

Российский государственный гуманитарный университет  
Russian State University for the Humanities



RGGU BULLETIN

№ 16(59)/10

Scientific Journal

Series “Philological Sciences. Linguistics” /  
Journal of Language Relationship  
Issue 4 (2010)

Moscow 2010

ВЕСТНИК РГГУ

№ 16(59)/10

Научный журнал

Серия «Филологические науки. Языкознание» /  
«Вопросы языкового родства»  
Выпуск 4 (2010)

Москва 2010

УДК 81(05)  
ББК 81я5

Главный редактор

Е.И. Пивовар

Заместитель главного редактора

Д.П. Бак

Ответственный секретарь

Б.Г. Власов

Главный художник

В.В. Сурков

Редакционный совет серии

«Филологические науки. Языкознание» / «Вопросы языкового родства»

Председатель Вяч. Вс. Иванов (Москва – Лос-Анджелес)

М. Е. Алексеев (Москва)

В. Блажек (Брно)

У. Бэкстер (Анн Арбор)

В. Ф. Выдрин (Санкт-Петербург)

М. Гелл-Манн (Санта Фе)

А. Б. Долгопольский (Хайфа)

Ф. Кортландт (Лейден)

А. Лубоцкий (Лейден)

Редакционная коллегия серии:

В. А. Дыбо (главный редактор)

Г. С. Старостин (заместитель главного редактора)

Т. А. Михайлова (ответственный секретарь)

К. В. Бабаев

С. Г. Болотов

А. В. Дыбо

А. С. Касьян

О. А. Мудрак

В. Е. Чернов

ISSN 1998-6769

© Российский государственный  
гуманитарный университет, 2010



# Table of Contents / Содержание

Table of Contents / Содержание . . . . .	vii
Contributors / Сведения об авторах . . . . .	ix
Note for Contributors / Будущим авторам . . . . .	x
Editors' Note. A Happy Jubilee to Aaron Dolgopolsky . . . . .	xi
[От редакции. А. Б. Долгопольскому — 80 лет]	

## Articles / Статьи

<i>Kirill Babaev</i> . Reconstructing Benue-Congo person marking II . . . . .	1
[К. В. Бабаев. Реконструкция личных показателей прабенуэ-конго, ч. II]	
<i>Alexei Kassian, George Starostin, Anna Dybo, Vasiliy Chernov</i> . The Swadesh wordlist. An attempt at semantic specification . . . . .	46
[А. С. Касьян, Г. С. Старостин, А. В. Дыбо, В. Е. Чернов. Уточнение семантики для элементов списка Сводеша]	
<i>Ilia Peiros</i> . Uto-Aztec comparative linguistics and etymological databases . . . . .	90
[И. И. Пейрос. Сравнительное изучение юто-ацтекских языков и этимологические базы данных]	
<i>Paul Sidwell</i> . The Austroasiatic central riverine hypothesis . . . . .	117
[П. Сидвелл. Гипотеза «центрально-речной» прародины австроазиатской семьи]	
<i>Gábor Takács</i> . The Afro-Asiatic background of West Rift *ʎ, *ɥ, *ʕ, *h, and *ʔ . . . . .	135
[Г. Такач. Афризийское происхождение западно-рифтских фонем *ʎ, *ɥ, *ʕ, *h и *ʔ]	
<i>Mikhail Zhivlov</i> . Studies in Uralic vocalism I: A more economical solution for the reconstruction of the Proto-Permic vowel system . . . . .	167
[М. А. Живлов. Исследования в области уральского вокализма I: Экономное решение проблемы реконструкции прапермской системы гласных]	

## Discussion Articles / Дискуссионные статьи

<i>V. Д. Соловьев</i> . Типологическая схожесть языков как метод изучения языковой эволюции. . . . .	177
[V. Solovyov. Using typological similarity of languages as a method for studying the evolution of language]	

## Book reviews / Рецензии

A. FOURNET & A. R. VOMHARD. The Indo-European Elements in Hurrian, 2010 ( <i>Alexei Kassian / А. С. Касьян</i> ) . . . . .	199
Аспекты компаративистики 4 [Aspects of Comparative Linguistics 4], 2009 ( <i>И. С. Якубович / Илья Якубович</i> ) . . . . .	207

**Periodic reviews / Периодика**

The Journal of Indo-European Studies. Vol. 36, 2008  
(Т. А. Михайлова / *Tatyana Mikhailova*) . . . . . 212

**Reports / Хроника**

Пятое традиционное чтение памяти С. А. Старостина,  
Москва, РГГУ, 2—3 апреля 2010 г. (Е. В. Коровина / *Evgeniya Korovina*) . . . . . 217  
[The 5<sup>th</sup> Traditional Conference in Memory of S. A. Starostin, Moscow, RSUH, April 2–3, 2010]

## Сведения об авторах

*Бабаев Кирилл Владимирович* — канд. филол. наук, доцент Института восточных культур и античности РГГУ, [babaev@yandex.ru](mailto:babaev@yandex.ru)

*Дыбо Анна Владимировна* — доктор филол. наук, чл.-кор. РАН, в.н.с. Центра компаративистики ИВКА РГГУ, зав. отделом урало-алтайских языков Института языкознания РАН (Москва), [adybo@rinet.ru](mailto:adybo@rinet.ru)

*Живлов Михаил Александрович* — канд. филол. наук, научный сотрудник отдела урало-алтайских языков Института языкознания РАН (Москва), [zhivlov@yandex.ru](mailto:zhivlov@yandex.ru)

*Кассьян Алексей Сергеевич* — канд. филол. наук, преп. Центра компаративистики ИВКА РГГУ (Москва), [a.kassian@gmail.com](mailto:a.kassian@gmail.com)

*Коровина Евгения Владимировна* — студент Центра компаративистики ИВКА РГГУ (Москва)

*Михайлова Татьяна Андреевна* — доктор филол. наук, проф. кафедры германской и кельтской филологии филологического факультета МГУ (Москва), [tamih.msu@mail.ru](mailto:tamih.msu@mail.ru)

*Пейрос Илья Иосифович* — доктор филол. наук, Институт Санта Фе (Нью-Мексико, США), [ilia@santafe.edu](mailto:ilia@santafe.edu)

*Сидвелл Пол* — руководитель мон-кхмерского отдела Центра исследований по вычислительной лингвистике, Бангкок; внештатный научный сотрудник Австралийского национального университета, Канберра, [paul.sidwell@anu.edu.au](mailto:paul.sidwell@anu.edu.au)

*Соловьев Валерий Дмитриевич* — доктор физ.-мат. наук, зам. директора Института информатики Казанского государственного университета, [maki.solovyev@mail.ru](mailto:maki.solovyev@mail.ru)

*Старостин Георгий Сергеевич* — канд. филол. наук, зав. кафедрой истории и филологии Дальнего Востока ИВКА РГГУ (Москва), [gstarst@rinet.ru](mailto:gstarst@rinet.ru)

*Такач Габор* — научный сотрудник отдела египтологии Будапештского университета, Венгрия, [gabtak@datatrans.hu](mailto:gabtak@datatrans.hu)

*Чернов Василий Евгеньевич* — программист, ООО «Яндекс» (Москва), [vechernov@yandex-team.ru](mailto:vechernov@yandex-team.ru)

*Якубович Илья Сергеевич* — кандидат филол. наук, научный сотрудник Института мировой культуры МГУ (Москва), [sogdiana783@gmail.com](mailto:sogdiana783@gmail.com)

## Contributors

*Kirill V. Babaev* — candidate of sciences (Philology), Associate Prof., Russian State University for the Humanities (Moscow), [babaev@yandex.ru](mailto:babaev@yandex.ru)

*Vasily Ye. Chernov* — programmer, “Yandex” (Moscow), [vechernov@yandex-team.ru](mailto:vechernov@yandex-team.ru)

*Anna V. Dybo* — doctor of sciences (Philology), corresponding member of the Russian Academy of Sciences, lead researcher, Center for Comparative Linguistics, Russian State University for the Humanities; Head of Department of Uralo-Altaic Studies, Institute of Linguistics, Russian Academy of Sciences (Moscow), [adybo@rinet.ru](mailto:adybo@rinet.ru)

*Alexei Kassian* — candidate of sciences (Philology), researcher, Center for Comparative Linguistics, Russian State University for the Humanities (Moscow), [a.kassian@gmail.com](mailto:a.kassian@gmail.com)

*Eugenia V. Korovina* — student, Center for Comparative Linguistics, Russian State University for the Humanities (Moscow)

*Tatyana A. Mikhailova* — doctor of sciences (Philology), professor, Department of Germanic and Celtic Philology, Faculty of Philology, Moscow State University (Moscow), [tamih.msu@mail.ru](mailto:tamih.msu@mail.ru)

*Ilya Peiros* — doctor of sciences (Philology), visiting researcher, Institute of Santa Fe (New Mexico, USA), [ilia@santafe.edu](mailto:ilia@santafe.edu)

*Paul Sidwell* — Paul Sidwell — Director of Mon-Khmer projects, Centre for Research in Computational Linguistics (Bangkok) and Visiting Research Fellow, Australian National University (Canberra).

