ЦЕНТР ПОДДЕРЖКИ АРМЕНИСТИЧЕСКИХ ИССЛЕДОВАНИЙ
Москва

АРМЯНСКИЙ гуманитарный вестник

2/3-II 2009

Ереван
"Зангак-97"
2009

УДК 001:030
ББК 72я2
А 839

А 839 Армянский гуманитарный вестник.—
Ер.: "Зангак-97", 2009.— 120 стр.

Հայագիտական ուսումնասիրություններ
ՊԱՐԲԵՐԱԳԻՐ

BULLETIN
of Armenian Studies
(periodical journal)

ISBN 978—99941—1—646—1

СОДЕРЖАНИЕ

Гуманитарный проект
Аспрам Нждеян, Ашота Балаяна, Арега Баяндура
Москва-Ереван-Степанакерт

В работе над номером принимали участие
Владимир Плунгян
Виктория Хуршудян
Михаил Даниэль

Оформление обложки
Мкртыч Матевосян

Дизайн-макет
Григорий Арутюнян

С редакцией можно связаться по электронной почте:
aregbayandur@mail.ru

С электронной версией журнала можно ознакомиться на сайте:
www.eanc.net, "Восточноармянский национальный корпус"

ПРЕДИСЛОВИЕ

В настоящем номере АРГУВ печатаются материалы семинара «Армянская корпусная лингвистика», организованного в Горисе (Армения) с 19 по 21 августа 2007 г. проектом «Восточноармянский 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.ruscorpora.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.

Восточноармянский 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 восточноармянского 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 Хуршудян, Арег Баяндур

Участники семинара

Анаид Донабедян (INALCO, Париж), *Lemmatization of the Western Armenian with: the state of affairs [Лемматизация западноармянского с помощью программы Nooj]*

Джос Вайтенберг (Лейденский университет, Лейден), *The LALT (Leiden Armenian Linguistic Textbase) project and Corpus Linguistics [Проект LALT и корпусная лингвистика]*

Давид Гюрджинян (ЕГЛУ, Ереван), *Թվի կարգի ներկայացումը հայերենի կորպուսում [Представление категории числа в армянском корпусе]*

Нарине Экекян (ЕГЛУ, Ереван), *Տառային հապավումների ներկայացումը հայկական կորպուսում [Представление буквенных аббревиатур в армянском корпусе]*

Лиана Овсепян (ЕГУ, Ереван), Сусанна Тиоян (ЕГЛУ, Ереван), Роберт Урутян (ИЯ ААН), *Հայերենի անվանական և բայական բառափոփոխություն [Именное и глагольное словоизменения в армянском языке]*

Михаил Даниэль (МГУ, ВАНК, СТ, Москва), Дмитрий Левонян (ВАНК, СТ, Москва), Владимир Плунгян (ИЯ РАН, МГУ, ВАНК), Алексей Поляков (ВАНК, СТ, Москва), Сергей Рубаков (ВАНК, СТ, Москва), Виктория Хуршудян (ВАНК, СТ, Москва), *Introduction to Eastern Armenian National Corpus [Восточноармянский национальный корпус]*

Вартушка Самуелян, Анаид Донабедян (INALCO, Париж), *Online database of Armenian dialects [Онлайн база данных армянских диалектов]*

Гаяне Геворгян (ИЯ ААН), *Բարբառային նյութի մշակման և համակարգչայնացման սկզբունքներն ու առանձնահատկությունները [Принципы и особенности обработки и компьютеризации диалектного материала]*

Лилиан Восганян (ICAR - CNRS Université Lyon2, Лион), *Dialectes arméniens en contact en situation de diaspora: le recours au corpus [Армянские диалекты в контакте в диаспоре: на корпусном материале]*

Виктория Хуршудян (ВАНК, СТ, Москва), Вера Подлесская (РГГУ, Москва), Андрей Кибрик (ИЯ РАН, Москва), *Segmentation and Other Oral Discourse Transcription Issues in Large Corpora [Сегментация и иные проблемы транскрибирования устного дискурса в больших корпусах]*

Шеф Барбиерс (Институт Миртенса, Амстердам & Университет Утрехта), *European Dialect Syntax: Methods, Storage, On-line retrieval and Results. [Синтаксис европейских диалектов: методы, хранение, онлайн поиск и результаты]*

Нина Сумбатова (РГГУ, Москва), *Проблемы исследования языков даргинской группы*

Лилит Варданян (Университет Пизы, Пиза), *Expedition to Armenian forests: any unique tree species growing there? [Экспедиция в армянские леса: растут ли там уникальные виды деревьев?]*

Эмилия Нерсисянц (Университет Тегерана, Тегеран), *Language as Action: Analyzing Performatives in a Corpus of Commentaries/ Monolinguals' and Bilinguals Narratives: Further Analysis of a Corpus [Язык как действие: анализ перформативов на материале корпуса комментариев / Нарративы монолингвов и билингвов: дальнейший анализ корпуса]*

М. Даниэль, Д. Левонян, В. Плунгян,
А. Поляков, С. Рубаков, В. Хуршудян

ВОСТОЧНОАРМЯНСКИЙ НАЦИОНАЛЬНЫЙ КОРПУС
www.eanc.net

M. Daniel, D. Levonian, V. Plungian, A. Polyakov, S. Rubakov, V. Khurshudian
EASTERN ARMENIAN NATIONAL CORPUS

Eastern Armenian National Corpus is a comprehensive linguistic database of texts in Standard Eastern Armenian. EANC is an annotated corpus of about 110 mln tokens provided with a powerful search engine for making complex lexical morphological queries. The corpus is primarily intended for linguists and Armenian language teachers but can also be used in typological, anthropological and historical studies as well as by anyone interested in Eastern Armenian and its history.

Восточноармянский национальный корпус (ВАНК) - это открытый интернет-ресурс на базе представительной совокупности текстов на современном восточноармянском языке, предлагающий гибкую функциональность для поиска примеров различных языковых форм. ВАНК, включающий около 110 млн. словоупотреблений, предназначен, в первую очередь, для использования лингвистами-арменистами, во вторую - преподавателями армянского языка, но может представлять ценность и для лингвистов-типологов, историков, культурологов, публицистов и всех тех, кто работает с армянским языковым материалом или интересуется армянским языком и его историей.

Что такое корпус и чем он отличается от электронной библиотеки

Первый вопрос, на который следует ответить - что такое электронный корпус языка и чем он отличается от электронных библиотек (таких, например, как доступные на популярных сайтах www.armenianhouse.org или www.hayeren.hayastan.com). На первый взгляд, и корпус, и электронная библиотека представляют собой просто электронную коллекцию текстов. Различие между этими типами ресурсов целиком определяется их предназначением. Корпус предназначен для поиска языкового материала, а библиотека - для чтения текстов. Все остальные различия, определяющие функциональность, интерфейс и другие аспекты реализации этих двух типов электронных ресурсов, являются логическими следствиями различия в их предназначении.

Библиотека должна предоставить пользователю электронные версии текстов в полном объеме. Главная задача разработчиков библиотеки - удобство отображения текстов на экране и навигации, т.е. перемещения по фрагментам текстов и между текстами (система оглавлений, перелистывания и т.п.). Основная проблема,

стоящая перед разработчиками электронных библиотек - это авторские права на размещаемые тексты. Некоторые русскоязычные ресурсы библиотечного профиля прекратили свое существование именно вследствие того, что размещение в интернете текстов в полном виде нарушало права интеллектуальной собственности.

Перед разработчиками корпуса проблема авторского права не стоит, так как они предоставляют своим пользователям лишь краткие фрагменты текстов - обычно одно предложение, иногда, как в случае ВАНК, с возможностью просмотра непосредственного контекста (несколько предложений вокруг найденного). Непосредственный контекст бывает необходим для изучения некоторых языковых явлений: например, при исследовании функций анафорических местоимений, кореферентного опущения и т.п. С другой стороны, корпус должен удовлетворять определенным требованиям, которые ставят достаточно сложные задачи (а) методики сбора материала, (б) его лингвистической обработки и (в) эффективной реализации поисковых алгоритмов.

Методика сбора материала. Корпус призван представить язык во всей его полноте, поэтому он должен покрыть как можно большее разнообразие типов текстов и текстовых жанров. Именно поэтому разработчики Британского национального корпуса, одного из первых электронных корпусов, приняли в свое время решение о том, что крупные литературные произведения будут представлены в корпусе лишь фрагментами. В противном случае для английского языка с его огромным литературным наследством либо классические литературные тексты вытеснили бы на периферию другие жанры, либо объем корпуса вышел бы за рамки алгоритмически допустимого (на то время). Ясно, что такого рода решения для электронных библиотек недопустимы. Кроме того, электронная библиотека не обязана представлять тексты второстепенных и третьестепенных авторов, а также тексты документальных или иных "скучных" жанров (от официальных документов до медицинских предписаний), а для корпуса включение таких текстов, по крайней мере в некотором объеме, желательно, так как они представляют собой одну из форм существования письменного языка. Специализированные корпуса (например, корпус газетных текстов) по определению ограничены одним жанром или типом текстов, но в рамках этого жанра они все равно стремятся к максимально полному охвату языкового материала (от аналитических изданий до желтых листков); электронная же библиотека может представить полное собрание сочинений одного крупного автора, но игнорировать другого, менее популярного, по выбору составителей библиотеки и в соответствии с интересами ее аудитории. Иными словами, состав библиотеки определяется тем, что интересно читателю, а состав корпуса - стремлением к возможно более полному покрытию языкового разнообразия.

Существует несколько принятых, но не вполне формализованных характеристик, описывающих электронную коллекцию. текстов, на которую опирается корпус. *Репрезентативным* называется корпус, представляющий языковые выражения (словоформы, лексемы, грамматические категории, словосочетания) во всем разнообразии их узуса - например, в разных жанрах

письменных текстов. *Сбалансированным* называется корпус, в котором соотношение между объемом текстов различных жанров не является произвольным, а опирается на какие-то содержательные критерии. *Полным* называется корпус, который представляет большую или значительную часть существующих письменных текстов данного языка. ВАНК является первым и пока единственным репрезентативным, сбалансированным и полным электронным корпусом современного восточноармянского языка. Такие корпуса иногда называются "национальными" - в том смысле, что они максимально полно представляют языковое наследие данной письменной традиции (Британский национальный корпус, Чешский национальный корпус, Национальный корпус русского языка и др.).

Строго говоря, ВАНК является первым и единственным электронным корпусом современного восточноармянского языка вообще. В сети размещено лишь несколько интернет-корпусов древнеармянского языка (это, например, открытые сайты проектов www.sd-editions.com/LALT/home.html и http://titus.uni-frankfurt.de/indexe.htm).

Лингвистическая обработка. Будучи в первую очередь инструментом лингвистического исследования, корпуса требуют достаточно глубокого лингвистического анализа языка. Простейший поиск - поиск слов в том виде, в котором они встречаются в тексте или поиск фрагментов словоформ - обеспечивается любым текстовым редактором и не требует специализированного корпусного обеспечения. Но функциональность корпуса шире, он ищет не только определенные словоформы, но и лексемы в любой форме (*գնալ* ***gnal*** 'идти'), и одинаковые грамматические формы разных лексем (например, императивы множественного числа от непереходных глаголов) - см. в приложении описание функциональности ВАНК. Для этого необходимо программное обеспечение, которое умеет осуществлять грамматический анализ словоформ - так называемые лемматизаторы, которые словоформе W сопоставляют исходную форму лексемы L (это позволяет находить словоформу W по запросу на поиск лексемы L) и ее грамматические и лексические характеристики G1, G2 ... Gn, L1, L2 ... Ln (это позволяет находить словоформу W по запросу на поиск грамматической или лексической категории или их сочетания). Создание программы-лемматизатора является сложной задачей не только для плохо описанных языков, но и для языков с богатой грамматической традицией. Подробнее об этом смотри ниже в разделе о грамматическом словнике ВАНК.

Эффективные поисковые алгоритмы. Корпус должен обрабатывать запрос и предоставлять пользователю интересующие его примеры в разумное время (время отклика). Большинство лингвистически содержательных запросов обрабатываются ВАНК в течение нескольких секунд. Время отклика является техническим требованием к ресурсу. То, насколько сложные алгоритмы потребуются для удовлетворения этого требования, зависит, во-первых, от объема корпуса и, во-вторых, от степени гибкости предполагаемой поисковой функциональности.

Реализация поиска на многомиллионных корпусах с функциональностью, аналогичной функциональности ВАНК, является очень сложной алгоритмической задачей, сопоставимой по сложности (и отчасти аналогичной) поисковым интернет-системам типа Google.

Как видно из вышесказанного, составители электронных библиотек и составители корпусов сталкиваются с проблемами разного рода; причем последние должны решать большее число научных, методологических и алгоритмических задач. Конечно, и тем, и другим на начальном этапе приходится осуществлять одну и ту же, технически сложную и затратную работу по оцифровке бумажных версий текстов - сканирование и автоматическое распознавание либо набор с последующей корректурой. В случае ВАНК было оцифровано более 70 млн. словоупотреблений; остальной объем (чуть менее 40 млн., из которых около 35 млн. составляет электронная пресса) был загружен с открытых интернет-сайтов.

В связи с обсуждением электронных библиотек отметим, что на сайте ВАНК открыт специальный раздел - Электронная библиотека ВАНК, в которой в полнотекстовом доступе размещено более ста произведений классической армянской литературы. Все размещенные в библиотеке произведения были оцифрованы в рамках проекта, поэтому значительных пересечений с другими общедоступными ресурсами (уже упомянутыми выше www.armenianhouse.org и пр.) нет. От других электронных библиотек библиотека ВАНК отличается наличием лексической и морфологической разметки, которой может пользоваться, например, студент, изучающий армянский язык. Если щелкнуть мышкой по словоформе, на экране появляется исходная форма лексемы, характеризующие данную словоформу словоизменительные признаки, для большинства словоформ английские переводы, а также некоторая другая лингвистическая информация.

Состав корпуса ВАНК

Как уже было сказано, чтобы быть эффективным инструментом лингвистических исследований, корпус должен представлять языковое выражение во всем спектре его употреблений и контекстов и отражать относительную частотность этих употреблений, то есть быть репрезентативным и сбалансированным. Несмотря на то, что языковая репрезентативность корпуса - характеристика трудноформализуемая, ясно, что корпус должен по крайней мере представлять жанровое разнообразие языка и что наиболее важные жанры - проза, поэзия, публицистика - должны быть представлены в нем достаточно полно. Корпус, который содержит только публицистические, только научные или только художественные тексты, или корпус, в котором все эти жанры есть, но публицистики, скажем, очень мало, не может полностью удовлетворить лингвиста. Конечно, полнота представленности - понятие также относительное, и поэтических текстов в любом национальном корпусе всегда мало (в ВАНК - менее четырех миллионов словоупотреблений, или около 3%). В ВАНК вошли тексты разных

жанров, не только проза, поэзия и пресса, но и официальные, научные, религиозные тексты. При этом корпус позволяет ограничивать поиск отдельными жанрами и группами жанров, которые интересуют пользователя в данный момент.

Таблица 1. Распределение письменных текстов по основным жанрам

Письменные тексты	**словоупотребления**	**% от ВАНК**
пресса	47 265 000	43,0%
проза	37 029 000	33,7%
наука	13 750 000	12,5%
другие нехудожественные	4 681 000	4,3%
поэзия	3 627 000	3,3%
Всего письменных текстов	**106 352 000**	**96,8%**

Важнейшим аспектом корпусной лингвистики являются микроисторические исследования, ориентированные на "быстрые" языковые изменения. Для таких исследований корпус должен иметь временную координату, по которой можно отслеживать изменения значений лексемы или граммемы, отмирание старых и появление новых конструкций. ВАНК покрывает весь новый период истории армянской литературы с самого начала ашхарабара в том состоянии языка, в котором он представлен, например, в произведениях Абовяна. Временные характеристики текстов также можно использовать при выборе подкорпуса - например, работать только с текстами двадцатого или второй половины двадцатого века.

Отправляя запрос и получая ответ в виде совокупности контекстов, пользователь ставит себя в зависимость от того, какие жанры и периоды больше, а какие меньше представлены в корпусе. Если, например, научная литература представлена в корпусе непропорционально широко, у исследователя может сложиться неверное представление о частотных характеристиках того или иного языкового феномена. Принято говорить о необходимости баланса разных жанров в составе корпуса. Установленных универсальных пропорций между разными жанрами в корпусной лингвистике не существует. Более того, кажется, что их и не может существовать - культуры могут отличаться по тому, какую роль в их письменном языке играет художественная литература, а какую публицистика, какую проза, а какую поэзия, и т.д. Чаще всего к проблеме баланса подходят именно с такой "культурологической" точки зрения; методов формальной оценки лингвистической сбалансированности корпуса пока, насколько нам известно, не существует. Культурологическая же оценка обычно бывает вполне приблизительна - считается, что пресса и художественная литература должны быть представлены сравнимыми объемами текстов, поэзии вполне может быть заметно меньше, чем

прозы, научной литературы может быть меньше, чем художественной, и так далее. Отметим, что такие пропорции достаточно точно отражают степень доступности текстов того или иного типа.

Однако на пропорциональную представленность в корпусе разных жанров порой накладываются внешние ограничения, в разной степени очевидные и в разной степени неизбежные. Важный тип текстов, почти не представленный в ВАНК - это эпистолярные тексты. Письма писателей и выдающихся людей в некотором объеме в корпусе представлены, а вот переписка простых людей, так сказать, "бытовая" эпистолярика в корпусе отсутствует. Несмотря на то, что доля бытовой письменной корреспонденции в общем объеме написанных на армянском языке письменных текстов значительна, включить такие документы в корпус затруднительно по чисто техническим причинам (трудности с доступом к рукописным документам и с их обработкой).

Интересным и пока слабо охваченным жанром письменных текстов является электронная коммуникация: электронная почта, смс, instant messengers, блоги. Эти тексты представляют интерес, поскольку они представляют новый тип коммуникации, отчасти промежуточный по своим характеристикам между письменным и устным. Основное затруднение заключается в том, что до самого последнего времени в текстах такого типа на армянском языке использовались в основном косвенные и нестандартизованные способы передачи армянской письменности. В марте 2009 г. в корпус был включен первый массив текстов этого рода - армяноязычный форум forum.am-kayq.com.

Но основная сложность заключается в том, что в идеале каждый жанр должен быть относительно равномерно распределен по годам; колебания, если таковые имеются, должны отражать какую-то содержательную характеристику письменной традиции. В качестве примера действия препятствующих такой временной сбалансированности технических ограничений можно привести распределение прессы ВАНК по годам. Значительная часть прессы относится к пост-советскому периоду за счет широкой представленности текстов открытых периодических интернет-изданий (в ВАНК - около 35 млн. словоупотреблений), доступный объем которых практически неограничен, в то время как пресса советского и досоветского периода существует только в виде печатных изданий. Эта проблема была отчасти преодолена после осуществления совместного с Национальной библиотекой Армении проекта, в рамках которого были отсканированы, распознаны и включены в корпус избранные выпуски 60 периодических изданий общим объемом более 12 млн. словоупотреблений. Архив периодики относительно равномерно покрывает всю историю существования армянской прессы, начиная от 70-х годов 19-го века по поздний советский период. Но полученная из электронных архивов пресса последних лет все равно дает всплеск, лишь отчасти мотивированный экспансией интернет-периодики. Поэтому баланс прессы по времени (а также соотношение публицистических и иных текстов в последней декаде) остаются не идеальными.

Ср. таблицу:

Таблица 2. Распределение жанров по времени
(объем в тысячах словоупотреблений и проценты от общего числа словоупотреблений в декаде)

период	**проза**		**поэзия**		**нехудожественные**		**пресса**		**итого за период**
		% за период		**% за период**		**% за период**		**% за период**	
до 1870	292	64%	4	1%	n/a	0%	161	35%	456
1870 - 1879	515	53%	49	5%	250	26%	150	16%	963
1880 - 1889	1431	74%	4	0%	48	3%	447	23%	1930
1890 - 1899	802	100%	n/a	0%	n/a	0%	n/a	0%	802
1900 - 1909	736	36%	84	4%	253	12%	955	47%	2029
1910 - 1919	452	60%	62	8%	n/a	0%	246	32%	759
1920 - 1929	740	44%	297	18%	44	3%	599	36%	1680
1930 - 1939	2211	57%	28	1%	243	6%	1410	36%	3892
1940 - 1949	923	46%	139	7%	199	10%	733	37%	1993
1950 - 1959	2408	47%	785	15%	463	9%	1422	28%	5078
1960 - 1969	4014	57%	479	7%	426	6%	2176	31%	7095
1970 - 1979	5885	48%	122	1%	4355	36%	1899	15%	12262
1980 - 1989	3984	34%	69	1%	5936	50%	1861	16%	11850
1990 - 1999	1227	37%	79	2%	1325	40%	650	20%	3281
2000 - 2008	879	2%	37	0%	3993	10%	34553	88%	39462
недатированные	10531	82%	1391	11%	897	7%	3	0%	12821
Итого	**37029**	**35%**	**3627**	**3%**	**18431**	**17%**	**47265**	**44%**	**106352**

Для армянской литературной традиции двадцатого века очень важны переводы русской и западной литературы. В корпус включен значительный объем (около 12 млн. словоупотреблений) переводных текстов более 100 авторов, в основном художественная литература, причем ВАНК дает возможность ограничить поиск только переводными (или, соответственно, только оригинальными) текстами. К переводным текстам отнесены, в том числе, переводы на современный армянский

классических древних авторов - Хоренаци, Нарекаци, Даврижеци и некоторых других.

Оригинальная восточноармянская художественная литература представлена произведениями 241 автора и составляет более 35 млн. словоупотреблений (около трети всего корпуса). Для сравнения можно сказать, что объем "Самвела" - менее 150 тыс. словоупотреблений, а полное собрание сочинений Раффи содержит менее полутора миллионов словоупотреблений. В этой связи излишне говорить, что ВАНК покрывает не только все тексты школьной программы, но практически все сколько-нибудь известные произведения армянской литературы. С другой стороны, ВАНК является не собранием критических филологических изданий, а лингвистическим инструментом. Поэтому в корпусе могут отсутствовать периферийные тексты классических авторов и иные случайные лакуны, отсутствуют сноски и комментарии, остается довольно много опечаток.

В целом ВАНК является одним из крупнейших языковых корпусов в мире (ср. краткие описания других корпусов на http://www.linguistlist.org/sp/texts.html).

Устный корпус ВАНК

Кроме различных письменных текстов в состав ВАНК входит большой подкорпус устной речи. Устные тексты также диверсифицированы и включают публичную устную речь (в основном расшифровки телепередач - около 1,7 млн. словоупотреблений), спонтанную устную речь (около 1 млн.), а также более мелкие и более специальные жанры - электронная коммуникация (http://forum.am-kayq.com - около 400 тыс. словоупотреблений), речь кино (расшифровка речи актеров - около 200 тыс. словоупотреблений), так называемая стимулированная речь (расшифровки психолингвистических экспериментов - около 70 тыс.), всего почти 3,5 млн. Отметим, что, хотя о балансе и о полноте устного корпуса говорить вообще сложно - множество устных текстов, в отличие от множества письменных текстов, принципиально открыто - основной тип устной речи, спонтанные диалоги, представлены в корпусе достаточно широко. Все устные тексты были записаны и расшифрованы в рамках реализации проекта ВАНК.

Армянский язык, таким образом, вошел в число очень немногих языков, для которых существуют обширные корпуса устной речи; в основном это крупные европейские языки - например, английский, итальянский, русский. Для сравнения скажем также, что, хотя общий объем устных текстов в Национальном корпусе русского языка составляет почти 6 млн., что заметно больше, чем в ВАНК, объем спонтанных устных текстов, которые, с нашей точки зрения, представляют стихию устной речи наиболее непосредственно, в ВАНК более чем вдвое превышает соответствующий подкорпус НКРЯ.

Устная речь - это динамичная и живая субстанция, практически не поддающаяся регулированию или иному целенаправленному воздействию извне и в диахроническом плане опережающая современное ей состояние письменного языка. В ней отражается большое число культурных влияний, социальных изменений, она является котлом из самых разных языковых процессов, а иногда и

разных языков; погружена в прагматику ситуации и потому гораздо менее эксплицитна, более эллиптична; гораздо менее отрефлектирована и потому во всех отношениях менее нормативна, не только и не столько на уровне лексики, сколько на уровне грамматики и особенно синтаксиса. Некоторые ученые считают, что только такая субстанция и является единственным достойным объектом лингвистического исследования, у других устные данные вызывают неприятие. Это верно для лингвистической традиции вообще и для арменистики в частности. Использование разговорных, ненормативных грамматических форм, иностранных, в основном русских или английских слов, жаргонизмов - все это расценивается не просто как отклонение от литературной и письменной нормы, но как "неправильность" языка, который поэтому не является достойным объектом изучения (те же доводы часто приводятся против изучения текстов электронной коммуникации). С первым трудно спорить, но и согласиться со вторым никак нельзя.

Все эти претензии, на самом деле, апеллируют не к недостаткам, а к языковым особенностям устной речи, которая имеет собственную, иногда кардинально отличную от литературной норму, действительно широко использует переключение кода, действительно обладает собственным синтаксисом и т.п. Именно для исследования этих особенностей устной речи и формируются подобные корпуса. Такие исследования обычно относятся не к сфере чистой лингвистики, а к смежным областям - социолингвистике (переключение кода), психолингвистике (особенности построения высказывания, речевые ошибки и т.п.), и в последнее время начинают проводиться на материале всех европейских языков. Примеры исследований такого рода в арменистике см. [Хуршудян 2006; Хуршудян, Подлесская 2006]. Конечно, устная речь значительно отличается от письменной. Но это не значит, что устная речь неправильна - просто она живет по своим собственным законам. Стихия устной речи позволяет изучать тенденции языкового развития. Знание этих тенденций оказывается важным не только для чистой лингвистики и социолингвистики, но и, например, для языкового планирования: осуществление языковой политики в отрыве от живых языковых процессов никогда не бывает вполне успешным. Поэтому мы считаем устный подкорпус важнейшей частью проекта ВАНК и хотели бы привлечь к нему особое внимание исследовательского сообщества.

Отсюда вытекает ответ и на другой вопрос о сбалансированности корпуса - как определить правильное соотношение между количеством письменных и устных текстов? Из вышесказанного ясно, что это вопрос бессодержательный. Объем письменного и устного подкорпусов могут находиться в произвольном отношении между собой, так как по сути это два разных способа существования языка, две грамматики, два корпуса.

Разметка корпуса

Как уже было сказано, важнейшее отличие корпуса от любой другой совокупности текстов - наличие грамматической и иной информации, внесенной в

него разработчиками. Такая информация обычно называется *разметкой* (markup). Неразмеченный корпус представляет для лингвиста определенную ценность, но лингвистическая ценность размеченного корпуса, снабженного эффективной системой поиска, на много порядков выше, а корпус, не сопровождаемый лингвистическим аппаратом, в принципе не может обладать эффективной (с точки зрения исследователя-лингвиста) системой поиска.

Технология внесения в корпус лингвистической информации может быть очень различна - от полностью ручной до полностью автоматической. Ручная или частично ручная разметка используется при документации диалектов или малых и вымирающих языков или в специальных проектах с малым объемом текстов, требующих специфической ручной обработки - например, при создании корпусов детской речи и в других тонких психолингвистических исследованиях. В случае национальных корпусов такая технология неприменима просто по причине их большого объема. Тексты ВАНК были обработаны программой лемматизатором, которая сопоставляла каждой словоформе исходную форму и набор лексических и грамматических помет. То, как эта информация записывается, не так важно (в частности, она может храниться и отдельно от самих текстов). Примерное представление о разметке дает следующий пример:

կայծակը
կայծակ (N, inanim) – *неодуш. сущ., исходная форма: կայծակ*
{sg nom def} – *грамматические признаки: единственное число, именительный падеж, определенная форма*
lightning – *английский перевод*

Как видно, в записи используются сжатые, условные и отчасти технические обозначения - N вместо существительного, sg вместо единственного числа, nom вместо именительного падежа, def вместо определенной формы: подробнее описание помет см. на сайте в разделе "Разметка ВАНК (список помет ВАНК)". Наличие разметки позволяет поисковому компоненту корпуса быстро находить интересующие пользователя словоформы и показывать контексты, в которых они встретились. Для эффективного и быстрого поиска происходит индексация корпуса, при которой сам корпус преобразуется в базу данных - но все это скрыто от пользователя, принципиальное отличие корпуса от электронной библиотеки заключается именно в наличии лексико-грамматической информации как таковой.

Не все словоформы корпуса разбираются программой-лемматизатором. Во многих случаях это связано с опечатками или ошибками распознавания исходных текстов. Есть и более содержательные ошибки - лемматизатор не в состоянии обработать авторскую орфографическую игру, при которой искажается правильное написание слова; вкрапления из грабара и западноармянского, иностранные слова (переключение кода) и т.п. Наконец, некоторые редкие лексемы могут быть не включены в словник ВАНК, им может быть приписан ошибочный словоизменительный тип или не указан допустимый словоизменительный вариант. На настоящий момент вообще не разбирается 7% словоупотреблений ВАНК; еще

около 3% составляют цифровые вхождения и вхождения использующие цифры других алфавитов. Это не значит, что остальные 90% разобраны правильно - вхождению может быть приписан неправильный разбор - однако такой статистикой мы не располагаем.

Еще одним важным компонентом ВАНК является наличие библиографической разметки, включающей жанр и название текста, время написания или издания, имя автора и некоторые другие сведения. Эта информация важна с двух точек зрения. Во-первых, пользователь корпуса, работая с примерами, должен знать, к какому жанру принадлежит данный текст, а также кем и особенно когда он был написан, так как от этого существенно зависит интерпретация языковых данных. Во-вторых, ВАНК предоставляет пользователю возможность ограничить поиск языкового материала заданным им подкорпусом. Подкорпус можно ограничить разными основаниями - жанр, время написания, автор и т.п. Понятно, что для обеспечения такой функциональности необходимо приписать текстам базовую библиографическую (метатекстовую) разметку. Оговоримся, что сведения о времени написания носят приблизительный характер и нуждаются в доработке - для многих произведений период написания определен лишь очень приблизительно, с точностью до пятидесяти лет. В значительной степени это связано с недоступностью сводных библиографических источников. Жанровая разметка и разметка по авторам произведений проведена достаточно последовательно и хорошо.

Грамматический словарь

Человек, читающий текст на иностранном языке, должен открыть словарь, найти возможные исходные формы и принять решение о том, формой какой лексемы является интересующая его словоформа. Примерно так работает и лингвистический аппарат корпуса. Для того чтобы произвести лемматизацию, т.е. внести лексическую и грамматическую разметку, программное обеспечение корпуса должно уметь устанавливать соответствие между словоформой в тексте и исходной формой лексемы. Для этого необходим уже кратко упоминавшийся выше грамматический словарь, в котором приводятся не только исходные формы (леммы) слов, но и их парадигматические типы, исчерпывающим образом определяющие их формальную парадигму.

В основу грамматического словаря ВАНК положен словник объемом около 80 тыс. слов. Этот словник является компиляцией многих источников - в первую очередь словаря Е. Г. Галстян [Галстян 1985] и части словаря Э. Б. Агаяна [Агаян 1976], но также словаря аббревиатур Д. С. Гюрджиняна и Н. А. Экекян [Гюрджинян, Экекян 2007], словаря географических названий А. Гргеаряна и Н. Арутюнян [Гргеарян, Арутюнян 1987-1989], различных имен собственных и пр.

Однако компиляция словника - это лишь часть работы, и весьма небольшая. Если программа не имеет никакой информации о словоизменении, она не сможет определить, что словоформа *գիտի gini* является именительным падежом от лексемы

գինի gini 'вино', а не родительным падежом от лексемы *գին gin* 'цена'. Только наличие в словнике словоизменительных помет позволяет понять, что родительный падеж от *գին gin* образуется с редукцией гласного – *գնի gni.* Составление грамматического словаря - трудо- и времяемкая работа, которая ранее на армянском материале, насколько нам известно, не проводилась; проект, который решал небольшую часть этой задачи - словарь форм множественного числа [Гюрджинян 2005]. Для сравнения скажем, что первый и достаточно полный грамматический словарь русского языка, содержащий свыше ста тысяч лексем, появился уже более тридцати лет назад [Зализняк 1977].

Несмотря на богатую традицию грамматического описания армянского языка, в готовом виде получить из какого-либо источника классификацию именных или глагольных парадигм, пригодную для автоматического анализа текстов, невозможно. Дело не в том, что какие-то словоизменительные типы остались неописанными в традиционной арменистике (хотя в ВАНК и встречаются типы, отсутствующие в традиционных грамматиках, но присутствующие в разговорном языке и потому отраженно представленные в литературе - например, *տղի tġi – տղու tġu*). Важнее то, что задачи автоматической обработки текстов имеют приоритеты, отличные от приоритетов теоретического и практического грамматического описания. Для построения алгоритма лемматизации армянских именных словоформ мы выделили более 50 формальных типов именного словоизменения. С лингвистически содержательной точки зрения многие из этих типов могут быть сведены друг к другу или выведены из фонотактики слова. Практически каждому набору падежей соответствует два словоизменительных типа в зависимости от выбора показателя множественного числа *–եր -er/-ներ -ner*, что "несодержательно" удваивает общее число типов. Сам выбор показателя множественности в статистически подавляющем большинстве случаев определяется количеством слогов в исходной форме. Некоторые типы отличаются друг от друга только наличием беглой гласной в корне. Однако для удобства автоматической обработки текстов вместо того, чтобы отягощать морфологическую модель различными содержательными правилами, гораздо проще приписать каждому существительному один из пятидесяти словоизменительных типов в готовом виде. Ясно, что в научных работах задача описания словоизменения таким образом не ставится. Полный список *всех* различных парадигматических типов приведен на сайте проекта в разделе "Разметка".

С практической точки зрения работа над словником была организована по принципу итераций - после того, как на основании различных источников был создан, размечен и выверен исходный вариант словника ВАНК, была проведена первоначальная обработка массива текстов. После этой обработки просматривались наиболее частотные формы, которые лемматизатор разобрать не смог - в основном это были либо словоформы лексем, отсутствовавших в словнике ВАНК, либо последствия неправильно приписанной словоизменительной информации (в

частности, отсутствие факультативного словоизменительного варианта). После этого словник ВАНК поправлялся и исправлялся. Такие итерации время от времени повторяются, после чего лексико-морфологическая разметка корпуса обновляется.

Морфологическая модель

В основе любого алгоритма морфологического анализа текста лежит то или иное лингвистическое описание. Необходимо принять решения об интерпретации спорных форм, иногда относительно периферийных, иногда центральных. В ходе исследования лингвист часто сталкивается с неоднозначностью языковых данных (не говоря уже о различиях в теоретических установках) и с возможностью их различной интерпретации; одни и те же формы могут интерпретироваться как словоизменение или словообразование, одни исследователи выделяют один набор частей речи, другие - другой и т.п. При разработке системы автоматического морфологического анализа нужно выбрать одно из возможных решений. Выбор решения прямо или косвенно сказывается на пользователе корпуса - ему приходится адаптироваться под используемую в корпусе модель морфологии и принимать во внимание интерпретации, которые могу казаться ему спорными или неверными по сути.

Примером вопроса о центральных, но спорных категориях, является проблема различения в армянском языке дательного и родительного падежа. Большая часть описаний различает эти формы, однако нельзя не признать, что грамматисты испытывают здесь серьезное давление классической грамматической традиции. Если же обратиться к внутренним языковым свидетельствам, то оказывается, что у имен дательный падеж отличается от родительного только своей способностью присоединять определенный артикль: первый обычно имеет показатель определенности, а второй никогда с ним не сочетается. Не вполне ясно, является ли это достаточным основанием для выделения двух разных падежей. Почти с теми же основаниями в армянском можно было бы выделять, например, специальную счетную форму - употребление формы единственного числа в контексте числительного очевидно имеет здесь особую функцию, которая с трудом может быть выведена из семантики единственности.

Важным доводом в пользу различения падежей могло бы служить то, что в армянском языке есть именная часть речи, у которой форма дательного падежа отличается от формы, которая используется в приименной позиции – это личные местоимения. Однако такие формы личных местоимений можно в принципе считать отдельными лексемами – притяжательными местоимениями. Таким образом, вполне оправдана точка зрения, согласно которой в армянском языке есть единый падеж, выполняющий одновременно функции дательного и винительного падежей; а поскольку он маркирует еще и прямое дополнение, его можно было бы называть просто косвенным падежом, аналогичным обликвусу (oblique), широко представленному во многих языках с бедной падежной системой. Однако в морфологической модели ВАНК принята другая, более традиционная точка зрения, согласно которой притяжательные формы являются элементами падежной

парадигмы местоимений, а у имен выделяются две разные падежные категории, датив и генитив, что не в последнюю очередь обусловлено стремлением сохранить единство структуры парадигмы для всех именных частей речи. Впрочем, лемматизатор ВАНК не может различить родительный и дательный падеж у существительных без артикля (они могут быть различены только в контексте, который лемматизатор не учитывает), поэтому де факто в таких случаях обе грамматические пометы приписываются одновременно: gen/dat.

Примерами периферийных форм, которые встречаются в корпусе, но почти не обсуждаются в традиционных описаниях, являются реляционная форма имени или форма ассоциативной множественности. Реляционная форма имени, или релятив – это относительно редкая форма, более характерная для устной речи, но изредка встречающаяся и в письменных текстах. С морфологической точки зрения релятив – это именная основа, к которой присоединен показатель родительного падежа, затем определенный артикль, а затем либо просто определенный артикль, либо, реже, показатель дательного падежа и определенный артикль или показатель родительного или одного из периферийных (пространственных) падежей.

1. *սեղանինը*
   seġan-i-n-ə
   стол-GEN-DEF-DEF
   ‘то, что на столе’

2. *սեղանիինը*
   seġan-i-n-in-ə
   стол-GEN-DEF-DAT-DEF
   ‘тому, что на столе’

3. *սեղանինով*
   seġan-i-n-ov
   стол-GEN-DEF-INST
   ‘тем, что на столе’

С функциональной точки зрения релятив является субстантивацией формы генитива и может присоединять именные грамматические показатели, например, падеж. Значение такой формы можно описать как 'принадлежащий / имеющий отношение к N', где N – это производящая именная основа. Имеются основания полагать, что артикль, следующий за показателем генитива, с функциональной точки зрения не может считаться артиклем, так как он выступает здесь в роли не артикля, а номинализатора или, в другой терминологии, субстантиватора. Поэтому реляционные формы имени в ВАНК считаются субстантивированными формами генитива.

Формы ассоциативной множественности - это формы, которые обозначают группу лиц, ассоциированную с неким главным членом этой группы, обозначенным производящей именной основой. Форма более характерна для разговорной и диалектной речи. В качестве примеров можно привести, например, *Վարդանանք* ‘Вартан и его группа (его друзья, родственники)’, *Շուշանենք* ‘Шушан и ее группа (ее друзья, родственники)’. Морфологически к формам ассоциативной множественности примыкают местоименные формы вида *մերոնք* ‘наша группа’, хотя с семантической точки зрения их интерпретация как форм ассоциативной множественности несколько проблематична.

Наконец, пользователю могут показаться непривычными или произвольными некоторые решения, относящиеся к сфере грамматической терминологии. Эти решения были направлены в первую очередь на сближение арменистической традиции с современной типологической практикой описания языков. Так, ВАНК

относит к различным частям речи причастия и деепричастия, в названиях падежей используется латинская номенклатура, деепричастие называется более распространенным в теоретической литературе термином конверб и т.п. Пожалуй, самым значимым терминологическим решением является использование вместо термина пассива термина медиопассив. Несмотря на то, что в русской армяноведческой традиции принято говорить о пассиве, на самом деле эта категория в армянском языке гораздо шире - она похожа на русскую возвратность, а в типологической номенклатуре соответствует термину медиальный залог, или медий.

Итак, при работе с ВАНК пользователю отчасти приходится вставать на теоретические позиции лежащей в его основании лингвистической модели. В целом, при известном навыке, это не должно создавать у пользователя серьезных проблем - нужно лишь желание работать с корпусом и готовность принять некоторые "правила игры". Конечно, вопрос не в смене научных взглядов, а только в умении и стремлении пользоваться корпусом как инструментом исследования. Кроме того, речь идет о словоизменительной морфологии и различных лексических пометах, в рамках которой разнообразие интерпретаций значительно более ограничено, чем, например, в сфере синтаксиса. Подробно различные технические решения и конвенции описаны на сайте ВАНК в разделе о разметке.

О грамматической омонимии и контекстной информации

Понимание естественного языка - это многокомпонентная деятельность, предполагающая работу сразу многих языковых модулей. Человек, читающий текст, воспринимает слово не вне, а внутри контекста - у него не возникает проблем различить значения одной и той же словоформы *գրում grum* в таких контекстах как:

4. *«Հառա՛ջ, ի փառս հայրենիքի», — հպարտորեն* ***գրում*** *է նա օրագրում:*
«Вперед, во славу родины!», — гордо пишет он в дневнике. (С. Цвейг «Звездные часы человечества»)

5. *Գիտնական այր դարձավ, սակայն հոգեւոր ոլորտի հմայքը չկորցրեց նաեւ իր* ***գրում:***
Он стал ученым, однако в своих произведениях продолжал отдавать должное духовной сфере. (газета «Азг», 23.01.2004)

Значительная часть лемматизаторов – в том числе лемматизатор ВАНК - устроена иначе. Форма рассматривается в отрыве от контекста, так что словоформе, которая может иметь несколько разных интерпретаций, приписываются все возможные грамматические интерпретации вне зависимости от окружающего контекста. В данном случае в обоих случаях приписывается и разбор словоформы

как локатива от не очень частотного существительного *գիր gir* 'письменность, письменное творчество', и как конверба (деепричастия) несовершенного вида от крайне частотного глагола *գրել grel* 'писать'. При поиске форм, допускающих омонимичные разборы, пользователь должен быть готов к тому, что в результатах нужные ему языковые данные окажутся перемешанными с поисковым "шумом". В приведенном выше случае, если пользователь хочет найти формы местного падежа от лексемы *գիր gir*, шумом окажется подавляющее большинство результатов поиска, так как в значении конверба форма *grum* встречается гораздо чаще, чем в значении падежной формы. Долю шума в результатах поиска можно несколько сократить, используя контекстный поиск. Например, если указать, что перед формой *գրում grum* идет прилагательное, форма родительного падежа или другой приименной атрибут, вероятность нахождения контекстов с локативностью несколько возрастает. Но при таком поиске мы упускаем те контексты, в которых *գրում grum* является формой имени, но не имеет при себе определения.

