому в конце предисловия приведен список всех участников и названия представленных ими докладов. Надеемся, что эти материалы, равно как, разумеется, и включенные в номер статьи, помогут читателю составить общее представление об обсуждавшейся проблематике, и будут способствовать расширению круга пользователей Восточноармянского национального корпуса.

Владимир Плунгян, Виктория Хуришудян, Арег Баяндур