*Valery D. Solovyov* — doctor of sciences (Physics and Mathematics), vice-Director of the Institute of Informatics of Kazan State University, [maki.solovyev@mail.ru](mailto:maki.solovyev@mail.ru)

*George Starostin* — candidate of sciences (Philology), Head of Department of the history and philology of the Far East, Institute of Eastern Cultures and Antiquity, RSUH (Moscow), [gstarst@rinet.ru](mailto:gstarst@rinet.ru)

*Gabor Takács* — researcher, Department of Egyptology, Eotvos Lorand University, Budapest, Hungary, [gabtak@datatrans.hu](mailto:gabtak@datatrans.hu)

*Ilya S. Yakubovich* — candidate of sciences (Philology), research associate, Institute of World Cultures, Moscow State University, [sogdiana783@gmail.com](mailto:sogdiana783@gmail.com)

*Mikhail A. Zhivlov* — candidate of sciences (Philology), researcher, Department of Uralo-Altaic Studies, Institute of Linguistics, Russian Academy of Sciences (Moscow), [zhivlov@yandex.ru](mailto:zhivlov@yandex.ru)

## Note for Contributors

*Journal of Language Relationship* welcomes submissions from everyone specializing in comparative-historical linguistics and related disciplines, in the form of original articles as well as reviews of recent publications. All such submissions should be sent to the managing editor:

G. Starostin  
Institute of Oriental Cultures and Antiquity  
Russian State University for the Humanities  
125267 Moscow, Russia  
Miuskaya Square, 6  
E-mail: [journal@nostratic.ru](mailto:journal@nostratic.ru)

Articles are published preferably in English or Russian, although publication of texts in other major European languages (French, German, etc.) is possible. Each article should be accompanied with an abstract (not exceeding 300 words).

For more detailed guidelines on article submission and editorial policies, please see our Website at: <http://journal.nostratic.ru> or address the editorial staff directly at [journal@nostratic.ru](mailto:journal@nostratic.ru).

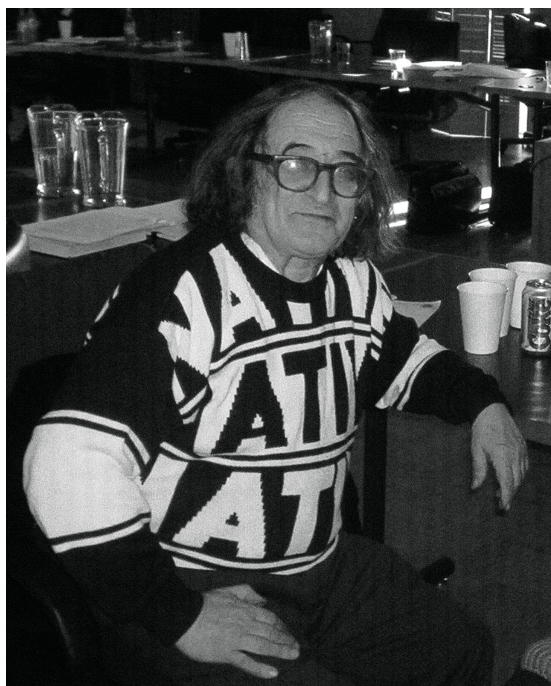
## Будущим авторам

Журнал *Вопросы языкового родства* принимает заявки на публикацию оригинальных научных статей, а также рецензий, от всех, кто специализируется в области сравнительно-исторического языкознания и смежных дисциплин. Рукописи можно высылать непосредственно заместителю главного редактора по адресу:

125267 Москва  
Миусская площадь, д. 6  
Российский государственный гуманитарный университет  
Институт восточных культур и античности  
Г. Старостину  
E-mail: [journal@nostratic.ru](mailto:journal@nostratic.ru)

Предпочтительные языки публикации — английский или русский, хотя возможна также публикация статей на других европейских языках (французский, немецкий и т. п.). К каждой статье обязательно прикладывается резюме (не более 300 слов).

Подробнее о требованиях к оформлению рукописи, редакционной политике журнала и т. п. Вы можете узнать на нашем сайте по адресу: <http://journal.nostratic.ru>, или же непосредственно обратившись к редакции по электронной почте ([journal@nostratic.ru](mailto:journal@nostratic.ru)).



## Editors' Note. A Happy Jubilee to Aaron Dolgopolsky

The editors of *Journal of Language Relationship* would like to use this opportunity to congratulate Aaron B. Dolgopolsky (b. November 18, 1930), one of the leading figures in long-range comparison and comparative linguistics as a whole, with an important jubilee date.

A. Dolgopolsky is generally known as one of the forefathers of the Nostratic theory — deep genetic relationship between Afro-Asiatic, Kartvelian, Indo-European, Uralic, Altaic, and Dravidian languages — of which he was one of the early pioneers along with the late V. M. Illich-Svitych and V. A. Dybo; more than fifty years of his continuous research in this field have recently culminated in the online publication of the voluminous *Nostratic Dictionary* (University of Cambridge, 2008, preliminary publication) that contains more than 2,800 etymologies, a veritable treasure grove of comparative data on early prehistoric relationships between the major language families of Eurasia. However, A. Prof. Dolgopolsky's research has also been equally valuable in more narrow areas of comparative linguistics, such as Indo-European, Cushitic, and Semitic studies; his monograph on the historical phonology of Hebrew in the general context of its Semitic and Afro-Asiatic origins (*From Proto-Semitic to Hebrew*, 1999) has been widely recognized as one of the most insightful works on this issue.

In achieving all of these results, Prof. Dolgopolsky has had to overcome numerous obstacles — skepticism and rejection by the “official science” while in the Soviet Union; forced emigration; in recent years, serious health problems — but we are happy to know that, despite all this, he is still going strong and continuing work on a further improved version of the *Nostratic Dictionary*. On behalf of all the editors, advisory board, and (we dare to hope) most readers of *JLR*, we wish him all the best and hope that his immense erudition, expertise, and talent will continue to guide us all for years to come.

We are happy to enclose the biography of A. Dolgopolsky and to list his major papers in Russian below.

## От редакции. А. Б. Долгопольскому — 80 лет

Один из крупнейших современных специалистов-компаративистов, Арон Борисович Долгопольский родился 18 ноября 1930 г. в Москве. Образование получил в Московском государственном педагогическом институте иностранных языков им. Мориса Тореза (ныне Лингвистический университет), на испанском отделении (окончил в 1954 г.). Защитил кандидатскую диссертацию по словообразованию в романских языках (1958). Научный сотрудник Института русского языка (1962–66), затем Африканского сектора Институ-