Аналогично обстоит ситуация и с поиском просто по грамматическим показателям. Если искать формы конверба несовершенного вида от любого глагола, в поиске будут попадаться формы, омонимичные с ними, но являющиеся формами местного падежа (тот же *գրում grum* в значении 'в (своих) книгах'). Как и при поиске словоформ, можно воспользоваться контекстным поиском: ввести такие условия, которые делают одну интерпретацию более вероятной, чем другую. В данном случае можно указать, что непосредственно перед или после искомой формы стоит глагол-связка. Тогда мы исключим заметную долю ненужных употреблений тех форм на *-ում -um*, которые имеют омонимию типа V, cvb ↔ N, loc. Но шум все равно остается, так как форма местного падежа вполне может стоять рядом со связкой, а часть нужных нам контекстов при этом теряется, так как не всегда конверб находится в непосредственном контакте с глаголом-связкой.

Еще один способ сократить шум - выбрать специальную опцию поиска, исключающую из результатов словоформы с неединственным разбором. В этом случае, если мы ищем словоформы лексемы *գիր gir*, мы найдем только именные формы, а если мы ищем словоформы глагола *գրել grel* - только глагольные. Но тогда словоформа *գրում grum* вообще не попадет в результаты поиска, так как у нее есть два разбора - а ведь она составляет значительную часть словоупотреблений глагола *գրել grel*; если исключить вообще все омонимичные разборы, то останется лишь около двадцати процентов всех контекстов с лексемой *գրել grel*.

Наконец, поиск без учета контекста имеет еще одно важное, системное грамматическое последствие. В глагольной парадигме современного армянского языка преобладают аналитические глагольные формы, состоящие из сочетания нефинитной глагольной формы (мы называем ее конвербом) со вспомогательным глаголом (связкой). Однако, поскольку лемматизатор не умеет устанавливать связи между словоформами предложения, он не может выделить из контекста вспомогательную конструкцию: формы конвербов и связки в любом случае

получают независимые разборы. Иными словами, лемматизатор полностью игнорирует морфосинтаксис. Как и в предыдущих случаях, для поиска собственно аналитических конструкций можно использовать косвенные возможности контекстных запросов.

Отсутствие контекстной информации также, и даже в еще большей степени, сказывается на семантической информации об употреблении тех или иных форм. И в результатах поиска, и в библиотеке в окошке подсветки отображается несколько вариантов английского перевода данного слова. При этом переводы никак не соотнесены с контекстом - пользователю приходится самому выбирать нужный, опираясь на смысл всего предложения. При поиске по переводам возможен точно такой же шум, что и при поиске по грамматическим признакам, лемме или словоформе. Так, поиск по значению 'pit' (яма) даст многочисленные контексты со словоформой *հոր hor*, которая в большем числе случаев, конечно, является родительным или дательным падежом от *հայր hayr* 'отец'. Точно так же лемматизатор не в состоянии определить, в каком значении употреблена та или иная граммема - например, различить реципиентные и локативные употребления дательного падежа или пассивное и реципрокальное употребление медиопассива. В отличие от межлексемной омонимии словоформ (типа обсуждаемых выше *գիր gir* → *գրում grum* ← *գրել grel*) варианты значения лексемы или граммемы трудно различить даже используя контекстный поиск.

Таким образом, при работе с корпусом надо иметь в виду, что в результатах поиска по некоторым запросам может содержаться большое количество ненужных пользователю примеров. Это связано с тем, что лемматизатор анализирует словоформу в отрыве от контекста. Если причина появления "лишних" контекстов на первый взгляд неясна, можно навести на найденную словоформу мышку и посмотреть, какие разборы ей приписаны. После этого в некоторых случаях удается сузить запрос исключением омонимичных разборов или использованием контекстных ограничений, повышающих вероятность искомого явления и/или снижающих объем шума. Однако, как мы видели выше, некоторая вероятность появления шума остается, а некоторая часть нужных примеров при этом может теряться. Пользователь корпуса часто вынужден просматривать найденные контексты, отбрасывая ненужные и сохраняя нужные, правильные результаты поиска.

Целевая аудитория ВАНК

Как уже говорилось, аудиторией проекта является в первую очередь сообщество арменистов, работающих с лексикой и грамматикой ашхарабара, а также решающие сравнительные задачи специалисты по западноармянскому языку или грабару. С точки зрения представленности различных форм и временных срезов, ВАНК покрывает значительную часть всего разнообразия языкового материала и ограничен почти только рамками логически невозможных (несовременные устные тексты) или крайне труднодоступных (жанр частной переписки) типов текстов. Как

было сказано, корпус позволяет искать не только определенные словоформы, но и лексемы, и грамматические категории, а также сочетания нескольких слов с определенными признаками в одном контексте (подробнее см. приложение о функциональности). Такого рода функциональность лучше всего приспособлена для изучения лексики и морфологической семантики (значений и правил употребления тех или иных грамматических форм) языка, в определенной степени морфосинтаксиса (например, для изучения глагольных управлений). Синтаксические явления, такие как структура именной группы, стратегии релятивизации, механизмы поддержания референции могут изучаться лишь косвенно, и исследователь-синтаксист должен быть готов применять обходные методы получения релевантных контекстов и/или вручную отсеивать значительное количество поискового "шума".

Благодаря включению переводов и глоссированного формата выдачи корпусом могут пользоваться исследователи, не владеющие или не вполне владеющие армянским языком - арменисты, которые только начинают изучать армянский язык, а также лингвисты-типологи, вообще не специализирующиеся в армянской филологии. Студент-лингвист, таким образом, может проводить собственные микроисследования, пользуясь корпусом точно так же, как и опытный исследователь-арменист, но при необходимости обращаясь к грамматическим разборам и переводам незнакомых форм. Студент-нелингвист сможет познакомиться с особенностями значения тех или иных лексем через исследование контекстов их употребления. В принципе, корпус поможет ему составить представление о функционировании тех или иных грамматических форм, но на практике такая постановка вопроса в большей степени свойственна людям с лингвистическим фундаментом.

В целевую аудиторию корпуса, как мы надеемся, могут войти школьные и вузовские преподаватели армянского языка и преподаватели армянского языка как иностранного. Использование корпусов в преподавании - вполне активная, а количественно едва ли не доминирующая отрасль корпусной лингвистики. В последнее время развернута активная пропаганда такого использования Национального корпуса русского языка - осенью прошлого года прошла посвященная этой теме конференция в Москве, по материалам конференции издан сборник статей [Добрушина ред. 2007], имеется также серия публикаций Н. Р. Добрушиной (например, [Добрушина 2005]), начинается разработка образовательного портала корпуса.

Бурное развитие этого направления носит, конечно, вполне прагматический характер (в случае языков, имеющих государственный статус, число изучающих такие языки студентов, по-видимому, всегда больше, чем академических исследователей) и поэтому пропорционально объему преподавания. Для армянского языка этот объем меньше, чем для японского, английского или русского. Но в том объеме, в каком армянский язык преподается, ВАНК мог бы послужить преподавателям важным подспорьем. Он дает возможность работать с живым языковым материалом и отойти от традиционных методов обучения, опирающихся на закрытый и ограниченный объем признанной литературной

классики и жестко нормативного языка. Из корпуса мы узнаем, как на языке говорят, как его в действительности используют.

Здесь естественно также упомянуть о той сфере использования корпусов, которая лежит в "серой" зоне между филологами и преподавателями языка - нормативной лингвистике. Как таковая эта отрасль не принадлежит ни к какому направлению академического лингвистического исследования и относится скорее к общественно-политической, чем научной сфере. Тем не менее нельзя не признать, что государственное регулирование языковых практик, по крайней мере в сферах, смежных с официальной, является социальной необходимостью и нуждается во внимании со стороны лингвистов. Как можно использовать корпус в языковом планировании? Понимая, что корпус ни в коем случае не является образцом нормы, еще раз повторим, что все-таки именно корпус, представляя действительный языковой узус, может и должен становиться основой для работы над нормой. Языковые реформы, которые оторваны от живых языковых процессов, обречены на фиаско, как видно сегодня на примере бесплодных попыток достичь консенсуса по реформе русской орфографии, а если таковые реформы принимаются, то в конечном итоге они будут отторгнуты языковой стихией В здоровом социуме лингвистический произвол в языковом реформировании невозможен, так как языковой процесс не поддается законодательному регулированию. И здесь важную роль может сыграть как сам ВАНК, фиксирующий языковые сдвиги на протяжении более чем полутора веков, так и устный корпус ВАНК, демонстрирующий живые языковые процессы и обладающий значительным (по сравнению с устными корпусами многих других языков мира) объемом почти 3,5 млн. словоупотреблений.

Для представителей других специальностей - историков, социологов, культурологов и др. - корпус может представлять интерес лишь постольку, поскольку они в своих исследованиях обращаются к языковому материалу (что происходит относительно редко). Речь идет о том, как социальные факторы или исторические процессы отражаются в языке, то есть о своего рода исторической социолингвистике. Социолингвисты в основном работают с современным состоянием языка и составляют собственные микрокорпуса, ориентированные на частные задачи. А вот на вопросы об узусе того или иного социального значимого концепта, о том, когда он впервые упоминается в письменных текстах, как распространяется, как отмирает, как меняется его наполнение, ВАНК сможет ответить достаточно однозначно. Здесь гибкая грамматическая и контекстная функциональность поиска оказывается излишней (достаточно поиска по лексемам), зато на первый план выступает репрезентативность корпуса и особенно включение значительного архива армянской периодики, более менее равномерно покрывающего весь период ее существования.

Наконец, часть потенциальной аудитории составляют люди, для которых обращение к корпусу вызвано не профессиональной необходимостью, а личным интересом к языковому материалу. Языковая рефлексия, рассуждения об узусе, значении тех или иных слов характерны для интеллигентного человека вообще. Поэтому при разработке интерфейса ВАНК мы пытались сделать его

функциональность по возможности интуитивной и прозрачной, а разъяснения о том, как искать лингвистическую информацию, максимально неспециальными и свободными от лингвистической терминологии. Пользователь корпуса - нелингвист может искать редкие слова (*եղերդ eġerd* 'цикорий'), слова, в значении которых он сомневается (*բազմապատկիչ bazmapatkič‘* 'множитель'), формы, которые кажутся ему неправильными, но которые он встретил в живой речи или в тексте или запрещаемые нормой формы, которые кажутся ему естественными или допустимыми (*տղու tġu* GEN от 'мальчик'). Привлечение такого рода пользователей требует особых усилий по популяризации корпуса: пользователь - нелингвист, даже если он случайно попадет на сайт корпуса, не сразу поймет, какую пользу он может извлечь из этого инструмента.

Краткий обзор истории проекта и его перспективы

Проект ВАНК был запущен в январе 2006 г. по инициативе группы московских исследователей и компании CorpusTechnologies. Летом 2007 г. открыт портал www.eanc.net, на котором размещен первый релиз корпуса. Второй релиз был размещен на том же портале весной 2008 г. Второй релиз отличался от первого объемом поискового корпуса (около 90 млн. словоупотреблений вместо 60 млн. в первом релизе), открытием полнотекстового доступа к более чем ста произведениям армянской классики (Электронная библиотека ВАНК), а также относительно немногочисленными, но важными функциональными расширениями: в разметку добавлены переводы лексем; добавлен специальный "глоссированный" формат отображения текстов, предназначенный для пользователей, не владеющих или недостаточно хорошо владеющих армянским языком; появилась возможность поиска специфической для армянского языка так называемой "внутренней" пунктуации. На момент сдачи статьи в печать (март 2009 г.) готовится к запуску третий релиз, который будет отличаться от второго в первую очередь объемом (около 110 млн. словоупотреблений).

Как кажется, по полноте и репрезентативности литературного языка ВАНК приблизился к некоторому качественному порогу, преодоление которого не только затруднительно, но и не очень осмысленно. Конечно, можно бесконечно продолжать добавлять в корпус те или иные не попавшие в него художественные произведения или публицистику, но качественных улучшений это уже не принесет. Говоря нестрого, для литературного восточноармянского языка корпус ВАНК является вполне репрезентативным. Устный корпус ВАНК не только можно, но и нужно расширять (теоретически до бесконечности), но уже сейчас он входит в число самых крупных устных корпусов в мире и является единственным устным корпусом такого объема для "среднего" языка. Возможное осмысленное расширение проекта - включение принципиально новых текстов на армянском языке в широком смысле этого слова - литературных западноармянских, средне- и древнеармянских текстов. В настоящий момент на базе проекта ВАНК силами нескольких аспирантов в Ереване осуществляется сбор тестовых диалектных корпусов. Ш. Асильбекян работает в Арцваберде, Г. Мкртчян в Шенаване, С.

Давтян в Гусане; для каждого из диалектов планируется собрать около 100 тыс. словоупотреблений.

В некоторых мелких доработках нуждается грамматическая модель, на которую опирается корпус – например, включение в разметку периферийных и пока не анализируемых грамматических категорий (например, усеченные звательные формы). Технически было бы важно оптимизировать некоторые типы запросов: например, запросы с отрицанием, обработка которых на настоящий момент занимает значительное время. Характерный для армянского языка высокий уровень грамматической омонимии наталкивает на мысль о необходимости работы по снятию омонимии или хотя бы о создании подкорпуса со снятой омонимией (ср. подкорпус с вручную снятой омонимией в Национальном корпусе русского языка). Создание синтаксической модели и внесение в корпус синтаксической разметки могло бы резко увеличить число академических и прикладных областей применимости корпуса. Однако добавление автоматического синтаксического парсера требует огромной теоретической разработки и не может быть реализовано в обозримом будущем.

Источники:

1. Агаян Э. Б. 1976. Արդի հայերենի բացատրական բառարան [Толковый словарь современного армянского языка]. Т. 1-2. Ереван.
2. Галстян Е. Г. (ред.) 1985. Հայ-ռուսերեն բառարան [Армяно-русский словарь]. Ереван.
3. Гргеарян А. К., Арутюнян Н. М. 1987-1989. Աշխարհագրական անունների բառարան [Словарь географических названий]. Ереван.
4. Гюрджинян Д. С. 2005. Անուն խոսքի մասերի թվի կարգը արդի հայերենում. Քերականական բառարան-տեղեկատու [Категория числа имен в современном армянском. Словарь-справочник]. Ереван.
5. Гюрджинян Д. С., Экекян Н. А. 2007. Հայերենում գործածվող տառային հապավումների բառարան [Словарь. Инициальные аббревиатуры в армянском языке]. Ереван.
6. Добрушина Н. Р. 2005. Как использовать Национальный корпус русского языка в образовании? // Национальный корпус русского языка: 2003 – 2005. Результаты и перспективы. М. – 308-330.
7. Добрушина Н. Р. (ред.) 2007. Национальный корпус русского языка и проблемы гуманитарного образования. Теис.
8. Зализняк А. А. 1977. Грамматический словарь русского языка. Словоизменение. М.
9. Хуршудян В. Г., Подлесская В. И. 2006. Армянское *ban* как дискурсивный маркер речевого сбоя // Армянский гуманитарный вестник № 1. Ереван. – 21-42.
10. Хуршудян В. Г. 2006. Средства выражения хезитации в устном армянском дискурсе в типологической перспективе. Диссертация ... кандидата фил. наук. М.: РГГУ.
11. www.armenianhouse.org – армянская электронная библиотека
12. www.eanc.net - Восточноармянский национальный корпус
13. http://forum.am-kayq.com – армянский форум
14. www.hayeren.hayastan.com – армянский образовательный портал
15. http://www.linguistlist.org/sp/texts.html – Страница ресурса «Linguistlist», посвященная электронным корпусным ресурсам.
16. www.ruscorpora.ru - Национальный корпус русского языка
17. www.sd-editions.com/LALT/home.html – Leiden Armenian Lexical Textbase
18. http://titus.uni-frankfurt.de/indexe.htm – Thesaurus Indogermanischer Text- und Sprachmaterialien

Приложение

Краткая характеристика функциональности ВАНК

ВАНК ориентирован в первую очередь на лексические и грамматические запросы. Синтаксические запросы можно делать лишь опосредовано, так как корпус не имеет синтаксической разметки. По поисковой функциональности ВАНК очень близок Национальному корпусу русского языка (www.ruscorpora.ru), который в значительной степени служил его прототипом. В рамках настоящей приложения поисковая функциональность может быть описана лишь в самых общих чертах, для более подробного ознакомления мы приглашаем читателя зайти на сайт www.eanc.net. Итак, возможны следующие типы поиска.

1. Поиск словоформы или лексемы. ВАНК позволяет искать как вхождения конкретной словоформы (например, *մարդու mardu*), так и вхождения всех словоформ определенной лексемы (например, словоформ *մարդ mard*, *մարդու mardu*, *մարդիկ mardik* и т.д. от лексемы *մարդ mard*).

2. Поиск по переводу. Вхождения лексем можно искать по их английским переводным эквивалентам.

3. Поиск по грамматическим признакам. ВАНК позволяет искать все словоформы, обладающие определенной грамматической характеристикой или набором грамматических характеристик (например, имперфективный конверб в пассиве). Грамматические признаки можно искать как вне зависимости от того, в какой лексеме они встретились, так и вместе с лексемой. При поиске можно учитывать словоизменительный тип словоформы. Грамматический запрос может быть определен как логическая конъюнкция или дизъюнкция нескольких категорий или совмещать конъюнкцию и дизъюнкцию в одной логической формуле.

4. Дополнительные признаки. Кроме этих, собственно лингвистических параметров поиска, можно использовать дополнительные графематические и иные параметры, иногда позволяющие эффективно сузить запрос. Так, можно искать только словоупотребления в начале или в конце предложения, накладывать определенные ограничения на регистр (написание с первой прописной или со всеми прописными), указывать знаки препинания слева и справа от вхождения и т.п.

5. Контекстные запросы. ВАНК позволяет искать сочетания конкретных словоформ, лексем или словоформ, определенных грамматическими признаками. Иными словами, все описанные выше "точечные" запросы на поиск одного слова можно сочетать в контекстные запросы, которые ищут сочетания этих точечных запросов в одном контексте. Расстояние между вхождениями можно изменять, меняя интервал допустимых расстояний. ВАНК позволяет искать вхождения не в одном, а в соседних (и далее) предложениях.

Выбор подкорпуса. Любой запрос, который может быть применен к ВАНК, может быть также применен и к определенному пользователем подкорпусу ВАНК. Окно подкорпуса состоит из следующих зон: авторы и произведения, период, жанр

(с достаточно подробной классификацией жанров и типов текстов), проза/поэзия, оригинальные / переводные тексты, детская / общая литература. Эти признаки можно использовать одновременно, например, ограничивая поиск стихотворными текстами первой половины двадцатого века и т.п.

*Отображение результатов***.** ВАНК позволяет осуществлять сортировку контекстов по целому ряду параметров: лемма (исходная форма), словоформа, словоформа слева от вхождения, автор, название произведения, автор, год создания, жанр. При этом ВАНК поддерживает четыре формата отображения найденной информации.

1. Полный (по умолчанию): каждый контекст сопровождается базовыми библиографическими сведениями (автор, название, год создания).

2. Краткий: библиографические сведения приводятся только в окне расширенного контекста.

3. KWIC (Key Words In Context): принятый в корпусных интернет-ресурсах способ отображения контекстов таким образом, чтобы они были визуально выровнены друг относительно друга по вхождению.

4. Глоссированный: этот формат предназначен в первую очередь для лингвистов-типологов и людей, начинающих изучать армянский язык. Данное отображение текста близко к так называемому морфологическому глоссированию (interlinear morphological glossing), используемому в типологических публикациях и описаниях малых языков, но без разбиения на морфемы и поморфемного перевода. Для всех словоформ в виде столбца, расположенного непосредственно под словоформой, выводится лексико-грамматический анализ, который в других типах выдачи доступен только при наведении мыши. В первой строчке столбца содержатся исходная форма и лексические признаки (например, частеречная характеристика). Во второй строке приводятся грамматические (словоизменительные) признаки словоформы. Если лексеме приписаны переводы, они даются в третьей строчке. Разные разборы одной словоформы визуально отделены друг от друга.

Каждый контекст представлен в окне выдачи одним предложением; слова-вхождения при этом выделены цветом. При контексте приводятся базовые библиографические характеристики: автор, название, год создания, для прессы также номер или дата выпуска. ВАНК позволяет расширить контекст найденного предложения. По умолчанию на экран выводятся три предложения – то предложение, в котором обнаружены искомые вхождения, а также одно предложение до него и одно предложение после него. Расширяя контекст, можно увеличивать размер контекста вплоть до девяти предложений (четыре предложения до и четыре предложения после того предложения, в котором обнаружено вхождение).

Результаты можно отображать как в армянском алфавите, так и в латинской транслитерации. Используемая в ВАНК транслитерация в основном следует международной армяноведческой традиции Хюбшманна-Мейе, адаптированной

под Unicode. Транслитерация используется в том числе при отображении имени автора и названия произведения.

При выдаче результатов запроса в верхней части экрана отображается общая информация об отклике и исходном запросе, в том числе:

1. Число вхождений (в случае больших контекстных запросов – примерная оценка их общего числа в корпусе).
2. Число документов (в случае больших контекстных запросов – примерная оценка числа документов, в которых они могут встретиться).
3. Выбранные критерии сортировки и размер подкорпуса, по которому осуществлялся поиск.

Ниже приводятся некоторые примеры микрозадач, которые могут решаться при помощи корпуса. Соответствующая каждой задаче последовательность действий (что нужно делать, чтобы ее решить) подробно описана на сайте в разделе "Как искать (примеры запросов)".

1. Найти вхождения словоформы *տղու tġu*
2. Найти все вхождения всех словоформ существительного *էրէրդ eġerd*
3. Найти формы императива множественного числа от медиопассивных глаголов
4. Найти формы императива глагола *անձրևել anjrevel*
5. Найти все словоформы лексем, заканчивающихся на *-թյուն -t‘yun* и начинающихся на *ան- an-*
6. Сравнить использование глагола *պրծացնել prcac‘nel* в 19-м веке с его использованием в 20-м веке и сегодня.
7. Сравнить использование глагола *պրծացնել prcac‘nel* в поэзии 19-го века с его использованием в поэзии 20-го века.

Статистика использования словоформ. В готовящемся на настоящий момент к запуску релизе ВАНК добавлена новая функциональность - интерфейс, с помощью которого пользователь может получить не только общую информацию о количестве вхождений словоформы в корпусе, но и подробную картину ее распределения по основным жанрам и декадам. Ниже в качестве примера приведена таблица со статистикой употребления словоформы *աստծո astco* (родительный падеж единственного числа от нерегулярно склоняющегося существительного *աստված astvac* 'бог') по данным корпуса. Кроме абсолютного числа употреблений словоформы, на сайте приводятся еще две характеристики: WPM (число вхождений на миллион), которая позволяет получить представление о частотности словоформы с учетом объема корпуса, а также ее ранг (логарифм отношения частотности самой частотной словоформы к частотности данной словоформы), которая показывает, насколько данная словоформа менее частотна,

чем самая частая словоформа данного сегмента. Для краткости приводится только таблица значений показателя WPM, значения которого легче всего интерпретировать.

Словоформа: *ши ин бп.* Число вхождений: 9,195 Ранг: 386.9 WPM: 548.

	Художественные	**Нехудожественные**	**Пресса**	**Устные**	**Всего по декаде**
(1800-1859)	18	n/a	0	n/a	11
(1860-1869)	95	n/a	0	n/a	61
(1870-1879)	108	40	0	n/a	74
(1880-1889)	90	124	0	n/a	70
(1890-1899)	72	n/a	n/a	n/a	72
(1900-1909)	56	39	0	n/a	28
(1910-1919)	179	n/a	0	n/a	121
(1920-1929)	14	0	3	n/a	10
(1930-1939)	75	8	4	n/a	45
(1940-1949)	77	60	0	n/a	47
(1950-1959)	64	39	8	187	47
(1960-1969)	258	45	14	318	171
(1970-1979)	126	45	6	55	79
(1980-1989)	224	24	11	182	91
(1990-1999)	182	57	91	32	112
(2000-2009)	169	127	100	61	59
недатированные	171	17	388	456	161
Всего по жанру	**151**	**55**	**38**	**68**	

Դավիթ Գյուրջինյան
Երևանի Վ.Բրյուսովի անվան պետական լեզվաբանական համալսարան

ԹՎԻ ԿԱՐԳԻ ՆԵՐԿԱՅԱՑՈՒՄԸ ՀԱՅԵՐԵՆԻ ԿՈՐՊՈՒՍՈՒՄ

Ժամանակակից արևելահայերենի կորպուսում անուն խոսքի մասերից գոյականների և ածականների, մասամբ նաև դերանունների և թվականների մի քանի խմբեր թվի կարգի ներկայացման հետ կապված որոշ դժվարություններ կարող են ունենալ:

Որո՞նք են այդ խմբերը, որո՞նք են այդ դժվարությունները և դրանց հաղթահարման ուղիները,- ահա այս և հարակից այլ հարցերի պատասխանը պիտի փորձենք տալ ստորև:

1. Միավանկ բառեր: Գրական արևելահայերենում միավանկ գոյականները հոգնակի թվի կազմության ժամանակ ստանում են ***-եր*** վերջավորությունը, ընդ որում արմատական ***ի*** և ***ու*** ձայնավորները շեշտազրկվելով սովորաբար հնչյունափոխվում են ***ը***-ի, ինչպես՝ ***քիծ - բծեր, լինդ - լնդեր, գունդ - գնդեր, քուրդ - բրդեր, թումբ - թմբեր:*** Բայց ահա մի շարք միավանկ բառերի հոգնակին կազմելիս ***ի*** և ***ու*** ձայնավորների հնչյունափոխություն տեղի չի ունենում, ինչպես՝ ***կիրճ - կիրճեր, ծիլ - ծիլեր, քուրք - քուրքեր, պուպ - պուպեր, ռումբ - ռումբեր*** և այլն: Այստեղ որևէ խնդիր կարծես թե չունենք:

Փոխառյալ բառերի ***ի*** և ***ու*** ձայնավորները որպես կանոն չեն հնչյունափոխվում՝ ***բիմ - բիմեր, պիկ*** «սրածայր լեռնագագաթ» - ***պիկեր,*** խսկց. ***փիր*** «հրաձգարան» ***- փիրեր, բինտ*** «վիրակապ» ***- բինտեր:*** Սրանց հնչյունափոխությունը բարբառային է և մերժելի, ինչպես՝ ***բ[ը]ինտեր, փ[ը]րեր:*** Այդպես էլ՝ ***փիկ - փկեր*** (փոխանակ՝ ***փիկեր)***, ***լինգ - լնգեր*** (փոխանակ՝ ***լինգեր)*** և այլն:

Որոշ դեպքերում միավանկ բառերի հոգնակիի զուգաձևություն կա. նրանք հանդես են գալիս հնչյունափոխված և անհնչյունափոխ տարբերակներով: Այսպես՝ ***քուրմ - քրմեր // քուրմեր, սուրբ - սրբեր // սուրբեր, խուց - խցեր // խուցեր, ճիպք - ճպքեր // ճիպքեր, ցլեր // ցուլեր*** և այլն:

Նախընտրելի են առաջին ձևերը, սակայն ժամանակակից արևելահայերենի ներկա փուլում որոշ բառերի դեպքում չհնչյունափոխելու միտում է նկատվում, որը հաճախ արտալեզվական գործոններով է բացատրվում: Օրինակ՝ ***սուրբեր*** ձևը կազմելու մեջ ոչ միայն արևմտահայերենի ազդեցությունը կա, այլև այս ձևը նախընտրողներին թվում է, թե իրենք այդպես «չեն աղավաղում» բառը, նրա արտաքին ձևավորումը:

Այլ դեպքերում դեր է խաղում ոմանց ռուսական կրթությունը. նրանք ևս փորձում են «անխաթար պահել» բառը (***ու - ը*** հնչյունափոխությունից խուսափելու պատճառը թերևս այն է, որ ռուսերենն ***ը*** հնչյուն չունի, որն էլ դժվարացնում է արտասանությունը):

Կարելի է այսպես ասել. մի կողմից արևմտահայերենի և այլ գործոնների ազդեցությամբ արմատական ձայնավորը չհնչյունափոխելու միտում կա, իսկ մյուս դեպքում խոսակցական լեզվի և որոշ բարբառներ ազդեցությամբ հնչյունափոխվում են անգամ փոխառյալ բառերը:

Մեր կարծիքով՝ ձևաբանական նորմով անթույլատրելի ձևերը արևելահայերենի կորպուսում պիտի արձանագրել համապատասխան նշումով, միաժամանակ հղում տալ կանոնական ձևին: Անշուշտ, այս կամ այն ձևի գործածությունն իրողություն է, բայց կորպուսից օգտվողին թերևս կարելի է հուշել, օրինակ, որ հանդիպած թե՛ ***քուրմեր***, թե՛ ***քրմեր*** ձևերը նորմատիվ են, իսկ ***քեռեր, գառեր*** ձևերը՝ ոչ:

Մեզ համար հստակ չէ, թե հայերենի կորպուսում առհասարակ զուգաձևությունների խնդիրն ինչպես է լուծվելու: Մեր կարծիքով՝ սոսկ արձանագրելը դեռևս մեկ քայլ է:

Բառավերջում ***ու*** ունեցող բառերից զուգաձևություն ունի ***քու***-ն՝ ***քվեր*** և ***քուեր***:

Այս խմբի բառերի մեջ առանձնանում է ***ալիչ***-ը: Սրա հոգնակին լինում է ***ալիչեր***, բայց ավելի շատ գործածվում է գրաբարյան քարացած ձևը՝ ***ալիչիք:***

Մի շարք միավանկ բառերի հոգնակին կազմելիս վերականգնվում է գրաբարում նրանց ունեցած ***և*** վերջնահնչյունը ***(քեռ, գառ, դուռ, եզ, բոռ, լեռ, ծոռ, ծունկ, հարս, ձուկ, մատ, մուկ, նուռ):***

Նույն ձևով է կազմվում բարբառային ***կուռ*** բառի հոգնակին՝ ***կռևեր:*** Ժողովրդական լեզվին բնորոշ ***ձեռ*** և ***ոտ*** բառերի հոգնակին լինում է ***ձեռներ, ոտներ:*** Գրական լեզվի համար վերջիններս ընդունելի չեն:

Ոդներ և ***ոտնգներ*** ձևերն այժմ անգործածական են. նրանց փոխարեն հանդես են գալիս ***ոդեր*** և ***ոտնգեր // ոնգեր*** ձևերը:

Երբ բառի վերջնաբաղադրիչը գրաբարում ***և*** վերջնահնչյուն ունեցած միավանկ բառ է, հոգնակին կրկին ***-ևեր*** վերջավորությամբ է կազմվում, ինչպես՝ ***շնաձուկ - շնաձկևեր, ուղեբեռ - ուղեբեռևեր, քաղաքաղուռ - քաղաքաղռևեր, դաշտամուկ - դաշտամկևեր:*** Կան բացառություններ՝ ***ջրահարս - ջրահարսեր, հավերժահարս - հավերժահարսեր:*** Իհարկե, լինում են նաև ***-ևեր***-ով կազմելու դեպքեր ***(ջրահարսևեր, հավերժահարսևեր.*** հմմտ. ***հարսևեր)***, սակայն դրանք տարածված ձևեր չեն:

Այս տեսակ որոշ բառերում զուգաձևություն է առկա, ինչպես՝ ***ցուցամատ - ցուցամատեր // ցուցամատևեր:***

Միավանկ գոյականներից միակը ***ռուս*** բառն է, որն ստանում է ***-ևեր*** վերջավորությունը՝ ***ռուսևեր (ռսևեր*** ձևը խոսակցական է և մերժելի):

2. Միավանկ վերջնաբաղադրիչով բաղադրյալ բառեր: Միավանկ վերջնաբաղադրիչով գոյականների հոգնակիի կազմության ժամանակ հանդես է գալիս ***-եր*** մասնիկը, երբ այդ բաղադրիչը գոյական է և բաղադրության մեջ պահպանում է իր հիմնական՝ առարկայական նշանակությունը, ինչպես՝ ***ցուցափեղկ – ցուցափեղկեր, արորավայր - արորավայրեր, զորամաս - զորամասեր:***

Հորաքույրեր, մորաքույրեր ձևերի փոխարեն խոսակցական լեզվում գործածում են ***–ևեր***-ով կազմվածները, որոնք կանոնական չեն: Հմմտ. ***բուժքույրեր, արբայաքույրեր:***

Երբ միավանկ վերջնաբաղադրիչը բայարմատ է և հանդես է գալիս բայական նշանակությամբ (գործողություն կատարողի, տվյալ վիճակի մեջ եղողի), հոգնակին կազմվում է ***-ևեր*** մասնիկով: Օրինակ՝ ***փայտահատ - փայտահատևեր, նորածին - նորածիևեր***, ***ժամացույց - ժամացույցևեր*** և այլն:

Երբեմն միավանկ վերջնաբաղադրիչը բայի իմաստ ստացած գոյական է լինում և հոգնակին կազմելիս կրկին ***–ներ*** վերջավորություն է ստանում: Օրինակ՝ ***մեծապուճ*** *(=մեծ պուճ ունեցող, այսինքն՝ հարուստ, ունևոր)* – ***մեծապուճներ***:

Այս բառերը կա՛մ գոյական, կա՛մ ածական են (և կամ նույն բառը համատեղում է երկու արժեքներն էլ), ինչպես՝ ***ականապես - ականապեսներ, բանապատի - բանապատիներ, գնդացիր - գնդացիրներ, դավադիր - դավադիրներ, պատմագիր - պատմագիրներ, հետուպատգույց - հետուպատգույցներ, չարամիտ - չարամիտներ, ավանդապատի - ավանդապատիներ***:

Գոյական միավանկ վերջնաբաղադրիչով ածականների հոգնակին առհասարակ կազմվում է ***-ներ*** վերջավորությամբ, ինչպես՝ ***չարասիրտ, բարդրաքիթ (մարդիկ) - չարասիրտներ(ը), բարդրաքիթներ(ը), զրահապատ (մեքենաներ) - զրահապատներ(ը)***:

Մի շարք դեպքերում միևնույն վերջնաբաղադրիչն է, բայց հոգնակին կազմվում է տարբեր վերջավորություններով, քանի որ այդ բաղադրիչը մերթ գոյականական, մերթ էլ բայական նշանակություն են արտահայտում: Օրինակ՝ ***դեղաչափ*** «դեղի այն չափը, որ նախատեսված է միանգամից ընդունելու համար», «տվյալ բաղադրության մեջ մտնող յուրաքանչյուր նյութի չափը» ***- դեղաչափեր, հողաչափ*** «հողաբաժանության համար օգտագործելի հողերը չափող և սահմանագծող մասնագետ» ***- հողաչափներ:*** Այդպես էլ՝ ***ասպատգույց - ասպատգույցներ, կողմնացույց - կողմնացույցներ, խնդրագիր - խնդրագրեր, հեքիաթագիր - հեքիաթագիրներ, առանձնապուճ - առանձնապուճներ, մեծապուճ - մեծապուճներ*** և այլն:

Երբ միավանկ վերջնաբաղադրիչի իմաստը հստակ չի գիտակցվում կամ այլ իմաստով է գործածվում (և կամ սերտաճած է լինում նախորդ բաղադրիչին), բառը սովորաբար ստանում է ***-ներ*** վերջավորությունը: Օրինակ՝ ***քղանցք - քղանցքներ, ընձուղտ - ընձուղտներ, տեսադեպ - տեսադեպեր, տաքդեղ - տաքդեղներ*** և այլն: Այդպես էլ՝ ***դիպանցք - դիպանցքեր, հորապանցք - հորապանցքեր,*** բայց՝ ***լեռնանցք - լեռնանցքներ, միջանցք - միջանցքներ, նրբանցք - նրբանցքներ:***

Որոշակի է, որ մի շարք բառերի (ավելի ճիշտ՝ վերջնաբաղադրիչների) պարագայում հստակ կանոնակարգում չունենք, և բառարանները հոգնակիի տարբեր ձևեր են ներկայացնում: Օրինակ՝ ***լեռնանցք*** բառի համար մի դեպքում տրվում է ***լեռնանցքեր*** ձևը, մյուս դեպքում՝ ***լեռնանցքներ:***

Սակայն որոշ իրողություններ կարելի է հստակ արձանագրել. այսօր առավելաբար ***–ներ*** վերջավորությամբ են կազմում ***-ահիշ*** և ***-պետ*** վերջնաբաղադրիչներով բաղադրությունների հոգնակին, ինչպես՝ ***ախտանիշներ, որոշմանիշներ, խորհրդանիշներ, հայտանիշներ, հատկանիշներ, չափանիշներ, ցուցանիշներ, բանապետներ, գնդապետներ, հազարապետներ, համայնքապետներ, մարզպետներ, միապետներ, նահանգապետներ, վարչապետներ, քաղաքապետներ:*** Իհարկե, ***գյուղապետեր, պահեստապետեր*** և նման ձևերը միանգամայն ճիշտ են, սակայն ցանկալի է նույնական մոտեցում կիրառելը: Առհասարակ այս կարգի բառերի հոգնակին ***–ներ*** մասնիկով կազմելու միտում կա:

Ժամանակակից հայերենի կորպուսում, բնականաբար, արձանագրվելու են զույգ ձևերն էլ, սակայն առաջնայնության, վերաբերմունք դրսևորելու խնդիր, այնուամենայնիվ, թվում է, թե կա:

3. Մեկուկեսվանկանի բառեր: Այն երկվանկ բառերը, որոնք վերջանում են գաղտնավանկով, հոգնակին կազմելիս ստանում են ***-եր*** մասնիկը, ինչպես՝ ***կայ/ ր /ր*** – ***կայսրեր:*** Սրանք այսպես կոչված մեկուկեսվանկանի բառերն են (գոյական և ածական)՝ ***դուստր - դուստրեր, արկղ - արկղեր, ասպղ - ասպղեր, մանր – մանրեր, կարծր – կարծրեր, քաղցր – քաղցրեր:***

Երբ բառի վերջնաբաղադրիչը այս կարգի բառ է, հոգնակին կրկին կազմվում է ***-եր*** վերջավորությամբ, ինչպես՝ ***փոմսարկղ - փոմսարկղեր, դրամարկղ - դրամարկղեր, գիսասպղ - գիսասպղեր, արբայադուստր - արբայադուստրեր, գիշանգղ - գիշանգղեր:***

Հայերենագետներից ոմանք ճիշտ են համարում ***փոմսարկղևեր, դրամարկղևեր*** և նման ձևերը:

Մեկուկեսվանկանի որոշ բառեր հատկապես խոսակցական լեզվում հանդես են բերում զուգաձևություններ: Հոգնակիի զուգաձևություններ նկատվում են նաև դասական գրողների լեզվում, տարբեր բառարաններում, ինչպես՝ ***անգղեր // անգղևեր, կայսրեր // կայսրևեր:***

Ի դեպ, ***–ևեր*** մասնիկով որոշ ձևեր ընդհանրանալու միտում ունեն: Կարելի է ասել, որ, օրինակ, ***գամփռևեր*** ձևը կայունացել է:

Մկ, ասպ, սպ, սթ, սք, սփ, զբ, զգ, շպ բաղաձայններով սկսվող և դրանցից առաջ թույլ ***ը*** ունեցող երկվանկ բառերը ևս ստանում են ***–ևեր*** վերջավորությունը, ինչպես՝ ***սրինք -սրինքևեր, սպա - սպաևեր*** և այլն: Այդպես էլ որոշ փոխառյալ բառերի դեպքում՝ ***բլոկ - բլոկևեր, շրաբ - շրաբևեր:*** Մամուլի և հեռուստատեսության լեզվում հանդիպում են ***բլոկեր, շրաբեր*** ձևերը (խոսակցականում՝ ***սպաեր***), որոնք կարելի է որակել իբրև սխալ:

Դպրոցական ուսուցման մեջ գրեթե կայունացած այս ձևերը կորպուսում առանց ***անհանգ.*** կամ նման այլ նշումի արձանագրելով, մեր կարծիքով, վնաս կհասցնենք հայերեն խոսքի մշակույթին:

4. Բազմիմաստ բառեր, որոնց հոգնակին տարբեր ձևերով է կազմվում: Երբեմն միևնույն բառն է այս կամ այն իմաստի թելադրանքով տարբեր վերջավորություններ ստանում: Օրինակ՝ ***պարսավագիր*** բառը «հրապարակախոսական երկ, որ բնութագրվում է որևէ անձի, հասարակական երևույթի սուր ծաղրով. պամֆլետ» (այսինքն՝ *պարսավական գիր՝ գրություն*) իմաստով գործածվելիս հոգնակի թվում ստանում է ***-եր*** վերջավորությունը, իսկ «պամֆլետի հեղինակ» և «մեկի դեմ պարսավական բան գրող անձ» (այսինքն՝ *պարսավագիր գրող*) իմաստներով գործածվելու ժամանակ՝ ***-ևեր՝ պարսավագիրևեր:***

Կամ էլ՝ ***արտագիր*** «սրտի աշխատանքը գրառող սարք» - ***արտագիրևեր*** և ***արտագիր*** «սրտի աշխատանքի գրառումը՝ պատկերը՝ գծագիրը» - ***արտագրեր, ձախշազարդ*** «ձախշերով արված զարդ» (գոյական է) - ***ձախշազարդեր*** և ***ձախշազարդ*** «ձախշերով զարդարված» (ածական է) - ***ձախշազարդևեր: Ջրանցք*** բառի հոգնակին լինում է ***ջրանցքևեր***, երբ գործածվում է իր հիմնական՝ «արհեստական մեծ հուն ջրի համար» նշանակությամբ, խիստ հազվադեպ, մասնավորապես տեխնիկական գրականության մեջ, կարող է հանդիպել ***ջրանցքեր*** ձևը, երբ գործածվում է երկրորդ՝ «ջրի անցք՝ ծակ» նշանակությամբ:

Մի շարք անհոգնական բառեր հոգնակի թիվ ունենում են միայն այս կամ այն բառային նշանակությամբ գործածվելիս: Այսպես, ***առաքելություն*** բառն անհոգնական է, բայց իբրև «ներկայացուցչություն» նշանակությամբ