та языкознания (1966—1976), где изучал кушитские языки. Уже в конце 50-х годов (1957—1958), занимаясь изучением креолизации в языках мира, Арон Борисович заинтересовался проблематикой дальнего родства. Будучи в начале своей научной деятельности далек от классического сравнительно-исторического языкознания, но зато хорошо знаком с математическими лингвистами и их деятельностью (в Институте Мориса Тореза процветала знаменитая Лаборатория машинного перевода под руководством В. Ю. Розенцвейга, где в разное время работали А. К. Жолковский, Вяч. Вс. Иванов, Н. Н. Леонтьева, Ю. С. Мартемьянов, И. А. Мельчук и другие известные структуралисты) и обладая богатым опытом переводческой деятельности, Арон Борисович в начале 1960-х годов, размышляя над сущностью переводческих трансформаций, осуществил несколько блестящих пионерских исследований о применении трансформационно-переводного метода в сфере сопоставительной лексикологии (4) и о контекстных факторах выбора граммы вида при переводе на русский язык (7). Однако вскоре его интересы целиком поглощаются компаративистикой. Арон Борисович поначалу полагал, что доказательство «ностратического родства» (по тогдашней его терминологии, «сиби́ро-европейского»; в отличие от В. М. Иллич-Свитыча, он включал тогда в него чукотско-камчатскую семью и не включал дравидийскую) невозможно с помощью сравнительного метода, но может быть обосновано математической статистикой. На эту тему он написал несколько остроумных статей (см. 1961 (5), 1964 (8) = 1986 (107), 1965 (14)), пока в 1964 г. не познакомился со славистом и индоевропеистом В. М. Иллич-Свитычем, который к этому времени уже второй год серьезно занимался созданием этимологического корпуса ностратической семьи языков и построением их сравнительно-исторической фонетики. Знакомство их состоялось в Центрально-справочном отделе Библиотеки им. В. И. Ленина (хранилище словарей), с помощью В. А. Дыбо, который хорошо знал обоих лингвистов. Со свойственной ему страстью, Долгопольский немедленно влился в работу по этому направлению и вскоре достиг в ней значительных успехов. Кроме того, он параллельно с В. М. Иллич-Свитычем занялся реконструкцией морфологии в ностратическом. Поскольку оба лингвиста полагали к этому моменту наименее разработанной афразийскую (семито-хамитскую) реконструкцию, они разделили работу в этой области: Иллич-Свитыч взял на себя чадские языки, Долгопольский — кушитские. После безвременной гибели В. М. Иллич-Свитыча в 1966 г. Долгопольский вместе с В. А. Дыбо принял деятельное участие в издании трудов Иллич-Свитыча по ностратике. Кроме того, он продолжал уточнение фонетической реконструкции ностратического консонантизма. Арон Борисович принял решение до выхода основной части работы Иллич-Свитыча не публиковать собственных крупных работ по ностратике, и в этот период у него выходят только статьи по частным вопросам ностратической реконструкции (1969 (30), 1971 (39), 1972 (45, 46), 1974 (72), 1984 (105)). При этом значительное внимание он уделял популяризации идей ностратики, читая лекции, давая интервью и публикуя статьи в научно-популярных журналах (прежде всего «Знание — сила»). Большой вклад в популяризацию лингвистических знаний внесли его многочисленные энциклопедические статьи в БСЭ о языках мира (особенно об афразийских и америндских языках) и письменностях, в которых проявилась эрудиция и разносторонность таланта Арона Борисовича<sup>1</sup>, а также многолетняя плодотворная работа по лингвистическому консультированию БСЭ 3-го издания (1969—1978) в части, касающейся языков мира, их генеалогической группировки и письменностей мира.

В 1972 г., после проведения в Институте славяноведения и балканистике в Москве конференции по сравнительно-исторической грамматике индоевропейских языков (12—14 декабря), был организован Ностратический семинар имени В. М. Иллич-Свитыча. Основное ядро его составили молодые лингвисты В. А. Терентьев, О. В. Столбова, Е. А. Хелимский, В. Э. Орёл, И. И. Пейрос, С. А. Старостин — студенты и выпускники Отделения структурной и прикладной лингвистики филологического факультета МГУ, многие из которых еще в школьном возрасте слушали лекции Арона Борисовича по проблемам сравнительно-исторического языкознания и дальнего родства языков, читавшиеся с 1966 по 1969 г. в рамках лингвистического кружка для школьников, функционировавшего при Отделении. Семинар заседал в первые годы по квартирам участников, но с 1975 г. получил официальную локализацию — в Институте славяноведения и балканистики. В 1974—1976 гг. А. Б. Долгопольский прочел здесь цикл докладов, посвященных проблеме ностратической прародины и уточнению корпуса ностратических сопоставлений.

В 1976 г. А. Б. Долгопольский по семейным обстоятельствам эмигрировал в Израиль. К сожалению, его эмиграция произошла как раз в тот момент, когда развитие ностратических исследований в Москве двига-

<sup>1</sup> Порой в этих статьях можно найти и остроумные общелингвистические положения: например, в статье «Знаки препинания» (написанной совместно с А. А. Леонтьевым) среди знаков препинания перечислен пробел (эта справедливая общетеоретическая идея, высказанная авторами как бы мимоходом, представляет собой полезное новшество).



лось настолько успешно, что привлекло внимание чиновников Академии наук и попало в список «основных достижений народного хозяйства», что создало новые возможности для развития исследований в СССР<sup>2</sup>. Ностратический семинар при Институте славяноведения и балканистики получил официальный статус, организовывал конференции и экспедиции; в феврале 1978 г. после доклада В. А. Дыбо на заседании Президиума РАН ностратические исследования были окончательно включены в состав направлений советской лингвистики и поставлены в план Института славяноведения и балканистики. Институт даже принял на стажировку по ностратике специалиста из братской страны — венгерку И.Хегедюш, впоследствии одного из ведущих европейских ностратистов. Члены Семинара продолжали работать в первую очередь над уточнением реконструкций отдельных семей внутри и вне ностратической макросемьи (но в контексте ностратической теории), а также над расширением и уточнением круга ностратических сопоставлений. А. Б. Долгопольский после переезда в Хайфу должен был заняться в первую очередь более традиционными областями сравнительно-исторического языкознания. В университете Хайфы он был загружен курсами по семитологии, прежде всего по исторической грамматике иврита в семитологическом контексте. Соответственно этому, он работал над продвижением семитской и афразийской реконструкций (1977 (101), 1982 (103), 1983 (104)). В частности, ему удалось применить методы Московской акцентологической школы к ряду проблем семитской реконструкции, например, к проблеме сеголатных имен (1978 (102), 1986 (106)). Несомненно, работы Арона Борисовича по семитологии выводят сравнительно-историческое языкознание в этой области на новый уровень — ср., хотя бы его историческую фонетику иврита (1999(135)).

В 1988 году, в связи с успехами перестройки, В. В. Шеворошкину вместе с американскими коллегами удалось организовать конференцию по проблемам дальнего родства в Мичиганском университете в Анн-Арборе и вывезти на нее практически весь наличный на тот момент состав Ностратического семинара. Там и состоялась, после десятилетней разлуки, встреча московских лингвистов с Долгопольским и Шеворошкиным, и первая (для большей части москвичей это был вообще первый выезд за пределы СССР) встреча восточноевропейских и западных специалистов по дальнему родству.

Последние два десятилетия Арон Борисович продолжает свои исследования по афразистике и одновременно развивает ностратику, в основном, по двум направлениям: сложные места морфологической реконструкции и полный ностратический словарь. В 2008 г. предварительная версия ностратического словаря А. Б. Долгопольского была опубликована в электронном виде (2008(143)). Потрясает прежде всего объем проведенной работы — более 2800 этимологических статей с полным приведением материала и библиографических ссылок. Последние пять лет Арон Борисович воздерживается от публикаций, целиком по-

<sup>2</sup> Любопытно в этой связи посмотреть, как выполнялось в случае с Ароном Борисовичем общее правило о запрете на упоминание фамилии ученого в печати, выходявшей после его отъезда. Запрет этот осуществлялся обычно на уровне работы издательского редактора с автором или ответственным редактором книги. Редактор заявлял, что с фамилией книга не может быть выпущена. Автор, посопротивившись, находил такую форму ссылки, чтобы фамилия не упоминалась, но «все» всем было ясно. Общая инерционность, свойственная советской системе, по-видимому, сказалась здесь в обоих направлениях. Книги, вышедшие в конце 1976 года — «Опыт сравнения ностратических языков — 2» и «Тайны древних письмен» (под ред. И. М. Дьяконова) — еще содержат и работы Долгопольского, и правильно оформленные ссылки с его фамилией (хотя уехавший годом раньше — и с более громким скандалом — В. В. Шеворошкин фигурирует в «Тайнах» уже только под псевдонимом «А. Ф. Деянов»). Эта же ситуация продолжает существовать в 1978 и в 1979 г.: в сборнике тезисов докладов конференции «Ностратические языки и ностратическое языкознание» (М., Институт славяноведения и балканистики, 1977) в тезисах Н. С. Пановой «Дополнения к ностратическим этимологиям по группе марги», с. 56, дважды упомянут А. Б. Долгопольский. В книге «Этимология 77» (М., «Наука», 1979) имеется статья В. А. Терентьева «Ностратические этимологии», где на с. 163 стоит ссылка на работу Долгопольского 1964 г. Ухудшилось положение к 1981 г.: два выпуска «Сравнительно-исторического словаря афразийских языков» (авторы: А.Г.Белова, И. М. Дьяконов, А.Ю.Милитарев, В.Я.Порхомовский, О. В. Столбова) ссылаются на «Сравнительно-историческую фонетику кушитских языков» по заглавию, без упоминания фамилии автора. В 3-м выпуске «Опыта сравнения ностратических языков» (М., «Наука», 1984) В. А. Дыбо в предисловии два раза по этому поводу отсылается к предыдущему выпуску 1976 г., где Долгопольский еще свободно цитирован. В статье В. А. Дыбо и В. А. Терентьева «Ностратическая семья и ее временная локализация» в материалах конференции «Лингвистическая реконструкция и древнейшая история Востока» (М., Институт востоковедения, 1984) дается ссылка на место публикации без фамилии и заглавия статьи. Семитологи же и дальше сохранили изобретенный ими метод ссылки, который мы видим уже и в самый разгар перестройки, в 1987 г., в книге «Африканское историческое языкознание. Проблемы реконструкции» (М., «Наука», Главная редакция Восточной литературы, 1987): на с. 29 и на с. 293 стоит название «Сравнительно-исторической фонетики кушитских языков» без фамилии автора. Однако уже в издании «Proceedings of the 9th International congress of Ethiopian studies — 1986», (М., Наука, 1988) в статье О.В. Столбовой «Cushitic and West Chadic: isoglosses and cognate sets» в библиографии стоит фамилия Долгопольского. И в 1989 г. в материалах конференции «Лингвистическая реконструкция и древнейшая история Востока» (М., Институт востоковедения) работы Арона Борисовича цитируются под его собственным именем — следовательно, запрет был уже снят.

свящая свое время выверке словаря. Сыграла свою роль и тяжелая болезнь, перенесенная ученым, вследствие которой он не рискует летать самолетом и отказывается от научных поездок.

Сейчас, однако, Арон Борисович творчески активен и всегда готов к научному общению с друзьями и учениками. Пожелаем же ученому к 80-летию продолжения его «хорошей полосы», завершения Ностратического словаря и новых ярких открытий!

Нижеследующий список трудов А. Б. Долгопольского содержит как публикации, так и неопубликованные рукописи; опубликованные работы пронумерованы арабскими цифрами, рукописи — латинскими буквами<sup>3</sup>.

## Статьи и монографии

### 1955

- (1) Долгопольский А. Б., Против ошибочной концепции «гибридных» языков. (О креольских наречиях) // Уч. зап. 1 Моск. гос. пед. института иностр. языков», 1955, т. 7.

### 1958

- (2) Теория вероятности и установление языкового родства // Тезисы докладов конференции по машинному переводу (15—21 мая 1958 года). М., 1958. С. 38—39.
- (3) Из истории развития типов отглагольных имен деятеля от латыни к романским языкам. (К проблеме развития словообразовательных типов). Автореф. дисс. ... к.филол.н. М., 1958.

### 1960

- (4) Изучение лексики с точки зрения трансформационно-переводного анализа плана содержания в языке // Тезисы докладов на VI пленарном заседании Комиссии, посвящённом современной проблематике лексикологии и семасиологии, 19—20 окт. 1960 г. М., 1960. С. 35—42.

### 1961

- (5) Статистическое изучение сохраняемости лексики // Тезисы докладов Межвузовской конференции по применению структурных и статистических методов исследования словарного состава языка. М., 1961. С. 87—90.

### 1963

- (6) *Están emparentadas las lenguas de Europa, Asia y Africa del Norte?* // *Tlatoani* (México) 1963, 17.
- (7) Категория вида в русском языке и вероятностный характер связи означающего с означаемым // Шаумян С. К. (отв. ред.) Проблемы структурной лингвистики 1963. М.: Изд-во АН СССР, 1963. С. 266—281, с табл.

### 1964

- (8) Гипотеза древнейшего родства языков Северной Евразии с вероятностной точки зрения // Вопросы языкознания 1964, № 2. С. 53—63.
- (9) Гипотеза древнейшего родства языков Северной Евразии (Проблемы фонетических соответствий) // VII Международный конгресс антропологических и этнографических наук (август 1964). М., 1964. С. 1—22.
- (10) Методы реконструкции общеиндоевропейского языка и внеиндоевропейские сопоставления // Проблемы сравнительной грамматики индоевропейских языков (тезисы докладов). М., Изд-во МГУ, 1964. С. 27—30.

### 1965

- (11) Методы реконструкции общеиндоевропейского языка и сибироевропейская гипотеза // Этимология 1964. М., Наука, 1965. С. 259—270.
- (12) [Хроника:] VII Конгресс антропологических и этнографических наук (Москва, август 1964) // Вопросы языкознания, 1965, № 1. С. 132—136.

---

<sup>3</sup> Библиография составлена при участии С. А. Крылова.



- (13) [перевод и примечания] *Сводеш М.* Лингвистические связи Америки и Евразии // *Этимология* 1964. М., Наука, 1965. С. 259—70; 311—322.
- (14) Сохраняемость лексики, универсалии и ареальная типология // *Лингвистическая типология и восточные языки*. М., Наука, 1965. С. 189—195.
- (a) Полевые записи Мача Оромо. М., 1965—68. Ркп.

## 1966

- (15) Материалы по сравнительно-исторической фонетике кушитских языков: губные и денгальные смычные в начальном положении // *Языки Африки* (ред. Б.А.Успенский). М., Наука, 1966. С. 35—88.
- (16) Ностратические основы с сочетанием двух шумных согласных // *Проблемы славянских этимологических исследований в связи с общей проблематикой этимологии*. М., 1966. С. 48—50.
- (17) В. М. Иллич-Свитыч (некролог) // *Известия АН СССР, серия литературы и языка*. 25/6, 1966. С. 563—564. (совм. с: В. А. Дыбо)
- (18) Языки — братья, дедушки и племянники // *Знание — сила*, 1966, № 2. С. 10—14.
- (19) Как говорили шесть тысяч лет тому назад // *Знание — сила*, 1966, № 7. С. 26—29.
- (20) О языковедах и языковедении // *Кондратов А. М. Звуки и знаки*. М.: Знание, 1966, с. 3.
- (21) Письмена планеты // *Наука и жизнь*, 1966, № 4. С. 88—91.

## 1967

- (22) От Сахары до Камчатки языки ищут родственников // *Знание — сила* 1967, № 1. С. 43—46.
- (23) В поисках дальнего родства // *Русская речь* 1967, № 6. С. 95—109, с табл...
- (24) Структура семитохамитского корня в сравнительно-историческом освещении // *Проблемы языкознания*. М., Наука, 1967. С. 278—282.
- (25) Cushitic roots with the initial \*m ~ \*b // *II International congress of africanists. Papers presented by the USSR delegation. Moscow 1967*. С. 3—12
- (26) Пособие по устному переводу с испанского языка на русский. Сост. Г. Я. Туровер, И. А. Триста, А. Б. Долгопольский. М., Высшая школа. 1967. 262 стр.

## 1968

- (27) Древние корни и древние люди // *Русская речь* 1968/2, 96—108.
- (28) Корректирующие примечания к статье В.М. Иллич-Свитыча «Соответствия смычных в ностратических языках» // *Этимология* 1966. М., Наука, 1968. С. 401—404.

## 1969

- (29) Ностратические основы с сочетанием шумных согласных // *Этимология* 1967. М., Наука, 1969. С. 296—313.
- (30) Языки Африки и аргумент метели // *Знание — Сила* 1969, № 11: 26—28.
- (31) Австралийские языки // *Большая Советская Энциклопедия*. 3-е изд. Т. 1, с. 75. М., 1969.
- (32) Австрические языки // *Большая Советская Энциклопедия*. 3-е изд. Т. 1, с. 93. М., 1969.
- (33) Агавские языки // *Большая Советская Энциклопедия*. 3-е изд. Т. 1, с. 173. М., 1969.

## 1970

- (34) A Long-Range Comparison of Some Languages of Northern Eurasia // *VII Международный конгресс антропологических и этнографических наук, т. 5 / VII Intern. Congress of Anthropological and Ethnographical Sciences, Vol. 5*. Москва / Moscow 1970: 620—634.
- (b) Заметки по юкагирскому языку. Москва, 1969—70. Ркп.
- (c) Полевые записи верхнеколымского диалекта юкагирского языка. (1970-е). Ркп.
- (35) Атапасские языки // *Большая Советская Энциклопедия*. 3-е изд. Т. 2, с. 368. М., 1970.
- (36) Афар-Сахо языки // *Большая Советская Энциклопедия*. 3-е изд. Т. 2, с. 420. М., 1970.
- (37) Ацтекский язык // *Большая Советская Энциклопедия*. 3-е изд. Т. 2, с. 470. М., 1970.
- (38) Багирми [язык народа багирми] // *Большая Советская Энциклопедия*. 3-е изд. Т. 2, с. 512. М., 1970.
- (39) Бедауйе язык // *Большая Советская Энциклопедия*. 3-е изд. Т. 3, с. 76. М., 1970.

## 1971

- (40) Ностратические этимологии и происхождение глагольных формантов // *Этимология* 1968. М., 1971. С. 237—242.
- (41) Вакашские языки // *Большая Советская Энциклопедия*. 3-е изд. Т. 4, с. 239. М., 1971.
- (42) Галла язык // *Большая Советская Энциклопедия*. 3-е изд. Т. 6, с. 67. М., 1971.