դիվանագիտական տերմին կարող է գործածվել հոգնակի թվով: ***Ոսկի*** բառն իբրև քիմիական տարրի անվանում անհոգնական է, բայց «ոսկեդրամ» կամ «ոսկյա զարդերի ընդհանուր անվանում» նշանակությամբ կարող է հոգնակի թվով գործածվել:

Արդյոք պիտի բավարարվե՞լ միայն ***հոգն.*** նշումով, թե՞ տվյալ բառագործածության համար համապատասխան իմաստը ևս նշել:

5. Համանուն բառեր, որոնց հոգնակին տարբեր ձևերով է կազմվում: Եզակի դեպքերում տարբեր մասնիկներ են ստանում համանուն բառերը, ինչպես՝ ***մատ*** «մարդու ձեռքի և ոտքի թաթի հինգ վերջավորություններից յուրաքանչյուրը» **-** ***մատներ, մատ*** «խաղողի վազի ոստ» **-** ***մատեր:*** Կամ՝ ***կոշկակար*** «կոշիկ կարող արհեստավոր» **-** ***կոշկակարներ, կոշկակար*** «կոշիկի կար» **-** ***կոշկակարեր, գիշերապահ*** «գիշերային պահակ՝ պահապան» - ***գիշերապահներ, գիշերապահ*** «գիշերվա պահ, գիշեր ժամանակ» - ***գիշերապահեր:*** Այս պարագայում բառիմաստների ներկայացումը կարևոր է:

6. Ժողովրդախոսակցական և հնացած ձևեր: ***-Ավցի, -եցի, -ցի*** վերջածանցներով գոյականների հոգնակին երբեմն կազմվում է ***-ք*** մասնիկով, ինչպես՝ ***ախալքալաքցի - ախալքալաքցիք, ֆրանսիացի - ֆրանսիացիք, լոռեցի - լոռեցիք, փնեցի - փնեցիք:*** Սրանք բնորոշ են ժողովրդախոսակցական լեզվին և արդի գրական լեզվում հանձնարարելի չեն (չնայած որ հանդիպում են նաև գեղարվեստական գրականության և մամուլի լեզվում):

–Ք մասնիկով հոգնակիի որոշ ձևեր բարբառային են, ինչպես՝ ***ճամփա - ճամփեք, փեսա - փեսեք, սապաւա - սապաւեք:***

Գրական լեզվում մերժելի են նաև ***-երք*** մասնիկով ***գղերք, աղջկերք*** ձևերը: Սակայն հիշյալ ձևերը գեղարվեստական գրականության մեջ կարող են գործածվել ժողովրդայնության երանգով:

-Ք-ով որոշ ձևեր էլ հնացած են, ինչպես՝ ***որդի - որդիք, փղա - փղայք, ևախևի - ևախևիք, եղբայր - եղբայրք:*** Գեղարվեստական գրականության լեզվում սրանք գործածվում են իբրև քերականական հնաբանություններ:

Վերոհիշյալ բոլոր բառերի հոգնակին արդի գրական արևելահայերենում որպես կանոն կազմվում է ***-ներ*** մասնիկով: Արդ, տարբերակված մոտեցում արդյոք անհրաժե՞շտ է: Եթե կորպուսն ընդգրկում է Րաֆֆի, Պռոշյան և ուրիշներ, հնացած, ժողովրդական ձևերի առկայությունն անխուսափելի է, և դա էլ իր հետ մի շարք հարցեր է բերում: Հետևել ակադեմիական և այլ բառարաններին և կիրառե՞լ ***հնց.***, ***ժող.*** և այլ նշումներ:

-Ք մասնիկով է կազմվում ***այլ*** անորոշ դերանվան հոգնակին՝ ***այլք:*** Շատ հաճախ կազմում են միանգամայն սխալ ***այլոք*** ձևը (***այլք***-ի սեռականի ***այլոց*** ճիշտ ձևի ազդեցությամբ), որի գրավոր տասնյակ գործածություններ կարող ենք բերել: Այս ձևը եթե անգամ սպրդի հայերենի կորպուսի մեջ, ապա ***հոգն.*** նշումի կողքին անշուշտ պիտի ավելացվի ***սխալ***:

Մի շարք դերանուններ ունեն հոգնակի ուղղականի զուգաձևություններ՝ ***-ներ*** և ***-ք*** մասնիկներով: Այսպես, ***այդպիսի - այդպիսիները // այդպիսիք, այնպիսի - այնպիսիները // այնպիսիք, նույնպիսի - նույնպիսիները // նույնպիսիք, ինչպիսի - ինչպիսիները // ինչպիսիք, որպիսի - որպիսիները // որպիսիք,*** ընդ որում ***-ք*** մասնիկով ձևերը չեն հոլովվում: Այս զուգաձևությունների համար փոխադարձ

հղումներ են հարկավոր, որովհետև գործածություններն են տարբեր (ըստ տեքստի ոճի և նախադասությունների կառուցվածքի)։

7. Անհոգնական բառեր։ Արդի հայերենում ոչ հաշվարկելի գոյականները սովորաբար միայն եզակի թվով են գործածվում։ Դրանց թվին են պատկանում՝բնության մեջ մեկ՝ միակ առարկաների անունները ***(լուսին, արևելք),*** վերացական գոյականները ***(զղջում, բերկրանք, խաղաղություն),*** հավաքական գոյականները (մասնավորապես ***-եղեն***, ***-ություն*** վերջածանցներով կազմվածները՝ ***տեսականի, սպիտակեղեն, երիտասարդություն),*** գիտությունների, ուսմունքների, տնտեսաճների, արվեստի ուղղությունների անունները ***(հանրահաշիվ, քրիստոնեություն, պլատոնականություն), -ություն*** վերջածանցով այն բառերը, որոնք արհեստ, մասնագիտություն, տնտեսության ճյուղ են նշանակում ***(դարբնություն, հողագործություն, գինեգործություն, հանքարդյունաբերություն),*** տարբեր խաղերի, մարզաձևերի անունները ***(շախմատ, լող, վազք),*** նյութերի անունները, քիմիական տարրերի անունները, ինչպես՝ ***(օդ, արճիճ, պողպատ, ազոտ),*** լեզուների անունները ***(արաբերեն, հայերեն, հունարեն)***։ Վերջիններս հոգնակի թվով կարող են գործածվել միայն փոխաբերական կիրառության դեպքում։

Համոզված ենք, որ համապատասխան բառաճնի համար ոճական նշում պիտի արվի, այլապես հատկապես հայերենի ոչ քաջածանոթ մեկը սխալ եզրակացություն պիտի անի։

-Ություն և ***-ում*** ածանցներով կազմված բայանուն գոյականները, որոնք եթե բանձրացական նշանակություն են ունենում կամ ցույց են տալիս տվյալ հատկանիշի մեկից ավելի դրսևորումը, գործածվում են հոգնակի թվով, ինչպես՝ ***հուզում - հուզումներ, կամայականություն - կամայականություններ, ցանկություն - ցանկություններ, հիշողություն - հիշողություններ, գիտություն - գիտություններ*** և այլն։

Սխալ են ***ամանեղեններ, խմորեղեններ, դեղորայքներ, տեսականիներ*** և նման ձևերը։

8. Անեզական կամ առավելաբար հոգնակիով գործածվող բառեր։ Հասարակ անուններից կարելի է նշել միայն մեկ-երկուսը, որոնք կա՛մ միայն հոգնակի թվով են գործածվում, կա՛մ էլ հիմնականում հոգնակիով են հանդիպում (չնայած որ եզակին ունեն)։ Այն պատճառով, որ բազմակիության կամ երկակիության իմաստ են արտահայտում։ Օրինակ՝ ***ձոթեր, կուլիսներ, բարթիչներ, խոպոպներ, խռիկներ:***

Գանգուրներ գոյականը եզակիով չի գործածվում. առանց հոգնակիակերտ մասնիկի ձևն ածական է։

Արդի արևելահայերեն գիտական խոսքում, գերազանցապես հոգնակի թվով են գործածվում բույսերի և կենդանիների տիպերի, դասերի, կարգերի, տեսակների, որոշ հանքատեսակների, քիմիական նյութերի խմբերի անվանումները, ինչպես՝ ***շուշանազգիներ, ամբակավորներ, բառափաննմաններ, սպալակտիդներ, ոիբոսոմներ:***

Նկատելի է, որ նմանատիպ բազմաթիվ բառեր գործածվում են առանց հոգնակիակերտ վերջավորությունների, ինչպես՝ ***երկկենցաղ - երկկենցաղներ, պարկավոր - պարկավորներ, վարդազգի - վարդազգիներ:*** Այդ ձևերը, սակայն, գոյականներ չեն, այլ ածականներ, որոնք գոյականների են վերածվում ***-ներ***

մասնիկի միջոցով: -***Ներ***-ը փաստորեն բառակազմական-քերականական մասնիկ է: Բառարաններն էլ այդ բառերը տալիս են հոգնակիի ձևով:

Այդպես է նաև որոշ ազգերի (հատկապես անհետացած) անունների պարագայում, օրինակ՝ ***արմեններ, մարեր, հոներ, խեթեր, գալլեր, գոթեր*** և այլն:

Որոշ բառեր գրեթե միշտ հոգնակիի ձևով են հանդես գալիս այս կամ այն իմաստի ձեռքբերման և դրա հետևանքով խոսքիմասային փոխանցման դեպքում: Օրինակ՝ ***կանաչներ*** «բնապահպաններ» (***կանաչ*** ածականից), ***կարմիրներ*** «սոցիալիստական հեղափոխության ուժեր», ***աջեր, ձախեր:*** Ակնհայտ է, որ խոսքիմասային անցում է տեղի ունեցել, և ածականի հոգնակին ձեռք է բերել գոյականի արժեք: Եթե նշում ենք ***գոյական, հոգնակի թիվ***, ստացվում է, որ եզակին ևս գոյական է, այնինչ իրականում այդպես չէ: Այս դեպքում ճիշտ կլինի նշել ***անեզական գոյական***, բայց իմաստի և ձևաբանական արժեքի արձանագրումը կարևոր է: Մյուս ելքն այն է, որ որակական ածականների կողքին նշվեն հոգնակիի հնարավոր ձևերը (չնայած որ այդ դեպքում բառային արժեք ստացածների հարցում կորուս կունենանք);

-Վնք, -ենք, -ոնք, -ունք վերջածանցներով բառերը, ինչպես՝ ***պապոնք, պապոնք*** հավաքականության իմաստ են արտահայտում և անեզական են: Այս ածանցներով որոշ ձևեր ժողովրդախոսակցական կամ բարբառային են, ինչպես՝ ***խնամոնք, քեռոնք, աներաոնք*** կամ էլ, ասենք, ***դպրավորենք:*** Կրկին նույն հարցադրումն է՝ բարբառային կամ այլ շերտերի նկատմամբ որևէ վերաբերմունք պիտի դրսևորվի՞, թե՞ ոչ: Օրակարգային է նաև հատուկ անուններից կազմվածների հարցը. ***Վարդանանք*** բառի կողքին ունենք ***Վարդանենք, Գրիգորենք*** և այլն:

9. Ոճական հոգնակի: Մի շարք դեպքերում անհոգնական գոյականները գործածվում են հոգնակի թվով: Դրանք ոճական և իմաստային այլևայլ դրսևորումներ ունեն:

Նյութերի անունները սովորաբար հոգնակի թիվ չեն ունենում: Այս բառերը հոգնակի թվով գործածվում են այն ժամանակ, երբ ցույց են տալիս դրանց տարբեր տեսակները կամ միևնույն տեսակը՝ հաշվարկելի ձևով: Օրինակ՝ ***գինիներ*** *(հայկական, ֆրանսիական, քաղցր, կիսաքաղցր, տրփիպ):* ***Ջրեր*** հոգնակին կարող է նշանակել թե՛ ջրի զանազան տեսակներ և թե՛ ջրի մեծ քանակություն:

Երևույթների անունների հոգնակին ցույց է տալիս տվյալ երևույթի հաճախակի կրկնությունը, պարբերականությունը՝ ***անձրևներ, շոգեր, ցրտեր:***

Հոգնակի թիվ չունեն շաբաթվա օրերի (շաբաթվա կտրվածքով), ամիսների, տարվա եղանակների անունները (մեկ տարվա շրջանակներում), ինչպես՝ ***հինգշաբթի, սեպտեմբեր, գարուն:*** Տևական ժամանակի ընթացքում, բնականաբար, թվարկված իրողությունները կրկնվում են, և մասնավորապես շաբաթվա օրերի և տարվա եղանակների անունները կարող են գործածվել հոգնակի թվով:

Ոճական է նաև թվականների գործածությունը հոգնակիով: Քերականական այդ ձևերը ցույց են տալիս՝ ա) թվանշանի անունը, բ) տվյալ թիվն ունեցող առարկա, գ) առարկաներ, որոնք տվյալ թվային կարգն ունեն: Եվ կարող ենք բազմաթիվ վկայություններ բերել, որտեղ լինեն ***երեքներ, հարյուրներ, հազարներ, միլիոններ, չորորդները*** և նման այլ ձևեր: Հարց է, այս դեպքում որևէ նշում պիտի արվի՞, թե՞ ոչ:

Անուն խոսքի մասերի թվի կարգի հետ կապված խնդիրներն իրականում շատ ավելին են, եթե քննենք մասնավոր դեպքերը: Սակայն արծարծված հարցերը սկզբունքային են, անհետաձգելի և ամենակարևորը՝ լուծելի:

ԳՐԱԿԱՆՈՒԹՅՈՒՆ

1. Աբրահամյան Ս.Գ. և ուրիշներ, Ժամանակակից հայոց լեզու, հ. 2, Երևան, 1974:
2. Աղաջանյան Զ.Խ., Չևաբանական նորմ և խոսքի մշակույթի հարցեր, Երևան, 2007:
3. Աճառյան Հր., Լիակատար քերականություն հայոց լեզվի, հ.3, Երևան, 1957:
4. Ասատրյան Մ.Ե., Ժամանակակից հայոց լեզվի ձևաբանության հարցեր, հ. Ա, Երևան,
5. 1970, հ. Բ, Երևան, 1973:
6. Բեղիրյան Պ.Ս., Հայ լեզուն և մեր խոսքը, Երևան, 2007:
7. Գյուրջինյան Դ.Ս., Անուն խոսքի մասերի թվի կարգը արդի հայերենում. քերականական
8. բառարան-տեղեկատու, Երևան, 2005:
9. Խլղաթյան Ֆ.Հ., Ժամանակակակից հայոց լեզու, պր. Գ, Երևան, 2006:
10. Մարգարյան Ա.Ս., Հայոց լեզվի քերականություն. ձևաբանություն, Երևան, 2004:
11. Սկրտչյան Ռ.Խ., Ժամանակակից հայերենի ձևաբանական ոճաբանություն, Երևան,
12. 1992:
13. Պետրոսյան Հ.Զ., Գոյականի թվի կարգը հայերենում, Երևան, 1972:
14. Ջահուկյան Գ.Բ., Ժամանակակից հայերենի տեսության հիմունքները, Երևան,1974:
15. Ջահուկյան Գ.Բ., Խլղաթյան Ֆ.Հ., Հայոց լեզու 9, Երևան, 1998:
16. www. eanc.net

Գայանե Գևորգյան

ԲԱՐԲԱՌԱՅԻՆ ՆՅՈՒԹԻ ՄՇԱԿՄԱՆ ԵՎ ՀԱՄԱԿԱՐԳՉԱՅՆԱՑՄԱՆ ՍԿԶԲՈՒՆՔՆԵՐՆ ՈՒ ԱՌԱՆՁՆԱՀԱՏԿՈՒԹՅՈՒՆՆԵՐԸ

Լեզվի պատմության համար խիստ կարևոր է բարբառների փոխառնչությունների գիտական քննությունը: Թեև հայերենի բարբառների քննության և ուսումնասիրության փորձերը սկսվել են անցյալ դարի կեսերից, և այդ ժամանակից սկսած՝ դրանք ենթարկվել են միահատկանիշ և բազմահատկանիշ դասակարգումների, սակայն բարբառային միավորների միջև գոյություն ունեցող բարդ փոխհարաբերությունների պարզաբանման, հստակեցման և ճշգրտման խնդիրներն առ այսօր մնում են չլուծված:

Բարբառագիտական հարցերի համակողմանի քննությունը հնարավոր է միայն բարբառագիտական ատլասի առկայության պայմաններում, որի ստեղծումը ճշանակալից երևույթ է ցանկացած լեզվի ուսումնասիրության պատմության մեջ: Վաղուց կազմված են շատ լեզուների բարբառագիտական ատլասներ: Հայերենի քննությունը այդ բնագավառում բավականին հետ է մնացել,

Հայերենի բարբառագիտական ատլասի /այսուհետև՝ ՀԲԱ/ կազմման գործընթացը հայ իրականության մեջ սկսվել է 1975 թվականից:

Քանի որ ատլասի ստեղծման համար անհրաժեշտ է կատարվելիք աշխատանքների մի քանի փուլ, դրանք սկսել են իրականացվել փուլ առ փուլ՝ հարկ եղած դեպքում նաև կիրառելով դրանց միաժամանակյա կատարումը:

Առաջին փուլի հիմնական խնդիրը հնչյունական, բառային ու քերականական այն հիմնական հատկանիշների ընտրությունն էր, որոնցով հետագայում պիտի որոշվեին հայերենի բարբառների՝ միմյանց նկատմամբ ունեցած ընդհանրություններն ու տարբերությունները:

Բարբառային տարբեր տարածքներին հատուկ առավել բնորոշ զուգաբանությունների /իզոգլոս/ ընտրությունը հնարավորություն տվեց 1977 թվականին հրատարակել «Հայերենի բարբառագիտական ատլասի նյութերի հավաքման ծրագիր» /այսուհետև՝ Ծրագիր/ խիստ արժեքավոր աշխատանքը, որը հիմնականում կազմել է հայոց լեզվի պատմության և բարբառագիտության երախտավոր, հանգուցյալ Հ. Մուրադյանը: «Քերականական զուգաբանություններ» բաժնի կազմությանը նրան օգնել են բանասիրական գիտությունների թեկնածուներ Ա. Վ. Գրիգորյանը, Դ. Մ. Կոստանդյանը և Ա. Ն. Հանեյանը:

Ծրագրի հրատարակումից հետո մշակվել են բարբառային նյութի հավաքման և գրանցման որոշակի սկզբունքներ:

Լեզվաբանական աշխարհագրության մեջ ընդունված է այն տեսակետը, որ բարբառային երևույթները կենտրոնից դեպի ծայրամասերը շարժվում են հավասարաչափ. «If we travel from village to village in a particular direction, we notice linguistic differences which distinguish one village from another. Sometimes these differences will be larger, sometimes smaller, but they will be cumulative. The further we get from our starting point, the larger the differences will become”.[1]

1 J. K. Chambers, P. Trudgill, Dialectology, Cambridge, 1980, p. 6:

Առաջնորդվելով այս դրույթով՝ սովորաբար ատլասի համար ընտրվում է բնակավայրերի մի այնպիսի ամբողջություն, որի վրա նշված կետերը միմյանցից գտնվում են համաչափ հեռավորության վրա: Նման պարագայում անհրաժեշտ չէ լրացնել բոլոր բնակավայրերի խոսվածքները: Սակայն հիշյալ սկզբունքը ընդունելի է ոչ բոլոր լեզուների համար: Այն կարող է կիրառվել միայն այնպիսի բարբառների դեպքում, որոնց կրողները տեղաշարժերի չեն ենթարկվել:

Ատլասի բնակավայրերի ընտրության վերոհիշյալ մոտեցումը ընդունելի չէ հայերենի բարբառների համար, քանի որ հայ ժողովրդի բազմաթիվ գաղթերն ու տեղահանությունները փոխել են հայերենի բարբառների տեղադրության նախնական պատկերը. «Հայերենի ատլասի համար ամենաճիշտ սկզբունքը՝ անխտիր բոլոր հայախոս բնակավայրերը հարցման ենթարկելն է, անկախ այն բանից, թե այդ բնակավայրերը կմտնեն քարտեզների մեջ, թե ոչ: Այդ ձևով նյութերի հավաքումը մի կողմից հնարավորություն կտա հայտնաբերելու իրենց բնօրրանից կտրված ու այլուր տեղափոխված բարբառային միավորները, որոնք քարտեզագրման ժամանակ կտեղադրվեն իրենց պատմական միջավայրում, մյուս կողմից՝ համատարած հարցումից հետո դժվար չի լինի քարտեզների համար ընտրել ավելի տիպական բնակավայրերը և դրանով իսկ ապահովել բնակավայրերի ցանցի անհրաժեշտ խտությունը»:[2]

Նկատի ունենալով վերոհիշյալը՝ հարցման ենթակա ելակետային միավորը համարվել է ոչ թե բարբառը. այլ կոնկրետ բնակավայրի խոսվածքը, և նյութերը սկսել են հավաքվել բոլոր հայաբնակ վայրերից:

Ատլասի համար նյութերի հավաքման աշխատանքները սկսվել են 1977թ: Դրանք կատարվել են գիտարշավների միջոցով, որոնց ընթացքում ուսումնասիրվել են տվյալ տարածքի բոլոր բնակավայրերը: Բարբառային միավորները գրանցվել են ըստ Ծրագրի:

Ծրագիրը կազմված է բառային, քերականական և հնչյունական զուգաբանություններն ընդգրկող բաժիններից: Առաջին բաժինը՝ «Բառային զուգաբանություններ», կազմված է երկու մասից՝ «Հասկացական-բառանվանողական զուգաբանություններ» և «Բառիմաստային զուգաբանություններ»:

«Հասկացական-բառանվանողական զուգաբանությունների» խնդիրն է պարզել բառային նույն միավորի դիմաց տարբեր վայրերում գործառող տարբերակները: Այն ընդգրկում է 478 հոդված՝ 1-478: Հասկացությունները ներկայացված են ըստ իմաստային դաշտերի՝ բնության երևույթներ, բույսեր, գործիքներ, հագուստ, զբաղմունք և այլն: Յուրաքանչյուր հոդված կազմված է տվյալ հասկացության գրական անվանումից, բառաշարքից, որի մեջ մտնում են այդ հասկացության մինչ այդ հայտնի բարբառային տարբերակները, և հասկացության իմաստը պարզաբանող օրինակից: Այսպես՝

288. Դառն - դառը, աղու, լեղի:

Կորիզը դառն է:[3]

Հոդվածում բերվող բառաշարքը, սակայն, սպառիչ չէ, և գրանցումների ընթացքում արձանագրվել են նորանոր տարբերակներ:

Բառային զուգաբանությունների երկրորդ՝ «Բառիմաստային զուգաբանություններ» հատվածի խնդիրն է պարզել բառային նույն միավորի իմաստային

2 Տես Հայերենի բարբառագիտական ատլաս, մաս 1, Երևան, 1982, էջ 21:
3 Հայերենի բարբառագիտական ատլասի նյութերի հավաքման ծրագիր, Երևան, 1977, էջ 52:

տարբերությունները բարբառախոս վայրերում: Այն բաղկացած է 80 հոդվածից /479-559/:

Հիշյալ բաժնում նախապես տրվում է զուգաբանությունը, որից հետո նշվում են տվյալ հասկացության բարբառային իմաստները, այսպես՝

517. Դեղ – 1. Բուժման միջոց: 2. Թույն: 3. Սկնդեղ: 4. Քթախոտ: 5. Աչքի ծարիր: 6. Պաների մակարդ:[4]

Ծրագրի երկրորդ բաժինը քերականական զուգաբանություններն են, որն ընդգրկում է ձևաբանական, մասամբ նաև՝ շարահյուսական հատկանիշներ: Այն կազմված է 120 հոդվածից՝ 560-680: Այս հատվածում փորձ է արվում պարզել, թե տվյալ քերականական իմաստի արտահայտման համար ձևական ինչ միջոցներ են գործառում բարբառախոս վայրերը: Այդ պատճառով նախ տրվում է քերականական իմաստը և դրսևորման ձևերը, որից հետո բերվում են օրինակներ, այսպես՝

577. –ություն ածանցով բառերի հոլովումը՝

ա/ ներքին ա հոլովմամբ /մեծութուն – մեծութան, լավութուն – լավութան/

բ/ ի հոլովմամբ /մեծութէնի, լավութէնի/

գ/ ու հոլովմամբ / մեծութունու, լավութունու/

աղքատության, բարձրության, եղբայրության, երկարության, լայնության, կարճության, հաստության, մեծության, չարության, քաջության…[5]

Ծրագրի երրորդ բաժնի՝ «Հնչյունական զուգաբանությունների» խնդիրը պատմական հնչյունափոխության երևույթների բացահայտումն է: Այս բաժնի զուգաբանությունները, որոնց թիվը հասնում է հարյուրի՝ ընդգրկելով 668-778 համարները, ունեն նախորդ բաժնի զուգաբանությունների կառուցվածքը:

Նյութերի հավաքումը ըստ Ծրագրի կատարվել է ամեն մի բնակավայրի համար առանձին մեկ տետրի մեջ: Եթե մի քանի վայրեր ունեցել են նույն խոսվածքը, ապա դրանք ներկայացվել են մեկ տետրում՝ նշելով այդ բնակավայրերը:

Տետրի սկզբում տրվում են տեղեկություններ գրանցվող վայրի, բնակիչների, խոսվածքի, գրանցման աղբյուրի՝ բարբառախոսների, ինչպես նաև խոսվածքը գրանցողի վերաբերյալ:

Առ այսօր ըստ Ծրագրի լրացված է մոտ 500 բարբառային միավոր Հայաստանի արևելյան և արևմտյան տարածքներից: Ինչպես արևելյան, այնպես էլ արևմտյան վայրերի նյութերը հավաքված են Հայաստանի Հանրապետության տարածքում: Արևմտյան տարածքի բարբառային միավորները գրանցվել են 1915 թվականից հետո Հայաստան եկած մարդկանց օգնությամբ, որոնք հաճախ տվյալ բնակավայրի միակ բնակիչն են եղել, հետևաբար բարբառային միավորի միակ կրողը:

Այժմ բարբառային նյութի հավաքման սկզբունքների, հավաստիության և գրանցման ժամանակ առկա դժվարությունների մասին:

Նյութը գրանցելիս կարևոր պայման է բարբառախոսների ճիշտ ընտրությունը, որոնց տեղեկություններից է կախված տվյալների հավաստիությունը: Փորձը ցույց է տվել, որ հավաստի տվյալներ հաղորդվել են նրանց կողմից, ովքեր երկար ժամանակ ապրել և շարունակում էին ապրել մայրենի խոսվածքը կրողների միջավայրում: Դժվարություններ էին առաջանում հատկապես արևմտահայ

4Հայերենի բարբառագիտական ատլասի նյութերի հավաքման ծրագիր,էջ 73:
5Հայերենի բարբառագիտական ատլասի նյութերի հավաքման ծրագիր, էջ 81:

տարածքի խոսվածքները լրացնելիս: Արևմտյան Հայաստանից գաղթած բարբառախոսները, երկար ժամանակ կտրված լինելով հարազատ միջավայրից, չէին կարողանում պատասխանել այս կամ այն հոդվածին վերաբերող հարցերին, հաճախ նաև չէին կողմորոշվում, թե իրենց կողմից գործածվող ձևերից որն է պատկանում մայրենի խոսվածքին: Ճիշտ է, ծրագրում ընտրված լեզվական հատկանիշները համակարգային բնույթ ունեն և դժվար մոռացվող են, բայց գրանցումների ընթացքում քիչ չեն պատահել նման դեպքեր: Այսպիսի դեպքերում այդ բացը լրացվել է շրջակա գյուղերի խոսվածքների տվյալներով, որոնք կամ արդեն գրանցված էին կամ էլ ծանոթ էին գրանցողին, որը կարողացել է կողմնորոշել տեղեկատու անձին՝ նրան հիշեցնելով շրջակա վայրերում գործածվող ձևերը: Հավելենք, որ նման մեթոդով աշխատելու դեպքում բարբառախոսները արագ կողմորոշվում էին և հիշում մայրենի խոսվածքին հարազատ բարբառային տարբերակները:

Այս պարագայում պակաս կարևոր չէ նաև նյութը գրանցողը, որի մասնագիտական պատրաստվածությունից ևս կախված է գրանցվող նյութի հավաստիությունն ու լիակատարությունը: Այդ պատճառով էլ նյութի հավաքումը հիմնականում կատարվել է մասնագետ-բարբառագետների կողմից, որոնք գիտական շահագրգռվածություն են հանդես բերում գրանցվող տվյալների նկատմամբ:

Դժվարություններից մեկն էլ այն էր, որ հաճախ բարբառախոսներին հարկ էր լինում համոզել սկսած աշխատանքը հաջողությամբ ավարտելու համար, մանավանդ երբ բարբառախոսների ընտրության հնարավորություն չէր լինում: Դա բացատրվում է նրանով, որ ծրագիրն ընդարձակ է, և մեկ խոսվածքի գրանցումը տևում է 6-7 օր, հետևաբար միշտ չէ, որ բարբառախոսները, որոնց տարիքը տատանվում էր 70-100-ի միջև, սիրով համաձայնում էին տեղեկություններ հաղորդել: Նման դեպքերում պարզ ու հասկանալի բացատրվում էր գործի կարևորությունը, ընգծվում էր այն կարևոր փաստը, որ չհրաժարվելու դեպում կփրկվի գոնե իրենց հարազատ գյուղի լեզուն և կհանձնվի սերունդներին, բացի այդ, ասվում էր, որ մենք ոչ թե երկար պատմություններ ենք ուզում, այլ ուզում ենք իմանալ՝ ինչպես են ասում իրենց գյուղում այս կամ այն ուտելիքին, բույսին, գործիքին: Այս բոլորից հետո նրանք աստիճանաբար տեղի էին տալիս և համաձայնում սկսել գործը: Դրան նպաստում էր նաև Ծրագրի կառուցվածքը, քանի որ այն սկսվում է «Բառային զուգաբանություններ» բաժնով, որը «…լինելով «Ծրագրի» առավել հետաքրքիր ու հեշտ ընկալելի բաժինը, շարժում է տեղեկատու անձանց հետաքրքրասիրությունը, նրանց հոգեկան աշխարհում արթնացնում է առանձնակի ապրումներ, հուշեր ու տպավորություններ, պարգևում է լեզվազգացողության հրճվալից պահեր և արվող գործի նկատմամբ առաջացնում է որոշակի ոգևորություն»:[6]

Թեև նյութերի հավաքումը դեռ ավարտված չէ, սակայն ծրագրի կառուցվածքը հնարավորություն է տալիս սկսել գրանցված նյութի մշակումը և համակարգչայնացումը: Այդ աշխատանքը կատարելու համար մեր կողմից ընտրվել է File maker ծրագրով: Մինչև նյութի համակարգչայնացումը այն պետք է մշակվի և դասակարգվի: Սա երկարատև աշխատանք է և նյութի լավ իմացություն է պահանջում: Սակայն քանի որ Ծրագրում, որի համաձայն հավաքվել է նյութը, լեզվական հատկանիշները ներկայացված են զուգաբանությունների ձևով, այդ փաստը դյուրացնում է նյութերի հավաքման և մշակման աշխատանքները: Ակնհայտ է, որ լեզվական հատկանիշները զուգաբանությունների տեսքով

6 Տես Հայերենի բարբառագիտական ատլաս, մաս 2, Երևան, 1985, էջ 17:

ներկայացնելուն անպայման նախորդել է գիտական որոշակի աշխատանք: Ասվածը հիմնավորելու համար ներկայացնենք 626 թվահամարն ունեցող զուգաբանությունը:

626. Սահմանական եղանակի ներկա ժամանակի դրսևորումը՝

1. անկատար դերբայով և օժանդակ բայի ներկայի ձևերով /գրում եմ, գրելիս եմ, նստման ամ/:
2. հին հայերենի ներկայի ձևերով /գրեմ, կարդամ/
3. հին հայերենի ներկայի ձևերի և բառ-մասնիկի զուգորդմամբ՝

ա/ կու, կա, կը, գու, գա, գը /կա գրեմ, գրեմ կը/

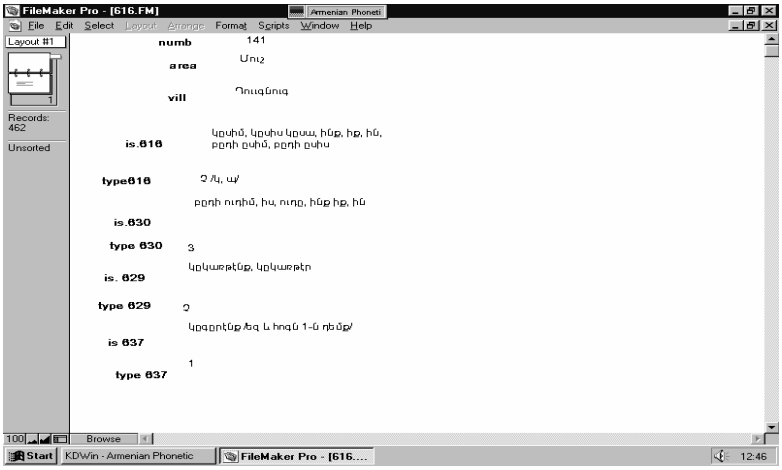
բ/ ահա, հայ, ա, ը…/հայ պանիմ, գարթամ ա.../

գրում եմ	գրում ենք	մնում եմ	մնում ենք
գրում ես	գրում եք	մնում ես	մնում եք
գրում է	գրում են	մնում է մնում են /Ծրագիր, էջ 97/	

Ինչպես նկատելի է, այս զուգաբանության դեպքում ամփոփվել են այն բոլոր կազմությունները, որոնք տարբեր խոսվածքներում գործածվում են սահմանական եղանակի ներկա ժամանակի իմաստով, այսինքն՝ կատարված է տվյալ հատկանիշի գիտական դասակարգում, որը օգնում է բարբառագետ-մասնագետին ավելի հեշտ և հստակ կողմորոշվել հսկայածավալ նյութի վրա աշխատելիս: Այս զուգաբանությունը մշակելիս տարբեր վայրերից հավաքված նյութը նախ կխմբավորվի ըստ դերբայով, եղանականիշ մասնիկով կամ գրաբարածև կազմություններով ներկա ժամանակը կազելու հատկանիշի, որից հետո դերբայականզմ ձևերը կբաժանվեն 3 ըստ զուգաբանության մեջ տրված տարբերակների, իսկ մասնիկավոր ձևերը՝ 2 ենթախմբերի: Հիշյալ երեք մեծ խմբերը համակարգչայնացման փուլում կարտահայտվեն տարբեր թվերով կամ տառերով, իսկ նրանց ենթախմբերը՝ այդ տառերի կամ թվերի փոփոխված տարբերակներով՝ 1.1, 1.2, 1,3 և այլն: Դրանք հետագայում քարտեզի վրա կփոխարինվեն պայմանական նշաններով:

Հավելենք, որ սա համակարգչայնացման ամենակարևոր փուլն է, քանի որ սրանով է պայմանավորված բարբառային զուգաբանությունների ճշգրիտ քարտեզագրումը:

Կարևոր է նաև համակարգչային դաշտերի ճիշտ կազմումը: Այն պետք է արտացոլի տվյալ զուգաբանության նյութն իր բոլոր ստորաբաժանումներով և պարունակի անհրաժեշտ տվյալներ այդ զուգաբանությունն ունեցող բարբառային միավորի վերաբերյալ: Այդ պատճառով էլ նախապես որոշվել է համակարգչային դաշտերի կառուցվածքը: Բայական զուգաբանությունների համար դաշտերը կազմվել են հետևյալ սկզբունքով՝ տետրի համարը նշող սյունակ, այնուհետև՝ տվյալ խոսվածքի շրջանի և գյուղի անվանումները, տվյալ հատկանիշի թվահամարը, վերջում՝ տեսակը: Վերջինս կարելի է ներկայացնել թվերով կամ տառերով: Մենք նպատակահարմար ենք գտել դրանք ներկայացնել թվերով:



Ինչպես երևում է լուսանկարից համակարգչային դաշտը կազմված է բայական մի քանի հատկանիշի համար: Վերցնենք 637 թվահամարը /տես դաշտ 1/: Ծրագրում այն ներկայացվում է հետևյալ ձևով. «Բայի դիմային վերջավորությունների նույնացում՝ ա/ անցյալի եզակի և հոգնակի առաջին դեմքում /ես գրեր եմ, մենք գրեր ենք/

բ/ հոգնակի առաջին և երկրորդ դեմքերում /մենք գրել ենք, դուք գրել եք/

գ/ անցյալի եզակի թվի բոլոր դեմքերում /ես կգրեր, դու կգրեր, նա կգրեր/

գրում էի	գրում էինք	գրում ենք	գրում եք	խմում էի
գրեցի	գրեցինք	գրելու ենք	գրելու եք	խմում էիր
գրել էի	գրել էինք	կգրենք կգրեք	խմում էր	
գրեի	գրեինք			

Ինչպես երևում է, այս զուգաբանության համար առանձնացված են 3 խմբեր, որոնցից յուրաքանչյուրը կներկայացվի որևէ թվահամարով, այսպես՝ ա կետի հատկանիշը կունենա 1, բ-ն՝ 2 և գ-ն՝ 3 թվահամարները:

Ծրագրի այս, ինչպես և մյուս դասակարգումները վերջնական չեն, և բարբառային նյութի մշակման ընթացքում հնարավոր է նոր խմբերի առանձնացում: Վերոբերյալ զուգաբանության համար, օրինակ, առանձնացվել են նաև հետևյալ խմբերը՝

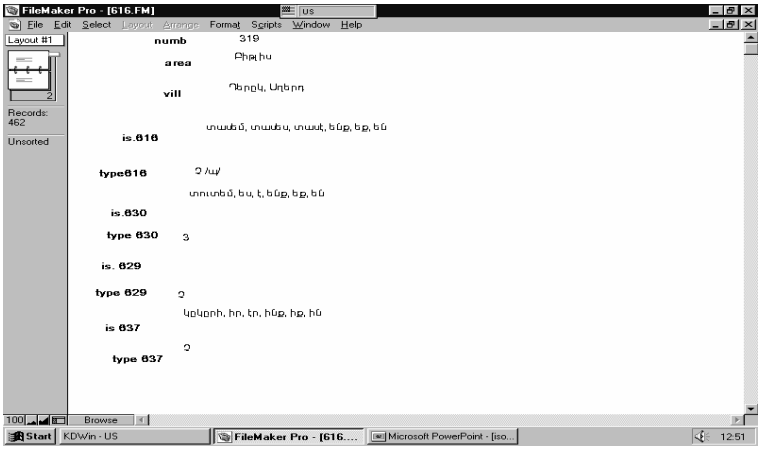
դ/ եզակի և հոգնակի 1-ին դեմքում, ինչպես նաև եզակի 2-րդ և 3-րդ դեմքերում /ես կգրէնք, մենք կգրէնք, ինչպես նաև՝ դու կըգրէր, նա կըգրէր/

ե/ եզակի 2-րդ և 3-րդ դեմքերում, ինչպես նաև հոգնակի 1-ին և 2-րդ դեմքերում /ես գըխօսէմ, դու գըխօսէր, նա գըխօսէր, մենք գըխօսէկ, դուք գըխօսէկ, նրանք գըխօսէն/

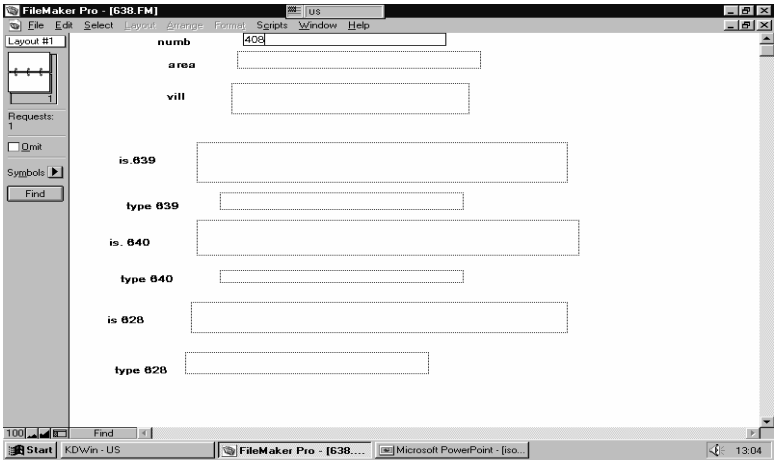
զ/ եզակի թվի 2-րդ և 3-րդ դեմքերում և եզակի 1-ին, հոգնակի 1-ին և 2-րդ դեմքերում /ես գըխօսէկ, դու գըխօսէր, նա գըխօսէր, մենք գըխօսէնկ, դուք գըխօսէկ, նրանք գըխօսէն/:

է/ եզակի 2-րդ և 3-րդ դեմքերում /ես կըգըրէ, դու կըգըրէր, նա կըգըրէր/:

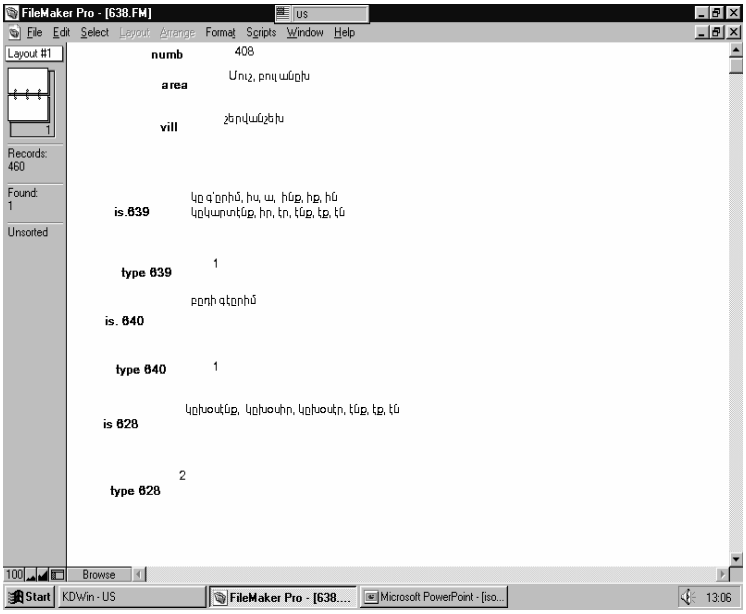
Քնականաբար նշված տարբերակներից յուրաքանչյուրը կստանա նոր թվահամար՝ համապատասխանաբար 4. 5. 6 և այլն: Հավելենք, որ սա միակ զուգաբանությունը չէ, որի դեպքում Ծրագրում նշված դասակարգումը լրացվել է նոր խմբերով:



Հետագայում այդ թվերը քարտեզների վրա կդառնան նշաններ, որոնց տարբերությամբ կտարորոշվեն տվյալ զուգաբանության տիպերը կամ տեսակները:



Բարբառային նյութի այս ձևով համակարգչայնացումը հնարավորություն է տալիս օգտվել դրանից և ստանալ ցանկացած հարցի պատասխան: Այսպես՝ ծրագիրն ունի Find հրահանգը, որի օգնությամբ կարելի է պարզել քերականական այս կամ այն հատկանիշի հետ կապված որևէ հարց, ասենք՝ տարածման սահմանները, տիպերը:



Առ այսօր մեր կողմից վերոնշյալ սկզբունքով համակարգչայնված են 460 բարբառային միավորների բայական զուգաբանությունները, որոնք ծրագրում 612-662 համարներն են ընդգրկում:

Կարևոր ենք համարում նշել, որ բարբառային նյութի համակարգչայնացման աշխատանքները նախաձեռնելիս մեզ օգնություն է ցույց տրվել Լեյդենի համալսարանի պրոֆեսոր, հայագետ Յ. Վայտենբերգը: Նրա օժանդակությամբ ենք ուսումնասիրել File Maker ծրագիրը, մշակել բարբառային նյութի համակարգչայնացման սկզբունքները:

Բարբառային նյութը, նրա մշակումը, հետագայում նաև քարտեզագրումը կարևորվում է տվյալների բազմազանությամբ, բարբառ - գրական եզր դրսևորումների յուրօրինակությամբ, նաև կորստյան վտանգի առջև կանգնած արևմտահայ բարբառների գոնե պահպանված վիճակի ամրագրմամբ և բարբառային հսկայածավալ նյութը անդառնալի կորստից փրկելու փաստով:

Նոր տվյալների օգնությամբ կճշտվեն բարբառային միավորների փոխհարաբերությունները, հայերենի բարբառների դասակարգման սկզբունքները, կբացահայտվեն բարբառային կարևոր հատկանիշների ընդհանրություններն ու տարբերությունները: Գիտական շրջանառության մեջ կդրվեն նոր տվյալներ, եզակի փաստեր, որոնք կնպաստեն հայերենի բարբառների ուսումնասիրությանը, հնարավորություն կտան պարզելու բարբառախոս վայրերի պատմական ընդհանրությունները, ավելի ստույգ ու ճշգրիտ մեկնաբանելու լեզվական փոխազդեցությունները:

Լեզվաբանական նշանակությունից բացի՝ բարբառային նյութը ունի նաև պատմական և մշակութային արժեք: Նրա օգնությամբ հնարավոր կլինի լուսաբանել ոչ միայն բանասիրական, լեզվաբանական, այլև պատմագիտական, ազգագրական և բազմաթիվ այլ հարցեր:

Ստորև որպես հավելված բերվում է այն բարբառային միավորների բնակավայրերի ցանկը, որոնց տվյալները համակարգչայնացվել են: Բնակավայրերը բերվում են համաձայն ՀՀԳԱԱ լեզվի ինստիտուտում Հայերենի բարբառագիտական ատլասի հավաքման ծրագրով կուտակված նյութերի՛ ըստ վարչական բաժանումների: Բնակավայրից հետո փակագծում տրվում է վարչական շրջանը կամ գավառը: Արևմտահայ տարածքի բնակավայրերի համար նկատի է առնվում 19-րդ դարի վերջի և 20-րդ դարի սկզբի վարչաշխարհագրական այն բաժանումը, որն ունեցել են վերոնշյալ հատվածում տեղակայված գավառներն ու բնակավայրերը:

Քեսաբ
Տիգրանակերտ
Շապին Գարահիսար
Ակն
Երզնկա
Խարբերդ
Վան
Դիադին
Փարպի /Աշտարակ/
Արփի /Սասուն/
Ներքին Խաթունարխ /Էջմիածին/
Ռամիս /Նախիջևան/
Փառակա /Նախիջևան/
Ազա /Նախիջևան/
Խանաբադ /Ղարաբաղ,
Ստեփանակերտ/
Ենգիջա /Եղեգնաձոր/
Շատախ
Գոպալ /Տարոն, Խնուս/
Արտաշատ
Օշական /Աշտարակ/
Կարճևան /Մեղրի/
Ագուլիս
Երևան /կենտրոն/
Դաշբուլաղ /Ղարաբաղ,
Ստեփանակերտ/
Ազնաբերդ /Նախիջևան/
Թուսկուլու /Վարդենիս/
Ջուղա, Ծղալթբիլա /Ախալցխա/
Չմշկածագ
Փալդրվան /Կարս/
Մամռտանց /Մոկս/
Սպիտակ
Շահումյան /Դաշքեսան/
Կրխու /Սասուն, Մոտկան/
Փիջունք /Բիջունք/ /Խութ/
Սկրագոմ /Մուշ/
Զանգի /Ապարան/
Սելիքգյուղ, Հարթավան, Արագած,
Արայի /Ապարան/
Ապարան
Ճենիկ /Համշեն/
Հավարիկ
Գորիս
Կոփ /Մուշ, Բուլանուխ/
Ղզըլղոչ, Ճըջըխլար /Կարս/
Մուղնի /Աշտարակ/
Գոպալ /Տարոն, Խնուս/
Քիշմիշբափա /Իրան/
Բանանց /Դաշքեսան/
Լեռնապար /Արագած/
Վաղուհաս /Ղարաբաղ/
Ծաղկավան /Իջևան/
Կարմիր /Բայազետ/
Մասիս /Արտաշատ/
Տոսու, Վերին Հուրուք /Սպարկերտ/
Ղալայչի, Արագած, Ատմա
/Ապարան/
Ընգուզնակ, Սաղդուս /Սասուն/
Ջորաղբյուր /Աբովյան/
Խաչան /Բերկրի/
Սիփան /Արծկե/
Հադրութ /Ղարաբաղ/
Մալաթիա
Վերին Հոռաթաղ / Ղարաբաղ,
Մարտակերտ/
Սեյսուլան / Ղարաբաղ,
Մարտակերտ/
Խոժոռնի /Վրաստան, Մառնեուլ/
Խանլար /Աղբրեջան/
Սառնաղբյուր /Անի/
Մուժումբար /Իրան/
Բուղաշեն /Ախալքալաք/
Մեծ Շեն / Ղարաբաղ,
Մարտակերտ/
Հին Ջուղա /Նախիջևան/
Նավուր /Շամշադին/
Վարդաբլուր /Ստեփանավան/
Վաչիան /Ախալքալաք/
Սեյրավար /Խոյ/
Այնթափ /Մասիս/
Գյումուր /Նախիջևան/
Ղասմաշեն /Քեյվան/
Թաղվու /Խութ/
Դիլիֆ /Բոգդանովկա/
Քխսաս /Բիթլիս/
Ախալքալաք
Տանակերտ /Նախիջևան/
Սամսոն /Համշեն/
Քարաղաշտ /Վան, Ոստան/
Բլուր /Սուրմալու/
Ուլյա Նորաշեն, Շարուրի
/Նախիջևան/
Բադամլու, Հորս /Եղեգնաձոր/
Ռոդոսթո
Խաչ /Էրզրում/
Քոփ /Սասուն/
Բազմաշեն /Խարբերդ/
Կտրած քար /Արճեշ/
Արարատ /Արարատ/
Փոգա /Բոգդանովկա/
Իրցանք /Սասուն/
Ջրվեժ
Գետահովիտ /Իջևան/
Գոմս /Վան, Թիմար/
Արճակ /Վան/
Պռոշնենց /Բիթլիս, Խիզան/
Բժգներդ /Աղբակ/
Դռնաշեն /Վան, Թիմար/
Հադգոն /Մուշ/
Հասպստան /Աղբակ/
Քյօչանի /Վան, Թիմար/
Հօրմօրուր /Վան/
Դերըկ /Բիթլիս, Սղերդ/
Ջենիս /Վան, Խոշաբ/
Բուզլուխ /Շահումյան/
Վանք /Վան, Աղբակ/
Գործոթ/ Վան, Բերկրի/
Սխկյա, Մխկներ /Վան/
Մարզիյան /Իրան,Գափլա/
Ղավրավա /Իրան, Ներքին
Բուրվարի/
Սևան, Ջարանց /Վան, Արճակ/
Նամագերդ /Իրան, Փերիա/
Սնձնուտ /Մուշ/
Լիմ /Վան, Արճակ/
Չիգյան /Իրան, Փերիա/
Ույքիլիսա /Ալաշկերտ/
Հուռթուկ /Վան, Խոշաբ/
Հնդստան /Վան, Խոշաբ/
Կիրովաբադ
Ճանիկ /Համշեն/
Մարաշ
Հացիկ /Մուշ/
Ներքին և Վերին Խոտանան
/Ղափան/
Ջիարեթ /Մուշ/
Թաղիաբադ /Իրան, Միանդաբ/
Նոր գեղ /Վան, Հայոց ձոր/
Արևշատ /Արտաշատ/
Դաջան /Իրան, Փերիա/
Վերին Խոգյան /Իրան, Փերիա/
Աբուլ /Ախալքալաք/
Մշկեղ /Սասուն, Փսանք/
Բոլորմակ /Վան, Արճեշ/
Ազնավուլ /Իրան, Փերիա/
Ղարաքյարիս /Իրան, Փոքր
Բուրվարի/
Բոլորան /Իրան, Փերիա/
Տուսու /Բիթլիս, Սպարկերտ/
Հաջիաբադ /Իրան, Սուլթանաբադ/
Վարդխաղ /Մուշ/
Հարդիֆ /Էրզրում, Քղի/
Ջարա /Մեբաստիա/
Վինան /Ղարադաղ/
Մարաղա

Հարությունագոմեր /Ղարաբաղ/
Կոճողոտ /Ղարաբաղ/
Ծաղկաշեն /Ապարան/
Չիբուխլու /Համշեն, Օրդու/
Խանագահ /Իրան, Ղարադաղ/
Մարդակերտ /Ղարաբաղ/
Հառեն, Արջրա /Արծկե/
Առնջկուլա /Արծկե/
Իշխանաձոր /Սասուն/
Ծայծակ /Արճեշ/
Միջին թաղ /Խոտրջուր/
Վարդենիս /Ապարան/
Օհանավան /Աշտարակ/
Գելիգուզան /Սասուն/
Վերին Ճամբարակ /Կրասնոսելսկ/
Մուժումբար /Թավրիզ/
Ղափան
Հախվերիս /Իգդիր/
Շեն /խութ/
Թաղավանք /Խութ/
Էգնախոջա /Մանազկերտ/
Բյուրական /Աշտարակ/
Ուրգութուն /Ղարադաղ/
Արագած /Էջմիածին/
Անտառուտ /Աշտարակ/
Բաղնոս /Մանազկերտ/
Ղաբուռ /Վրաստան, Ճալկա/
Քասախ /Ապարան/
Աչաջուր /Իջևան/
Արծափ /Կոգովիտ/
Նորք /Երևան/
Մամզարա /Բոգդանովկա/
Սեբաստիա
Սևերեկ
Կոշ /Աշտարակ/
Եմիշճան /Ղարաբաղ, Մարտունի/
Չովդար /Դաշքեսան/
Եղվարդ
Մարտունի /Շամքոր/
Փարադաշ /Նախիջևան/
Օրջոնիկիձե /Կրասնոսելսկ/
Տումի /Ղարաբաղ, Հադրութ/
Հին Բաշքենդ /Կրասնոսելսկ/
Օրջոնիկիձե /Շամքոր/
Հաջըն
Քյալղապա/Համշեն, Օրդու/
Գյումրի
Սիրբար /Մուշ/
Գետք /Ախուրյան/
Կոթի /Նոյեմբերյան/
Տրապիզոն
Սարդու /Ղարադաղ/
Քարահունջ /Գորիս/
Խնածախ /Գորիս/
Խնձորեսկ /Գորիս/
Վերիշեն, Բռուն /Գորիս/
Կորնիձոր /Գորիս/
Վերին կարմիր աղբյուր
/Շամշադին/
Թալիշ /Ղարաբաղ, Մարտակերտ/
Կազի /Ալաշկերտ/
Բանթաղ /Ալաշկերտ/
Ջեյթուն
Եկմալ, Լաթար /Մուշ/
Ազատավան /Արտաշատ/
Սառնաղբյուր /Անի/
Ծովազարդ /Քայազետ/
Լճաշեն /Սևան/
Բուրաստան /Արտաշատ/
Գառնի
Ղզլչախչախ /Կարս/
Գետաշեն /Խանլար/
Իքի Աղաջ /Ուրմիա/
Տեղեր /Աշտարակ/
Շուռուտ /Նախիջևան/
Սվջող /Խարբերդ/
Կյաղ /Գալ/ /Նախիջևան/
Բիստ /Նախիջևան/
Հադան /Իրան, Փերիա/
Չղդրան /Ղարաբաղ, Մարտակերտ/
Հասանղայա-Մարալյան սարով
/Ղարաբաղ, Մարտակերտ/
Չափար /Ղարաբաղ, Մարտակերտ/
Մեհտիշեն /Ասկերան/
Հայադի /Ղարաբաղ, Մարտակերտ/
Ղազարահող /Ղարաբաղ/
Պողոսագոմեր /Ղարաբաղ/
Ճարտար /Ղարաբաղ/
Ջագլիկ /Ղարաբաղ/
Հաթերք /Ղարաբաղ/
Մարաղա /Ղարաբաղ/
Մսրվանիս /Նախիջևան/
Կաթնաղբյուր /Աբովյան/
Քանաքեռ
Խոսրովավան, Իշխու Բասեն/
Հազարեն /Վան/
Շուրիշկան /Փերիա/
Մեծ Սամսար /Ախալքալաք/
Աստապատ /Նախիջևան/
Առինջ /Աբովյան/
Շահումյան /Վրաստան/
Էլար /Աբովյան/
Պտղնի /Աբովյան/
Լիճք /Մարտունի/
Հիթինք /Սասուն/
Միլակերտ /Փերիա/
Բերկրի /Վան/
Կողք /Իգդիր/
Առաջածոր /Ղարաբաղ,
Մարտակերտ/
Ջանյաթաղ /Ղարաբաղ,
Մարտակերտ/
Մադաղիս /Ղարաբաղ,
Մարտակերտ/
Ղազանչի /Ղարաբաղ,
Մարտակերտ/
Քոլատակ, Թբղլվի, Գառնաքար
/Ղարաբաղ, Մարտակերտ/
Տռնաշեն, Սողոմոնագոմեր, Գետի
գոմեր, Դաստակերտ /Ղարաբաղ,
Մարտակերտ/
Մադինա /Ալաշկերտ/
Ունիե Բալղուղ /Համշեն/
Մրգաշատ /Հոկտեմբերյան/
Կարբի /Աշտարակ/
Լիզ /Մուշ, Բուլանուխ/
Խոսպ /Վան, Հայոց ձոր/
Արտամետ /Վան/
Մերաբադ /Իրան, Դովլաթաբադ/
Սանգիբարան /Իրան, Փերիա/
Ազնա, Քալլավա /Իրան, Քյազազ/
Գուլիզարդ /Իրան, Քյազազ/
Կարակըզի, Բադիլբոյ /Խոյ/
Գետազատ /Արտաշատ/
Փայաջուկ /Խոյ/
Խեք /Վան/
Արծափ /Կոգովիտ/
Ալիջան /Սուրմալու/
Խարիդ /Խիզան/
Հովտամեջ /Էջմիածին/
Առատաշեն /Էջմիածին/
Հազարջիրիք /իրան, Փերիա/
Բոլդաջի /Իրան, Չալմահալ/
Աղբուլաղ /Իրան, Չալմահալ/
Վերին քրդեր /Իրան, Քյարվանդ/
Դեյնով /Իրան, Վերին Բուրվարի/
Յաղափլու /Վարդաշեն/
Ռստամ-Գեդուկ /Մանազկերտ/
Կոճերեր /Արծկե/
Մցու /Բիթլիս, Մոտկան/
Նիչ /Բիթլիս, Մոտկան/
Աղավնատուն, Հայթաղ /Էջմիածին/
Ծաղկալանջ /Էջմիածին/
Ամբերդ, Դոդս /Էջմիածին/
Ծաղկունք, Շահումյան /Էջմիածին/
Նիգավան /Ապարան/
Քուչակ /Ապարան/
Քարագլուխ /Ղարաբաղ, Հադրութ/
Ղարաղշլաղ /Շուշի/
Բեժանո /Ախալքալաք/
Չինչին /Շամշադին/
Եխնաթափա /Ալաշկերտ/
Դամիա /Վրաստան, Մառնեուլ/
Տադվերան /Բասեն/
Դաշբուռուն /Սուրմալու/
Ասլանբեկ
Շեն /Բիթլիս, Հիզան/
Նուխի
Աթարբեկյան /Էջմիածին/
Մրցանք /Բիթլիս, Մոտկան/
Վարդենիս /Մուշ/
Մոլալու /Մարտունի/
Գյարմավ /Շատախ/
Կոռ /Բիթլիս, Մոտկան/
Ջուջան /Դիադին/
Կաղզվան /Կարս/
Մոլլաղամար /Սուրմալու/
408 Շերվանշեխ /Մուշ, Բուլանուխ/
Կողակ /Մուշ, Բուլանուխ/
Օրմոնքոյ /Համշեն, Ճենիկ/
Նուխի
Չլկանի /Ալաշկերտ/
Ղազուկլաղ /Սուրմալու/
Ծպնի /Կարս/
Բուլանուխ /Կարս/
Լղարակ /Շամխոր/
Սաղյան /Շամախի/
Խորզենդ /Իրան, Վերին Բուրվարի/
Ղարղուն /Իրան, Փերիա/
Փանիկ /Սուրմալու/
Վելի Բաբա /Բասեն/
Խարկեն /Արճեշ/

Տաթև, Սևարանց /Գորիս/
Եղեսիա
Գարմունջ /Եղեսիա/
Ուզ /Սիսիան/
Գնիշիկ /Եղեգնաձոր/
Աղավնաձոր /Եղեգնաձոր/
Ծոկու /Կարկառ/
Հազզո /Սասուն/
Օլավերդ /Ախալքալաք/
Հոսներ /Սասուն/
Գանձա /Բոգդանովկա/
Խոտ /Գորիս/
Վաղատուր /Գորիս/
Շինուհայր /Գորիս/
Բասեն /Գոմաձոր/
Հին Թաղլար / Ղարաբաղ, Հադրութ/
Թազագյուղ /Մարտունի/
Շնող /Թումանյան/
Նյուզգեր /Շամխոր/
Երևանի մայլա /Բյուրական/
Արաբկիր
Ծաղկիս /Կարին/
Սանահին /Թումանյան/
Դսեղ /Թումանյան/
Թեղուտ /Թումանյան/
Շամուտ /Թումանյան/
Ճոճկան /Թումանյան/
Հաղբատ /Թումանյան/
Բարսում /Շամքոր/
Դուգնուգ /Մուշ/
Վերնաշեն /Եղեգնաձոր/
Ղզլքիլիսա /Ծալկա/
Նորս /Նախիջևան/
Ջովունի /Ապարան/
Թազաղենդ /Ղարադաղ/
Խանիշեն /Շամախի/
Մաղրասա /Շամախի/
Ցղնա /Նախիջևան/
Արփա /Եղեգնաձոր
Աշտարակ
Շամիրամ /Քաղեշ/
Ալիճագրակ /Բասեն/
Եղեգիս /Եղեգնաձոր/
Մարմետ /Վան, Թիմար/
Քոշկ /Էրզրում, Այնթափ/
Դարման, Ծվստան /Վան/
Չոլախլու /Կարս/
Արմտլու /Բասեն/
Մազրա /Ալաշկերտ
Ուզունքիլիսա /Կարս/
Կողպանց /Վան/
Շուշանց /Վան/
Կյուսենց /Վան/
Լեսկ /Վան/
Չանախչի /Արարատ/
Արամուս /Աբովյան/
Ավրակ /Վան/
Զառ /Աբովյան/
Ակունք /Աբովյան/
Մառնիկ /Սասուն/
Գեղաշեն /Աբովյան/
Կապուտան /Աբովյան/
Խավենց /Վան/
Շատախ /Տղացպար/
Գյաբոլներ /Մուշ, Բուլանուխ/
Խտրբեկ /Սվեդիա/
Կամարիս /Աբովյան/
Բերդ /Շամշադին/
Խասկալ /Նիկոմեդիա/
Պարտիզակ
Քյուրագդարա /Կարս/
Հալնճի /Բիթլիս/
Հաղին /Մոկս/
Խալենց /Մոկս/
Կճավ /Մոկս/
Հարպերտ /Գավաշ/
Կուրթան /Ստեփանավան/
Յօնջալու, Իրիցու /Ալաշկերտ/
Յոնջալու /Մուշ, Բուլանուխ/
Խուռնավուլ /Սեբաստիա/
Ֆուռնոս Ուլնիա /Զեյթուն/
Ապարանց, Սեղ /Մոկս/
Քռեր /Վան, Խոշաբ/
Կոմք /Սասուն
Բալահովիտ /Աբովյան/
Սրվտրկի /Շատախ/
Կվերս /Շատախ/
Ավան /Աբովյան/
Հաբուսի /Խարբերդ/
Արմաշ /Նիկոմեդիա/
Մանդան /Արճակ/
Ադափազար /Նիկոմեդիա/
Կյաղ /Նախիջևան/
Ալյուր /Վան/
Արմշատ /Շատախ/
Էվջիլար /Սուրմալու/
Ոտնչոր /Մուշ, Բուլանուխ/
Բոստաբյանդ /Մուշ/
Դուման /Խնուս/
Նոր շեն /Ղարաբաղ, Մարտունի/
Մարգարա /Հոկտեմբերյան/
Թոխաթ /Եվդոկիա/
Գյուրջի /Իրան, Գափլա/
Նալբանդյան /Հոկտեմբերյան/
Ջորագլուխ /Ապարան/
Կուլաբ /Սուրմալու/
Նալբանդ /Սպիտակ/
Տոլորս /Սիսիան/
Մուսուն /Բայազետ/
Գյուզակ /Վան, Բերկրի/
Բոլնիս-Խաչեն /Բոլնիս/
Կաղարծի /Ղարաբաղ, Մարտունի/
Շահնազար /Կալինինո/
Սինագան /Չալմահալ/
Լիվասյան, Աղբուլաղ /Իրան, Չալմահալ/
Խաբլջող /Սասուն/
Տնգետ /Սասուն/
Նախիջևան
Արծվիկ /Սասուն/
Ծղնակ /Մուշ/
Ռաբաթ /Սասուն/
Ակունք /Թալին/
Սեմալ /Սասուն/
Խան /Սասուն, Փսանք/
Նաջարան /Բալու/
Խաստուր /Ալաշկերտ/
Նախիջևան /Կարս/
Կարծախ /Ախալքալաք/
Փիան /Մուշ, Բուլանուխ/
Մանջալիկ /Սեբաստիա/
Դենդիլ /Սեբաստիա/
Բեյլան /Անտիոք/
Մավրակ /Կարս/
Թալին
Նալբանդյան /Հոկտեմբերյան/
Կարմրաշեն /Թալին/
Խրմանջուղ /Ղարաբաղ/

Նարինե Հեքեքյան
Երևանի Վ.Բրյուսովի անվան պետական լեզվաբանական համալսարան

ՏԱՌԱՅԻՆ ՀԱՊԱՎՈՒՄՆԵՐԻ ՆԵՐԿԱՅԱՑՈՒՄԸ ՀԱՅԵՐԵՆԻ ԿՈՐՊՈՒՍՈՒՄ

Հապավումները, իբրև բառակազմական յուրահատուկ գործընթացի արդյունք, լեզվական անժխտելի իրողություն են և արդի լեզուների, նաև գրական արևելահայերենի բառապաշարի ուրույն ու բավականաչափ հարուստ շերտը: Ներկայումս սա լեզվի բառային կազմի թերևս ամենից արագ փոփոխվող ու համալրվող շերտն է: Բառակազմության այս տեսակի ծավալվումը թելադրված է կյանքի մերօրյա ռիթմով և տնտեսման սկզբունքով: Մի կողմից հասարակական-քաղաքական, տնտեսական, գիտական ու մշակութային կյանքում տեղի ունեցող անընդհատ փոփոխություններին արձագանքելու, միաժամանակ բառային հնարավորինս կարճ միավորներ ստեղծելու լեզվական անհրաժեշտությունը, մյուս կողմից՝ նյութի, տեղի, գրողի կամ խոսողի ժամանակի, ինչու չէ՝ նաև ջանքերի խնայողությունը այն գործոններն են, որոնք խթանում են հապավման, իբրև բառակազմական եղանակի, աշխուժացումը մեր օրերում: Սրանում պակաս կարևոր չէ նաև հապավումների ստեղծման ու գործածության ոլորտը. դրանք մեծապես գործածվում են զանգվածային լրատվության միջոցներում (պարբերական մամուլ, ռադիո, հեռուստատեսություն, ինտերնետ), որոնք տեղեկատվական մեր դարաշրջանի շարժիչ ուժն են, վարչական ու գիտական խոսքում (ներպետական և միջպետական փաստաթղթեր, պաշտոնական գրագրություն, գիտական ուսումնասիրություններ և այլն), և խոսքի հենց այս տեսակներն են այսօր ամենից շատ գրվողն ու ընթերցվողը: Բավական է նշել, որ վերջին տասնամյակի ընթացքում հարյուրավոր նորաստեղծ հապավումներ բառացիորեն ողողել են պարբերական մամուլի, տեղեկատու գրականության էջերը: Ու թեև հապավումների, մասնավորապես տառային հապավումների այսչափ ծավալումը խրախուսելի երևույթ չէ, սակայն փաստ է, որը չի կարելի անտեսել: Հապավումների հավաքումը և ուսումնասիրությունն անհրաժեշտ ու կարևոր են ոչ միայն լեզվաբանական, այլ նաև պատմական ու մշակութաբանական տեսանկյունից, քանի որ դրանք արտացոլում են պատմական որոշակի ժամանակաշրջանում երկրում տեղի ունեցող հասարակական-քաղաքական, տնտեսական զարգացումները, գիտության, կրթության, մշակույթի վիճակը և այլն: Ուստի գիտական և գործնական արժեք ունի հապավումների ներկայացման խնդիրը ինչպես բառարաններում, այնպես էլ արևելահայերենի կորպուսում:

Հայերենի բացատրական և ուղղագրական բառարաններում[1] հիմնականում ընդգրկված են մասային և մասաբառային հապավումներ (օրինակ՝ *գիմկոմհասարիպ, ուսմասվար, դասվար, կոլտնտեսություն, բաներիպղպրոց, կենտկոմ, կուսթեկնածու* և այլն), իսկ տառային հապավումները ներկայացված են եզակի նմուշներով, ինչպես՝ *բուհ, ևէպ, զագս, գէս* (ՀԷԿ-ի փոխառյալ տարբերակն

[1] Ժամանակակից հայոց լեզվի բացատրական բառարան, հհ. 1-4, Եր., 1969-1980, Էդ.Աղայան, Արդի հայերենի բացատրական բառարան, հհ. 1-2, Եր., 1976, Հ.Բարսեղյան, Հայերենի ուղղագրական-ուղղախոսական, տերմինաբանական բառարան, Եր., 1973:

է)[2], որոնք լեզվաբանական գրականության մեջ գործածվում են իբրև հասարակ գոյականի արժեք ստացած հապավումների դասական օրինակներ[3]:

Տառային հապավումներն իրենց յուրահատուկ կազմությամբ առանձնանում են հապավումների շարքում: Նրանք որոշակիորեն տարբերվում են հապավումների մյուս տեսակներից, որոնք կառուցվածքով նման են բաղադրյալ բառերին, իմաստային տեսակետից համեմատաբար ավելի դյուրընկալելի են և թեքման հարացույցում դրսևորում են նույն օրինաչափությունները, ինչ բաղադրյալ բառերը (օր՝. *ուսմասվար* -ներ, -ի, *զինկոմիսարիատ* -ներ, -ի): Բացի այդ, տառային հապավումներն իրենց գերակշիռ մասով հատուկ անունների արժեք ունեն, մինչդեռ հապավման մյուս տեսակները մեծ մասամբ հասարակ գոյականներ են: Թերևս այս վերջին հանգամանքով է պայմանավորված այն իրողությունը, որ հիշյալ բառարաններում մասային և մասաբառային հապավումներին ավելի շատ տեղ է հատկացված, քան տառայիններին:

Անժխտելի է, որ տառային հապավումները, որոնք գլխավորապես գոյականի արժեքով բաղադրյալ խիստ յուրահատուկ կազմություններ են, պետք է արտացոլվեն բառարաններում: Այլ հարց է՝ ի՛նչ չափով և ինչպե՛ս: Տառային հապավումների ընդգրկման ծավալի խնդիրը պայմանավորված է նախ լեզվական այդ միավորների նկատմամբ ունեցած մոտեցումներով, ապա նաև հապավումների ընտրության սկզբունքներով ու չափանիշներով:

Թերևս ճիշտ է ընդհանուր բառարանների մեջ այն տառային հապավումների ընդգրկումը, որոնք հասարակ գոյականների (հազվադեպ՝ ածականների) արժեք ունեն և առավել տարածված են ու կայունացած, ինչպես, ասենք՝ *ԱՀԿ* (ավտոմատ հեռախոսակայան), *ՀԷԿ* (հիդրոէլեկտրակայան), *ՋԷԿ* (ջերմային էլեկտրակայան), *ԱԷԿ* (ատոմային էլեկտրակայան), *ճՏՊ* (ճանապարհատրանսպորտային պատահար), օգգ (օգտակար գործողության գործակից), *ՔՊ* (քաղաքացիական պաշտպանություն), *ՏՏ* (տեղեկատվական տեխնոլոգիաներ), *ՀԻԱՀ* (ձեռքբերովի իմունային անբավարարության համախտանիշ), *ՌՕ* (ռազմածովային) և այլն, ինչու չէ՝ նաև առավել գործածական մասնագիտական որոշ հապավումներ՝ համապատասխան նշումով, ինչպես՝ *ՀՆԱ* (համախառն ներքին արդյունք), ՖԱԴ (ֆիզիկաաշխարհագրական դիրք) և այլն:

Սակայն բնական է, որ նույնիսկ նման մոտեցման դեպքում հայերենի հապավումների մի զգալի շերտ մնում է չներկայացված: Ակնհայտ է, որ այս բացը լավագույնս լրացվում է հապավումների բառարաններով: Խորհրդային շրջանում մեզանում եղել է հապավումների մեկ բառարան՝ Ալ.Մարգարյանի հեղինակած բառարանը[4], որի մեջ ընդգրկված են այդ ժամանակաշրջանի իրականության անվանող 4000 հապավումներ, որոնցից միայն 470-նն են տառային:

Հայաստանում անկախության հռչակումից հետո պետական, հասարակական-քաղաքական կյանքում, միջազգային հարաբերություններում տեղի ունեցած խոշոր տեղաշարժերի հետևանքով կրկին աշխուժացում նկատվեց բառապաշարի այս շերտում, ստեղծվեցին հարյուրավոր նոր տառային հապավումներ, կատարվեցին

[2] Նշենք, որ չնայած հիշյալ օրինակների բառարանային արձանագրմանը, դրանց նկատմամբ վերաբերմունքի հարցում միասնականություն չկա: Օր.՝ *զագս*-ը նշված ուղղագրական բառարանում ներկայացված է իբրև հապավում, իսկ բացատրականներում՝ ոչ, *դասվար*-ը Էդ.Աղայանի բացատրական բառարանում ներկայացված է որպես հապավում, ուղղագրական և ակադեմիական բացատրական բառարաններում՝ ոչ և այլն:

[3] Էդ.Աղայանի բացատրական բառարանում տեղ է գտել նաև *կենկրտում* հապավման՝ խոսակցական լեզվում ժամանակին գործածվող *ցեկա* (ռուս. ЦК) օտար համարժեքը՝ առանց որևէ նշումի:

[4] Ալ. Մարգարյան, Հայոց լեզվի հապավումների բառարան, Եր., 1979:

փոխառություններ: Միայն մի քանի տարվա ընթացքում պարբերական մամուլի էջերում ստեղծված ու գործածված տառային հապավումները բավականաչափ նյութ ընձեռեցին առանձին բառարան-տեղեկատուի[5] (շուրջ 1300 միավոր) համար, որի լույսընծայումից հետո ընդամենը 6-7 տարվա ընթացքում հապավումների թիվը առնվազն կիսով չափ ավելացավ, և տառային հապավումների մի նոր բառարանի[6] (շուրջ 2000 միավոր) ստեղծման կարիք զգացվեց: Սակայն մի բան ակնհայտ է. այսօր տառային հապավումների շերտում փոփոխություններն այնքան արագ են տեղի ունենում, որ տպագիր բառարանն իր լույսընծայումից կարճ ժամանակ անց արդեն «հնացած» կարող է համարվել:

Ժամանակակից տեխնոլոգիաներն այս իմաստով ավելի մեծ հնարավորություններ են ընձեռում. էլեկտրոնային բառարաններն օգտագործման համար ավելի հարմար են, կարող են ավելի արագ արձագանքել լեզվի բառապաշարում տեղի ունեցող տեղաշարժերին, մշտապես թարմացվել և ավելի հարուստ բառամթերք, այդ թվում նաև տառային հապավումներ ընդգրկել:

Կորպուսում ևս արևելահայերենի բառապաշարն իր բոլոր շերտերով, այդ թվում և տառային հապավումներով պետք է հնարավորինս լիակատար արտացոլվի: Դրա հետ կապված՝ որոշ խնդիրներ են առաջ գալիս, որոնց կփորձենք անդրադառնալ մեր հետագա շարադրանքում:

1. **Ո՛ր հապավումները և ի՛նչ ծավալով պետք է ներկայացվեն կորպուսում**

Քանի որ կորպուսում լեզվի բառապաշարը պետք է ներկայացվի հնարավորինս ամբողջական ձևով և իր բոլոր շերտերով, ուրեմն տառային հապավումներն էլ նրանում պետք է իրենց համապատասխան արտացոլումը գտնեն: Կարծում ենք՝ դրանք կարելի է ներկայացնել հետևյալ շերտերով (իհարկե, տարբեր համամասնությամբ).

ա) *համագործածական հապավումներ*

Սրանք տարբեր ոլորտներին բնորոշ, ակտիվ գործածություն ունեցող հապավումներն են, որոնք հանդիպում են մամուլի էջերում, գրքերում, ազդագրերում, ցուցանակներում, անգամ պիտակների վրա և թե՛ հատուկ, թե՛ հասարակ անունների արժեք ունեն:

Բնական է, որ տառային այս հապավումները կորպուսում ամենամեծ թիվը պետք է կազմեն:

բ) *տերմինային տարբեր համակարգերի առավել տարածված հապավումներ*

Այս դեպքում նկատի չունենք համընդհանուր գործածություն ստացած և այդ պատճառով նեղ մասնագիտական արժեքը կորցրած այնպիսի հապավումներ, ինչպիսիք են, ասենք, *ԱԱՀ, ՃԻԱՀ, ՄԻԱՎ, ԴՆԹ, ՌՃՈՒ, ՌՕՈՒ* և այլն: Խոսքը վերաբերում է տնտեսագիտական (*ՀԱԱ* - համախառն ազգային արդյունք / համախառն արտաքին ակտիվներ), *ՀԱՊ* - համախառն արտաքին պարտավորություններ, *ՀՓԻ* - հատուկ փոխառության իրավունք, *ՆԱԱ* - ներքին ազգային ակտիվներ և այլն), աշխարհագրական (*ՏԱԴ* -տնտեսաաշխարհագրական դիրք, *ՔԱԴ* - քաղաքաաշխարհագրական դիրք ևն), գիտատեխնիկական (*ԳԲԹ* - գերբարձր թույլատրելիություն, *ՕՀՄ* - օպերատիվ հիշող սարք և այլն), ռազմական (*ՌՏԿ* - ռադիոտեղորոշումային կայանք, *ՌԱՏ* - ռազմական ավտոտեխնիկա, *ՋՀԿ* - զենիթահրետանային կայանք և այլն), բժշկական (*ԿԱՀ* -

[5] Դ.Գյուրջինյան, Ն.Հեքեքյան, Հայերենի տառային հապավումների բառարան-տեղեկատու, Եր., 2000:
[6] Դ.Գյուրջինյան, Ն.Հեքեքյան, Հայերենում գործածվող հապավումների բառարան, Եր., 2007:

կենսաբանորեն ակտիվ հավելումներ և այլն) տերմինների արժեք ունեցող փոքրաթիվ հապավումների:

գ) *հնացած և պատմական հապավումներ*

Այսօր կյանքի տարբեր բնագավառներում տեղի ունեցող արագ զարգացումները հաճախակի փոփոխություններ են առաջ բերում տառային հապավումների շերտում: Կյանքի թելադրանքով ակտիվ գործածությունից դուրս են գալիս որոշ հապավումներ, ստեղծվում կամ փոխառվում են նորերը: Այս շարունակական գործընթացի հետևանքով երբեմն մի քանի տարվա կյանք ունեցող հապավումները վերածվում են պատմական (օր՝. *ԽՍՀՄ/ՍՍՀՄ, ՀԽՍՀ/ՀՍՍՀ, ԲԺՀ, ԼՂԻՄ* և այլն) կամ հնացած (օր.՝ *ՀՓԸ* - Հայաստանի թատերական ընկերություն, *ՀՄԻ* - Հայկական մանկավարժական ինստիտուտ և այլն) միավորների, որոնք, ճիշտ է, ակտիվ գործածություն չունեն, սակայն արժեքավոր են նրանով, որ պատմական որոշակի իրողությունների անվանումներ են, որոշակի ժամանակաշրջանի մասին տեղեկությունների կրողներ: Ուստի սրանց ներկայացումը կորպուսում պակաս կարևոր չէ:

դ) *փոխառյալ հապավումներ*

Տառային հապավումների մեջ զգալի թիվ են կազմում փոխառյալ միավորները: Խորհրդային շրջանում դրանց զգալի մասը փոխառված կամ պատճենված էր ռուսերենից կամ նրա միջոցով՝ այլ լեզուներից, ինչպես՝ *ՄԱԳԱՏԷ, ՅՈՒՆԵՍԿՕ, ՆԱՏՕ* և այլն: Մեր օրերում օրեցօր զարգացող ու ամրապնդվող միջազգային քաղաքական, տնտեսական, մշակութային հարաբերությունների պայմաններում աշխուժացել է տառային հապավումների ուղղակի փոխառությունը նաև եվրոպական լեզուներից, ինչպես՝ *ՅՈՒՆԻԴՕ, ՏԱՍԻՄ, ՏՐԱՍԵԿԱ* և այլն: Կարծում ենք, որ տառային հապավումների այս շերտը ևս պետք է իր արժանի արտացոլումը գտնի արևելահայերենի կորպուսում:

ե) *խոսակցական ոճում գործածվող հապավումներ*

Հապավումները բնորոշ են գրավոր խոսքի առանձին ոճերի և բանավոր խոսքում գրեթե չեն գործածվում: Չնայած դրան՝ մեկ-երկու տառային հապավում ստեղծվել է հենց բանավոր խոսքում, ինչպես՝ *խծք, բոմժ* (ի դեպ, վերջինը ռուսերեն կազմություն է): Սրանք այն եզակի օրինակներից են, երբ հապավումը բանավոր խոսքից թափանցում է գրավոր խոսք: Մամուլի էջերում, ուղեցուցակներում, հայտարարություններում և այլուր, սակայն, հանդիպում են նաև հապավումներ, որոնք օտարաբանություններ լինելով՝ բնորոշ են առավելապես առօրյա խոսքին, ինչպես՝ *ԳՈՒՄ, ԺԷԿ, ԳԱԻ և այլն*: Քանի որ կորպուսում լեզուն ներկայացվում է բոլոր շերտերով, այս փոքրաթիվ հապավումները ևս պետք է ներկայացվեն նրանում:

Տառային հապավումների խոսքաշարային արձանագրումը համապատասխան բնագրային օրինակներով, անշուշտ, կարևոր է, սակայն բավարա՞ր է արդյոք: Կարծում ենք, որ դրանք միայն *հպվ.* նշումով առանձնացնելը չպետք է վերջնական քայլը լինի: Լրացուցիչ նշումները (*պատմ., հնց., մանգ., խսկց.*) կարող են շատ օգտակար լինել կորպուսից օգտվողների համար, որոնք կկարողանան հեշտորեն և ավելի ճիշտ կողմնորոշվել իրենց անհրաժեշտ բնագրային նյութի ընտրության ժամանակ:

2. **Համանունություն**

Տառային հապավումների կազմության յուրահատկությամբ է պայմանավորված նրանց մեջ համանունության առաջացումը: Քանի որ տառային հապավումներն ստեղծվում են բառակապակցության բաղադրիչների սկզբնատառերով, բնական է, որ նրանց մեջ ավելի մեծ է գրությամբ և արտասանությամբ նույնական միավորների ստեղծման հավանականությունը: Մյուս կողմից էլ դրան նպաստում է այն, որ բառակազմության այս եղանակը խիստ գործուն է, հապավումների ստեղծման գործընթացը՝ անընդհատ: Հետևաբար համանունների թիվը գնալով ավելի ու ավելի է աճում, ընդ որում դրանք միայն զույգերով չեն հանդես գալիս. նույնիսկ քառանդամ, հնգանդամ համանունական շարքեր են արձանագրված: Սա, իհարկե, խրախուսելի երևույթ չէ, քանի որ բարդացնում է տառային հապավումները հասկանալու, դրանք մեկնաբանելու առանց այն էլ դժվարին խնդիրը, սակայն իրողություն է:

Համանուն տառային հապավումները ճիշտ հասկանալու և մեկնաբանելու գործում անշուշտ մեծ դեր ունի խոսքաշարը: Նրանում գործածված առանձին բառեր կարող են հուշել, թե համանուն հապավումներից ո՛ր մեկն է գործածված տվյալ համատեքստում: Հմմտ., օրինակ, *ՄԱԿ-ի Անվտանգության խորհրդի անդամները ԱՄՆ-ին կոչ են արել մարել իր պարտքը Միավորված ազգերին (ՀՀ,*1.04.00*)* և *Նա հավատարեց, որ իբրև քաղաքական կառույց՝ ՄԱԿ համբակցությունը նախագիծը մշակելիս առաջնորդվել է իր արժեքային համակարգով....* (ՀՀ, 14.09.06): Կասկած չկա, որ առաջին դեպքում խոսքը Միավորված ազգերի կազմակերպության մասին է (*Անվտանգության խորհուրդ, Միավորված ազգեր*), իսկ երկրորդ դեպքում՝ Միավորված աշխատանքային կուսակցության (*համբակցություն*):

Սակայն դեպքեր էլ կան, երբ անգամ նախադասությունը չի կարող կողմնորոշել համանուն հապավման ճիշտ մեկնաբանության մեջ: Շփոթ կարող է առաջանալ այն դեպքում, երբ համանուն են պետությունների, քաղաքական կուսակցությունների անուն հապավումները (օրինակ՝ *ԲՀ* - Բելառուսի / Բուլղարիայի Հանրապետություն, *ԼՀ* - Լեհաստանի / Լիբանանի Հանրապետություն, *ԱԺԿ* - Ազգային / Արցախի ժողովրդավարական կուսակցություն և այլն), ինչպես նաև այնպիսի հապավումներ, որոնց հիմքում ընկած բառակապակցությունները տարբերվում մեկ բաղադրիչով (օրինակ՝ *ԿԽ* - կենտրոնական/ կառավարման խորհուրդ, *ԱՄԿ* - Առողջապահության / Աշխատանքի միջազգային կազմակերպություն, *ՀԱԸ* - Հայաստանի աստվածաշնչային ընկերություն և Հայաստանի ավետարանչական ընկերակցություն և այլն): Օր.՝ *Ինչ վերաբերում է ԱԺԿ համամասնական ցուցակին, այն արդեն պատրաստ է* (ՀԺ, 02.03.07) նախադասությունից դժվար է կռահել, թե խոսքը ո՛ր կուսակցության մասին է: Նման դեպքերում, իհարկե, ավելի ծավալուն խոսքաշար է հարկավոր:

Բացի այդ, նշենք, որ հապավումների համանունական շարքում հանդիպում են թե՛ համագործածական, թե՛ սահմանափակ գործածություն ունեցող, այսինքն՝ հնացած, պատմական, երբեմն հազվադեպ կամ դիպվածային հապավումներ, որոնք, բնականաբար, նույն հարթության վրա ներկայացվել չեն կարող:

Ուշագրավ երևույթ է նաև հապավումների և միավանկ գոյականների համանունությունը, ինչպես՝ *ԿԱՊ* (Կոլեկտիվ անվտանգության պայմանագիր), *ՀԱՄԿ* (Հայաստանի ազգային սկաուտական կազմակերպություն), *ԱԽՏ* (Ասիախաղաղօվկիանոսյան տարածաշրջան) և այլն:

Կարծում ենք, որ կորպուսում համանուն տառային հապավումների ներկայացման հետ կապված խնդիրները հանգում են հետևյալին. ինչպե՞ս են դրանք մատուցվելու կորպուսից օգտվողին. միայն հաջորդաբար ներկայացվող բնագրային համապատասխան օրինակներո՞վ, թե՞ նաև հարակից մեկնաբանություններով ու նշումներով:

Տառային հապավումների, հատկապես նորակազմությունների առատությունն այսօր լուրջ դժվարություններ է հարուցում այն առումով, որ դրանցից շատերը կարող են պարզապես չհասկացվել, իսկ համանունների դեպքում՝ երբեմն նաև շփոթվել: Ուստի այս կազմությունների յուրահատկությունը հաշվի առնելով՝ գուցե արժե մտածել դրանց բացված տարբերակները (համապատասխան բառակապակցությունները) որևէ կերպ ներկայացնելու մասին:

3. **Տարբերակայնություն**

Հապավումների ստեղծումն ու գործածությունն այսօր մի տեսակ հախուռն և, կարելի է ասել, երբեմն նաև կամայական բնույթ ունի: Այդ պատճառով էլ հաճախ բախվում ենք լեզվական այնպիսի երևույթի, ինչպիսին տարբերակայնությունն է: Այն դրսևորվում է հետևյալ ձևերով.