## 1972

- (43) Гуайкуру [группа языков] // *Большая Советская Энциклопедия*. 3-е изд. Т. 7, с. 419—420. М., 1972.
- (44) Материалы по сравнительно-исторической фонетике кушитских языков: велярный звонкий в анлауте // *Проблемы африканского языкознания: Типология, компаративистика, описание языков*. М., 1972. С. 197—216.
- (45) Ностратические корни с сочетанием латерального и звонкого ларингала // *Этимология* 1970. М., 1972. С. 356—69.
- (46) Опыт реконструкции общеностратической грамматической системы. А. О системе служебных и местоименных слов. Б. О ностратическом синтаксисе: реконструкция порядка слов // *Конференция по сравнительно-исторической грамматике индоевропейских языков (12—14 декабря)*. Предварительные материалы. М., Наука, Институт славяноведения и балканистики АН СССР, 1972. С. 32—34.
- (47) Реконструкция чукотско-корякских корней и ностратические этимологии. // *Конференция по сравнительно-исторической грамматике индоевропейских языков (12—14 декабря)*. Предварительные материалы. М., Наука, Институт славяноведения и балканистики АН СССР, 1972. С. 27—30. (совм. с: А. Н. Головастиков)
- (48) Центральночадские фонемы и этимология центрально-чадских корней в свете данных внешнего сравнения // *Конференция по сравнительно-исторической грамматике индоевропейских языков (12—14 декабря)*. Предварительные материалы. М., Наука, Институт славяноведения и балканистики АН СССР, 1972. С. 64—65. (совм. с: Н. С. Панова, В. Я. Порхомовский)
- (49) Долгопольский А. Б. О происхождении личных окончаний глаголов в восточносидамских и ираквских языках. стр. 103—112. // *Africana (Африканский этнографический сборник) IX*. (Труды Института этнографии им. Н.Н. Миклухо-Маклая, т. 100). Л., 1972. С. 103—112.
- (50) La permutation des \**m* et \**b* initiaux dans les langues couchitiques // *Congrès International des africanistes, 2<sup>ème</sup> session*. Paris: Présence africaine 1972, 225—234.
- (51) Какие языки родственны европейским? // *Наука и человечество: доступно и точно о главном в мировой науке: Международный ежегодник*. 1971—72. М., 1972. С. 106—119.
- (52) Еврейское письмо // *Большая Советская Энциклопедия*. 3-е изд. Т. 9, с. 13—14. М., 1972 (совм. с: Э. М. Фалькович).
- (53) Древнеюжноаравийский язык // *Большая Советская Энциклопедия*. 3-е изд. Т. 8, с. 490. М., 1972.
- (54) Знаки препинания // *Большая Советская Энциклопедия*. 3-е изд. Т. 9, с. 550—551. М., 1972 (совм. с: А. А. Леонтьев).
- (55) Ибибио [язык] // *Большая Советская Энциклопедия*. 3-е изд. Т. 9, с. 616. М., 1972.
- (56) Иврит // *Большая Советская Энциклопедия*. 3-е изд. Т. 10, с. 23. М., 1972.
- (57) Индейские языки // *Большая Советская Энциклопедия*. 3-е изд. Т. 10, с. 182—183. М., 1972.
- (58) Ираквские языки // *Большая Советская Энциклопедия*. 3-е изд. Т. 10, с. 403. М., 1972.

## 1973

- (59) Сравнительно-историческая фонетика кушитских языков. М., Наука, 1973.
- (60) Материалы по лексике языка хадиа. Именные части речи // *Бесписьменные и младописьменные языки Африки*. М., Наука, 1973. С. 57—82.
- (61) Вклад В. М. Иллич-Свитыча в сравнительно-историческую грамматику индоевропейских и ностратических языков // *Советское славяноведение*. 1973, №5, с. 82 — 91. (совм. с: В. А. Дыбо, А. А. Зализняк)
- (62) Boreisch — Ursprache Eurasiens // *Bild der Wissenschaft* 1973, 1140—1146.
- (63) О так называемых границах применения методов сравнительно-исторического языкознания и о лингвистических свидетельствах родства языков. // *Генетические и ареальные связи языков Азии и Африки*. Тезисы докладов. М., Институт востоковедения, 1973. С. 109—112.
- (64) Карибские языки // *Большая Советская Энциклопедия*. 3-е изд. Т. 11, с. 425. М., 1973.

- (65) Китайско-тибетские языки // Большая Советская Энциклопедия. 3-е изд. Т. 12, с. 245. М., 1973.  
 (66) Классификация языков // Большая Советская Энциклопедия. 3-е изд. Т. 12, с. 270. М., 1973 (совм. с: М. А. Журиной).  
 (67) Конго-кордофанские языки // Большая Советская Энциклопедия. 3-е изд. Т. 12, с. 622. М., 1973.  
 (68) Кордофанские языки // Большая Советская Энциклопедия. 3-е изд. Т. 13, с. 150. М., 1973.  
 (69) Креольские языки // Большая Советская Энциклопедия. 3-е изд. Т. 13, с. 386. М., 1973.  
 (70) Кушитские языки // Большая Советская Энциклопедия. 3-е изд. Т. 14, с. 64. М., 1973.  
 (71) Ливийский язык // Большая Советская Энциклопедия. 3-е изд. Т. 14, с. 419. М., 1973.

## 1974

- (72) О ностратической системе аффрикат и сибилантов: корни с фонемой \*ʒ // Этимология 1972. М., 1974. С. 163—176.  
 (73) Майнхоф, Мейнхоф (Meinhof) Карл // Большая Советская Энциклопедия. 3-е изд. Т. 15, с. 221. М., 1974 (совм. с: Р. А. Агеева)  
 (74) Майя письмо // Большая Советская Энциклопедия. 3-е изд. Т. 15, с. 226. М., 1974  
 (75) Майя-соке языки // Большая Советская Энциклопедия. 3-е изд. Т. 15, с. 226—227. М., 1974  
 (76) Мальтийский язык // Большая Советская Энциклопедия. 3-е изд. Т. 15, с. 303. М., 1974.  
 (77) Масаи язык // Большая Советская Энциклопедия. 3-е изд. Т. 15, с. 435. М., 1974  
 (78) Матако-Мака языки // Большая Советская Энциклопедия. 3-е изд. Т. 15, с. 464. М., 1974  
 (79) Матрес лекционис // Большая Советская Энциклопедия. 3-е изд. Т. 15, с. 507. М., 1974.  
 (80) Мериотский язык // Большая Советская Энциклопедия. 3-е изд. Т. 16, с. 94. М., 1974.  
 (81) Мискито-матагальпские языки // Большая Советская Энциклопедия. 3-е изд. Т. 16, с. 328. М., 1974  
 (82) На-дене языки // Большая Советская Энциклопедия. 3-е изд. Т. 17, с. 207—208. М., 1974.  
 (83) Нилотские языки // Большая Советская Энциклопедия. 3-е изд. Т. 18, с. 20. М., 1974  
 (84) Ностратические языки // Большая Советская Энциклопедия. 3-е изд. Т. 18, с. 139. М., 1974.  
 (85) Огамическое письмо // Большая Советская Энциклопедия. 3-е изд. Т. 18, с. 283. М., 1974

## 1975

- (86) Языки и проблемы прародины // Знание — Сила 1975/6, 15—19.  
 (87) Contributions to the Afroasiatic Comparative Word List // *Proceedings of the Sixth Conference on African Linguistics* (Ohio State University, Columbus, April 1975). Columbus (Ohio): Ohio State University 1975, 42—43.  
 (88) Палеонтология лингвистическая // Большая Советская Энциклопедия. 3-е изд. Т. 19, с. 113. М., 1975  
 (89) Пиктский язык // Большая Советская Энциклопедия. 3-е изд. Т. 19, с. 530. М., 1975  
 (90) Языкознание (в статье «Румыния») // Большая Советская Энциклопедия. 3-е изд. Т. 22, с. 382. М., 1975  
 (91) Саамский язык // Большая Советская Энциклопедия. 3-е изд. Т. 22, с. 472. М., 1975  
 (92) Сапаро языки // Большая Советская Энциклопедия. 3-е изд. Т. 22, с. 582. М., 1975.

## 1976

- (93) Семья языков // Большая Советская Энциклопедия. 3-е изд. Т. 23, с. 245. М., 1976.  
 (94) Нумидийское (восточноливийское) письмо Северной Африки // Тайны древних письмен: проблемы дешифровки. Под ред. И. М. Дьяконова. М., Прогресс, 1976. С. 384—401.  
 (95) Комментарий к переводу статьи Ж. Марси «Ливийские двуязычные надписи Северной Африки» — там же, 413—414  
 (96) Комментарий к переводу статьи Ж. Феврие «Что мы знаем о ливийском языке» — там же, 421  
 (97) Комментарий к переводу статьи О. Рёсслера «Язык Нумидии» — там же, 442—443  
 (98) Сомали язык // Большая Советская Энциклопедия. 3-е изд. Т. 24, кн. 1, с. 174. М., 1976  
 (99) Сравнительно-историческое языкознание // Большая Советская Энциклопедия. 3-е изд. Т. 24, кн. 1, с. 357—358. М., 1976  
 (100) Тасманийские языки // Большая Советская Энциклопедия. 3-е изд. Т. 25, с. 285. М., 1976.