ա) *հապավումների շարադասական տարբերակներ*

Սրանք բնորոշ են առավելաբար պատճենված հապավումներին: Ստեղծվում են աղբյուր կամ միջնորդ լեզվի բառակապակցության բաղադրիչների շարադասության պահպանմամբ կամ հայերենին բնորոշ շարադասությամբ: Հմմտ., օրինակ, *ԵՎԶԲ* (Եվրոպական վերակառուցման և զարգացման բանկ) և *ՎԶԵԲ* (Վերակառուցման և զարգացման եվրոպական բանկ), *ՄՀՀ* (Միջազգային արժութային հիմնադրամ) - *ԱՄՀ* (Արժույթի միջազգային հիմնադրամ), *ՄՀՖ* (Մշակույթի հայկական ֆոնդ) - *ՀՄՖ* (Հայկական մշակույթի ֆոնդ) և այլն:

Սրանց մեջ կան կայունացած և չկայունացած, ընդունելի և մերժելի տարբերակներ, և դրանց նկատմամբ վերաբերմունքի դրսևորման անհրաժեշտություն է ծագում, այլապես հավասարության նշան կդրվի մերժելի և ընդունելի ձևերի միջև:

բ) *պաշտոնական և ոչ պաշտոնական հապավումներ*

Գաղտնիք չէ, որ տառային հապավումներն իրենց գերակշիռ մասով հատուկ գոյականների արժեք ունեն: Դրանք մեծ մասամբ՝

– պետությունների, նրանց պետական և կառավարման բարձրագույն և այլ մարմինների *(ՀՀ, ԼՂՀ, ՌԴ, ԱՄՆ, ՌՀ, ՎՀ, ԱԺ, ԱԳՆ, ՆԳՆ, ԱԱՆ, ԿԳՆ* և այլն),

– կուսակցությունների, կազմակերպությունների, միությունների, հիմնադրամների (*ԱԺՄ, ՀՀԿ, ՀՌԱԿ, ԵԿՄ, ՀԲԸՄ, ՀՕՖ* և այլն),

– համաշխարհային, միջազգային և տարածաշրջանային կազմակերպությունների (*ԵՄ, ԵԽ, ԵԱՀԿ, ՄԽՎ, ՆԱՏՕ, ՄՕԿ, ՅՈՒՆԻՍԵՖ* և այլն),

– կրթական և մշակութային օջախների (*ԵՊՀ, ՀՖՆ, ԵԴԹ, ՊՀԹ* և այլն) անվանումներ են:

Հայտնի է, որ կուսակցությունները, կազմակերպությունները, հիմնարկությունները և զանազան այլ մարմիններ գրանցվելիս իրենց լրիվ անվան հետ հաճախ ստանում են նաև պաշտոնական կրճատ անվանումը, որը ոչ այլինչ է, եթե ոչ նրա անվան հապավված տարբերակը: Սակայն շատ հաճախ հատկապես մամուլի էջերում հանդիպում են հապավումների կամայականորեն ստեղծված տարբերակներ՝

առանձին բաղադրիչների կրճատումով կամ աղավաղումով: Օրինակ՝ Երևանի պետական լեզվաբանական համալսարանը պաշտոնապես ընդունված *ԵՊԼՀ*[7] ձևին զուգահեռ ներկայացվում է նաև *ԵՊԼՀ, ԵԼՊՀ*[8] տարբերակներով, Մարդու իրավունքների եվրոպական դատարանը *(ՄԻԵԴ)*՝ *ԵԴ* տարբերակով, Հայաստանի պետական ճարտարագիտական համալսարանը *(ՀՊՃՀ)*՝ *ԵՊՃՀ* (Երևանի պետական ճարտարագիտական համալսարան) ձևով: Հմմտ. նաև *ԵՔԸՀ - ԵՔԸՀ* (Երևանի քաղաքային ընտրական հանձնաժողով), *ՀՊՏՀ* (Հայաստանի պետական տնտեսագիտական համալսարան) - *ԵՏՀ* (Երևանի տնտեսագիտական համալսարան) և այլն: Այս վերջին օրինակում տարբերակայնությունն առաջացել է ոչ միայն տեղայնացնող բաղադրիչի փոփոխման, այլ նաև կարգավիճակը նշող բաղադրիչի (*պետական*) բացակայության հետևանքով: Այս հիմքով տարբերվող հապավումներն էլ քիչ չեն: Հմմտ. *ԵրՄՄԳՀԻ* (Երևանի մաթեմատիկական մեքենաների գիտահետազոտական ինստիտուտ) - *ՄՄԳՀԻ* (Մաթեմատիկական մեքենաների գիտահետազոտական ինստիտուտ), *ՀԳԱԱ* (Հայաստանի գիտությունների ազգային ակադեմիա) - *ՀԳԱ* (Հայաստանի գիտությունների ակադեմիա), *ԳԱԱ* (գիտությունների ազգային ակադեմիա)[9] և այլն: Սրանք բոլորն էլ կարող են ընդունելի տարբերակներ լինել, ուստի հարց է ծագում՝ անհրաժեշտ են արդյոք փոխադարձ հղումները և ի՛նչ չափով:

գ) *հայերեն կամ համարժեք օտար բաղադրիչով տարբերակային հապավումներ*

Երբեմն հապավումների տարբերակայնությունը հետևանք է համապատասխան բառակապացության մեջ փոխառյալ բաղադրիչի կամ հայերեն համարժեքի գործածության. հմմտ, օրինակ, *ՀԹԱ* (չբացահայտված թռչող առարկա) և *ՀԹՕ* (~ օբյեկտ), *ԳՄԿ* (Գաղթականների միջազգային կազմակերպություն) և *ՄՄԿ* (Միգրացիայի ~), *ՀԱՊ* (Հավաքական անվտանգության պայմանագիր) և *ԿԱՊ* (Կոլեկտիվ ~) և այլն: Նման տարբերակների դեպքում նախապատվությունը այս կամ այն ձևին տալը դժվար է, ուստի գուցե արժե պարզապես փոխադարձ հղումներ անել:

դ) *փոխառյալ հապավումներ և դրանց հայերեն համարժեքներ*

Հայաստանի ինտեգրումը միջազգային հանրությանը մի կողմից առաջ է բերում օտարալեզու տառային հապավումների մեծ հոսք, մյուս կողմից էլ խթանում դրանց հայերեն ընդունելի տարբերակների ստեղծումն ու կիրառությունը: Միջազգային զանազան հաստատություններ անվանող հապավումները հաճախ սկզբնապես փոխառվում են, ապա ստեղծվում են դրանց հայերեն համարժեքները, և սրանք երբեմն գործածվում են տարբերակներով: Օրինակ՝ *ՕՊԵՔ* (անգլ. Organization of Petroleum Exporting Countries) – *ՆԱԵԿ* (Նավթ արդյունահանող երկրների կազմակերպություն), *ԻՍՕ* (անգլ. International Standartization Organization) – *ՍՄԿ* (Ստանդարտացման միջազգային կազմակերպություն) և այլն:

Այսպիսի տարբերակների դեպքում ցանկալի է փոխադարձ հղումներ տալ *պե՛ս*-ով նախընտրելի, *հմմտ*-ով՝ ընդունելի ձևին:

ե) *տառադարձման և գրության զուգաչևությամբ տարբերակներ*

Երբեմն երկգրություն ունեցող հապավումներ են ստեղծվում ինչպես փոխառյալ հապավումները տարբեր կերպ տառադարձելու (*ՄԱԳԱՏԷ/ՄԱԳԱՏԵ,*

[7] Այս և նման կազմություններն պայմանականորեն ենք ընդգրկում տառային հապավումների խմբի մեջ:

[8] Սրանում առկա է նաև բաղադրիչների շարադասության տարբերություն՝ Երևանի լեզվաբանական պետական համալսարան:

[9] ԳԱ -ն, որպես ընդհանրական տարբերակ, համապատասխան տեղանունների հետ գործածվում է տարբեր երկրների գիտությունների ակադեմիաներն անվանելու համար, ինչպես՝ ՌԴ ԳԱ, Ուկրաինայի ԳԱ և այլն:

ՕՊԵՔ/ՕՊԵԿ), այնպես էլ առանձին հապավումների երկգրության (օր.՝ բուհ/ԲՈՒՀ, ՀԱՍԱՍ/«Համաս» և այլն) պատճառով։

Այս դեպքում ևս դրանք ներկայացնելու եղանակների խնդիր է առաջ գալիս։ Արդյոք երկգրություն ունեցող հապավումների բնագրային օրինակները կբացվեն առանձնաբար, նմանատիպ հապավումները միայն համապատասխան գրությամբ ներանցելու դեպքում, թե՞ միասնաբար, ցանկացած գրությամբ ներանցելու դեպքում։ Սրանով պայմանավորված՝ ինչպե՞ս պետք է դրսևորվի վերաբերմունքը դրանց նկատմամբ։

գ) *բառակապակցության ամբողջական և մասնակի հապավում*

Մի շարք դեպքերում ունենք բառակապակցության տառային ամբողջական հապավում և տառային հապավում՝ վերջին չհապավված բաղադրիչի հետ, ինչպես՝ *ՀԾԿՀ* և *ՀԾԿ հանձնաժողով* (հանրային ծառայությունները կարգավորող հանձնաժողով), *ԲԲԸ* և *ԲԲ ընկերություն* (բաց բաժնետիրական ընկերություն), *ԱԳՆ* և *ԱԳ նախարարություն* (Արտաքին գործերի նախարարություն) այլն։ Քանի որ նշված երկրորդ տարբերակներում հապավումը հանդես է գալիս ոչ թե անվանողական արժեքով, այլ որպես բառակապակցության հապավված բաղադրիչ, թերևս ճիշտ է սրանցից հղում տալ ամբողջական հապավումներին։

Հետաքրքրականն այն է, որ տարբեր է նաև այդ ոչ անվանողական արժեք ունեցող հապավումների կապակցելիության աստիճանը. եթե *ՍՊ* (սահմանափակ պատասխանատվությամբ), *ՓԲ* (փակ բաժնետիրական), *ԲԲ* (բաց բաժնետիրական) հապավումները կապակցվում են միայն *ընկերություն*, ՀՅ-ն՝ *դաշնակցություն* գոյականների հետ[10], ապա այլ է վիճակը *ԿԳ, ԿԱ, ՆԳ, ԱԳ* և նման այլ հապավումների դեպքում. սրանք կարող են կապակցվել մեկից ավելի գոյականների հետ (ասենք՝ *ԿԳ / ԱԳ / ՆԳ նախարարություն, նախարար, փոխնախարար, ԿԱ* (կոլեկտիվ անվտանգության) *պայմանագիր/համաձայնագիր, խորհուրդ, ուժեր* և այլն), ընդ որում առանձին դեպքերում դրանք համարժեք ամբողջական տարբերակ չունեն (օր. *ԿԳ նախարար, փոխնախարար*)։ Այստեղ արդեն խնդիր է առաջ գալիս. նմանատիպ հապավումների դեպքում պե՞տք են արդյոք ինչ-որ նշումներ կամ հղումներ, և ինչպե՞ս պետք է տրվեն դրանք։

4. **Տառանուններով կազմված հապավումներ**

Փոխառյալ տառային հապավումների մի առանձին ենթաշերտ են օտար տառանուններով կազմված հապավումները, ինչպես՝ *Սի-Էն-Էն, Բի-Բի-Սի, Էյչ-Էս-Բի-Սի* և այլն։ Ներկայացվելո՞ւ են արդյոք այս հապավումները և, եթե այո, ապա ինչպե՞ս։

5. **Քերականական հատկանիշների արտացոլումը**

Տառային հապավումներն իրենց խոսքիմասային պատկանելությամբ գերազանցապես գոյականի արժեք ունեցող կազմություններ են, սակայն, ինչպես արդեն նշվել է, կան նաև ածականի արժեքով գործածվող հապավումներ։ Այսպես՝ *ԵԱ* - երկարալիք, *ԿԱ* - կարճալիք[11], *ՈՒՄ* - ուլտրամանուշակագույն, *ԹԽՏ* -

[10] Թերևս այս պատճառով էլ նման գործածությունն արդարացված և խրախուսելի չէ, սակայն իրողություն է, որը, անկասկած, պետք է արձանագրվի։

[11] Սրանք, ի դեպ, բնորոշվում են խոսքիմասային տարարժեքությամբ և գործածվում են նաև գոյականի արժեքով՝ ԵԱ - երկար ալիք, ԿԱ - կարճ ալիք։

թռչնաբուծական, տավարաբուծական, խոզաբուծական, *ԳՀ* - գիտահետազոտական, *ՌԾ* - ռազմածովային, *ՄԴ* - սոցիալ-դեմոկրատական և այլն։

Գոյականի արժեք ունեցող տառային հապավումներին բնորոշ են որոշյալության, թվի և հոլովի քերականական կարգերը։ Բոլոր հապավումները, անկախ համապատասխան բառակապակցության վերջին բաղադրիչից, ենթարկվում են ընդհանրական *Ի* հոլովման։

Հայերենի առանձնահատկությունն այն է, որ հապավումները խոսքում հանդես են գալիս ինչպես ուղիղ, այնպես էլ թեք ձևերով[12], հարկ եղած սակավաթիվ դեպքերում նաև հոգնակիի կամ որոշյալ առման ձևով։ Օրինակ՝ *Կմասնակցեն չինական, ճապոնական ընկերություններ,* ՌԱՕ ԵԷՍ-ը (Ա,13.04. 00)։ *Գրեթե չեն կազմակերպվում հասարակական ցույցեր, հալածվում են անկախ* ՈԿԿ-ները (Առ, 30.09.99)։ *Բարչդրացրեց* ՋՕՄ-երի *աշխատանքի հարցը* (ՀՀ, 12.04.00)*։ 1960-ական թթ. ունիվերսալ* ԷՀՄ-ներից *բացի* ՍՄԳՀԻ-ում *ստեղծվել են մի շարք մասնագիտացված մեքենաներ* (ՀՀՀ, հ. 3, էջ 302)։ ՍՌՏ-ով *խորհրդային բանակի զինման գործում այլ ազգերի հետ մեկտեղ մեծ ներդրում ունեցան նաև հայազգի մասնագետները* (ՀԲ, 2000, 2, էջ 74)։

Այս առանձնահատկությունն էլ բոլորովին այլ մոտեցում է թելադրում արևելահայերենի կորպուսում հապավումների ներկայացմանը։ Հնարավոր կլինի՞ արդյոք այնտեղ գտնել հապավման համապատասխան ձևերի գործածության օրինակներ՝ որոնման ժամանակ նշելով նաև անհրաժեշտ քերականական հատկանիշները։

Այս և նախորդ հարցադրումները, թեև կարևորությամբ ու առաջնահերթությամբ տարբեր, սակայն սկզբունքային խնդիրներ են, որոնք արժանի ուշադրության և մանրակրկիտ ուսումնասիրության դեպքում, համոզված ենք, անհրաժեշտ լուծում կգտնեն։

Համառոտագրություն-հապավումներ

Ա - Ազգ
Առ - Առավոտ
ՀԲ - Հայոց բանակ
ՀՀ – Հայաստանի Հանրապետություն
ՀՀՀ – Հայկական համառոտ հանրագիտարան, հ.3, Երևան, 2000

[12] Ուղիղ և թեք ձևերի գործածության անմիօրինակություն կա միայն հատկացուցչի պաշտոնում հապավման գործածության մեջ։ Ըստ գործող կանոնի՝ տառային այն հապավումները, որոնց բաղադրիչներն արտասանվում են առանձնաբար՝ ուղեկցվելով չգրվող ը-ով, հատկացուցչի պաշտոնում պետք է գործածվեն ուղիղ ձևով (օր.՝ ՀՀ կառավարություն, ԵԽ անդամ), իսկ մյուս հապավումները՝ թեք ձևով՝ ի հոլովիչով (ՆԱՍԱ-ի աշխատակիցները, ՊՈԱԿ-ի հետ համատեղ)։ Այլ կիրառությունների դեպքում բոլոր հապավումները գործածվում են համապատասխան վերջավորություններով (օր.՝ ՀՀ-ի և ԼՂՀ-ի միջև, ԱՄՆ-ից)։ Այս մասին տե՛ս Տերմինաբանական և ուղղագրական տեղեկատու, Եր., 1988, էջ 311-312։

Լիանա Հովսեփյան,
Ռոբերտ Ուռուրյան, Սուսաննա Տիրյան

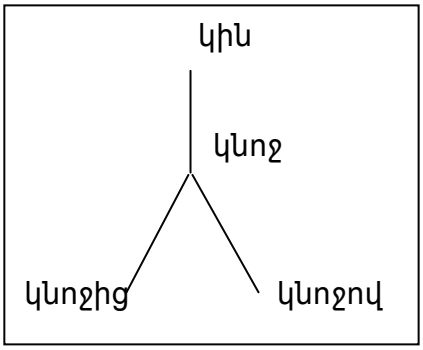
ՀԱՅԵՐԵՆԻ ԱՆՎԱՆԱԿԱՆ ԵՎ ԲԱՌԱՅԻՆ ԲԱՌԱՓՈՓՈԽՈՒԹՅՈՒՆ

Ժամանակակից հայերենում բառերի երկու կարգի փոփոխությունները, որոնք հայտնի են հոլովում և խոնարհում անվանումներով, բացահայտում է մի պարզ օրինաչափություն. հոլովվող բառերը ենթարկվում են մի կարգի փոփոխության, իսկ խոնարհվող բառերը՝ մեկ այլ: Իսկ խոսքի մեջ գոյականաբար օգտագործվող բոլոր բառերը թեքվում են գոյականների պես:

Բառերի այնպիսի փոփոխությունը, որին ենթարկվում են գոյական անունները և գոյականաբար օգտագործվող բառերը, կոչվում է անվանական բառափոփոխություն:

Բառափոփոխման գործընթացները նկարագրելու համար սովորաբար օգտվում են հիմք և վերջավորություն հասկացություններից: Այս կապակցությամբ ակադեմիկոս Գ. Ջահուկյանը գրում է. »Հիմք է կոչվում բառի յուրաքանչյուր իմաստակիր հատվածը, որից առանձնացվում է տվյալ եզրային՝ քերականական ձևույթը, թեև այդ հատվածը ինքնին կարող է նույնպես քերականական ձևույթներ պարունակել« (Ջահուկյան, Ժամանակակից հայերենի տեսության հիմունքները, Երևան, 1974, 159):

Այսինքն, բառափոփոխման յուրաքանչյուր փուլում որպես հիմք կարող է ծառայել մի այնպիսի ելակետային բառաձև, որն ինքնին բառափոփոխման իմաստով սերման արդյունք կարող է հանդիսանալ: Եթե խոսում ենք բառափոփոխման մասին ընդհանրապես, ապա պետք է ընդունել այդ գործընթացի մակարդակային բնույթը, երբ բառափոփոխությունը կարող է ընթանալ ինչպես վերևից ներքև, այնպես էլ հակառակը: Այսպես, օրինակ, **կին** բառաձևից նոր բառաքերականական իմաստների սերման սխեման հետևյալ տեսքը կունենա.



Սխեմայից պարզ երևում է, որ **կնոջ** բառաձևը սերման գործընթացում տեղ է գրավում միջին մակարդակում՝ այն կարող է ստացվել երկու ճանապարհով՝ ա) **կին** բառաձևից՝ **ոջ** երկրորդային ձևույթի կցումով և բ) **կնոջից** կամ **կնոջով**

բառաձևերից՝ համապատասխանաբար **ից** կամ կամ **ով** վերջավորությունների անջատումով։

Թվի կարգի ներկայացումը

Եթե նկատի առնենք անկախ գործածություն ունեցող տիպերը, ապա կարելի է նշել թվաձևային հատկյալ վերջավորությունները՝ **եր, ներ, իկ, այք, երք, ք, անք ենք, ոնք, ունք**, որոնք կցվում են հիմնային ձևույթներին։

Մի կողմ թողնելով անվանական թվի քերականական կարգի ծագման բացատրությունները՝ առաջնորդվենք Գ.Ջահուկյանի »Հայոց լեզվի զարգացումն ու կառուցվածքը« (Ջահուկյան 1969) և Է.Աղայանի »Ժամանակակից հայերենի հոլովումը և խոնարհումը« (Աղայան 1967) աշխատություններով ու բերենք հոգնակիակերտ երկրորդային ձևույթների գործածության բաշխումը՝ կախված ելակետային բառաձևից։

Ենթա-ձևույթ	Ընտրման պայմանները
եր	Միավանկ ելակետային բառաձևերի դեպքում (հանած իկ և այք ձևույթների համապատասխան ձևերը, տե՛ս ստորև)
ներ	Բազմավանկ ելակետային բառաձևերի դեպքում
իկ	**Մարդ** բառաձևի և նրանով վերջացող բարդությունների դեպքում (բայց՝ **ձևեմարդ - ձևեմարդեր**)
այք	**Կին, տիկին** և **պարոն** բառաձևերի դեպքում (բայց՝ տանտիկին – տանտիկիններ)
երք	**Տղա** և **աղջիկ** բառաձևերի դեպքում
անք	Հատուկ անուններ արտահայտող բառաձևերի դեպքում (**Վարդանանք**)
ենք	Հատուկ անուններ արտահայտող բառաձևերի դեպքում (**Հակոբենք**)
ունք	Հատուկ անուններ արտահայտող բառաձևերի դեպքում
ոնք	Միավանկ ազգակցական անունների (**պապ, տատ**), այլև ստացական դերանունների դեպքում (**մերոնք, ձերոնք**)
ք	Տեղանվամբ տվյալ տեղի բնակչին բնութագրելու դեպքում (**լոռեցիք, երևանցիք**)

Բերած աղյուսակից դժվար չէ հանգել այն եզրակացության, որ ժամանակակից հայերենի համար որպես նորմ ընդունվում է **եր-ներ** ձևույթներով հոգնակիացումը։ Այդ կապակցությամբ Մ.Աբեղյանը (Մ.Աբեղյան, Հայոց լեզվի տեսություն, Ե., 1965) գրում է »... նոր լեզվի մեջ հոգնակի թիվը միակերպ բոլոր հոլովների համար կազմվում է միավանկների վրա **եր** ավելանալով, իսկ բազմավանկների վրա՝ **ներ**, միայն պետք է նկատի ունենալ հետևյալ երկու պայմանը. 1. մի քանի բառեր հին լեզվում վերջանում են »ըն«, որը նոր լեզվի մեջ դուրս է ընկել. դրանց հոգնակի թվի մեջ այդ »ն« հնչյունը վերականգնվում է. օրինակ՝ մատ-մատներ, ձեռ-ձեռներ և այլն, 2. բարդ բազմավանկ բառերը, երբ երկրորդ բաղադրիչ մասը միավանկ է և պահում է իր նշանակությունը, ընդունում են **եր**. դասագրքեր, վառելափայտեր և այլն։ Բայց երբ մասը բազմարմատ կամ նոր նշանակությամբ մի բառ է կազմվում, ապա այս դեպքում ավելանում է **ներ**. մարդասերներ, մատենագիրներ և այլն« (337)։

Հոլովի կարգի ներկայացումը

Նախքան հոլովական համակարգի ներկայացումը, պարզենք հոլովական համակարգի ընտրության հարցը։ Բանն այն է, որ հոլովների քանակի ընտրությունը ժամանակակից հայերենի վիճելի հարցերից մեկն է։ Հայ քերականագիտության մեջ

ապացուցվում է հինգ հոլովի բավարար լինելը, մինչդեռ անհրաժեշտ է համարվում նաև ընտրել վեց, յոթ և նույնիսկ ութ հոլով: Այդ »լրացուցիչ« հոլովների առանձնացման համար,– գրում է Գ.Ջահուկյանը,– հիմք են ծառայում մասամբ ավանդույթը, մասամբ էլ այդ հոլովների շարահյուսական կիրառության առանձնահատկությունները: Հոլովների քանակի հարցում մենք առաջնորդվում ենք զուտ ձևային մոտեցումով՝ հարացուցային համակարգում ընդգրկելով այն նվազագույն քանակը, որը տարբերակվում է քերականական իմաստներով:

Ժամանակակից հայերենի համար կարելի է նշել սեռական հոլովական իմաստի 16 դրսևորում, որոնցից 12-ը՝ երկրորդային ձևույթների կցումով (օրինակ՝ **ա** (աղջկա), **ան** (գարնան), **անց** (մարդկանց), **ի** (քաղաքի, քաղաքների, աղջիկների և այլն), իսկ 4-ը՝ ներմարմնավորմամբ (օրինակ՝ **ա** (տան), **ե** (կայսեր), **ո** (հոր), **վա** (անվան):

Առաջին կարգի հոլովական ձևույթներ		Գործածության պայմանները		
		Ձևային	Իմաստային	Պատմական
ա		+	+	-
ի		-	-	-
ո		-	-	+
ու	ձի	+	-	-
	տալ	+	+	-
	մարդ	-	-	+
ան	գարուն	+	+	-
	խոստում	+	+	-
	մուկ	-	-	+
անց		+	+	-
յան		+	+	-
ոջ		-	+	-
վա		-	+	-
վան		-	+	-
ո		+	+	-
ե		+	+	-
ու > ա	մանկություն	+	-	-
	տուն	+	-	-
ու > վա		+	-	-
ք > ջ		+	+	-
յք > նց		+	+	-
իք > ոց		+/-	+	-

Սեռականի կազմության ժամանակ ընդհանուր առմամբ գոյություն չունի այս կամ այն գործոնով պայմանավորված կանոն և ամեն մի բառածևի սերման համար անհրաժեշտ է ցուցաբերել առանձին մոտեցում: Ասածը լուսաբանելու համար բերենք տարբեր լծորդության առաջին կարգի հոլովական ձևույթների գործածության ընթացքում առաջ եկող հերթագայությունների մի քանի օրինակ:

Առաջին կարգի հոլովական ձևույթներ	Հերթագայություն	Օրինակներ
ի	ի > ∅	սիրտ-սրտի
	ի > վ	կաշի-կաշվի
	յու > ∅	ալյուր-ալրի
	յու > ն	ձյուն-ձնի
	ու > ∅	թումբ-թմբի
	ու > վ	առու-առվի
ու	ի > ∅	ամուսին-ամուսնու
	ի > ∅	անկողին-անկողնու
ան	ա > ∅	ամառ-ամռան
	ե > ∅	ձմեռ-ձմռան
	ու > ∅	դուռ-դռան
վա	ի > ∅	ամիս-ամսվա
ոջ	ե > ի	տեր-տիրոջ
	ի > ∅	կին-կնոջ
	ու > ∅	սկեսուր-սկեսրոջ
	ույ > ∅	քույր-քրոջ
ա	ու > ա	տուն-տան
ո	այ > ո	հայր-հոր
	ե > ո	տատմեր-տատմոր
ո	ե > ի	սեր-սիրո
	յ > ∅	հույս-հուսո
	ու > ∅	սուգ-սգո
յան	ի > ու	հանգիստ-հանգստյան
	ու > ∅	ծնունդ-ծննդյան
ա	ի > ∅	աղջիկ-աղջկա
վա	ու > վա	անուն-անվան
ե	∅ > ե	կայսր-կայսեր

Երկրորդ կարգի հոլովական ձևույթներով արտահայտվում են ժամանակակից հայերենի բացառականի, գործիականի և ներգոյականի իմաստները: Ի տարբերություն առաջին կարգի հոլովական ձևույթների, երկրորդ կարգի հոլովական ձևույթների քանակը անհամեմատ քիչ է, որոնցից մի քանիսն ունեն գործառության որոշակի սահմանափակություններ:

Աղյուսակով ներկայացնենք երկրորդ կարգի հոլովական ձևերի կազմումը: Աղյուսակում ∅ նշանով տրված են այն դեպքերը, երբ սերման հիմքը առնված է ուղղականով, իսկ * և ∇ նշված են այն դեպքերը, երբ երկրորդ կարգի հոլովական ձևերի սերման հիմքը սեռականով առնված բառաձևն է, կամ վերջինից առաջին կարգի հոլովական ձևույթն անջատված հատվածը՝ համապատասխանաբար:

Ելակետային բառաձև	Առաջին կարգի ձևույթ	Հերթագայու-թյուն	ից	ով	ում	ուց	մբ
անկյուն	∅	ու → ա	∅	∅	∅		
արյուն	∅	ու → ա	∅	∅	∅		*
համեստություն	∅	ու → ա	∅	∅			*
նախադասություն	∅	ու → ա	∅	∅	∅		*
տուն	∅	ու → ա	∅	∅	∅		
դուստր	∅	∅ → ե	∅	∅			
մայր	∅	այ → ո	*	*			

աճուն	∅	ու → վա	∅	∅			*
հակոբենք	ց	ք → ց	*	*			
պապոնք	ց	ք → ց	*	*			
վարդանանք	ց	ք → ց	*	*			
կանայք	անց	այք → ∅	*	*			
մարդիկ	անց	ի → ∅	*	*			
տղերք	անց	ք → ∅	*	*			
գյուղացիք	ոց	իք → ∅					
աղջիկ	ա	ի → ∅	*	*			
գարուն	ան	ու → ∅	*	∅			
դուռ	ան	ու → ∅	∇	∇	∇		
թոռ	ան		∅	∅			
լեռ	ան		∅	∅	∅		
ծունկ	ան	ու → ∅	∇	∇	∇		
մուկ	ան	ու → ∅	∇	∇			
մանուկ	ան	ու → ∅	∅	∅			
որոշում	ան	ու → ∅	∅	∅			*
թեկնածու	ի	ու →	∅	∅			
կատու	ի	ու → վ	∇	∇			
ճու	ի	ու → վ	∇	∇	∇		
ջրեր	ի		∅	∅	∅		
ջուր	ի	ու → ∅	∇	∇	∇		
սեղաններ	ի		∅	∅	∅		
սեղան	ի		∅	∅	∅		
ընկեր	ի		∅	∅			
տեգր	ի		∅	∅			
հանգիստ	յան	ի → ∅	∇	∇	∇		
հույս	ո	յ → ∅	∅	∅	∅		
սեր	ո	ե → ի		*		*	
սուգ	ո	ու → ∅	∇	∇	∇		
ընկեր	ոջ		*	*			
կին	ոջ	ի → ∅	*	*			
տեգր	ոջ		∅	∅			
այսօր	վա		*	∅			
ամիս	վա	ի → ∅	*	∅	∇	∇	
ամիս	վա	ի → ∅	∇	∇	∇	∇	
ժամ	վա		∅	∅	∅		
տարի	վա	ի → ∅	*	∇	∇	∇	
մահ	վան		∅	∅	∅		
այգի	ու	ի → ∅		∇	∇	∇	
բանալի	ու	ի → ∅	∅	∅	∅	∇	

Հիմք ընդունելով քերականական տեսությունը՝ մշակվել են գործնական աղյուսակներ՝ վերլուծության ու սերման համապատասխան համար։

Ահա դրանք՝

Հայերենի անվանական հարացույցների ամփոփիչ աղյուսակ ձևաբանական վերլուծության համար

GNP

1. 2. #	Հոլովումներ Dec	Հիմքեր (Stem)			Հարացույցներ (Paradigm)				
		B	C	D	Եզակի թիվ		Հոգնակի թիվ		Հավաքական ClPl
					Sg1	Sg2	Pl1	Pl2	
3.	EInf ա	աղջիկ	աղջկ	-	1	0	2	0	0
4.	EInf ան	շարժում	շարժմ	շարժմամ	2	0	2	0	0
5.	EInf ան	բեռ	-	-	3	0	2	0	0
6.	EInf ան	գարուն	գարն	-	4	0	2	0	0
7.	EInf ան	դուռ	դռ	դռն	5	0	4	0	0
8.	EInf ան	ձուկ	ձկ	ձկն	6	0	4	0	0
9.	EInf ի	ծառ	-	-	7	0	1	0	0
10.	EInf ի	սենյակ	-	-	7	0	2	0	0
11.	EInf ի	անձ	-	-	8	0	1	7	0
12.	EInf ի	տղա	տղայ	-	9	0	2	0	0
13.	EInf ի	թուփ	թփ	-	10	0	3	0	0
14.	EInf ի	երկիր	երկր	-	10	0	4	0	0
15.	EInf ի	պարոն	-	-	8	0	11	2	0
16.	EInf յան	հանգիստ	հանգստ		11	10	0	0	0
17.	EInf յան	ծնունդ	ծննդ	-	12	0	2	0	0
18.	EInf ո	սեր	սիր	-	13	7	1	0	0
19.	EInf ո	հույս	հուս		14	0	1	2	0
20.	EInf ո	սուգ	սգ	-	15	0	1	3	0
21.	EInf ոջ	կին	կն	կան	16	0	9	0	0
22.	EInf ոջ	տիկին	տիկն	-	16	0	10	2	0
23.	EInf ոջ	ընկեր	-	-	17	0	2	0	0
24.	EInf ոջ	տեր	-	-	16	0	1	0	0
25.	EInf վա	օր	-	-	18	7	1	0	0
26.	EInf վա	ժամ	-	-	19	7	1	0	0
27.	EInf վա	ամիս	ամս	-	20	7	2	0	0
28.	EInf վա	ամառ	-	-	18	0	2	0	0
29.	EInf վան	աուուն	աո	-	21	0	2	0	0
30.	EInf վան	մահ	-	-	22	0	1	0	0
31.	EInf ու	ամուսին	ամուսն	-	23	0	2	0	0
32.	EInf ու	մարդ	-	-	24	0	8	0	0
33.	EInf ու	գյուղացի	գյուղաց	-	25	0	12	2	0
34.	EInf ու	այգի	այգ	-	26	0	2	0	0
35.	EInf ու	ձի			24	0	1	0	0
36.	IInf ա	տուն	տան	տն	27	0	5	0	0
37.	IInf ա	շուն	շան	շն	28	0	6	0	0
38.	IInf ա	անկյուն	անկյան	անկյամ	29	0	2	0	0
39.	IInf ա	էություն	էության	էությամ	30	0	2	0	0
40.	IInf ա	սյուն	սյան	-	31	0	1	0	0
41.	IInf ե	դուստր	դստեր	դստր	32	0	1	3	0
42.	IInf ե	կայսր	կայսեր	-	33	0	1	0	0
43.	IInf ո	հայր	հոր	-	34	0	1	0	0
44.	IInf ո	եղբայր	եղբոր	-	34	0	2	0	0

45.	EInf g	Վարդանանք	-	-	0	0	0	0	13
46.	EInf g	Հակոբեանք	-	-	0	0	0	0	14
47.	EInf g	պապոնք	-	-	0	0	0	0	15

Հայերենի ձևաբանական անվանական հարացույցներ

Եզակի թիվ

Աղյուսակ **NP.Sg**

1.	Հոլովներ (Case)						Հիմքեր (Stem)			Հոլովումներ (Dec)	
2. #	Ուղղ. Nom	Սեռ. Gen	Տր. Dat	Բաց. Abl	Գործ. Ins	Ներգ. Loc	B	C	D	Արտաքին թեքում EInf	Ներքին թեքում IInf
3.	**1**	**2**	**3**	**4**	**5**	**6**					
4.	**NP.Sg** Nom	**NP.Sg** Gen	**NP.Sg** Dat	**NP.Sg** Abl	**NP.Sg** Ins	**NP.Sg** Loc					
5.	B	C*ա	C*ան	C*անից	C*անով	*	աղջիկ	աղջկ	-		
6.	B	C*ան	C*անն C*անը	B*ից	B*ով D*ը	*	շարժում	շարժմ	շարժմամ	ա	-
7.	B	B*ան	B*անն B*անը	B*ից	B*ով	*	բեռ	-	-	ան	-
8.	B	C*ան	C*անն C*անը	B*ից C*ից	B*ով	*	գարուն	գարն	-	ան	-
9.	B	C*ան	C*անն C*անը	D*ից	D*ով	*	դուռ	դռ	դռն	ան	-
10.	B	C*ան	C*անն C*անը	C*ից D*ից	C*ով D*ով	*	ձուկ	ձկ	ձկն	ան	-
11.	B	B*ի	B*ին	B*ից	B*ով	B*ում	սենյակ	-	-	ան	-
12.	B	B*ի	B*ին	B*ից	B*ով	*	անձ	-	-	ի	-
13.	B	C*ի	C*ի	C*ից	C*ով	*	տղա	տղայ	-	ի	-
14.	B	C*ի	C*ի	C*ից	C*ով	C*ում	թուփ	թփ	-	ի	-
15.	B	C*յան	C*ի C*ին	C*ից	C*ով	*	հանգիստ	հանգստ		ի	-
16.	B	C*յան	C*յանր B*ին	C*ից B*ից	C*ով B*ով	*	ծնունդ	ծննդ	-	յան	-
17.	B	C*ո	C*ո C*ուն	C*ուց	C*ով	*	սեր	սիր	-	յան	-
18.	B	C*ո	B*ին	B*ից	B*ով C*ով	*	հույս	հուս		ո	-
19.	B	C*ո	C*ին	C*ից B*ից	C*ով B*ով	*	սուգ	սգ	-	ո	-
20.	B	C*ոջ	C*ոջն C*ոջը	C*ոջ	C*ոջ	*	կին	կն	կան	ո	-
21.	B	B *ոջ	B *ոջն B *ոջը	B *ոջից	B *ոջով B*ով	*	ընկեր	-	-	ոջ	-
22.	B	B*վա	B*վան	B*վանից B*ից	B*վանով B*ով	B*ում	օր	-	-	ոջ	-
23.	B	B*վա	B *ին	B*ից	B*ով	B*ում	ժամ	-	-	վա	-
24.	B	C*վա	C*վան	C*վանից C*ից	B*ով	B*ում	ամիս	ամս	-	վա	-
25.	B	C*վան	C*վանն C*վանը	B*ից	B*ով	*	անուն	ան	-	վա	-

26.	B	B*վան	B*վանն B*վանը	B*վանից B*ից	B*ով	*	մահ	-	-	վան	-
27.	B	C*ու	C*ու ն	C*ու ուց	C*ով	*	ամուսին	ամուսն	-	վան	-
28.	B	B*ու	B*ու ն	B*ուց	B*ով	*	մարդ	-	-	ու	-
29.	B	C*ու	C*ու*ն	C*ուց	B*ով	*	գյուղացի	գյուղաց	-	ու	-
30.	B	C *ու	C *ուն	C*ուց	B*ով C*ով	*ում	այգի	այգ	-	ու	-
31.	B	C	C* ն C* ր	C*ից D*ից	D*ով	B*ում	տուն	տան	տն	ու	-
32.	B	C	C* ն C* ր	D*ից	D*ով	*	շուն	շան	շն	-	ա
33.	B	C	C* ն C* ր	B*ից	B*ով D*ր	B*ում	անկյուն	անկյան	անկյամ	-	ա
34.	B	C	C* ն C* ր	B*ից	B*ով D*ր	*	էություն	էության	էությամ	-	ա
35.	B	C	C* ն C* ր	B*ից	B*ով	*	սյուն	սյան	-	-	ա
36.	B	C	C* ն C* ր	B*ից C*ից	B*ով D*ով	*	դուստր	դստեր	դստր	-	ե
37.	B	C	C* ն C* ր	B*ից	B*ով	*	կայսր	կայսեր	-	-	ե
38.	B	C	C* ն C* ր	C*ից	C*ով	*	հայր	հոր	-	-	ո

Հոգնակի թիվ
Աղյուսակ NP. Pl.

1.	Հոլովներ (Case)						Հիմքեր (Stem)		
2. #	Ուղղ. Nom	Սեռ. Gen	Տր. Dat	Բաց. Abl	Գործ. Ins	Ներգ. Loc	B	C	D
3.	**1**	**2**	**3**	**4**	**5**	**6**			
4.	**NP. Pl. Nom**	**NP. Pl. Gen**	**NP. Pl. Dat**	**NP. Pl. Abl**	**NP. Pl. Ins**	**NP. Pl. Loc**			
5.	B*եր	B*երի	B*երի(ն)	B*երից	B*երով	B*երում	ծառ		-
6.	B*ներ	B*ների	B*ների(ն)	B*ներից	B*ներով	B*ներում	սենյակ		-
7.	C*եր	C*երի	C*երի(ն)	C*երից	C*երով	C*երում	թուփ	թփ	-
8.	C*ներ	C*ների	C*ների(ն)	C*ներից	C*ներով	C*ներում	երկիր	երկր	-
9.	D*եր	D*երի	D*երի(ն)	D*երից	D*երով	D*երում	տուն	տան	տն
10.	D*եր	D*երի	D*երի(ն)	D*երից	D*երով	*	շուն	շան	շն
11.	B*ինք	B*անց	B*անց	B*անցից	B*անցով	*	անձ	-	-
12.	B*իկ	B*կանց	B*կանց	B*կանցից	B*անցով	*	մարդ	-	-
13.	D*այք	D*անց	D*անց	D*անցից	D*անցով	*	կին	կն	կան
14.	C*այք	C*անց	C*անց	C*անցից	C*անցով	*	տիկին	տիկն	-
15.	B*այք	*	*	*	*	*	պարոն	-	-
16.	B*ք	*	*	*	*	*	գյուղացի	գյուղաց	-
17.	B*անք	B*անց	B*անց	B*անցից	B*անցով	*	Վարդան	-	-
18.	B*ենք	B*ենց	B*ենց	B*ենցից	B*ենցով	*	Հակոբ	-	-
19.	B*ոնք	B*ոնց	B*ոնց	B*ոնցից	B*ոնցով	*	պապ	-	-

Հավաքական հոգնակի
Աղյուսակ NP.ClPl.

#	Հոլովներ (Case)						Հիմքեր (Stem)		
	Ուղղ. Nom	Սեռ. Gen	Տր. Dat	Բաց. Abl	Գործ. Ins	Ներգ.	B	C	D D
	1	**2**	**3**	**4**	**5**	**6**			
0	NP.ClPl. Nom	NP.ClPl. Gen	NP.ClPl. Dat	NP.ClPl. Abl	NP.ClPl. Ins	NP.ClPl. Loc			
1.	B*անք	B*անց	B*անց	B*անցից	B*անցով	*	Վարդան	-	-
2.	B*ենք	B*ենց	B*ենց	B*ենցից	B*ենցով	*	Հակոբ	-	-
3.	B*ոնք	B*ոնց	B*ոնց	B*ոնցից	B*ոնցով	*	պապ	-	-

Հայերենի բայական ձևակազմություն

Անվանական ու բայական հարացուցային միավորների կազմությունը անվիճելիորեն ընդհանուր բազմաթիվ կողմեր ունի, սակայն գոյություն ունեն նաև էական տարբերություններ, որոնք բխում են են ավանական և բայական բառաձևերի սերման սկզբունքից:

Այս կապակցությամբ Գ.Ջահուկյանը գրում է. »Եթե հոլովման դեպքում հնարավոր է միայն շարժում տեղում և դեպի աջ, ապա խոնարհման դեպքում հնարավոր է նաև շարժում դեպի ետ, որ դժվարացնում է կաղապարի գործողության ավտոմատացման հնարավորությունը: Այնտեղ գործ ունենք միայն հերթագայման և վերջնամասնիկավորման, այստեղ՝ նաև միջնամասնիկավորման և նախամասնիկավորման հետ« (Գ.Ջահուկյան, Ժամանակակից հայերենի տեսության հիմունքներ, Երևան, 1974, 223-224):

Հավելենք մի առանձնահատկություն ևս. եթե ժամանակակից հայերենի անվանական բառաձևերը, աննշան բացառությամբ, ունեն համադրական ձև, ապա բայական բառաձևերի կառուցվածքը տարասեռ է: Այստեղ առկա են և՛ համադրական, և՛ հարադրական ձևեր, որոնց կազմումը նոր բարդություններ է առաջացնում ձևային քերականության մշակման հարցում:

Աչքաթող չպետք է անել ժխտական ձևերի կազմության հարցը ևս, քանի որ այն պահանջ է դնում լրացուցիչ մասնակաղապարի. այստեղ ելակետ է ընդունվում ստորոգական բառերի հաստատական ձևերը, որոնցից կազմվում են նրանց ժխտականները:

Ելակետային բառաձևի ընտրությունը բայական բառաձևերի սերման հանգուցային հարցերից է: Եթե անվանական ձևերի սերման դեպքում այն ունի պարզ լուծում՝ իբրև հարացույցի ելակետային բառաձև ընտրվում է բառարանային ձևը, ապա բայական բառաձևերի համար հարցը միանշանակ լուծում չունի:

Ելակետային ձևի ընտրության հարցում կարևոր տեղ է հատկացվում անորոշ դերբային: »Սակայն, իբրև ելակետ, անորոշ դերբայի ընտրության դեպքում,– գրում է Գ.Ջահուկյանը,– ստեղծվում է մի այլ դժվարություն. հիմք ընդունելով **գր-ել, կարդ-ալ** տիպի ձևերը (արմատ + ել, արմատ + ալ կաղապարները)՝ քերականները ստիպված են **փախ-չ-ել, մոտ-ենա-ալ, իմ-անա-ալ** ձևերի դեպքում ընդունել, որ ձևաբանական առումով **գրել** և **կարդալ** ձևերը ելակետային կաղապարներ ընդունվել չեն կարող, և որ **փախչել, մոտենալ, իմանալ** ձևերը անկախ կաղապարներ են. հակառակ դեպքում մենք ստիպված ենք խոնարհման դեպքում

ընդունել միջմասնիկավորման եղանակ և բարդացնել խոսքի մոդելավորման (կաղապարման) գործողությունը ետադարձ շարժումով« (Գ.Ջահուկյան 1974, 223)։

Գործնական նպատակներով որպես բայի ուղիղ, ելակետային ձև կարելի է ընդունել անորոշ դերբայի ձևը՝ ***գրել, կարդալ,*** որը հավասարեցվում է գոյականների ուղիղ ձևին։ Ասվածը հաստատվում է անորոշ դերբայի կերպաժամանակային առանձնահատկությամբ. այն ո՛չ առարկայի վիճակ է արտահայտում, ո՛չ էլ գործողության կերպ. այն պարզապես անվանում է գործողությունը։ Հարցի այսպիսի լուծումը, իհարկե, կապված է նաև դժվարությունների հետ, հատկապես այնպիսիների, որոնց մասին նշում է Գ.Ջահուկյանը, թե հիմք ընդունելով արմատ-ել, արմատ-ալ կաղապարները քերականներն ստիպված են *փախ-չ-ել, մոտ-ենա-ալ, իմ-ան-ալ* ձևերի համար կազմել նոր անկախ կաղապարներ։

Բերենք հայերենի բայական հարացույցների ամփոփիչ աղյուսակը.

GVP

1.	Բայեր (հիմք)	**Դերբայ**	**Սահմանական**				**Ըղձա-կան**	**Հրամայական**	**Հիմքեր**	
			Անցյալ կատարյալ		**Ներկա, Անցյալ անկա-տար**					
2.	**StemB**	**Part**	**Ind Aor1**	**Ind Aor2**	**Ind Pres**	**Ind Ipf**	**Opt**	**Imp**	**StemC**	**StemD**
3.	գրել (-ել)	1	1	0	0	0	1	1	գր	գրեց
4.	կարդալ (-ալ)	2	1	0	0	0	2	2	կարդ	կարդաց
5.	հեռանալ (-անալ)	2	2	0	0	0	2	3	հեռան	հեռաց
	մոտենալ (-ենալ)	2	2	0	0	0	2	3	մոտեն	մոտեց
6.	գտնել (-նել)	3	2	0	0	0	1	4	գտն	գտ
	թռչել (-չել)	3	2	0	0	0	1	4	թռչ	թռ
7.	հեռացնել (-ացնել)	3	3	0	0	0	1	5	հեռացն	հեռացր
	մոտեցնել (-եցնել)	3	3	0	0	0	1	5	մոտեցն	մոտեցր
	կորցնել (-ցնել)	3	3	0	0	0	1	5	կորցն	կորցր
	դարձնել (-ձնել)	3	3	0	0	0	1	5	դարձն	դարձր
8.	աճել	3	3	4	0	0	1	6	աճ	ար
9.	առնել	3	2	0	0	0	1	7	առն	առ
10.	արժել	6	0	0	1, 2	1	0	0	արժ	-
11.	ասել	5	1	0	0	0	1	8	աս	ասաց
12.	բերել	1	3	0	0	0	1	9	բեր	բերեց
13.	գալ	4	2	0	0	0	3	10	գա	եկ
14.	գիտեմ	0	0	0	1, 2	1	0	0	գիտ	-
15.	դառնալ	2	2	0	0	0	2	11	դառն	դարձ
16.	դնել	3	3	4	0	0	1	12	դն	դր

17.	ելնել	3	2	0	0	0	1	13	ելն	ել
18.	թողնել	3	3	4	0	0	1	14	թողն	թող
19.	լալ	4	3	4	0	0	3	15	լա	լաց
20.	լվանալ	2	1	0	0	0	2	16	լվան	լվաց
21.	կենալ	2	2	0	0	0	2	17	կեն	կեց/կաց
22.	տալ	4	3	4	0	0	3	18	տա	տվ
23.	տեսնել	3	2	0	0	0	1	14	տեսն	տես
24.	ունեմ	0	0	0	1	1	0	0	ուն	-
25.	լինել	3	2	0	0	0	1	4	լին	եղ

Այս աղյուսակի հիման վրա կազմվել է ժամանակակից հայերենի բայածների վերլուծության և սերման մասնակաղապարները։

Աղյուսակ **Part**

	Դերբայական ձևեր **Part**								
	Անորոշ	Անկատար	Ապակատար I	Ապակատար II	Զուգընթացական	Վաղակատար	Հարակատար	Ենթակայական	Ժխտական
0	**VP. Part.Inf**	**VP. PartIpf**	**VP. PartFut1**	**VP. PartFut2**	**VP. PartAdv**	**VP. PartPer**	**VP. PartRes**	**VP. PartSub**	**VP. PartNeg**
1	B	C*ում	B*ու	B*իք	B*իս	C*ել	C*ած	C*ող	C*ի
2	B	C*ում	B*ու	B*իք	B*իս	D*ել	D*ած	D*ող	C*ա
3	B	C*ում	B*ու	B*իք	B*իս	D*ել	D*ած	C*ող	C*ի
4	B	*	B*ու	B*իք	B*իս	D*ել	D*ած	D*ող	C
5	B	C*ում	B*ու	B*իք	B*իս	C*ել	C*ած	C*ող	C*ա
6	B	*	*	*	*	*	*	C*ող	*

Աղյուսակ **Ind Aor**

	Սահմանական եղանակ **Ind**					
	Անցյալ կատարյալ **Aor**					
	Եզակի **Sg**			Հոգնակի **Pl**		
	1 դեմք	2 դեմք	3 դեմք	1 դեմք	2 դեմք	3 դեմք
#	**IndAorSgPrs1**	**IndAorSgPrs2**	**IndAor.Sg.Prs3**	**Ind.Aor.Pl.Prs1**	**Ind.Aor.Pl.Prs2**	**Ind.Aor.Pl.Prs3**
1	D*ի	D*իր	D	D*ինք	D*իք	D*ին
2	D*ա	D*ար	D*ավ	D*անք	D*աք	D*ան
3	D*ի	D*իր	D*եց	D*ինք	D*իք	D*ին
4	D*եցի	D*եցիր	D*եց	D*եցինք	D*եցիք	D*եցին

Աղյուսակ **Ind Pres**

	Սահմանական եղանակ **Ind**					
	Ներկա **Pres**					
	Եզակի **Sg**			Հոգնակի **Pl**		
	1 դեմք	2 դեմք	3 դեմք	1 դեմք	2 դեմք	3 դեմք
0	**Ind Pres Sg Prs1**	**Ind Pres Sg Prs2**	**Ind Pres Sg Prs3**	**Ind Pres Pl Prs1**	**Ind Pres Pl Prs2**	**Ind Pres Pl Prs3**
1.	C*եմ	C*ես	C*է	C*ենք	C*եք	C*են
2.	C*եմ	C*ես	C*ի	C*ենք	C*եք	C*են

Աղյուսակ **Ind Ipf**

	Սահմանական եղանակ **Ind**					
	Անցյալ անկատար **Ipf**					
	Եզակի **Sg**			Հոգնակի **Pl**		
	1 դեմք	2 դեմք	3 դեմք	1 դեմք	2 դեմք	3 դեմք
0	**Ind Ipf Sg Prs1**	**Ind Ipf Sg Prs2**	**Ind Ipf Sg Prs3**	**Ind Ipf Pl Prs1**	**Ind Ipf Pl Prs2**	**Ind Ipf Pl Prs3**
1.	C*էի	C*էիր	C*էր	C*էինք	C*էիք	C*էին

Աղյուսակ **OptFut**

	Ըղձական Opt											
	Ըղձական ապակատար **OptFut**						Ըղձական անցյալ **OptPast**					
	Եզակի **Sg**			Հոգնակի **Pl**			Եզակի **Sg**			Հոգնակի **Pl**		
	1 դեմք	2 դեմք	3 դեմք	1 դեմք	2 դեմք	3 դեմք	1 դեմք	2 դեմք	3 դեմք	1 դեմք	2 դեմք	3 դեմք
0	**OptFut Sg Prs1**	**OptFut Sg Prs2**	**OptFut Sg Prs3**	**OptFut Pl Prs1**	**OptFut Pl Prs2**	**OptFut Pl Prs3**	**OptPast Sg Prs1**	**OptPast Sg Prs2**	**OptPast Sg Prs3**	**OptPast Pl Prs1**	**OptPast Pl Prs2**	**OptPast Pl Prs3**
1	C*եմ	C*ես	C*ի	C*ենք	C*եք	C*են	C*եի	C*եիր	C*եր	C*եինք	C*եիք	C*եին
2	C*ամ	C*աս	C*ա	C*անք	C*աք	C*ան	C*այի	C*այիր	C*ար	C*այինք	C*այիք	C*ային
3	C*մ	C*ս	C	C*նք	C*ք	C*ն	C*յի	C*յիր	C*ր	C*յինք	C*յիք	C*յին

Քայական հարացույց
Աղյուսակ **VPImp**

	Հրամայական եղանակ **Imp**						
	B	**Imp/ NegImp**		Բուն հրամայական **Imp**		Արգելական **NegImp**	
		Եզակի **Sg**	Հոգնակի **Pl**	Եզակի **Sg**	Հոգնակի **Pl**	Եզակի **Sg**	Հոգնակի **Pl**
		2 դեմք **Prs2**	2 դեմք **Prs2**	2 դեմք **Prs2**	2 դեմք **Prs2**	2 դեմք **Prs2**	2 դեմք **Prs2**
0		**Imp/NegImp Sg Prs2**	**Imp/ NegImp Pl Prs2**	**Imp Sg Prs2**	**Imp Pl Prs2**	**NegImp Sg Prs2**	**NegImp Pl Prs2**
1.	գրել	C*իր	C*եք		C*եցեք		
2.	կարդալ	C*ա	C*ացեք				C*աք

3.	մոտենալ	D*իր	D*եք			C*ա	
	մոտենալ					C*ար	
4.	գտնել, լինել	D*իր	D*եք			C*իր	C*եք
5.	իջեցնել	D*ու	D*եք				
6.	անել	D*ա	D*եք			C*իր	C*եք
7.	առնել, ուտել	D	D*եք			C*իր	C*եք
8.	ասել	C*ա	C*եք			C*իր	
	ասել		C*ացեք				
9.	բերել, տեսնել	D	D*եք				
10.	գալ		D*եք	արի		C	
	գալ		C*ք	D		C*ր	
11.	դառնալ	D*իր	D*եք				
12.	դնել	դիր	D*եք			C*իր	C*եք
13.	ելնել	D*իր	D*եք	D		C*իր	C*եք
14.	թողնել	D	C*եք			D*իր	D*եք
	թողնել					C*իր	
15.	լալ	D	D*եք			լա	
	լալ					լար	
	լալ					D*իր	
16.	լվանալ	լվա	D*եք				
17.	կենալ	կաց	D*եք				
	կենալ	D*իր					
18.	տալ	տուր	D*եք				

Բաղադրյալ ժամանակներ

Աղյուսակ **Ind Pres**

	Սահմանական եղանակ **Ind**					
	Ներկա **Pres**					
	Եզակի **Sg**			Հոգնակի **Pl**		
0	**Ind Pres Sg** Prs1	**Ind Pres Sg** Prs2	**Ind Pres Sg** Prs3	**Ind Pres Pl** Prs1	**Ind Pres Pl** Prs2	**Ind Pres Pl** Prs3
1	PartIpf *#*եմ	PartIpf *#*ես	PartIpf *#*է	PartIpf *#*ենք	PartIpf *#*եք	PartIpf *#*են
2	PartAdv*#*եմ	PartAdv*#*ես	PartAdv*#*է	PartAdv*#*ենք	PartAdv*#*եք	PartAdv*#*են

Աղյուսակ **Ind Per Pres**

	Սահմանական եղանակ **Ind**					
	Վաղակատար **Per**					
	Ներկա **Pres**					
	Եզակի **Sg**			Հոգնակի **Pl**		
0	**Ind Pres Per Sg** Prs1	**Ind Pres Per Sg** Prs2	**Ind Pres Per Sg** Prs3	**Ind Pres Per Pl** Prs1	**Ind Pres Per Pl** Prs2	**Ind Pres Per Pl** Prs2
1	PartPer*#*եմ	PartPer*#*ես	PartPer*#*է	PartPer*#*ենք	PartPer*#*եք	PartPer*#*են

Աղյուսակ **Ind Res Pres**

	Սահմանական եղանակ **Ind**					
	Հարակատար **Res**					
	Ներկա **Pres**					
	Եզակի **Sg**			Հոգնակի **Pl**		
0	**Ind Pres Res Sg** Prs1	**Ind Pres Res Sg** Prs2	**Ind Pres Res Sg** Prs3	**Ind Pres Res Pl** Prs1	**Ind Pres Res Pl** Prs2	**Ind Pres Res Pl** Prs3
1	PartRes*#*եմ	PartRes*#*ես	PartRes*#*է	PartRes*#*ենք	PartRes*#*եք	PartRes*#*են

Աղյուսակ **Ind Fut Pres**

	Սահմանական եղանակ **Ind**					
	Ապակատար **Fut**					
	Ներկա **Pres**					
	Եզակի **Sg**			Հոգնակի **Pl**		
0	**Ind Fut Pres Sg** Prs1	**Ind Fut Pres Sg** Prs2	**Ind Fut Pres Sg** Prs3	**Ind Fut Pres Pl** Prs1	**Ind Fut Pres Pl** Prs2	**Ind Fut Pres Pl** Prs3
1	PartFut1*#*եմ	PartFut1*#*ես	PartFut1*#*է	PartFut1*#*ենք	PartFut1*#*եք	PartFut1*#*են

Աղյուսակ **Ind Ipf**

	Սահմանական եղանակ **Ind**					
	Անցյալ անկատար Ipf					
	Եզակի **Sg**			Հոգնակի **Pl**		
0	**Ind** Ipf **Sg** Prs1	**Ind** Ipf **Sg** Prs2	**Ind** Ipf **Sg** Prs3	**Ind** Ipf **Pl** Prs1	**Ind** Ipf **Pl** Prs2	**Ind** Ipf **Pl** Prs3
1	PartIpf *#*էի	PartIpf *#*էիր	PartIpf *#*էր	PartIpf *#*էինք	PartIpf *#*էիք	PartIpf *#*էին
2	PartAdv*#*էի	PartAdv*#*էիր	PartAdv*#*էր	PartAdv*#*էինք	PartAdv*#*էիք	PartAdv*#*էին

Աղյուսակ **IndPastPer**

	Սահմանականեղանակ **Ind**					
	Վաղակատար**Per**					
	Անցյալ **Past**					
	Եզակի **Sg**			Հոգնակի **Pl**		
0	**Ind Past Per Sg** Prs1	**Ind Past Per Sg** Prs2	**Ind Past Per Sg** Prs3	**Ind Past Per Pl** Prs1	**Ind Past Per Pl** Prs2	**Ind Past Per Pl** Prs2
1	PartPer*#*էի	PartPer*#*էիր	PartPer*#*էր	PartPer*#*էինք	PartPer*#*էիք	PartPer*#*էին

Աղյուսակ **IndPastRes**

	Սահմանական եղանակ **Ind**					
	Հարակատար **Res**					
	Անցյալ **Past**					
	Եզակի **Sg**			Հոգնակի **Pl**		
0	**Ind Past Res Sg** Prs1	**Ind Past Res Sg** Prs2	**Ind Past Res Sg** Prs3	**Ind Past Res Pl** Prs1	**Ind Past Res Pl** Prs2	**Ind Past Res Pl** Prs3
1	PartRes*#*էի	PartRes*#*էիր	PartRes*#*էր	PartRes*#*էինք	PartRes*#*էիք	PartRes*#*էին

Աղյուսակ **Ind Fut Past**

	Սահմանական եղանակ **Ind**					
	Ապակատար **Fut**					
	Անցյալ **Past**					
	Եզակի **Sg**			Հոգնակի **Pl**		
0	**Ind Fut Past Sg** Prs1	**Ind Fut Past Sg** Prs2	**Ind Fut Past Sg** Prs3	**Ind Fut Past Pl** Prs1	**Ind Fut Past Pl** Prs2	**Ind Fut Past Pl** Prs3
1	PartFut1*#*էի	PartFut1*#*էիր	PartFut1*#*էր	PartFut1*#*էինք	PartFut1*#*էիք	PartFut1*#*էին

Emilia Nercissians

LANGUAGE AS ACTION: ANALYZING PERFORMATIVES IN A CORPUS OF COMMENTARIES/ MONOLINGUALS' AND BILINGUALS NARRATIVES: FURTHER ANALYSIS OF A CORPUS

The main inspiration for the analysis presented in this doublet has stemmed from the requirement to present one of the two proposed contributions, in which two very different corpora were to be investigated (or, more correctly put, revisited) with rather different motivations. It came to light that there were common themes, which could more clearly be elaborated via examination of both cases. The central idea for both studies has been the mining of what can be construed as the difference between monolingual and bilingual modes of language use. That concern already carries the investigator beyond the confines of mainstream linguistics drawing upon ethnography of speaking, sociolinguistics, pragmatics, and similar neighboring disciplines. A starting point is to abandon considering language as means of expressing logical propositions that can be true or false. Rather, following the traditions of speech act, communicative action, and ecological communication theories, linguistic utterances and expressions are first and foremost considered as sociocultural and environmental actions. Like other actions, they are not primarily communicating meanings, but are responding to drives, in pursuit of intensions, reacting to circumstances, and trying to exert control and influence the world in which they are constantly engaged. The emphasis thus shifts from conveying and ascribing meaning to sense making. The analysis carried out in this research, therefore, will not be confined to the level of locutions. Consideration of illocutions and perlocutions, rather, will pose the focal problematic. In previous studies (Nercissians, 2003), strategies of contextualization and identification were primary thematic for explication. Various aspects of those processes have been subjects of consecutive investigations that have already been reported, in each case associated with corresponding theoretical development. The present work intends to go further, where a return to mainstream linguistic will be within sight.

The so called "triple C" model for language use, elaborated more than a decade ago by the author (Nercissians, 1996, 1998), has motivated several studies about the communicative, cognitive, and cultural aspects of language use, especially in bilingual and multilingual settings, in conjunction with various projects and dissertations of colleagues and students. However, in more recent investigations (Nercissians, 2005), attention has shifted away from comparison of competences between language use in each sphere towards elicitation of motivations and intentional stances that give rise to those behaviors. When looked from the viewpoint of language as action, vernacular languages and cultures, it can be argued, encompass ideologies, modes of perception, and implicit knowledge that are often neglected in mainstream sociolinguistics. With the language come the knowledge, and historically shaped modes of perceptualization and conceptualization, which as a result of sociolinguistic exclusion are lost to the determent of all sides, as a result of not being leveraged and utilized.

On the other hand, there are many commonalities, universal attributes, and shared background assumptions, similarity of experiences, among members of different ethnicities living in a common polity that allows not only mutual understandability through similar interpretations, but also collaborative sense making and construction of shared cosmovisions, which escape the attention of those researchers who do not problematize what was taken for granted and never expressed linguistically. [Common references are cited at the end of the second contribution]

I. MONOLINGUALS' AND BILINGUALS NARRATIVES: FURTHER ANALYSIS OF A CORPUS

The corpora of verbal narratives chosen for further critical analysis was initially collected in conjunction with a thesis project carried out under the author's supervision in order to compare the linguistic competences of monolingual and bilingual schoolgirls. The study, despite careful selection of comparison metrics and experimental design, did not reveal significant differences substantiating any theoretical stance on bilingual communicative and cognitive abilities in educational linguistics. In contrast, a subsequent analysis of the same corpora, although having been collected for a different purpose, showed marked difference with respect to verbal explicitness verses context- embeddedness of the chosen codes by the subjects belonging to monolingual and bilingual groups asked to narrate a story illustrated by four consequent images (Nercissians, 1996, 1998). It was also argued that the difference need not be interpreted, as suggested by several well known theories on language acquisition and use of the working class, ethnocultural minorities, children, and primitives, in terms of educational and linguistic disadvantage and differential capabilities. An alternative interpretation was developed according to which contextualization strategies were viewed in terms of identity- oriented and solidarity-stressing linguistic behavior especially adopted by communities characterized by closeknit networks making reference to their common background. At the same time, contextualized narration can also be viewed as creative strategy for compensating for lexical and other limitations experienced by second language learners (Nercissians, 1998, 2000, 2003). The further analysis presented here intends to go beyond those premises and show the inadequacy of the very exophora- endophora distinction drawn by functionalist linguistics. The anaphoric relation in a narration can refer not to the exact and literal referent in some other part of the narration, but to the mental representation of the hearer(s). The question of discourse representation is also important. The narrative expresses the narrator's specific representation of the world, which is integrated with her more general world model. Hearers then are expected to build their own representation of what has been communicated by the narrator. To minimize possible mismatches, narrators are typically expected to take into consideration the developing representations of the audience especially, in one sided discourses like narrations, those features of their propensities they can depend upon, so that hearers are able to use in trying to identify intended referents. Analysis of the corpora shows that exophoric cannot always be the same as explicit. Decontextualized narratives can thus be conceptualized as discourses that are merely recontextualized to take into account the formality conventions of the dominant culture (Young, 1981). Alternatively, less formally organized minorities have, in addition to the exophoric and inter- narrative regularity features that can help hearers in interpreting intended meanings in

particular narrative fragments, important extra- narrative means in the form of socio- cultural knowledge and situational cues that can serve as reinforcement. Further steps are taken towards providing explanations beyond the confines of interest of pure linguistics through referral to speech act theory, discourse analysis, artificial intelligence, knowledge management, and poststructuralist textual analysis. The focus of attention is shifted towards the intentional and representative aspects of narration. A two dimensional model elaborated in the past studies of the author is restated in this new light. The extended model is then used for analyzing the differences between the monolingual and bilingual corpora from a critical standpoint.

THE NARRATIVES

The corpora under investigation was obtained by asking second grade schoolgirls in Educational zone eight in Tehran, where there is sizeable concentration of the Iranian Armenian community to narrate the story depicted in the four images shown in Fig. 1. The respondents were divided into two: monolingual (majority) and bilingual (minority) groups. The initial intention for collecting the data was to compare the linguistic competences of the two groups; therefore, no attempt had been made to select images that could facilitate other studies as well.

ANALYSIS

Initial investigations conducted in partial fulfillment of a thesis under the author's supervision revealed the Iranian Armenian community under investigation had linguistic competences in all components comparable with those of monolinguals. However, later research carried by the author showed that there were clear differences between the language use patterns of the two groups. Here are two typical examples of how the story was narrated by two respondents.

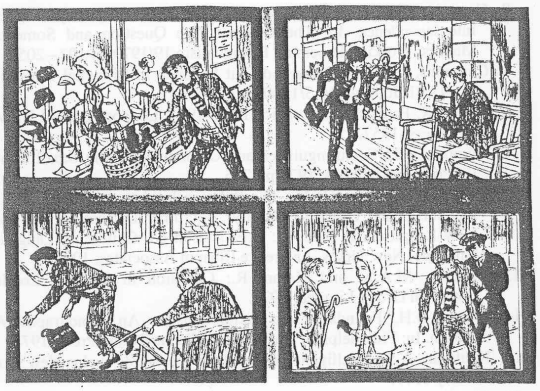


Fig. 1: The Story

"This lady has come across this store. She is watching the hats; and this man took that from her bag. And here, the woman found out that this man has taken her (hand)bag from her bag. And then the woman runs after him. And this man too, when sitting here, holds

his walking stick between that thief's feet. Then the bag falls from his hands. Here the police comes and catches him. And that man hands her bag. This woman thanks that man.

"A thief took a woman's bag and ran. A man was sitting on a chair. That man was able to detect that this is thief. A police came and arrested the thief; and the bag returned to the woman".

The examples show a very distinct divide between narrative strategies of the two groups. Half of the twenty two respondents from the group of monolinguals had narrated the story without any reference to the images, while only two, out of twenty two respondents in the bilingual group had used the "decontextualized" code and the rest had context embedded narratives. This can give rise to various language deficit or disadvantage theories. Context embedded codes may support high communicative competence, but low cognitive competence theories. It may also indicate tendancy towards use of restricted codes. Generally, context embeddedness has been said to mark primitiveness and immature communication. In discourse analysis, that narrative style can be interpreted as failure to follow the introduction then follow-up routines marking textuality and coherence. Models and analysis proposed by the author, however, furnished alternative deductions. Context embedded forms of narratives could also indicate more or less deliberate and conscious solidarity enhancing strategies on the one hand, and intelligent means of compensation for lexical or other linguistic difficulties on the other. Bilingual speakers, it is argued, are capable of using more limited linguistic repertoire without showing any communicative disadvantage, because they have compensatory measures through leveraging context more often. This line of argument carries us beyond the limits defined by the dominant functionalist paradigm prevailing in sociolinguistics towards more interactionist, constuctivist, and dialectic stances. Now it can be further argued that majority and minority subcultures do not converge in two dimensional status- solidarity plane. Identity oriented and solidarity stressing attitudes is to be encountered more often in minority settings, while majority attitudes often attach more significance to status and prestige dimension. Social creativity motivates bilinguals who belong to some ethnolinguistic minority towards stressing, enhancing, as well as leveraging the common identity and discounting the requirements of formal discourse. Consciousness towards higher prestige and social status, on the other hand, inclines majority group members, who have the dominant subculture, towards higher levels of abstraction and formality. Returning to the discourse technique of using indefinite clause when introducing someone or something, and definite clause as the person or thing or concept is being tracked throughout the story, the question might be asked whether not using that technique could necessarily be interpreted as linguistic deficit in conceptualization of indefinite categories, or at least utilization of discourse formalities for enhancing cohesion. A more critical analyst could hardly fail to observe that the narrators were given the task of telling the story from a set of only four images, the cohesion of which could itself be a subject of analysis. In effect, the subjects were asked to retell a story that has already been narrated pictorially. So their discourse genre could be considered as commentary rather than narration. Anaphoric references would thus not necessarily be a failure to observe the discourse formalities, if a consciousness of their lack in the original unfolding of the story were there. How much are the subjects conscious of the existing of diverse subcultures that could support differential interpretations of the story as told by the images, so that they would rather leverage what they assumed common background that existed outside of the text and the stories (the story narrated by them, as well as the story

narrated by the images)? And, one cannot fail to eventually pose the question: is telling the story the whole purpose of the speech act? Are there not many other things that need be said (done) during the discourse? One here can hardly fail to appreciate the notable commonality of stances. For example, none of the narrators have tried to tell the story from the thief's viewpoint (e.g. the thief being in dire need, all the people- the old man, the police, and the woman- acting so as to prevent his satisfying the need by transgressing property relations). Surely, it is not merely the images' use of the usual techniques of textuality that has excluded alternative interpretations (e.g. the woman's son wanted to buy something his mother disapproved of, but the old man and the police would not let him). The story, irrespective of the subculture of its narrator or addressee, both uses and reinforces dominant cultural elements like "women are week and need be protected", "property is right and should not be transgressed", "police serve people", "old men are respectable and wise", "there are many bad men around that are young and agile so that one always has to be on guard" etc. It is only within this network of common dominant ethical presuppositions, social norms, and behavioral patterns (running for escaping) that subcultural elements are allowed to show themselves. For example, a comparison reveals that minority narrators had more tendency towards descriptive generalizations (e.g description of what the woman is busy doing in front of the store, what the thief tells the police as he is being arrested, where the thief is taken afterwards, how the old man was alerted to the situation). This also demonstrates their higher level of consciousness of the need to interpret the discrete set of pictures. For example, making sense of the fact that the old man interferes and stops the free running of the young man would be very problematic, unless it is made clear that not only has the young man broken an important ethical norm, but also the fact is known to the old man as well. The bilingual narrators resolved that problem by telling that the old man had heard the woman's request for help, or had seen the act of stealing. In the case where the monolingual narrator had also appreciated the need for an explanation, she had sufficed to mention that the old man had detected that the running man was a thief (not bothering to narrate how). Another fine point to be mentioned is the part of the story narrating the old man's action. The author has elaborated elsewhere (Nercissians, 2003), how bilingual narrators have used contextualization strategies in order to compensate for their not using the word "walking stick". But further elicitation can also lead to the deduction that the contextualization strategy also helps in avoiding the use of possessive dexis. (For example, instead of saying "the old man extended his walking stick", it can be said that "the old man extended the wooden stick"). In all cases of the monolinguals' narrations the terms walking stick or hand stick has been mentioned together with a possessive pronoun. On the other hand, not only have the bilingual narrators often used contextualization for describing what was being extended, inhibiting the free run of the young man, but also in all but one case no possessive definite marker has been used for identifying that thing (stick).

CONCLUSIONS

In this analysis, it has been attempted to elucidate differential strategies adopted by majority and minority schoolgirls when narrating a story beyond what has been conjectured by the author in previous studies. Every step presents further critique of pathological views towards language use patterns. Instead of looking for grammars and

norms of correct use patterns and deviation from those norms indicating incompetence, dysfunctional behavior, and cognitive deficit, the researcher should try to explain why people leverage different strategies in selecting their actions. The key point here is to view language as an action that is more than communicating a mental state or a logical proposition. Language use also encompasses expression of viewpoints, attempts towards forging social bonds and links, and most of all endeavors seeking to reach certain goals. Of course, not all actions of ours are deliberate or volitive or intentional, accessible to our consciousness or awareness. We do not select actions for no reason even if we do it subconsciously. Stories tell not only what has happened, but also the associated value attachments that reflect the historical experiences of sociocultural groups. It is not merely the semantic aspects that need to be investigated but the semiotic and signifying moments too are in need of explication. Minorities, it is argued, tend to use specific procedures of contextualization that can help them take advantage of, stress and express alignment with, and in final analysis reinforce and reproduce common backgrounds. The more closeknit social networks they possess, the less dominant position the occupy in the greater polity, the more they will be likely to utilize this strategy. They are more conscious of prevailing sociocultural and sociolinguistic diversities and use discourse strategies that are more anaphoric. Their language use is geared towards identity stressing and solidarity enhancing discourse technologies. For example, they often take advantage of sympathetic circularity through use of expressions like “she does thus” or “he does that thing”. The language use patterns of the dominant group members, on the other hand, are often discernable as prestige and status seeking actions; not to be designated as decontextualized, but formal.

There are three further aspects that need be mentioned. The first is the notion of spatial deixis. The notions of text deixis and discourse deixis has been studied extensively. Shingo Imai (2003)Imai presents a comparative investigation of how space surrounding the speaker is demarcated via different languages. However, a microsociolinguistic and microethnopragmatic research has very different needs than pure linguistic research. Can it be that minority group members find it more useful to leverage “anchors” than majority group members irrespective of the language they use? Second is the concept of argument ellipsis, which is the subject of the research conducted by Shigeko Nariyama (2000) for Japanese. It elucidates the linguistic mechanisms with which to identify the referents of the ellipted arguments. And the third construct is designated as linguistically specific features of discourse, which according to Alida Anderson (2006) is an important facet of the competence children acquire in early childhood, which enables them to convey information and use language appropriately in different discourse situations. It can be argued that the less sociocultural distance there is between interlocutors, the more exophoric and internarrative regularity features and shared knowledge items and situational cues will tend to be leveraged to direct hearers/readers toward certain interpretations of the text.

II. LANGUAGE AS ACTION: ANALYZING PERFORMATIVES IN A CORPUS OF COMMENTARIES

The multilingual corpus revisited for further analysis consists of commentaries expressed via Armenian, Farsi (Persian), as well as several other languages; even via drawings and cartoons in a few cases, in 1993 by visitors of an art exhibition showing the paintings of S. Parajanov. The exhibition was held by authorities representing the Parajanov museum in Armenia, following invitation by Iranian authorities in movie industry, which, at that time, was in the process of policy shifts towards further liberalization (Nercissians, 1995, 2000). The theoretical frameworks for the analysis are elaborated in the first part of this paper. A focal point is to go beyond mere consideration of sentences and linguistic elements as constructs and premises used for encompassing truth values. Extending the models developed via speech act theory and communicative action approach, the commentaries are considered as intentional acts representing, in addition to content statements, communicative, cognitive and cultural goals and values. The conceptualization of text poses another question. Textual corpus involves more than written aspect of the language. Firstly, it must have textuality as problematized by functionalist linguistics. But more importantly, especially due to the action- oriented approach adopted in this paper, texts are not only writerly, but also readerly. The productivity aspect of the text is important for poststructural analysis (Young, 1981). Rather than being the product of communicative effort, it is the very theatre of production where the producer and reader of the text meet. Even the written and fixed text starts working whenever it is taken and subjected to doxa. It is, therefore, the purpose of the analysis presented in this paper to go beyond criticism in traditional as well as structuralist senses: trying to reveal the latent intention, to express what has not been stated, but by rereading what has been stated. The text incorporates many voices that need to be untangled. In the continuation of the paper the commentaries of the corpus under investigation are examined. An important aspect is the analysis of intertextuality, which refers to the ways texts derive their meaning from other texts. In addition to the expressed locutions, illocutions and perlocutions are also carefully taken into consideration. Among the internal procedures used for controlling each discourse are the concepts of major narratives and communities of discourse. Of particular interest in our study, is the elaboration of the commentaries in terms of an extended two dimensional model representing status and solidarity orientations. Finally, another most important component of our analysis of what is actually said in commentaries is the critical examination of what is never said. It is concluded that commentaries can be studied in an ethnomethodological sense as technologies people use to overcome the instability of commonsense knowledge creating shared meanings anew in each new encounter.

THE COMMENTARIES

The corpus under study consists of the comments written by the participants in an art exhibition displaying the works of Sergey Parajanov. The exhibition of the paintings took place in Tehran in 1993 by individuals that were involved in the movie industry and the Parajanov museum. The event, one of the first organized jointly by Institutions affiliated with

Iranian and Armenian states after the downfall of Soviet Union, brought the promise of more intense cultural relations between the two countries and a movement away from the cold war that had exerted severe limitations, both politically and culturally. The fact that Parajanov is considered as a major figure in Soviet dissident movement, and that his art style is also far from being simple and conventional is also significant. Among the participants, many were from the Iranian Armenian community; a community that had good reasons to resent the cold war tensions. Irrespective of which side they sympathized with, most Iranian Armenians would have welcomed closer ties with Armenia. Some even had relatives in that country to whom they could not even pay visit without undue difficulties, especially before the regime change in Iran (which happened before the independence of Armenia). The corpus is multilingual because participants were free to write in any language. There were also many non- Iranian participants, showing that the event was of international significance; and many commentaries have been written in English as well as other languages not spoken by the different ethnicities living in Iran as their vernacular tongues. Names also suggest that there were many foreign visitors. Finally, some participants had chosen non linguistic means (like drawing) to express themselves. Since the entries were results of voluntary action, the fact that someone had decided to express herself in the commentary book presupposed high level of motivation, although at least one writer had mentioned that his commentary has been written due to the "insistence of friends". Following the line of reasoning introduced in the companion paper, we start by considering the writing (or drawing) of comments as actions rather than mere linguistic sentences. So our attention is drawn to the intentional stances of the commentary writers rather than the meanings of what they have commented (Nercissians, 1996, 1998, 2003, 2005). They have entered their comments because they have felt a need to express themselves; and perhaps to offer their interpretations of the works exhibited. In trying to elaborate on the intended meanings or even, like we endeavor to do, the intensional stances of Parajanov himself, commentaries ought to "say for the first time what had, nonetheless, already been said, and tirelessly repeat what had, , however, never been said".