## 1977

- (101) Emphatic consonants in Semitic // *Israel Oriental Studies VII* (1977), 1—13.

1978

- (102) On phonemic stress in Proto-Semitic // *Israel Oriental Studies VIII* (1978), 1–12.

1982

- (103) Chadic – Semitic – Cushitic: Epenthetic  $-\gamma-$  in Sura in the light of Hamito-Semitic comparative linguistics // *The Chadic Languages in the Hamitosemitic-Nigrific Border Area*, ed. by Herrmann Jungraithmayr. Berlin: Reimer 1982, 32–46.

1983

- (104) Semitic and East Cushitic: sound correspondences and cognate sets // *Ethiopian Studies dedicated to Wolf Leslau*, ed. by St. Segert & A. Bodrogligetti. Wiesbaden: Harrassowitz 1983, 123–42.

1984

- (105) On Personal Pronouns in the Nostratic Languages // *Linguistica et Philologica. Gedenkschrift für Björn Collinder (1894–1983)*, ed. O. Gschwantler et al. Wien: Braumüller 1984, 65–112.

1986

- (106) Semitic nomina segolata in Ethiopic // *Ethiopian Studies: Proceedings of the Sixth International Conference of Ethiopian Studies*, ed. by Gideon Goldenberg. Rotterdam-Boston: Balkema 1986, 71–90.
- (107) A Probabilistic Hypothesis Concerning the Oldest Relationship Among the Language Families of Northern Eurasia // *Typology, Relationship and Time*, ed. by Tom L. Markey & V.V. Shevoroshkin. Ann Arbor: Karoma 1986, 27–50 [перевод статьи 1964 (5)].

1987

- (108) Cultural contacts of Proto-Indo-European and Proto-Indo-Iranian with neighbouring languages // *Folia Linguistica Historica VIII/1–2* (1987), 3–36.
- (109) South Cushitic lateral consonants as compared to Semitic and East Cushitic // *Proceedings of the Fourth International Hamito-Semitic Congress*, ed. by H. Jungraithmayr & W.W. Müller. Amsterdam-Philadelphia: Benjamins 1987, 195–214.

1988

- (110) On etymology of pronouns and classification of the Chadic languages // *FUCUS. A Semitic/Afrasian Gathering in Remembrance of Albert Ehrman*, ed. by Y. Arbeitman. Amsterdam-Philadelphia: Benjamins 1988, 201–20.
- (111) The Indo-European homeland and lexical contacts of Proto-Indo-European with other languages // *Mediterranean Language Review III* (1988), 7–27.
- (112) Semitic and East Cushitic: word-initial laryngeals // *Proceedings of the Eight International Conference of Ethiopian Studies (University of Addis Ababa 1984)*, ed. by Taddese Beyene, vol. I. Addis Ababa-Frankfurt am Main: Institute of Ethiopian Studies 1988, 629–637.

1989

- (113) Problems of Semitic comparative morphology: Mimation and nunation // *Societatis Linguisticae Europae Sodalitium Israëdense. Proceedings of the Sixth Annual Meeting (Tel Aviv 1988)*. Jerusalem 1989, III–VII.
- (114) Problems of Nostratic comparative phonology // *Reconstructing Languages and Cultures*, ed. by Vitaly Shevoroshkin. Bochum: Brockmeyer 1989, 90–98.
- (115) On lateral obstruents in Hamito-Semitic [abstracts] // *Reconstructing Languages and Cultures*, ed. by Vitaly Shevoroshkin. Bochum: Brockmeyer 1989, 99–103.
- (d) Происхождение алтайских восходящих дифтонгов в свете данных внешнего сравнения. Доклад на Международной конференции «Лингвистическая реконструкция и древнейшая история Востока (Москва, 1989). Ркп.
- (e) On the origin of the Altaic ascending diphthongs in the light of external comparison. Доклад на Международной конференции «Лингвистическая реконструкция и древнейшая история Востока (Москва, 1989). Ркп. [английская версия предыдущего номера].

- (f) Судьба ностратических гласных в индоевропейских языках. Тезисы сообщения на Международной конференции «Лингвистическая реконструкция и древнейшая история Востока». (Москва, 1989).
- (g) Imperfective of prefix-conjugated verbs in Cushitic, Semitic, and Berber: Origin and Development». Paper presented at the Conference of Cushitic and Omotic Studies (Turin 1989). Ркп.
- (h) On lateral obstruents in Hamito-Semitic. Haifa, 1989. Ркп.
- (i) Problems of Nostratic phonology. Haifa, 1989. Ркп.

## 1990

- (116) On Chadic correspondences of Sem. \*š // Proceedings of the Fifth International Hamito-Semitic Congress (Wien 1987), Vol. I, ed. by Hans Mukarovsky. Wien: Afro-Pub 1990, 213—226.
- (117) Language relationship and the history of mankind. Abstract // Evolution: From Moleculesto Culture, arranged by R. Dawkins & J. Diamond. Cold Spring Harbor (NY) 1990, 32.
- (118) [Барух Подольский] Практическая грамматика языка иврит. / Под ред. Арона Долгопольского. М., Интербук, 1990.
- (119) Современный иврит: Самоучитель: Начальный курс. / Сост. Ш. Блюм, Х. Рабин; подготовка и адаптация А. Долгопольского. М., Советский писатель. 1990.
- (j) О разграничении эпиглоттальных согласных и увулярных щелевых на ностратическом уровне. Хайфа, 1990. Ркп.

## 1991

- (120) Two problems of Semitic historical linguistics // Semitic Studies: in honor of Wolf Leslau on the occasion of his 85<sup>th</sup> birthday, ed. by Allan S. Kaye. Wiesbaden: Harrassowitz 1991, 328—339.
- (121) Кушитские языки // Языки Азии и Африки IV: Афразийские языки. Кн. 2. М., Главная редакция Восточной литературы, 1991. С. 5—147. (раздел книги написан в 1976 году)
- (k) Chadic interrogative pronouns in etymological context // Handout of a paper. Réunion du Groupe d'études tchadiques (Paris 1991). Ркп.

## 1992

- (122) A Hypothesis concerning Ancient Relationships Between Languages in Northern Eurasia // Nostratic, Dene-Caucasian, Austric and Amerind, ed. by Vitaly Shevoroshkin. Bochum: Brockmeyer 1992, 267—289 [перевод статьи 1964 (6)].
- (123) Nostratic etymologies and the origin of verbal formatives // Nostratic, Dene-Caucasian, Austric and Amerind, ed. by Vitaly Shevoroshkin. Bochum: Brockmeyer 1992, 290—297 [перевод статьи 1971 (28)].
- (124) The Nostratic vowels in Indo-European // Nostratic, Dene-Caucasian, Austric and Amerind, ed. by Vitaly Shevoroshkin. Bochum: Brockmeyer 1992, 298—331. [статья написана в 1976 г.].
- (l) Language relationship and the history of mankind. Paper presented at ZIF Conference on Biological and Cultural Aspects of Language Development, Jan. 20—22, 1992, Zentrum für interdisziplinäre Forschung, Univ. of Bielefeld, Germany). Ркп.

## 1993

- (125) More about the Indo-European homeland problem // Mediterranean Language Review VI-VII (1993), 251—72

## 1994

- (126) The Aramaic reflex of the Semitic glottalized lateral consonant // Rocznik Orientalistyczny XLIX (1994), 5—14.
- (127) Nostratic // Encyclopedia of Language and Linguistics, V, ed. by R.E. Asher: University Press 1994, 2838.
- (128) Hamito-Semitic names of body parts // Semitic and Cushitic Studies, ed. by Gideon Goldenberg & Shlomo Raz. Wiesbaden: Harrassowitz 1994, 267—87.
- (129) Семитские языки // Краткая еврейская энциклопедия VII. Иерусалим 1994, 744—747.
- (m) External relations of Afro-Asiatic // World Archaeological Congress: Theme 3: Language, Anthropology, and Archaeology (New Delhi, Dec 1994). Ркп.



- (n) Hamito-Semitic Etymologies. Paper presented at the 6th International Hamito-Semitic Congress, Moscow April 1994. Ркп.
- (o) Origin of Some Semitic Names of Body Parts. Haifa, 1994. Ркп.

1995

- (130) Судьба ностратических гласных в индоевропейском языке // Московский Лингвистический Журнал I (1995), 14—33.
- (131) Linguistic prehistory // Cambridge Archaeological Journal V/2 (1995), 268—71. (совм. с: Colin Renfrew, Theodora Bynon, Merritt Ruhlen, Peter Bellwood)

1996

- (p) Origin of Gender in Hamito-Semitic. Paper presented at the 8° Incontro di Linguistica Afro-Asiatica, Napoli, January 1996. Ркп.