ANALYSIS

One of the most important things in discourse analysis is to discuss not only what has been said, but also, and more importantly, what has not been said. The most important thing that strikes the reader is the large scale absence of what one would expect would constitute the main theme of commentaries: analysis of artistic expression and style of Parajanov. Of course most writers have written one or two short sentences on the beauty or excellence of his work, but have immediately shifted to discussions on how good it is to have such events organized, what they think constitutes the essence of Armenian or Oriental thinking, how they have seen others make do with simple things to produce works of art, etc. One commentary, for example, only mentions "long live the Armenian". Clearly, despite the fact that all those secondary topics could also be considered as discussions of various aspects of Parajanov's work, one can safely conclude that most commentary writers have used the occasion to express opinions that they would also express in any other occasion and in radically dissimilar settings. Another striking element in commentaries is the high frequency of calls for further explanation of Parajanov's work and criticism of the organizers' work because of their failure to provide

further explications. This is especially evident in those commentaries that can be inferred to have been written by Armenian participants either because they have chosen to write in Armenian or their names indicate their ethnic origin. It should be noted that in Iran (as perhaps in other countries too), artists especially resent being asked to explain their works. Ironically, this request is frequently made in artistic gatherings. Artists think their works are already the expression of their thoughts and the call for explanation amounts to the declaration of their failure to do the explanation through their works. One important aspect of the commentaries that has been analyzed by the author in past works is the collective construction of identity (Nercissians, 1999, 2003a). The use of pronouns, especially the first person plural "we" is very indicative. Examples like "we are anticipating more works of his…" or "we are proud of…" abound. More generally, the use of deictic expressions (the so called shifters or indexical expressions) can be viewed as performing important communicative functions as identification in terms of relations to the ongoing interactive context in which the discourse takes place. Examples are "I hope in the near future we will have such artists (end of comment)", or "It's good to see these exhibitions, I think they should be advertised more often so others can enjoy them". In some cases boundaries are made only to point out the need to bridge them. For example "we should thank our Armenian friends…" and "today, watching these artistic works as a compatriot, I am glad that I see artistic appreciation in the peoples' eyes" mark distances that need to be bridged. Again, in those cases where the Armenian origin of the commentary writer is easy to infer, one cannot escape noting how ethnic pride is emphasized (even with literally writing "we are proud of…" in several occasions) in so many entries. Also, in cases where the non Armenian origin of the writer could be inferred, one could read sentences like "… he, in my opinion, is the analogous of a complete human; same as there is in the poems of Hafez the great Iranian poet. I congratulate the people of Armenia and to all humanity for having such a genious…".

To proceed further, let us concentrate on two specific motivations that can be shown to be present in many commentaries. The first one is the attempt to use the occasion to emphasize the importance and to explicate the positive aspects of a common identity. That common identity can be the Armenian identity, the Irano Caucasian identity, the identity of the non Western, or sometimes other identities like the identity of people who appreciate art, the identity of political dissidents etc. That the occasion is only an excuse for this identity construction venture is in many cases not veiled. The second motivation, also evident in many commentaries, is to declare the writer's art loving attribute. Generally, appreciation of art is considered a matter of prestige in many developing countries like Iran. It can be a demonstration of affluence (the fact that one can allocate time for pursuit of intellectual fulfillment) and knowledge. However, the demonstration can often not be matched with real exhibition of mastery. Commenting on general themes like the beauty of Parajanov's work, his ability to use simple things to produce works of art, etc. relieves the commentary writer from the need to demonstrate his knowledge by commenting on the specific artistic virtues and stylistic distinctions of the art works, while maintaining the pretence that the commentary writer, unlike most other people who are not appreciative of works of art, is very interested and knowledgeable. The

commentary can then proceed with other general statements to become lengthier than other writings without necessarily being more informative or more expressive or even more directive (like asking for better lighting, more suitable frames, etc. like some commentary writers have done).

CONCLUSIONS

Discussion of commentaries is an important step towards investigation of ideological effects prevalent in a polity. The corpus represents an ensemble of discursive pieces, provoked or induced by primary works that exhibit cosmovisions, attitudes, illusions, wishes, as well as technologies for collective construction of those elements by their writers. More than any other discursive genre, they are actions having, in addition to locutions, specific illocutions and perlocutions. Original artworks of Parajanov, likewise are his commentaries encompassing the same ideological elements, and so is any analysis of the corpus of commentaries. Considering the status and solidarity dimensions elaborated in previous works, the two: demonstration of appreciation for art, and collective identity construction performatives of several commentaries can be categorized as special cases that can locate the commentaries in the two dimensional plane. It then becomes an obvious theme of investigation to see whether the comments, when so placed in accord to their perceived strengths along the two dimensions, cluster with respect to the ethnic origin or selected language (and more generally mode of expression) of their writers. Initial observations confirm that indeed there is an evident clustering. While both motivations can be said to be present, though with differential strengths, in almost all entries, identity oriented motivations are stronger in the case of Iranian Armenian writers while prestige based approaches can be seen most often in commentaries that have been written in Farsi (Persian). In both cases, psychoanalytic factors are clearly present. It is the perception of the lack or at least the absence of common cognition, most of all, which has been the main driving force for the endeavor to construct. It is being sought, through discursive methods, to reassure oneself of a positively valued place against what is conceptualized as the cultural and political dominance or economic and social pragmatism. The commentaries show that the majority of the visitors, who have taken time to write their thoughts in the book, were not professional drawers or art critics. They were, rather, members of Iranian Armenian community, who availed themselves of the opportunity to participate in the event that had to do with the Armenian part of their collective identity, members of the Iranian community at large, who wanted to participate in what promised to be intellectually rewarding, and tourists, who wanted to see something of the culture of this distant part of the world that was less conventional. One can understand how the feeling of being present in an important event, which on the one hand is detaching them for a short while from their daily social and economic activities, and on the other hand, makes them think about different cultures and their reflection in Parajanov’s artwork, furnishes a setting in which their conscious and unconscious ideas and concerns seek a route towards expression.

REFERENCES

A. Anderson, "Literate Language Feature Use in Preschool Age Children with Specific Language Impairment and Typically Developing Language". PhD Dissertation, Department of Special Education, University of Maryland, USA.

E. Nercissians (ed). 1998. **Bilingualism.** Tehran, Iran: Intelligent Systems Research Faculty, IPM.

E. Nercissians, 1995. "The Discourse Analysis of Painting Exhibition Visitors". **The Third Conference on Theoretical and Applied Linguistics**. Tehran, Iran. Feb. 24- 25.

E. Nercissians, 1999."Who is 'We'? Community Construction trough Commentaries". **Problems of Ethnography of the Caucasus**. Yerevan, Armenia. May 19- 21.

E. Nercissians, 2003a. "Regional Identity Construction". **The 3rd International Conference on Ethnic Problems of Caucasus**. Yerevan State University. Yerevan, Armenia. Nov. 17- 18.

E. Nercissians, 2005. "Vernacular Languages and Cultures in Rural Development: Theoretical Discourse and Some Examples". **LAGSUS II: Annual Conference.** Frankfurt, Germany, October 28- November 1.

E. Nercissians. 1996. "Cognitive and Affective Development of Bilinguals: A Critique of Dominant Social and Clinical Approaches". In H. Ghasemzadeh (ed). **Cognition and Affect: Clinical and Social Aspects.** Tehran, Iran: Farhangan Pub. pp. 47 – 68

E. Nercissians. 2000. "Context and Interpretation". In K. Badie, F. G. Walners, and E. Berger (eds). **Interpretative Processing and Environmental Fitting**. Vienna, Austria: Wilhelm Braumulker. pp. 119- 132.

E. Nercissians. 2003. **Topics in Biligualism from the Perspective of Social Sciences**. Tehran, Iran: Cultural Heritage of Iran.

R. Young (ed), 1981. **Untying the Text: A Post Structural Reader** New York: Rutledge and Reagan Paul.

S. Imai, 2003. "Spatial Deixis". PhD Dissertation, Department of Linguistics, State University of New York at Buffalo, USA.

S. Nariyama, 2000. "Referent Identification for Ellipted Arguments in Japanese". PhD Dissertation, Department of Linguistics and Applied Linguistics, The University of Melbourne, Australia.

L. Vardanyan,

EXPEDITION TO ARMENIAN FORESTS: ANY UNIQUE TREE SPECIES GROWING THERE?
Towards Treebanking of Modern Eastern Armenian

ABSTRACT

This paper is aimed at sketching out an outline about the research work in progress in frames of PhD 2006-2008 program "Linguistica generale, storica, applicata, computazionale e delle lingue moderne", carried out at the Department of Linguistics "Tristano Bolelli", University of Pisa.

We propose and support the practical and theoretical background for developing a sample of a syntactically annotated corpus (otherwise called treebank) for Modern Eastern Armenian. The acknowledged lack and need for descriptive studies on Armenian language resources stands for good reason to set and accomplish such task. Basing on the conviction that only the annotated linguistic data may gain their real value for linguistic observations and research, we will present different annotation levels and formats of corpora. We target the dependency structures description in the form of syntactic functions (manual) annotation of the naturally occurring sentences in a corpus of written Eastern Armenian. The project's current focus of study is building an annotation scheme alongside setting forth a syntactic tagset to be further applied on a morphologically analyzed corpus on surface syntax level.

1. CORPORA ANNOTATION

1.1. Introductory remarks

Study of syntax is one of the milestones of general linguistic science. Throughout many decades traditional grammarians of various languages have built theories of syntax – most of them following prescriptive frameworks of grammar. The word syntax and syntactic analysis may bring to some of us bad memories on an exercise-book corrected by school teacher with red-pen underlines on text we exposed so convinced in its correctness and naturalness. In fact, our school grammar books are indeed valuable pool of grammatical rules governing the structures of a language, yet they inevitably lack the coverage of all possibly occurring grammatical structures in a natural language adopted by its speakers, being the source of any language's syntax in general. Such knowledge can be gained only through the descriptive framework of studies that compile data-driven, empirical recourses and rules from studied material.

Current progress in information technologies and availability of growing amount of digitally processed and machine-readable textual data of a particular language is providing support for holding empirical research. Corpora for variety of languages serve best source for a particular language sample, yet they are of limited use. They need to be not only larger in order to be representative, but also carefully encoded and enriched with linguistic descriptions - annotations, since the latter make the corpora valuable linguistic resources for searches and computations.

The annotation process of corpora is layered according to levels of aimed linguistic description. The following layers of annotation are generally differentiated in annotated corpora:

- Morphological annotation (lemmatization, inflection, derivation, compounding)
- Morpho-syntactic annotation (POS tagging, contextual morphological disambiguation)
- Syntactic annotation (parsing, constituency or functional dependency structure representation)
- Semantic annotation (word-sense disambiguation, anaphora, coreference resolution, information structure)
- Discourse annotation (dialog turns, speech acts)

To keep in frames of the current project objectives, we will limit ourselves at describing the syntactic level annotation of linguistic data with further projection onto the Eastern Armenian morphologically analyzed textual data.

1.2. Syntactic annotation of corpora

Syntactic annotation is commonly interpreted through the parsing – the process of analyzing an input sentential segment in order to determine its grammatical structure with respect to a given formal grammar. It presents a sentence as a data structure, usually ***a tree***, which captures the implied hierarchy of its structure according to either constituency or grammatical relations.

There are developed various automatic parser programs for carrying out such task on major languages. Yet, natural language sentences are not easily parsed by programs, as there is substantial ambiguity in the structure of a sentence composed by human. Despite the rise and efforts of computational technologies to build such linguistic resources automatically, there are actually no projects of syntactic annotation of corpora, implemented in fully automated fashion. Human annotator's intervention is indispensable in developing a syntactically annotated corpus – a ***treebank***, from fully manual annotation up to post-editing of the parser's output.

1.3. General Guidelines of Treebanking

To obtain the objective of building a treebank, certain guidelines should be followed. The quality and usability depends much on the corpus size, domain selection, as well as the motivation for which such databank is being created. These may vary from applying and testing of a given linguistic theory, up to developing a new resource, independent of any particular linguistic theory. Some projects also have a precise application goal, like inducing some linguistic resource (lexicon, grammar) for training and evaluating parsers. Such efficient tools have been developed and applied to a variety of studied languages. Nevertheless, these tools cannot either be made independent of a particular language to be treated. In other words, English languages tools, for example, may not work for Armenian once one will take a task of automatic parsing Armenian textual data.

Another and perhaps major issue is the choice of the ***annotation format*** – the representation fitting to a formalism of syntactic structures. A detailed ***annotation scheme*** should be designed where the descriptive rules of the language's syntactic

structures are drawn and the consistency of the application of each rule is strongly followed up (that is, the same constructions should receive the same analysis every time they occur). A ***tagset*** – the list of symbols used for syntactic categories represented in the corpus, should be build for applying throughout the annotation process and depending on the annotation type. This list may vary in size from a dozen up to millions depending on a specific language.

Currently the choice of annotation type for building treebanks is mainly distinguished between two trends: the linguistic resource is annotated according to either constituent or functional structure. The first efforts in building treebanks were done in the phrase structure – ***constituency-based*** frameworks, represented as tree-shaped constructions.

In the recent years there has been wide interest towards the grammatical function annotation of treebanks. The dependency grammar formalism appears fitting for this purpose and many ***dependency-based*** treebanks have been constructed. In addition, grammatical function annotation has been added to some constituent-based treebanks. A hybrid scheme is also applied by some projects, where the constituents are minimal phrased, called chunks, linked by dependency relations.

The major issue playing a determinative role for annotation format choice is mostly depending on theory-specific considerations of the project, and mainly the language-specific features to be represented in tree structures.

Briefly, the constituency structure annotation is rather suitable for (relatively) fixed word order languages, such as English. This annotation format describes the phrase structure of the entire clauses. Here the context-free grammar formalisms, otherwise called also phrase-structure grammars (e.g. HPSG, LFG) support and play pivotal role in the annotation scheme. A prominent example of a treebank of English built in this fashion is the Penn Treebank (Marcus et al. 1993) of 3 mln tokens skeletally parsed - syntactically bracketed text. Another project following this trend in its start phase is NEGRA treebank for German newspaper textual material (Brants et al. 1999).

In the constituency-based schemes adopted by these projects, the sequence of tokens (word order) in a phrase structure is followed. The occurring non-local dependencies, discontinuous structures and elliptical material are represented with trace-fillers (e.g. adding empty-nodes). In a language where such phenomena are of quite frequent and moreover specific nature, the phrase structure models do not perform efficiently in their constituency annotation. While such representation is quite compatible for a language like English, the annotation of a language with a (relatively) free word order under such format is not always optimal. Such languages are apt to perform a range of features causing problems, such as discontinuous constituents (e.g. topicalization, split phrases), ellipsis, etc. The structural handling of freer word order assumes well-formed constraints on structures involving many trace-fillers, which makes the rules and tree structure overloaded and far transparent.

Here the grammar frameworks based on functional representations of syntactic structures come to support the annotation. The grammar formalism representing dependency relations is the theoretical backbone of such annotation scheme. The description of dependency relations between the nodes (words) of a sentence tree and attachment of the functional labels make a rather powerful formalism to deal with

annotation of free word order languages – a reasonable argument that is taken into consideration for further annotation choice of the Armenian language.

Perhaps the largest project representing this framework of treebanking is being implemented for Czech language – PDT (Prague Dependency Treebank) (Hajic 1998). This is a representative treebank that in its more than 10 years of development produced over 1.8 mln tokens (around 110 K sentences) of annotated material from newspaper, science, economics, literary and other domains. Worth to mention, that many aspects and experience of PDT project are followed up and inherited in taking the current project of Armenian treebank building. Other projects that have adopted dependency annotation schemata are implemented for Italian – ISST (Italian Syntactic-Semantic Treebank) (though the entire treebank exploits parallel layers of both constituent and dependency structure annotation), TUT (Turin University Treebank); for Dutch – the Alpino Treebank; the Dependency Treebank for Russian, etc.

2. FIRST STEPS TO TREEBANKING OF EASTERN ARMENIAN

Bearing in mind the guidelines described above, our project targets treebank building for Standard Eastern Armenian (henceforward referred to as Armenian in the frames of the project). A morphologically analyzed subcorpus from EANC (Eastern Armenian National Corpus) from press and fiction genres is targeted to be manually annotated. To take decisions upon choosing annotation format we deduce from the language specific requirements. From the typological point of view, Armenian is a pro-drop inflectional language showing agglutinative, synthetic and analytical features in word and phrase construction. The word order variation ranges from relatively fixed (mainly in NPs) up to very flexible inside a clause. The classical assumption about its basic word order is referred as SOV in neutral, unmarked sentence, though contextual occurrences of other combinations are rather frequent. Relevant to mention, that anyhow there is no data-driven evidence of the frequency of even basic order variation. In general many aspects of syntactic structures of Armenian language are not described and analyzed consistently enough. “There is a clear absence of a syntactic corpus study in previous work on Armenian standard varieties.” (Dum-Tragut 2002). A treebank as a linguistic resource may come useful to support such empiric research.

2.1 Annotation Scheme and Syntactic Tagset Development

Taking into account the syntactic characteristics of the target language and the past experience of treebankers, we were lead to choose the dependency representation framework for building an annotation scheme for Armenian sentences.

The core notion of dependency representation is the relational head-dependent structure. It can be graphically introduced by a tree – directed acyclic graph, which has one root and whose nodes have at least and most one parent node – the grammatical or technical (as the case may be) head.

In such dependency-based tree structure only the lexical words (nodes) are recognized, and the phrasal ones are omitted. All the lexical words together with the punctuation and other graphical symbols, i.e. every input token in a sentential segment is treated to be a node in the tree structure. No additional node insertion is allowed. There

are no empty categories, which makes dependency structures more optimal and human-readable for the annotator to work on. The nodes in the sentence are linked with edges representing the dependency relation types in the form of syntactic function attribute values (given in the syntactic tagset in the section 2.2), which are attached to the nodes for technical data representation, yet in fact belong to edges (which is, by the way, visually seen on the tree view of the annotation tool as shown on Fig. 1 below).

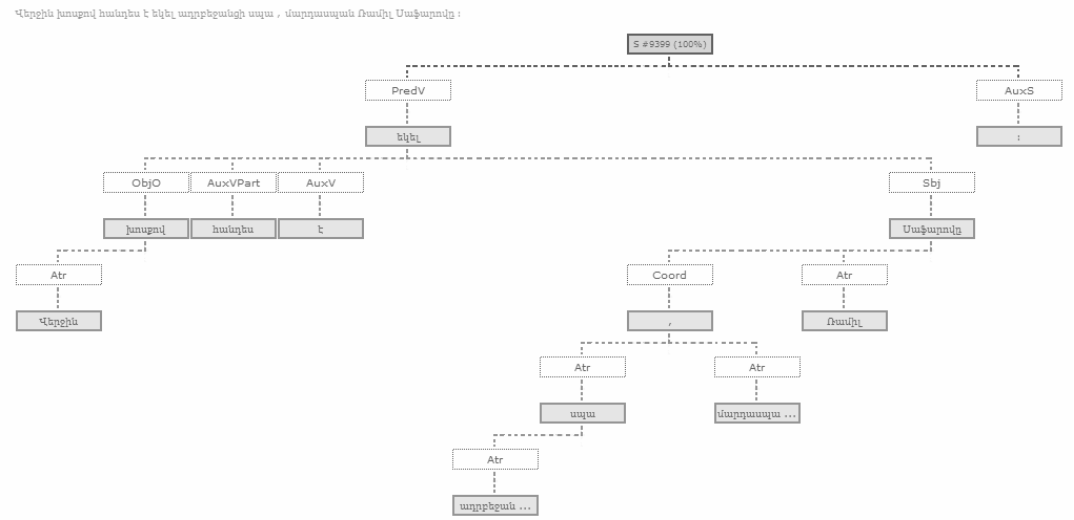


Fig. 1. Syntactic function and dependency representation in a tree structure

Following the basic assumption about the syntagmatic relationships, where two or more elements co-occur together, there is a dominant, principal element which is the primary determinant of the properties in such relationship. As regards to the sentence principal element, in our representation it is the predicate node taking the governing function in the entire utterance. Its valence (here it's worth to mention that whenever speaking of it, we always take into account the differentiation between the semantic and syntactic valence, and always refer to the latter), is determinant of the argument (again syntactic) structure of the whole sentence. Thus, in our representation, alongside with other arguments, the subject node is depending on the verb predicate (or other predication element). This very basic principle of annotation already differs from the traditional concept about sentence structure we find in Armenian grammar books that regard the subject as principal element together with the predicate. In general, wherever possible we follow the categorization and determination of grammatical relations and functions given in the textbook "Modern Standard Armenian" (Abrahamyan: 1981). Nevertheless, due to some inconsistency and incompatibility to formalism we come across in the grammar throughout developing the annotation scheme and tagset, we employ modifications to syntactic categories, either narrowing or broadening their delimitations, up to introducing new categories and structural representations.

Also, for the sake of formalism and computational approach, many nodes are labelled in a purely technical mode to represent structures and functions that are not linguistically competent. Due to the work being in preliminary stage and progress, many modifications and corrections are assumed to be done throughout their further exploitation in annotating the material.

One of the technical representations is the "dependence" of the predicate node on the sentence (<se>) node in the input file structure together with the sentence final punctuation mark. All the other elements are correspondingly suspended under this node in accordance with the valence frame of the governing verb or otherwise, being described in detail in the rules of annotation scheme. For keeping consistency, avoiding misinterpretations and clearing up dependency and syntactic function assignment, lexicons of verbs with their valence frame, as well as other relevant parts of speech with their assumed syntactic functions, are being compiled.

Another important non-linguistic dependency representation is drawn for linguistic phenomena like coordination and apposition, traditionally interpreted as equal sentence members. To represent such data, the members of coordination or apposition are represented as "depending" on the coordinating node (conjunction, punctuation), which itself hangs on the head node governing the coordination or apposition.

Many traditionally approached syntactic structures and functions are also getting more formal description and representation. The rules and tagset of syntactic functions are currently on the way of being formulated and populated, so that modifications and additions are in constant raise.

Below a brief list and description of the adopted syntactic tagset and functions are given. Once again it should be mentioned that the list is far from being complete and explanatory enough for many linguistic phenomena, moreover much more complex ones, which are unsurprisingly quite frequent in naturally occurring linguistic material to be annotated.

2.2 Syntactic Function Taglist and Brief Description

<Syn Fct> values	**Function. Description, examples of expressions**
PredV	**Predicate** The parent node of the sentence tree. Does not depend on other sentence member nodes, yet technically is suspended under the <se> node in data structure. Is expressed by simple finite verb forms or converbs that form finite predications. *Note:* The predicate of a subordinate clause does not take this function tag, but rather takes the function of the clause it represents.
PredN	**Predicative nominal element** This is the predicative part in the compound nominal predication; always hangs on the AuxV, if latter is expressed overtly in the sentence.

AuxV **Auxiliary verb 'Է'**

This is the function tag assigned to the copula when being a part of the main predication and always depends on PredV node. Otherwise stands as the governing node of the sentence with dependent compound nominal predicative. If not the previous two, this 'Է' takes the function AuxPart hanging on modal particle '(չ)պետք' in neighbourhood with subjunctive verb. The particle itself takes the function AuxMod being suspended under verb.

Note: The occurrence of '(չ)պետք Է' in a uniclausal occurrence with infinitive is represented as a construction of the case: $PredN_{[D]}$ $AuxV_{[H]}$

AuxVPart **Compound verbal particle**

This is the non-verbal element in compound analytic verb forms. Always is suspended under the main verb it "complements". Ex. '*ամն*$_{[D]}$ գալ$_{[H]}$' '*գլուխ*$_{[D]}$ բերել$_{[H]}$', etc. A lexicon of such particles and respective verbs is envisaged to be compiled to avoid the mislabelling of these elements with other functions, as well as the fact that in such forms the main verbs, having lost their main lexical meaning, change their argument structure too.

Sbj **Subject**

The direct argument of the predicate which can be missing from the overt structure of the sentence. Compliant to our dependency representation it is depending on the predicate or auxiliary verb.

ObjD **Direct object**

An obligatory direct argument depending on the verb according to its valence frame: in active voice utterances it usually corresponds to the semantic role of patient. Usually morphologically is expressed by a noun, pronoun, other nominalizations in nom/acc(dat in animate nouns) case marking. The infinitival part of compound predicate also takes this function. Ex.: 'ՈւԻզում$_{[H]}$ է *աշխատել*$_{[D]}$:'

ObjI **Indirect object**

An obligatory indirect argument depending on the verb according to its valence frame. Usually, but not always, the semantic role of recipient corresponds to this argument.

ObjO **Oblique object**

Adjunct or non-obligatory syntactic argument whose nature and behaviour are more describable in semantic terms than syntactic. They show different semantic relations to the verb and direct arguments and are likely to be most constrained in the semantic roles they may

individually express, hence no further subclassification is given between kinds of oblique objects. This element is likely to be marked by an adposition, in which case the latter is the governing node, or case affix, in which case the ObjO is hung directly on the verb.

Comp **Complement**

Arguable it may seem, as it does not have an exact match in traditional grammar, or more correctly, it is presented under various categories. This is a direct argument with predicative role complementing itself one of the arguments of the verb; either the subject or object. In general the nature of this element is of double dependency both on verb and on the argument it refers to. Yet again, following the principles of annotation of unidirectional dependence, in the presence of the verb in the sentence, it gets the tag Comp and gets suspended under the node of argument it concerns, or the preposition 'որպես', as the case may be. Ex., 'Պատերը[H] *սպիտակ*[D] ներկեցինք:' ' Նա[H] *հիվանդ*[D] է ձևանում:' 'Նա[H] ելույթ ունեցավ որպես *տնօրեն*[D]:'

Note: should not be confused with other functions like Obj, Adv, neither the overtly similar seeming apposition case with 'որպես', etc.

CompV **Verbal Complement**

When the governing node to which a Complement element refers is elided or missing at all, it gets the function tag CompV, and is suspended under the verb itself. Ex., '*հիվանդ*[D] է ձևանում[H]', 'ելույթ ունեցավ[H] որպես *տնօրեն*[D]'

CompAdv **Adverbial complement**

An obligatory argument dependent on the verb that is common to misinterpret with either adverbial or object: its usage is determined with the subcategorization frame of the verb. Depends on the verb it complements.

Note: a test to identify this function - the sentence is ill-formed, ungrammatical without it, unlike with adverbial: *Նա գտնվում է: *Գիրքը դնել:

Adv **Adverbial**

Optional arguments and adjuncts modifying the verbal elements. As the further subcategorization of these modifiers is according to their semantic roles, no further classification is given. Depends on the verb it modifies.

Atr **Attribute**

The traditional nominal modifying element. It depends on the nominal

it modifies. Typically is expressed by lexical categories as adjectives, quantifiers, demonstratives, numerals, but also may be expressed by nominals marked with nominative, ablative, instrumental, locative case.

AtrGen **Genitive Attribute**

The possessor attribute which is actually a semantically constrained subcategory of Atr. Nevertheless a syntactic function is categorized as its inflectional case marking is easy for look-up. As the Atr, it is hanging on the nominal it modifies.

Apos **Apposition**

Again a technical node for representing the apposition. Is expressed by the punctuation mark "bouth", while the members of the apposition themselves are suspended under this node:

Note: The case of apposition introduced with 'որպես', 'իբրև' should not be misinterpreted with Comp: 'Նա[H]՝ որպես *տնօրեն*[D], ելույթ ունեցավ:'

Coord **Coordination**

The coordination node technically dependent on the higher element to which its own dependents refer, thus having those as child nodes. Is expressed by punctuation marks -comma, period, or coordinating conjunction.

AuxSub **Conjunctive element**

Subordinate conjunction or punctuation (period or "bouth" otherwise not specified as main apposition node).

AuxA **Adposition**

Prepositions and postpositions: their position in the tree is governing the nominal they refer to, thus they are suspended where the nominal introduced by them would take place and correspondingly the nominal is hung on them.

AuxMod **Modals and emphasizing words**

These can be suspended under any element member, according to the function they take to emphasize a particular word. A list of these words is assumed to be compiled. Should not be misinterpreted with adverbials or other similar elements. In case such word refers, emphasizes the whole utterance, without any clear-cut emphasize on a particular member, it gets always suspended under the predicate element of the sentence.

AuxPart **Particles of multi-token words**

These are parts of analytical word constructions that have no syntactic

function; always suspended under their main lexical node. Ex., է in ‘(չ)պետք$_{[H]}$ է$_{[D]}$’,

AuxP **Punctuation**

All punctuation marks with otherwise not specified above function they take (quotes, brackets, sentence start dash), abbreviation period and other graphical symbols. Their suspension depends on the position they take and are to be more detailed explained in tagging manual.

AuxS **Colon : full stop**

The sentence final punctuation mark; always hangs on the root of the sentence tree together with the PredV or in the absence of the latter, another sentence governing node.

ExD **External dependency**

A technically added function which labels the node that misses its actual governor in the sentence; typically the ellipses are handled by this function tag. Another example of such function tag exploitation is for the attributes referring to the lexical part of relational nouns, which are then suspended on the entire wordform. Ex. ‘մեր *հարևան*$_{[D]}$ Արամինը$_{[D]}$’.

Voc **Vocative**

Depends on the verb of the imperative sentence.

2.3. Data representation and the annotation tool for manual tagging

In data structure and representation the XML markup standard is followed up. The source input data is shared by Eastern Armenian National Corpus (EANC), and its format already has the corresponding structure. The textual data is analyzed morphologically; that is, the lexical nodes already bear morpho-syntactic information in the form of tags with attributes with corresponding values. Nodes appear in their linear order without contextual disambiguation.

Fig 2. shows the file format serving as an input to be pre-processed with slight modifications and be prepared for syntactic annotation by the application that is currently under development to serve the objectives of manual annotation.

The Tool is currently being developed with following goals:

- entitles to import morphologically tagged data from XML source
- enables to perform basic updates to the imported XML structure, as well as imports punctuation tags otherwise left untagged in the source files
- integrates the imported files into a database for easy data handling
- allows visualisation and edition data organized in paragraphs and sentences
- allows to perform manual syntactical annotation of the tokens
- displays real time visual tree structure of the syntactically annotated data

- re-exports original format files after modifications and syntactically annotated file to XML structures
-

Fig 2. Original source file with morphological analysis

The syntactical information is introduced by adding to each token a <Syn> tag with attributes <fct> and <dep>, respectively marking the function of a node and pointing to its governor node. As mentioned before, the syntactic function actually belongs to the edge of the tree structure, relating the node to its governor.

The tool enhancing the annotation is web based application with SQL Server database use. The web application will allow easier later sharing of knowledge, and easy cooperation with other projects without requiring special computer configurations or tools to install.

The system is made of several building blocks described below. As the application is under development the following list may not be up to date at the time of reading this paper.

2.3.1. Importation Routine

This imports morphologically analyzed data files and performs basic updates to make the data compliant with the syntactical layer needs. The input files are structured into paragraphs, sentences, words and “gloss” tags representing the morphological layer. As the morphological disambiguation is not carried out, there may be several glosses per word. The data is analyzed in linear manner and bears only the morphological information for the lexical items. The punctuation marks and digits in the imported files are left out tagging.

The initial pre-processing operates a preliminary parser on the source file and results in integrating punctuation and other graphical symbols into the structure (as the latter also take functions in a sentence structure and will need to be positioned into the tree), also adding explicit numbering of paragraphs, sentences and ordering the word sequences. Once updated, the file is parsed into a relational database.

Files are organized into projects that are created on user needs (per project, per annotator teams, etc.). Each file is divided into "paragraphs", subsequently divided into "sentences" and then into "tokens" - <t>, which have come to substitute the original <w> tags and also include punctuation marks and digits.

A Token contains the character string <txt>, which in the case of a word carries the morphological information about it within zero (if the word has not been recognized during the morphological analysis), one or several "glosses", its sequential order <ord> in the sentence, its <type> with values for lexical words, punctuation, graphical symbols and digits, and after annotation, the syntactical layer information within an additional <syn> tag.

2.3.2. Data handling tools

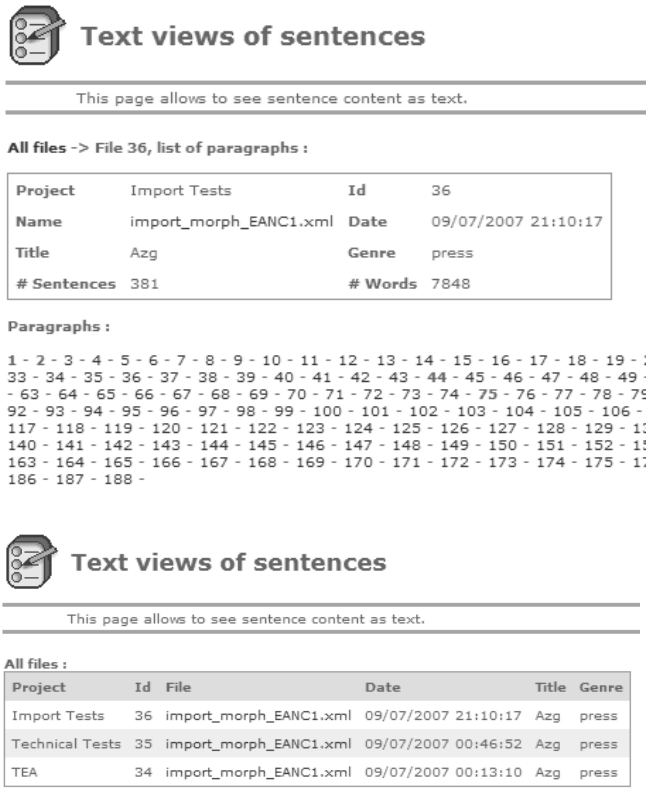


Fig 3. Navigaition tools

This block consists of two main types: the structure and text data navigation tools and, the data edition tools.

The *structure and text navigation tools* allow the annotator to navigate within the data by browsing projects, files, paragraphs and sentences (Fig. 3). This navigation entitles the user to work on XML files omitting all technical XML / database tags, making the files readable, shown as plain text (Fig. 4),with or without morphological and syntactical information.

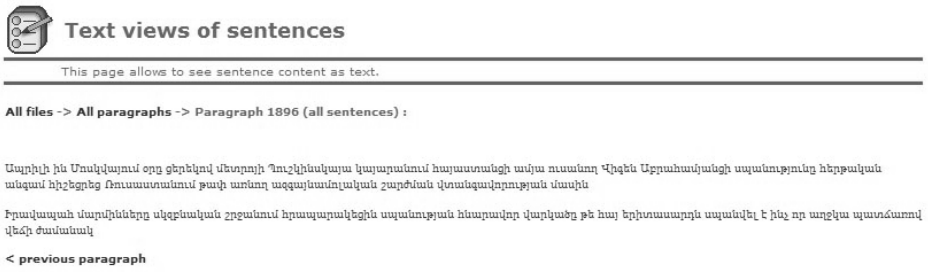


Fig. 4 . Text view of a file

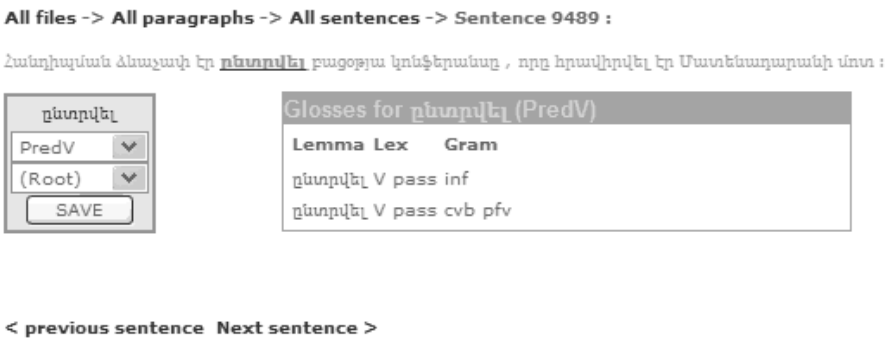


Fig. 5. The annotation tool with the annotation toolbox with the imported morphological information

The annotator's (manual) work can be performed from *data edition tools*. The application can work on token per token basis or full sentence annotation by means of annotation toolbox (to choose the syntactical function tag of a token and its governing node).

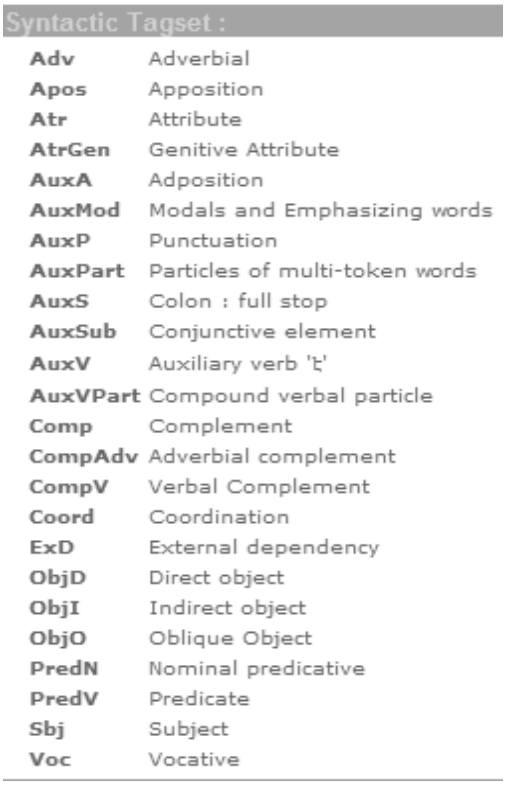


Fig. 6. Syntactic tagset display

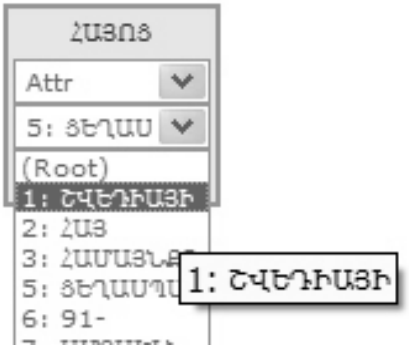


Fig. 7. Head selection

The list of syntactical function tags described in above sections (Fig. 6) is available in the first drop-down list of the annotation toolbox. These functions are stored into the database and can be managed by the user (add / edit / removed, if not used).

The second list (Fig. 7) displays the tokens in the sentence from which the annotator chooses the one, on which the token under analysis depends on, according to the annotation rules and principles.

The process of manual syntactical annotation allows creating the tree view of the sentence. The web based tree view component is developed in HTML, using <table> and simple images to be easily and widely usable, on all platforms, without software installation needed. The tree logic is all calculated on server, to produce light client application allowing use of the annotation tools on computers with lower resources or slow internet access.

Throughout the annotation of a sentence, the tree is being generated and can be browsed in real time (Fig. 8). This figure displays the tree being built, with the root being the <s>entence.

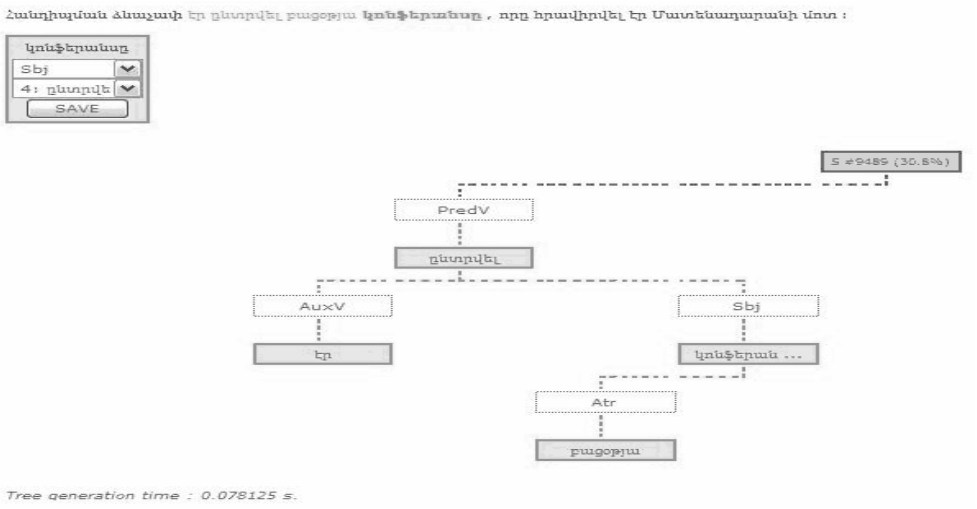


Fig 8. The tree is being built while the annotator works on the sentence

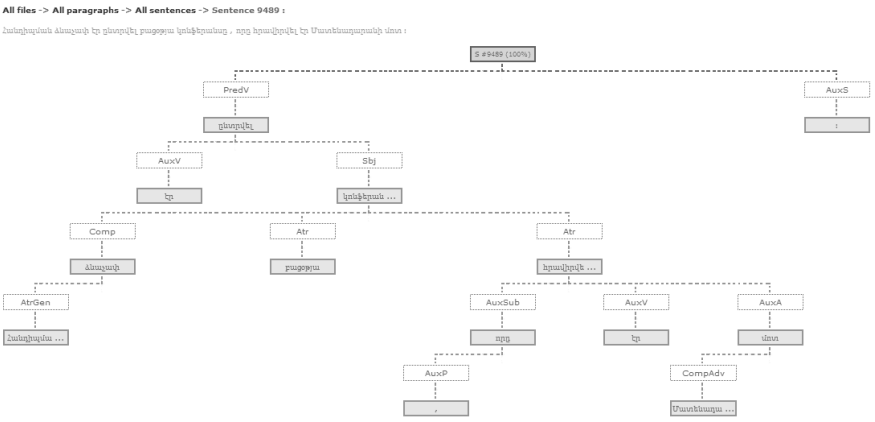


Fig 9. Dependency tree sample

The tree logic incorporates arches, that are made of images put in tables, with branches crossings and branches divisions etc., put in visual layout. Below an example of a fully annotated sentence tree is given:

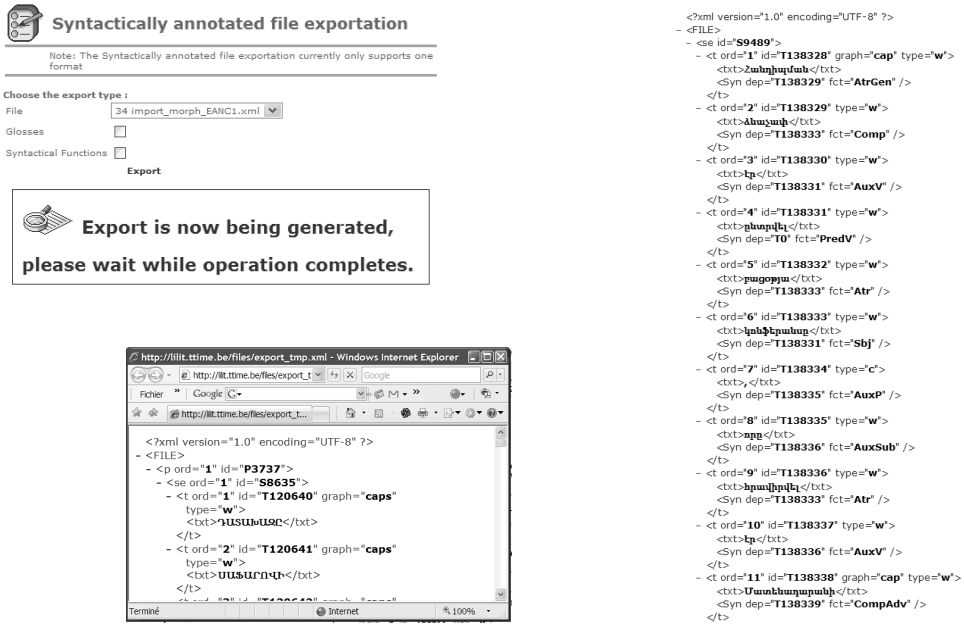


Fig 10. Syntactically annotated file exportation and exported file sample

2.3.3. Exportation Routine

The exportation routine, as the names says, allows recreating XML files in various supported formats, with or without morphological information and with or without syntactical information.

CONCLUSION

The annotation of linguistic data with integration of syntactic structure information has been widely exploited during the past decade in the general and field linguistics. These efforts have proven to be a very successful basis for the progress in language theory and processing technology, from which apparently the broadly spoken and studied languages have favoured, while the minor ones perform lacks in such research experience.

Armenian language represents a separate node in the Indo-European language family tree, thus lacks any close relatives from the genetic point of view. Inevitably it has inherited structures proper to other family languages, yet not surprisingly it may have developed quite individual features. In this stage of linguistic science when language technologies afford accumulation of valuable knowledge, it is worth to build resources serving for such needs.

Treebank building is a challenging task. It is time and labour consuming, yet is very promising for language resource building and field linguistic knowledge updating to match the current linguistics trends. Although we perform manual annotation in the frames of the project, as a preliminary stage, the work assumes to give challenge and prospects for possible development and training of automatic tools for processing

Armenian language resources. Further development and fine-grained annotation scheme and rules are envisaged to be carried out, as well as manual annotation of reasonable amount of data. We believe it may come useful in its later exploitation for linguistic research and NLP tools development purposes.

REFERENCES

Ann Abellié (ed.) 2003. *Treebanks: Building and Using Parsed Corpora*. Kluwer

Jurafsky D., Martin J.H. 2000. *Speech and Language Processing*. New Jersey

Manning C., Schutze H. 1999. *Foundations of Statistical Natural Language Processing.* London

Jan Hajic, Joakim Nivre. 2006. *TLT 2006. Proceedings of the Fifth Workshop on Treebanks and Linguistic Theories*. Prague

Jan Hajic. 1998. *Building a Syntactically Annotated Corpus: The Prague Dependency Treebank.* Prague

L.van der Beek (ed.) 2002. *The Alpino Dependency Treebank. Algorithms for Linguistic Processing NWO PIONIER Progress Report*. Groningen

Skut W. (ed.). 1997. *An Annotation Scheme for Free Word Order Languages*. In *Proceedings of the Fifth Conference on Applied Natural Language Processing (ANLP).* Washington, D.C.

Tuomo Kakkonen. 2005. *Dependency Treebanks: Methods, Annotation Schemes and Tools.* In *Proceedings of the 15th Nordic Conference of Computational Linguistics (NODALIDA 2005).* Joensuu

Boguslavsky (ed.). 2002. *Development of a dependency treebank for Russian and its possible applications in NLP*. In *Proceedings of the 3rd International Conference on Language Resources and Evaluation*. Las Palmas

C. Bosco, V. Lombardo. 2004. *Dependency and relational structure in treebank annotation*. In *Proceedings of Workshop on Recent Advances in Dependency Grammar at COLING'04*.Geneve

Kiril Simov. 2003. *HPSG-Based Annotation Scheme for Corpora Development and Parsing Evaluation. Proceedings of RANLP*. Borovets

J. Weitenberg, S. Simonian. 1993. *Computers in Armenian Philology*. Yerevan

Van Valin. 2001. *Syntax*. Cambridge

J. Dum-Tragut. 2002. *Word-order correlations and word order change: An applied-typological study on literary Armenian variant.* München

S. Abrahamyan. 1981. *Modern Standard Armenian.* Yerevan

V. Arakelyan. 1958-62. *Syntax of Modern Armenian, I-II.* Yerevan

N. Kozintseva. 1995. *Modern Eastern Armenian*, München

Н. Р. Сумбатова
Российский государственный гуманитарный университет, Москва

ПРОБЛЕМЫ ИССЛЕДОВАНИЯ ЯЗЫКОВ ДАРГИНСКОЙ ГРУППЫ

1. Общие сведения о даргинском языке

Традиция рассматривает даргинский как единый язык, составляющий отдельную подгруппу нахско-дагестанской семьи. Даргинский в основном распространен в восточной части Республики Дагестан, число говорящих на нем – около 520 тысяч человек (оценка, 2002 г.). Его отличительная особенность – наличие большого числа сильно расходящихся диалектов (см. ниже).

1.1. Грамматические особенности

Структура даргинского языка в основном типична для нахско-дагестанской семьи, однако вместе с тем он обладает рядом нетривиальных и даже уникальных грамматических свойств. Как и в большинстве восточнокавказских языков, имена существительные в даргинском имеют категории числа, падежа и именного класса. Классное согласование, помимо глаголов, демонстрируют некоторых прилагательные, местоимения и деривационные суффиксы, а также показатели эссива, в том числе в составе послелогов и локативных наречий. В диалектах даргинского языка представлены, как правило, от двух до четырех местных падежей и четыре–шесть локативных серий, среди которых нередко выделяется типологически редкая серия «нахождения перед ориентиром». Две формы латива (суперлатив и иллатив), формально относящиеся к локативным падежам, часто имеют свойства грамматических падежей, выполняя ряд функций, типичных для датива.

Глагол в даргинском языке образует чрезвычайно большое число видо-временных и модальных парадигм. Большая часть глагольных форм представляет собой комбинацию одной из нефинитных форм глагола либо одной из глагольных основ и потенциально отделяемой от глагола частицы, выражающей лицо, время, отрицание, вопрос и т. п. Система нефинитных форм включает простые и специализированные деепричастия, краткие и полные причастия и несколько отглагольных существительных. В некоторых диалектах имеются спрягаемые формы конъюнктива, соответствующие инфинитиву других диалектов. Даргинский язык принадлежат к числу немногих нахско-дагестанских языков, имеющих развитое личное согласование.

Как и другие нахско-дагестанские языки, даргинский – язык с преобладанием эргативной конструкции предложения, однако встречаются и другие конструкции – биабсолютивная, аффективная и антипассивная; впрочем, сфера их применения сравнительно узка.

1.2. Диалектное членение

Разные авторы насчитывают в даргинском языке от 11 до 40 диалектов, причем различия между большинством из них так велики, что практически исключают взаимопонимание носителей, а расхождение основных диалектов, по данным глоттохронологии (Ю. Б. Коряков, личное сообщение), относится к 3 в. до н. э. Следует ли из этого, что даргинские диалекты следует скорее считать даргинскими *языками* – вопрос, заслуживающий отдельного обсуждения. На наш взгляд, с собственно лингвистической точки зрения проблема языка vs. диалекта не может иметь четкого решения, да, собственно говоря, и не является существенной. Поэтому вопрос о том, следует ли рассматривать хотя бы часть идиомов даргинской группы как отдельные языки, не будет обсуждаться в данной работе. Мы будем придерживаться традиционного термина *диалект* или употреблять нейтральный термин *идиом*. Однако, как будет показано в разделах 1.3 и 1.4, то или иное решение этой проблемы оказывается весьма важным с социолингвистической точки зрения.

Традиционно (Абдуллаев 1954, Хайдаков 1985, Мусаев 1999; 2002 и др.) даргинские диалекты подразделяются на две группы – диалекты акушинского и цудахарского типа, или, соответственно, северные и южные. К северной группе относятся, в частности, акушинский, губденский, кадарский, урахинский, уркарахский, мегебский диалекты, к южной – цудахарский, кубачинский, аштинский, ицаринский, чирагский, кайтагский, мугинский, муиринский, сирхинский и др.

Это членение базируется на существенном различии в системе согласных: в южных диалектах шумные смычные и аффрикаты образуют четверки, состоящие из простого (глухого придыхательного) согласного, сильного (геминированного, непридыхательного), абруптивного и звонкого (например, t – t’ – tː – d; c – c’ – cː – ʒ); северные диалекты не имеют геминированных согласных, так что четверкам соответствуют тройки (t – t’ – d; c – c’ – ʒ)[1]. В большинстве случаев наблюдается также корреляция между наличием/отсутствием геминат и двумя морфологическими чертами: формой личного местоимения ‘я’ (в северных диалектах nu, в южных – du) и формой личных окончаний, используемых в базовых финитных формах (см. раздел 3.2).

Две группы диалектов не вполне равноценны: количество диалектов в составе южной группы больше, и различия между ними гораздо более существенны. В южной группе наблюдается также более высокая степень дробности диалектного членения: именно на юге даргиноязычного ареала, в горных районах, мы сталкиваемся с одноаульными идиомами, причем языки соседних аулов могут демонстрировать достаточно серьезные различия. По-видимому, выделение двух диалектных групп необоснованно с исторической точки зрения. Такой вывод

[1] В даргинском языке нет фрикативных абруптивов, поэтому фрикативные согласные в южных диалектах образуют тройки вида «глухой простой – глухой геминированный – звонкий» (χ – χː – ʁ), а в северных – пары «глухой – звонкий» (χ – χː – ʁ vs. χ – ʁ).

подтверждается предварительными лексикостатистическими подсчетами, сделанными Ю. Б. Коряковым (устное сообщение; см. тж. Коряков 2006), который разделяет идиомы даргинской группы на шесть подгрупп: 1) северную (акушинский, урахинский, мугинский, цудахарский, гапшиминско-бутринский, мюрего-губденский, кадарский, муиринский; 2) мегебский; 3) юго-западную (сирхинский, амухско-худуцкий, кункинский, санжи-ицаринский; 4) чирагский; 5) кайтагскую (верхне- и нижнекайтагский; 6) кубачи-аштинскую (кубачинский, аштинский).

Тем не менее можно сказать, что установившейся классификации идиомов даргинской группы до сих пор не существует.

1.3. Степень изученности

Первая большая работа по даргинскому языку – книга П. К. Услара «Хюркилинский язык» (Услар 1892), посвященная фонетике и морфологии диалекта, который сейчас, как правило, именуется урахинским. Характерно, что и даргинский язык в целом, и хюркилинский (урахинский) диалект Услар называет «языком». Традиция, в которой даргинский язык рассматривается как единое целое, в основном сложилась, очевидно, в двадцатые годы XX века – период советской политики «языкового строительства», когда на базе акушинского диалекта был создан даргинский литературный язык, была написана его первая грамматика (Жирков 1926) и создана первая письменность (на латинской основе – 1928 г.[2]).

К сожалению, подход к даргинскому как к единому языку оказался неблагоприятным для исследования языка. Проще говоря, начиная с 20-х гг. основная масса исследований была посвящена литературному языку или в лучшем случае акушинскому диалекту, на базе которого он создан, исследования же других диалектов рассматривались как область диалектологии и велись от случая к случаю.

В частности, литературному языку (помимо уже упомянутой грамматики Л. И. Жиркова) посвящена краткая, но содержательная грамматика (Абдуллаев 1954), работы М.-С. Мусаева (1999, 2002), Х. ван ден Берг (1999; 2001) и других исследователей. Материал диалектов хотя и привлекается во многих исследованиях, но чаще всего используется как фон, на котором выделяются особенности литературного языка. Имеются также собственно диалектологические исследования, но в них в основном приводятся разрозненные и несистематические сведения о тех или иных грамматических характеристиках некоторых диалектов, причем сам набор диалектов, привлекаемых для сравнения, разнится от случая к случаю. Почти единственным исключением до последнего времени были очень качественные работы А. А. Магометова, оставившего описания двух языков/диалектов даргинской группы – мегебского (1982) и особенно кубачинского (1963) – родного языка исследователя. В последние годы появились также описания ицаринского диалекта (Sumbatova, Mutalov 2003) и кайтагского (Темирбулатова 2004).

[2] В 1938 г. письменность на латинской основе была заменена кириллической.

Таким образом, на настоящий момент мы имеем описания следующих диалектов (в хронологическом порядке): урахинского, акушинского, кубачинского, мегебского, ицаринского, кайтагского. Следует заметить, что эти описания в основном сосредоточены на вопросах фонетики и морфологии; работы по синтаксису и грамматической семантике даргинского языка, тем более выполненные на современном теоретическом уровне, практически отсутствуют[3]. Однако для прочих диалектов мы не имеем даже и этого. Хотя в некоторых работах – в особенности в книге С. М. Гасановой (1970) и в статьях А. А. Магометова (1962, 1976, 1978 и др.) содержится довольно много информации, но эта информация, даже будучи систематизированной, не дает сколько-нибудь полного представления ни об одном из диалектов, помимо перечисленных – даже в области фонетики и морфологии.

Таким образом, описание фонетики и морфологии даргинских диалектов, помимо перечисленных выше, а также синтаксиса и грамматической семантики всех без исключения даргинских диалектов – задача, к решению которой исследователи практически еще не приступали.

1.4. Степень сохранности

Как большинство языков Кавказа, даргинский не относят к числу языков, находящихся под угрозой исчезновения. Действительно, данные последних лет указывают на значительный рост числа говорящих (365 тыс. по данным переписи 1989 г. против 520 в 2002 г.).

Однако оптимистические цифры являются результатом подхода к даргинскому как к единому языку. Растет число даргинцев в целом, однако большая часть из них говорит на нескольких наиболее крупных диалектах (акушинском, урахинском, цудахарском, губденском). При этом мы не знаем, что происходит с диалектами, число говорящих на которых не так велико. В связи с тем, что отток жителей из высокогорных аулов, наблюдающийся уже с шестидесятых – семидесятых годов прошлого века, в последние десять–пятнадцать лет очень ускорился, можно предполагать, что по крайней мере какие-то диалекты находятся под угрозой исчезновения. Конкретной информации у нас, к сожалению, недостаточно.

1.5. Необходимость изучения и документации даргинских диалектов

Таким образом, современное состояние даргинского языка и его изучения можно кратко суммировать так:

– даргинский имеет много сильно расходящихся диалектов: по-видимому, их не менее 18–20, однако точное число еще предстоит установить;

– глубина расхождения диалектов превышает две тысячи лет, лингвистические различия между ними огромны, взаимопонимание в общем случае невозможно;

[3] Синтаксический раздел имеется в кратком очерке грамматики ицаринского диалекта (Sumbatova, Mutalov 2003), но и это описание – вынужденное краткое из-за формата серии, в которой оно издано, – не дает о синтаксисе диалекта достаточного представления.
Кроме того, полезные сведения о синтаксисе чирагского диалекта содержатся в работе (Кибрик 1979/2003).

– общепринятой генетической классификации этих идиомов не существует; лексикостатистические подсчеты (основанные на неполном материале) дают результаты, весьма далекие от традиционной бинарной классификации, для которой, впрочем, не существует диахронического обоснования;

– мы располагаем сравнительно полными сведениями о шести диалектах, причем даже для этих шести это в основном сведения из области фонетики и морфологии; синтаксис и грамматическая семантика остаются практически неисследованными;

– сведения о прочих диалектах обрывочны и плохо сопоставимы друг с другом;

– даргинский язык в целом не находится под угрозой исчезновения, но это не относится к большинству диалектов: степень их сохранности неизвестна.

Таким образом, можно сказать что одна из самых интересных групп в составе нахско-дагестанской семье явно не получает достаточно внимания лингвистов. Исправить это положение хотя бы отчасти призван проект, о котором пойдет речь в следующем разделе.

2. Проект изучения даргинских диалектов

Как очевидно из вышесказанного, до составления хотя бы сравнительно кратких грамматических описаний всех идиомов даргинской группы еще очень далеко, и вряд ли это достижимо в обозримый период. Тем не менее, довольно существенное количество важных сведений можно собрать и систематизировать сравнительно быстро, если опираться на то, что уже известно о грамматических структурах языков даргинской группы. Это является одной из задач проекта исследования языков даргинской группы, работа над которым началась в текущем году (участники проекта – Н. Р. Сумбатова, Д. С. Ганенков, Р. О. Муталов, Н. В. Сердобольская, В. С. Хуршудян).

Основной метод сбора данных, предусмотренный проектом, – полевые исследования в местах проживания носителей. Помимо полевой работы, предполагается собрать и по возможности систематизировать все сведения о диалектах, содержащиеся в имеющихся публикациях, многие из которых выходили ограниченным тиражом в малоизвестных издательствах и недоступны большинству исследователей.

Проект предполагает целенаправленный сбор и систематизацию сведений о диалектах даргинского языка, до сих пор не ставших объектом последовательного изучения, и включает следующие виды исследований:

– сбор базового словаря – в основном по списку из работы (Кибрик, Кодзасов 1988);

– сбор, анализ, глоссирование сравнительно небольшого корпуса текстов, представляющих разные жанры устной речи;

– сбор материала по базовым грамматическим проблемам – на основании заранее разработанных анкет, специально ориентированных на языки даргинской группы;

– сбор базовых социолингвистических сведений (число носителей, проживающих в местах исконного обитания, степень сохранности языка).

«Грамматическая» часть проекта предполагает концентрацию исследований на тех аспектах языковой структуры, которые наиболее характерны для даргинских языков и, возможно, отличают их от других языков нахско-дагестанской семьи. Имеющиеся сведения о тех диалектах, описания которых уже существуют (см. раздел 1.3), позволяют ограничить пространство ожидаемых возможностей для большинства параметров языковой структуры и тем самым несколько упростить исследовательскую процедуру. Ниже, в разделе 3, мы пытаемся показать пример такого ограничения – параметры межъязыкового варьирования идиомов даргинской группы, связанные с выражением личного согласования.

3. Личное согласование в идиомах даргинской группы: параметры межъязыкового варьирования

Как уже отмечалось, даргинские языки принадлежат к числу немногих нахско-дагестанских языков, где представлена категория личного согласования[4]. Причем в отличие от языков, довольствующихся бинарным противопоставлением *локутор – нелокутор* (*первое/второе лицо* vs. *третье лицо*) или *первое лицо* vs. *непервое лицо* (как в лакском и ахвахском), даргинский, как правило, противопоставляет по меньшей мере три лично-числовых формы глагола.

Расхождения между известными нам диалектами в области лично-числового согласования чрезвычайно существенны, но в то же время пределы, в которых происходит такое варьирование, очерчены сравнительно четко. Рассмотрим отдельные составляющие этого явления более подробно.

3.1. Контроль личного согласования

(1) Потенциальные контролеры

Во всех даргинских языках лицо контролируется либо аргументом в абсолютивном падеже, либо эргативным агенсом[5]. При экспериенциальных глаголах (чаще всего при глаголе ‘хотеть, любить’) контролером согласования может быть экспериенцер в дативе или локативном падеже, выполняющем функции датива. Будем называть эти аргументы ядерными аргументами предиката. Маловероятно, что в каком-либо из даргинских диалектов найдутся другие типы контролеров.

(2) Правило выбора контролера

Правила выбора контролера в каждом конкретном случае варьируют чрезвычайно широко. В частности, засвидетельствованы следующие возможности:

1. личная иерархия 2 > 1 > 3;
2. личная иерархия 1, 2 > 3, преимущество абсолютива;
3. личная иерархия 1, 2 > 3, преимущество эргатива;
4. подлежащный контроль.

[4] Помимо даргинского, личное согласование представлено также в лакском, удинском, табасаранском, бацбийском и ахвахском языке.

[5] Можно сказать, что контролером является аргумент в абсолютивном падеже, но когда речь идет об эргативе, то необходимо указывать, что речь идет об эргативном агенсе, проявляющем свойства подлежащего: даргинский эргатив может использоваться также для выражения, например, инструмента и в этом случае никаких видов согласования контролировать не может.

В первом случае в качестве контролера личного согласования выбирается аргумент второго лица, если он представлен среди ядерных аргументов; если аргументов второго лица среди ядерных нет, но есть аргумент первого лица, то согласование указывает на присутствие первого лица; если оба ядерных аргумента третьего лица, то в третьем лице, соответственно, стоит и глагол (форму третьего лица можно считать немаркированной), ср. примеры (1a–e) из ицаринского диалекта:

1a) du-l u r-uc-ib-di
I-ERG you F-catch:PF-AOR-2
‘Я поймал тебя (F).’

1b) ul du ucib-di ‘Ты поймал меня (M).’ (глагол в форме второго лица)

1c) dul ʁuʕr bucib-da ‘Я поймал зайца.’ (глагол в форме первого лица)

1d) muradil du rucib-da ‘Мурад поймал меня (F).’ (глагол в форме первого лица)

1e) muradil ʁuʕr bucib-cab ‘Мурад поймал зайца.’ (глагол в форме третьего лица)

Во втором случае первое и второе лица не имеют преимущества друг перед другом при выборе контролера личного согласования: если из двух ядерных аргументов один первого, а другой – второго лица, то форма глагола указывает на лицо абсолютивного аргумента; если один из аргументов первого или второго лица, а другой – третьего, то форма глагола указывает на соответственно первое или второе лицо, см. примеры из акушинского диалекта (van den Berg 2001: 16):

2a) nu-ni ħu r-it-i-ri
me-ERG you:ABS F-beat-AOR-2
‘Я побил тебя (F).’

2b) ħu-ni nu r-it-i-ra ‘Ты побил меня (F).’ (глагол в форме первого лица)

2c) dudeš-li nu r-it-i-ra ‘Отец побил меня (F).’ (глагол в форме первого лица)

2d) nu-ni rursi r-it-i-ra ‘Я побил девочку.’ (глагол в форме первого лица)

2e) dudeš-li rursi r-it-ib ‘Отец побил девочку.’ (глагол в форме третьего лица)

Третий случай отличается от второго тем, что в случае, когда один из ядерных аргументов первого, а другой – второго лица, личное согласование идет по эргативу (примеры из работы Кибрик 1979/2003 – чирагский диалект):

3a) dicce ʕu r-iqqan-da
me-ERG you:ABS F-lead-AOR-1
‘Я веду тебя (F).’

3b) ʕicce du r-iqqan-de ‘Ты ведешь меня (F).’ (глагол в форме второго лица)

3c) dicce it r-iqqan-da ‘Я веду ее.’ (глагол в форме первого лица)

3d) ite du r-iqqan-da ‘Он(а) ведет меня (F).’ (глагол в форме первого лица)

3e) ite russe r-iqqle ‘Он(а) ведет девочку.’ (глагол в форме третьего лица)

Наконец, в последнем случае мы имеем дело с более привычным нам типом личного согласования: при переходных глаголах лицо контролируется эргативным агенсом, при непереходных – единственным ядерным аргументом в абсолютиве, ср. кубачинские примеры из (Магометов 1963):

4a) dudil u wītul-da ‘I am beating you.’ (глагол в форме первого лица)

4b) udil id wītul-de ‘You are beating him.’ (глагол в форме второго лица)

4c) udil du wītul-de ‘You are beating me.’ (глагол в форме второго лица)

4d) iddil du wītul-saw ‘He is beating me.’ (глагол в форме третьего лица)

4e) iddil id wītul-saw ‘He is beating him.’ (глагол в форме третьего лица)

Правило первого типа обнаруживается в ицаринском (Sumbatova, Mutalov 2003) и кайтагском (Магометов 1976) диалектах[6]. Второе правило зафиксировано в акушинском и урахинском диалектах, а также в литературном языке (van den Berg 1999; Uslar 1892; Magometov 1976); третье представлено в чирагском (изолированный диалект на крайнем юге ареала; см. Кибрик 2003: 483 – 485). Наконец, подлежащный контроль в основном характерен для кубачинского диалекта[7] (Магометов 1963), а также для изолированного мегебского (Магометов 1982).

Ни в одном из диалектов не отмечено случаев расщепления правила контроля в зависимости от, скажем, выбора, видо-временной формы. Случаи колебания, как правило, связаны с синтаксическими особенностями (степенью слитности) конструкции.

Легко заметить, что во всех известных случаях в контроле личного согласования участвует эргативный агенс (это в основном отличает правило личного согласования от согласования по именному классу, почти везде последовательно контролируемому абсолютивом). На наш взгляд, полностью абсолютивный контроль личного согласования в других диалектах также крайне маловероятен.

3.2. Морфология личного согласования: наборы показателей лица

Во всех ранее описанных даргинских диалектах имеется по нескольку наборов показателей личного согласования, причем материально они достаточно разнообразны. Но даже самый поверхностный анализ показывает, что все это разнообразие в абсолютном большинстве случаев сводится к трем базовым наборам лично-числовых показателей, которые мы соответственно назовем (1) клитическим набором; (2) «ирреальным» набором и (3) «оптативным» набором.

[6] Ицаринский, кайтагский, а также кункинский и амухский диалекты, где также действует это правило (см. раздел 4), относятся по традиционной классификации к диалектам цудахарского типа (южным) и расположены сравнительно недалеко друг от друга. Таким образом, выяснение того, является ли согласование по личной иерархии общей генетической характеристикой одной из подгрупп даргинских языков или, возможно, ареальной чертой, в данный момент не имеет ответа.

[7] А. А. Магометов описывает ряд случаев, когда в кубачинском языке при переходных глаголах контроль может передаваться абсолютивному пациенсу, см. Магометов 1976; 1963: 274 – 275.

Форма показателей клитического набора хорошо коррелирует с традиционным членением диалектов на северные и южные – в северных это *-ra* и *-ri* (или *-ra* и *-re*), в южных – *-da* и *-di* (*-da* и *-de*). Показатели данного набора противопоставляют второе лицо единственного числа (*-ri/-re, -di/-de*) граммеме, объединяющей второе лицо множественного числа и первое лицо независимо от числа (*-ra* или *-da*):

	Южные диалекты	
	SG	PL
1		-da
2	-di (-de)	

	Северные диалекты	
	SG	PL
1		-ra
2	-ri (-re)	

Показатели этого набора в той или иной степени сохраняют свойства клитик: в большинстве случаев они употребляются не только при глаголах, но и при предикатах других типов – именах, атрибутах, обстоятельственных группах (пример 6). В конструкциях с узким (аргументным фокусом) эти показатели отделяются от глагола и прикрепляются справа к фокусной группе (примеры 5b, 7):

Акушинский

5a) nu w-ak'-i–ra quli
I M-come:PF-AOR-1 home
'Я пришел домой'.

5b) nu–ra w-ak'-ib-si quli
I-1 M-come:PF-AOR-ATR home
'Это я пришел домой.' (Хайдаков 1985: 195–196)

Кункинский (полевые данные)

1) ča u-di ? – pat'imat(-da)
who you-2 Patimat(-1)
'Кто ты? – Я Патимат.'

7) du-de lal-ti et:ij peškeš-d-arq'-ib-il / peškeš-d-arq'-ib-ce
I-2 that-PL you:DAT present-NPL-do:PF-PRET-CONV / present-NPL-do:PF-PRET-ATR
'Это я тебе их подарил(а).'

Прочие наборы личных показателей содержат обычные суффиксы, однако отличаются как собственно набором морфем, так и типом выражаемой ими оппозиции. «Ирреальный» набор противопоставляет первое лицо второму, в то время как в «оптативном» наборе оппозиция устроена так же, как в клитическом.

Маркеры «ирреального» набора

	Южные диалекты		Северные диалекты		
	SG	PL		SG	PL
1	*-d*		1	*-s*	*-ħe*
2	*-tː*	*-tː-a, -tː-aja*	2	*-d*	*-d-a, -d-aja*

Маркеры «оптативного» набора (все диалекты)

	SG	PL
1	*-a*	
2	*-i / -e*	*-a, -aja*

В ряде диалектов встречаются и другие наборы показателей лица, однако, насколько можно судить по неполным данным, все они представляют собой комбинацию маркеров клитического и «ирреального» набора. Наличие в диалектах лично-числовых показателей, существенно отличающихся по форме и типу грамматической оппозиции от перечисленных здесь, маловероятно.

3.3. Выражение числа

В первом лице единственное и множественное число, как правило, не противопоставляются. В северных диалектах можно наблюдать наличие двух разных показателей для единственного и множественного числа первого лица, но только в «ирреальном» наборе показателей.

Для второго лица характерно противопоставление числовых форм. В клитическом и «оптативном» наборе для этого имеются разные личные показатели (которые тем самым являются лично-числовыми). Кроме того, имеется специальная морфема множественного числа -a, -aː или -aja, которая обнаруживается в большинстве форм, использующих «ирреальный» набор личных показателей. Существенно, что этот показатель сочетается исключительно с формами второго лица.

3.4. Распределение наборов личных показателей

Как уже отмечалось выше, наборы личных показателей распределены по разным видо-временным парадигмам. Проблема описания состоит, в частности, в том, что состав видо-временных парадигм отдельных диалектов практически не исследован (в частности, мы знаем о нем гораздо меньше, чем о категории личного согласования), хотя очевидно. что различия между диалектами по этому параметру чрезвычайно существенны. Поэтому здесь мы можем сформулировать только самые общие тенденции, которые, видимо, являются следствием разного происхождения личных показателей.

Показатели ирреального набора – единственные, для которых точно установлено их происхождение от личных местоимений (Troubetzkoy 1929; Nikolaev, Starostin 1994). Как правило, они сочетаются с формами, употребляемыми в зависимых

клаузах: условными (которые могут быть весьма многочисленными), уступительными, конъюнктивными. Кроме того, они могут захватывать область форм с хабитуальным или гномическим значением – в частности, почти всегда употребляются в так называемом «настоящем общем» времени[8] и/или хабитуальном прошедшем.

Клитические показатели могут выступать только в вершинных клаузах предложений. В частности, только они используются в предложениях с неглагольными предикатами, а также в составе аналитических форм и более слитных форм, восходящих к аналитическим.

«Оптативные» показатели выступают в формах оптатива и сходных с ними по семантике, иногда также в формах, характерных для аподозиса условных конструкций.

4. Заключение: новые данные (с. Кунки и Худуц)

Сопоставление имеющихся сведений о даргинских диалектах позволяет хотя бы вчерне ограничить пределы варьирования основных параметров языка. В предыдущем разделе мы показали пределы такого варьирования для параметров, связанных с категорией личного согласования.

Прежде чем перейти к собственно полевым исследованиям, следует провести систематизацию сведений, касающихся других областей грамматики, прежде всего таких, которые особенно характерны для даргинского языка: например, количество и состав локализаций в именных формах; формальный состав глагольной парадигмы; именные классы и классное согласование; базовые конструкции простого предложения (непереходная, эргативная, биабсолютивная, антипассивная) и т д.

Такая методика применялась в первой экспедиции, проводившейся в рамках описываемого проекта (апрель – май 2007 г.). В качестве объекта исследования были выбраны диалекты селений Кунки и Худуц. Оба селения расположены на юго-западе даргиноязычного ареала, в верховьях реки Уллучай. Жители Кунки говорят на идиоме, распространенном только в этом селении, речь жителей с. Худуц принято относить к амухскому диалекту. Оба диалекта относятся к южным (цудахарского типа) по традиционной классификации, к юго-западным по классификации Ю. Б. Корякова.

Хотя селения Кунки и Худуц находятся на расстоянии 6–7 км друг от друга, расхождения между их диалектами весьма существенны. Жители обоих селений хотя и пассивно, но достаточно хорошо владеют языком соседей, поэтому оказалось практически невозможно выяснить, насколько эти два диалекта взаимопонимаемы: жители Кунки и Худуца практически всегда понимают друг друга, но понимание основывается не на сходстве диалектов, а на владении языком соседей.

Сбор грамматической информации показал, что диалектные различия велики, но не выходят за пределы, обозначенные предварительной схемой. В частности, параметры. связанные с выражением личного согласования, выглядят следующим образом:

[8] В акушинском и некоторых других диалектах формы, морфологически параллельные настоящему общему, выражают значение будущего времени и также присоединяют «ирреальные» окончания».

(1) контролеры: пациенс, эргативный агенс, в редких случаях дативный экспериенцер (для каждого экспериенциального глагола возможность контроля согласования отмечается в словаре);

(2) правило контроля: в обоих диалектах действует иерархия 2 > 1 > 3 (как в кайтагском и ицаринском);

(3) морфология личного согласования: в обоих диалектах представлены уже знакомые нам наборы, однако тут есть некоторые расхождения. Кункинский диалект имеет в клитическом наборе показатели -da (первое лицо, второе лицо множественного числа) и -de (второе лицо единственного числа), худуцкий, соответственно – -da и -di. Аналогичные различия наблюдаются в «оптативном» наборе (Кунки: -a, -e; Худуц: -a, -i). В «ирреальном» наборе оба диалекта имеют обычные для южных идиомов показатели –d (первое лицо) и -tː (второе лицо). Однако в идиоме Худуца в результате оглушения и дегеминации согласных на конце слова оба окончания представлены в варианте -t во всех формах, где за ними не следуют никакие другие морфемы. В результате в парадигме настоящего общего времени различие между формами первого и второго лица нейтрализуется, в отличие от других форм, использующих данный набор окончаний, ср. učʼa-t ‘я (всегда) читаю’ и ‘ты (всегда) читаешь’, но učʼa-d-i ‘я (постоянно) читал’ vs. učʼa-tː-i ‘ты (постоянно) читал’. Нейтрализация различий первого и второго лица, насколько нам известно, до сих пор не отмечалась ни для одного даргинского диалекта;

(4) выражение множественного числа: множественное число выражается только во втором лице, в клитическом и оптативном наборе при помощи специального показателя лица/числа, в «ирреальном» имеется известный уже нам показатель множественного числа -a (в кункинском диалекте встречается также вариант -ja).

(5) распределение наборов личных показателей по видо-временным парадигмам подробно не исследовано (так как не полностью известен состав глагольной парадигмы), однако общие тенденции те же, что были сформулированы в разделе 3.4: в обоих диалектах клитический набор употребляется во всех видо-временных парадигмах, употребляемых в вершинных клаузах, за исключением настоящего общего и параллельного ему хабитуального прошедшего. Две последние формы, а также разнообразные формы условных наклонений, образованные от них уступительные формы и формы конъюнктива присоединяются окончания «ирреального» набора. «Оптативные» окончания пока обнаружены только в формах оптатива и прохибитива.

Помимо собственно лингвистических наблюдений. нам пришлось, к сожалению, сделать еще один – возможно, гораздо более важный – вывод: даргинские диалекты действительно, как мы и опасались, постепенно переходят в положение угрожаемых. В последние годы жители быстро покидают горные селения, а на равнине опасность нивелирования диалектных различий и даже полной утраты даргинского языка чрезвычайно велика. В частности, бывшие жители Худуца, проживающие в Махачкале, уже сейчас владеют родным языком только пассивно[9].

[9] Особенно далеко этот процесс зашел в с. Ицари, где еще в 1991 году (год нашего первого визита) проживало около 150 семей, а сейчас их не более 20 – 30, причем многие из оставшихся также планируют переселиться на равнину.

Таким образом, проблема описания малых даргинских диалектов перестает быть чисто научной задачей: как и многие другие языки и диалекты, они нуждаются в срочном исследовании и документации в связи с реальной угрозой вымирания.

СОКРАЩЕНИЯ

1, 2 – первое, второе лицо; ABS – абсолютив; AOR – аорист; ATR – атрибутивная (причастная) форма; CONV – деепричастие; DAT – датив; ERG – эргатив; F – женский именной класс; M – мужской именной класс; NPL – неличный именно класс, множественное число; PF – перфектив; PL – множественное число; PRET – претерит; SG – единственное число

ЛИТЕРАТУРА

Абдуллаев С. Н. *Грамматика даргинского языка (фонетика и морфология)*. Махачкала, 1954.

Гасанова С. М. *Очерки по даргинской диалектологии*. Махачкала, 1970

Жирков Л. И. *Грамматика даргинского языка*. М.: Центральное издательство народов С.С.С.Р., 1926.

Кибрик А. Е., Кодзасов С. В. *Сопоставительное изучение дагестанских языков. Глагол*. М., 1988.

Кибрик, А. Е. Материалы к типологии эргативности. Даргинский язык (чирагский диалект). В кн.: Кибрик, А. Е. *Константы и переменные языка*. СПб: Алетейя, 2003. С. 482–504 [первое издание: *Материалы к типологии эргативности, Предварительные публикации Института русского языка АН СССР*. Вып. 127. М., 1979].

Коряков, Ю. Б. *Атлас кавказских языков*. РАН, Институт языкознания. М.: Пилигрим, 2006.

Магометов, А. А. *Кубачинский язык*. Тбилиси: АН Грузинской ССР. 1963.

Магометов, А. А. Личное спряжение в даргинском языке сравнительно со спряжением в табасаранском и лакском языках. *Иберийско-кавказское языкознание XIII*, 1962. С. 313–342.

Магометов, А. А. Личные формы инфинитива в даргинском языке. *Иберийско-кавказское языкознание XX*, 1978. С. 264–279.

Магометов, А. А. *Мегебский диалект даргинского языка*. Тбилиси: Мецниереба, 1982.

Магометов, А. А. Субъектно-объектное согласование глагола в лакском и даргинском языках. *Ежегодник иберийско-кавказского языкознания III*, 1976. С. 203–218.

Мусаев, М.-С. М. *Даргинский язык*. М.: Academia, 2002.

Мусаев, М.-С. М. Даргинский язык. В кн.: *Языки мира. Кавказские языки*. М.: Academia, 2001. С. 357–369.

Темирбулатова С. М. *Хайдакский диалект даргинского языка*. Махачкала, 2004.

Услар, П. К. *Этнография Кавказа. Языкознание. V. Хюркилинский язык.* Тифлис: Управление кавказского учебного округа, 1892.

Хайдаков С. М. *Даргинский и мегебский языки (принципы словоизменения)*. М.: Наука, 1985.

Helmbrecht, Johannes. 1996. The syntax of personal agreement in East Caucasian languages. *Sprachtypologie und Universalienforschung 49*: 127–148.

Nikolaev, S. L. and Starostin, S. A. 1994. *A North Caucasian Etymological Dictionary*. Moscow: Asterisk.

Sumbatova, N. R., and R. O. Mutalov. 2003. *A Grammar of Icari Dargwa*. (Languages of the Worlds/Materials 92.) München: LINCOM Europa.

Troubetzkoy, N. S. 1929. Notes sur les désinences du verbe dans les langues tchétchénolesghiennes (caucasiques-orientales). *Bulletin de la Société de Linguistique de Paris*, t. 29 fasc.3 (No.88) : 153–171.

Van den Berg, Helma. 1999. Gender and person agreement in Aqusha Dargi. *Folia linguistica XXXIII*/1: 153–168

Van den Berg, Helma. 2001. Dargi Folktales. Oral Stories from the Caucasus and an Introduction to Dargi Grammar. Leiden: CNWS.

НАШИ АВТОРЫ

Лилит Варданян, Университет Пизы, Пиза, Италия

Гаяне Геворгян, Институт языка им. Ачаряна НАН, Ереван

Давид Гюрджинян, Ереванский государственный лингвистический университет им. В. Брюсова, Ереван

Михаил Даниэль, Московский государственный университет; Восточноармянский национальный корпус, Corpus Technologies, Москва

Дмитрий Левонян, Corpus Technologies, Svarog Capital, со-основатель проекта ВАНК, Москва

Эмилия Нерсисянц, Университет Тегерана, Тегеран, Исламская Республика Иран

Лиана Овсепян, Ереванский государственный университет, Ереван

Владимир Плунгян, Институт языкознания, РАН; Московский государственный университет; со-основатель проекта ВАНК, Москва

Алексей Поляков, Восточноармянский национальный корпус, Corpus Technologies, Москва

Сергей Рубаков, Восточноармянский национальный корпус, Corpus Technologies, Москва

Нина Сумбатова, Российский государственный гуманитарный университет, Москва

Сусанна Тиоян, Ереванский государственный лингвистический университет им. В. Брюсова, Ереван

Роберт Урутян, Институт языка им. Ачаряна НАН, Ереван

Виктория Хуршудян, Восточноармянский национальный корпус, Corpus Technologies, Москва

Нарине Экекян, Ереванский государственный лингвистический университет им. В. Брюсова, Ереван

OUR AUTHORS

Michael Daniel, Moscow State University; Eastern Armenian National Corpus, Corpus Technologies, Moscow

Gayane Gevorgian, Institute of Language after Acharyan, NAS, Yerevan

David Gyurjinian, Yerevan State Linguistic University after V. Brusov, Yerevan

Narine Hekekian, Yerevan State Linguistic University after V. Brusov, Yerevan

Liana Hovsepian, Yerevan State University, Yerevan

Victoria Khurshudian, Eastern Armenian National Corpus, Corpus Technologies, Moscow

Dmitri Levonian, Svarog Capital; Corpus Technologies, EANC co-founder, Moscow

Emilia Nercissians, University of Tehran, Tehran, Islamic Republic of Iran

Vladimir Plungian, Institute of Linguistics, RAS; Moscow State University; EANC co-founder, Moscow

Aleksey Polyakov, Eastern Armenian National Corpus, Corpus Technologies, Moscow

Sergey Rubakov, Eastern Armenian National Corpus, Corpus Technologies, Moscow

Nina Sumbatova, Russian State University for Humanities, Moscow

Susanna Tioian, Yerevan State Linguistic University after V. Brusov, Yerevan

Robert Urutian, Institute of Language after Acharyan, NAS, Yerevan

Lilit Vardanian, University of Pisa, Pisa, Italia

ЦЕНТР ПОДДЕРЖКИ АРМЕНИСТИЧЕСКИХ ИССЛЕДОВАНИЙ

Москва

•

Армянский гуманитарный вестник

Հայագիտական ուսումնասիրություններ պարբերագիր

Bulletin of Armenian Studies

2/3-II 2009

ISBN 978–99941–1–646–1

Печать офсетная. Формат 70x100 1/16. Бумага офсетная, 80 г.
Объем 7.5 п.л. Тираж 500 экз.

ИЗДАТЕЛЬСТВО «ЗАНГАК–97»

0051, Ереван, пр. Комитаса 49/2, тел.: (+37410) 23-25-28,
факс: (+37410) 23–25–95, эл. почта: info@zangak.am,
эл. сайт: www.zangak.am, www.book.am

Отпечатано в типографии издательства "Зангак-97"