1997

- (132) The Indo-European stops in the light of the long range relationship of Indo-European with Afroasiatic and some language families of Northern and Eastern Asia // IV Международная конференция по языкам Дальнего Востока, Юго-Восточной Азии и Западной Африки. Тезисы докладов. II. М., 1997. С. 109—112.
- (133) Postscript (to the paper by Heltzer, M.: *On the Origin of the Near Eastern Archaeological Amber*) // *Michmanim* 11 (1997), 37—38.

1998

- (134) *The Nostratic Macrofamily and Linguistic Palaeontology*. Cambridge: McDonald Institute for Archaeological Research 1998.
- (q) The Nostratic macrofamily. A short introduction // Symposium on the Nostratic Macrofamily. Cambridge [Precirculated papers] 1998.

1999

- (135) From Proto-Semitic to Hebrew. Phonology. Etymological approach in a Hamito-Semitic perspective. Milano: Centro Studi Camito-Semitici, 1999.
- (136) On the etymology of Hebrew // Historical, Epigraphical, and Biblical Studies in Honor of Prof. Michael Heltzer, ed. by Yitzak Avishur & Robert Deutsch. Jaffa 1999, 157—60.
- (137) On the origin of the Hebrew *nota accusativi* 'eṭ ~ 'eṭ and the *t*-accusative in Akkadian, Agaw, and Saho // Afroasiatica Tergestina. Papers from the 9th Italian meeting of Afro-Asiatic (Hamito-Semitic) Linguistics (Trieste, April 1998), ed. by Marcello Lamberti & Livia Tonelli. Padova: Unipress 1999, 43—46.
- (138) The Nostratic macrofamily: a short introduction // Nostratic: Examining a Linguistic Macrofamily, ed. by Colin Renfrew and D. Nettle. Cambridge: McDonald Institute for Archaeological Research 1999, 19—44.

2000

- (139) Sources of linguistic chronology // Time Depth in Historical Linguistics, ed. by Colin Renfrew, A. McJahon & Larry Trask. McDonald Institute for Archaeological Research 2000, 401—409.
- (r) Lexical convergence and long-range comparison of languages. Хэндаут к докладу на международной конференции «Проблемы изучения дальнего родства языков на рубеже третьего тысячелетия» (Москва, 29 мая — 2 июня 2000 г.). Ркп.

2001

- (140) Mah nēcappeh mi — millôn 'étimôlôgî shel ha-ṣáfáh ha-'ivrît? [What do we expect from a Hebrew etymological dictionary?] // *Ha-Civrît vë-'axyôtèha* I. Haifa 2001, 69—74.
- (s) Emphatic and plain voiceless consonants in Hamito-Semitic. Haifa, 2001 (Abstracts and handout for the 10<sup>th</sup> Meeting of Hamito-Semitic [Afoasiatic] Linguistics, Florence, April 2001). Ркп.
- (t) Lexical and grammatical convergence in Semitic. Abstracts of a lecture (Tel-Aviv Univ., 2001). Haifa, 2001. Ркп.

## 2002

- (u) Berber roots and grammar in the light of long-range comparison. Handout for the 2nd Bayreuth-Frankfurt Colloquium on Berber Linguistics (Frankfurt, July 2002). Ркп.

## 2005

- (141) Emphatic and Plain Voiceless Consonants in Hamito-Semitic in the Light of Internal and External Comparative Evidence // Proceedings of the 10<sup>th</sup> Meeting of Hamito-Semitic (Afroasiatic) Linguistics (Florence, April 2001), ed. by P. Fronzaroli & P. Marassini. Firenze: Dipartimento di Linguistica (Quaderni di Semitistica 25) 2005, 29–34.
- (142) Nostratic Grammar: Synthetic or Analytic? // *Orientalia et Classica*. Труды Института Восточных культур и античности, вып. VI: АСПЕКТЫ КОМПАРАТИВИСТИКИ 1. М., РГГУ, 2005. С. 13–38.

## 2008

- (143) *Nostratic Dictionary*. Cambridge: <<http://www.dspace.cam.ac.uk/handle/1810/196512>>

## Рецензии

## 1964

- (144) Naert, Pierre (1958) *La situation linguistique del' Aïnou. 1: Aïnou et indoeuropéen*, I. Lund: Gleerup (Lunds universitets årsskrift, N.F. Avd. 1., Bd.53, Nr. 4) // *Этимология* 1963. М., 1964. С. 293–299.

## 1975

- (145) Joki, Aulis (1973) *Uralier und Indogermanen. Die älteren Berührungen zwischen den uralischen und indogermanischen Sprachen*. Helsinki: MSFOu 151. // *Общественные науки за рубежом. Реферативный журнал, Серия 6: Языкознание* 1 (1975): 99–110.
- (146) Hymes Dell, H. ed. (1974) *Pidginization and Creolization of Languages*. Cambridge: University Press. — *Language in Society* 4 (1975), 243–247.

## 1986

- (147) Bomhard, Allan R. (1984) *Toward Proto-Nostratic: A New Approach to the Comparatison of Proto-Indo-European and Proto-Afroasiatic*. Amsterdam-Philadelphia: Benjamins. // *Bulletin de la Société de linguistique de Paris* 8/2 (1986), 91–97.
- (148) Bynon, James (ed. (1984): *Current Progress in Afro-Asiatic Linguistics. Papers of the third International Hamito-Semitic Congress* // *Bulletin of the School of Oriental and African Society* 49/2 (1986), 426–429.

## 2002

- (149) Takács, Gábor (1999) *Etymological Dictionary of Egyptian, Vol. I: A Phonological Introduction*, by Leiden-Boston-Köln: Brill 1999 // *Israel Oriental Society* XX (2002), 527–535.

## О юбиляре

Languages and their speakers in Ancient Eurasia. Dedicated to Professor Aharon Dolgopolsky on his 70<sup>th</sup> Birthday / Ed. by V.Shevoroshkin and P. Sidwell. AHL Studies in the Science and History of Language 5. Canberra, 2002.





## Reconstructing Benue-Congo person marking II

This paper is the second and last part of a comparative analysis of person marking systems in Benue-Congo (BC) languages, started in (Babaev 2008, available online for reference). The first part of the paper containing sections 1–2 gave an overview of the linguistic studies on the issue to date and presented a tentative reconstruction of person marking in the Proto-Bantoid language. In the second part of the paper, this work is continued by collecting data from all the other branches of BC and making the first step towards a reconstruction of the Proto-BC system of person marking.

*Keywords:* Niger-Congo, Benue-Congo, personal pronouns, comparative research, reconstruction, person marking.

The comparative outlook of person marking systems in the language families lying to the west of the Bantoid-speaking area is a challenge. These language stocks (the East BC families of Cross River, Plateau, Kainji and Jukunoid, and the West BC including Edoid, Nupoid, Defoid, Idomoid, Igboid and a few genetically isolated languages of Nigeria) are still far from being sufficiently studied or even described, and the amount of linguistic data for many of them remains quite scarce. In comparison with the Bantu family which has enjoyed much attention from comparative linguists within the last decades, there are very few papers researching the other subfamilies of BC from a comparative standpoint. This is especially true for studies in morphology, including person marking.

The aim here is therefore to make the very first step towards the comparative analysis and reconstruction of person markers in BC. This preliminary step will include compiling the data on as many languages as possible, systematising it into a single picture and making tentative conclusions about the general principles of person marking in the proto-language. By now, this seems the most ambitious objective we can achieve, so that the future, more complex work in this domain could have a basic foundation to move forward.

An important note here is that we had to omit the 3rd person markers in the present paper. First, in BC, they are mostly a part of the noun class system and therefore do not usually function as parts of the pronominal paradigms. Moreover, adding the 3rd person would have made the paper too huge. However, we will attract the 3rd person data where necessary in our analysis.

### 3. Cross River

The Cross River (CR) family, consisting of some 68 languages<sup>1</sup>, covers an area in south-eastern Nigeria (Korop and Usakade are also spoken in the border districts of Cameroon), resembling a long arch stretching from SW to NE, from the Niger delta region to the upper

---

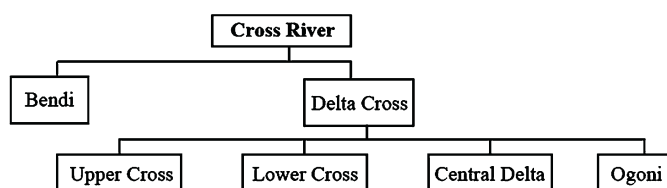
<sup>1</sup> Figures hereinafter according to Lewis 2009.

Cross river valley. Its southernmost tongues border the Ijoid languages, while central and northern dialects are squeezed between the Igboid area to the east and the Bantoid area to the west.

Williamson (1988: 104) gave her indication on the homeland and early migrations of the CR speakers from the BC home area, which she located at the confluence of the Niger and the Benue. According to her, the break-up of Bendi and Delta Cross subbranches of CR took place some 5,000 years ago, when the speakers of the latter branch moved to the south from the hills of Ogoja. On their way, CR speaking people settled along the banks of the river and around, and the migratory movement only stopped approximately 1,500 years ago in Central Delta, where the CR newcomers pressed the Ijo people.

Since Greenberg (1963) classified CR as a branch of BC, the internal structuring of the family has not been changed much, and the most recent family tree scheme (Williamson & Blench 2000: 33) looks generally as follows:

Diagram 1



Still, the family cannot be called fully homogeneous, and the discussions are still under way whether the Bendi branch belongs to CR or is a closer relative with Bantoid (Connell 1988). This is the main issue about the CR classification, since the languages of the second major branch, Delta Cross, seem to share a number of common items in both lexicon and grammar, though, again, very few lexical innovations and almost no phonetic innovations have been described even for Proto-Delta Cross.

The reconstructed consonant system of Proto-CR is presented following Dimmendaal (1978) and contains two basic sets of voiced / voiceless obstruents plus a special series of “fortis” obstruents, sometimes seen as geminates (Faraclas 1989: 386):

pp	tt	kk			j
p	t	k	k <sup>w</sup>	kp	(w)?
b	d		g <sup>w</sup>		
m	n	ŋ			(ɲ)?

Comparative works on CR include (Faraclas 1989) and (Connell 1988). In subbranch reconstructions, Dimmendaal (1978) and Sterk (1979) should be named for Upper Cross, Wolff (1969) and Alex (1989) for Central Delta, Williamson (1985) and Ikoro (1989) for Ogoni, Connell (1991, 1995) for Lower Cross. However, these works mainly deal with lexical and phonetic reconstruction, only briefly touching the morphology. The reconstruction of some elements of person marking are found in Bond & Anderson (2006) for Ogoni and Connell (ms.) for Lower Cross. Person markers for Proto-CR have not been subject to specific comparative research so far.

In most CR languages, there are four distinct series of markers denoting subject (predicative), object (predicative), possessive (nominal) and emphatic, also called “stressed” (independent pronouns). The subject markers are either prefixes or clitics — in the latter case, usually identical with the independent pronouns. Moreover, verb prefixes are often emphasised

by independent pronouns in a normal verb phrase. Object markers can be both independent and suffixed, and possessive markers are predominantly suffixed either to the noun or the class determiner in concord with the noun.

### 1st person singular (Table 3.1)

Some forms in 3.1 and subsequent charts have variable tone, depending on the tense or aspect of the predicate. For lack of space, we will unfortunately have to omit tonal distinctions in most cases if the segmental form does not change. Classificatory subgroup names are given in brackets following the name of the language.

The subject prefixed marker can be reconstructed as *\*N-*, a homorganic syllabic nasal. It might be suggested that it appeared due to the reduction of an earlier *\*mV-*: we see the same process in the system of noun class markers, where the Proto-BC pluralizer *\*ma-* > *\*N-* all over CR. When followed by a vowel, the initial consonant of the prefix is always *m-*. The vowel in unmarked forms is most often *-i-*. The *\*ma-* form (a long vowel is possible) is a contraction with a perfective marker. The subject prefix is lengthened in the negative in Delta Cross languages (cf. Iwara 2002: 106–107).

The independent pronoun in the majority of languages is derived from *\*àmì*, which confirms the reconstruction by Connell (ms.) for Lower Cross. A similar form is used as a possessive suffix, often lacking the final vowel: *\*-am*. In Kohumono *à-n-ùm* ‘I, me’, a class determiner *-n-* is infix between the association particle *a-* and the person marker. In Delta Cross, we see *\*mi-nV* which is also found in some Northwest Bantu independent and object series (Basaa A *mɛn*, Viya B *mɛnì*, etc.) as well as in other Bantoid languages (Tikar *mùn*, Yemba *mèŋ*) (Babaev 2008: 153–154).

The object marker for Proto-CR could be *\*mi* or *\*-m(i)*. Bond & Anderson (2006) reconstruct Proto-Ogoni *\*mI*, with a variable front vowel.

Therefore, the tentative reconstructed forms are:

<i>*N-</i> , <i>*mi</i>	subject
<i>*ma-</i>	subject (perfective)
<i>*m(i)</i>	object
<i>*àmì</i>	independent, possessive
<i>*minV</i>	independent (Proto-Delta Cross)

### 2nd person singular (Table 3.2)

The subject marker was certainly a vocalic one, though we cannot be sure whether *\*o* or *\*a* was the original vowel. The negative marker *-a-* lengthens the prefix in Ibibio and Legbo, the same as in the 1sg. The remaining three sets — object, possessive and independent — are mostly based on the proto-language root *\*wo* / *\*wu* / *\*we*.

The object marker can be reconstructed as *\*-o* / *\*wo* / *\*we*, the independent (and most probably possessive) as *\*òwò* which is *\*òfò* in Proto-Lower Cross. In Kana, one of the two possessive pronouns is cognate to the emphatic pronoun, the other one (denoting body parts only) is the normal object pronoun (Ikoro 1996), and it seems that Tee shows the same situation. That means the absence of a specific possessive set in Proto-Ogoni.

There are signs of the Proto-CR suffix *\*-n(V)* attached to non-subject markers (Bekwarra *íwò-n* ‘your’, Kukele *wě-ni* ‘you’, Ibibio *fiè-n* ‘you’). It is also in use in other person / number forms, and may serve as an indirect case marker.

Table 3.1

	subject	object	possessive	independent / emphatic
Bekwarra (Bendi) <sup>2</sup>	<i>N-</i> , <i>maa-</i> (perf.)	<i>m̄</i> (after CV verb), <i>mī</i>	<i>íyī</i>	<i>àmì</i>
Boki (Bendi) <sup>3</sup>	<i>me</i>	<i>mε</i>		
Kohumono (Delta, U) <sup>4</sup>	<i>N-</i>	<i>ànùṃ</i>	<i>-úm / -úṃ</i>	<i>ànùṃ</i>
Kukele (Delta, U) <sup>5</sup>	<i>mě</i>	<i>mě</i>	<i>-amě</i>	<i>měni</i>
Lokaa (Delta, U) <sup>6</sup>	<i>n-</i> , <i>nn-</i> (neg.)	<i>Ñ</i> , <i>mñ</i>	<i>-mì</i>	<i>àmí</i>
Mbembe (Delta, U) <sup>7</sup>	<i>N-</i>	<i>-m</i>	<i>-aam</i>	<i>kaam</i>
Ibibio (Delta, L) <sup>8</sup>	<i>N-</i>	<i>m̄n</i> , <i>-éḡ-</i>	<i>ṅm̄ / ñm̄</i>	<i>àmì</i>
Efik (Delta, L) <sup>9</sup>	<i>N-</i>	<i>m̄</i>	<i>(ò)m̄</i>	<i>àmì</i>
Obolo (Delta, L) <sup>10</sup>	<i>mi-</i> , <i>mù</i>	<i>mùḡ</i>	<i>mà</i>	<i>èmi</i>
Legbo (Delta, L) <sup>11</sup>	<i>N-</i> , <i>NN-</i> (neg.)	<i>-ń</i>		
Proto-Lower Cross <sup>12</sup>				<i>*àmì</i>
Ogbia (Delta, C) <sup>13</sup>	<i>mi / mu</i>	<i>məami</i>	<i>-ami / -ámí</i>	<i>əmi</i>
Abua (Delta, C) <sup>14</sup>	<i>mí</i>	<i>í-mí</i>	<i>a-mí</i>	<i>mí, mína</i>
Eleme (Delta, Ogoni) <sup>15</sup>	<i>Ñ-</i> , <i>rĩ-</i> (neg.) <i>ma-</i> (perf.)	<i>-mi</i>	<i>-àmi, ná</i>	<i>àmi</i>
Kana (Delta, Ogoni) <sup>16</sup>	<i>ń-</i> / <i>ṅ-</i> , <i>ńm</i> (neg.), <i>māā</i> (progr.), <i>māū-</i> (perf.)	<i>mē, mī</i>	<i>nà, ń-dā</i> (emph.), <i>mē</i> (body parts)	<i>ṅṅ, ṅ-dā</i> (emph.)
Tee (Delta, Ogoni) <sup>17</sup>	<i>m</i>	<i>me</i>	<i>nà, ńdāa</i>	<i>daà</i>

<sup>2</sup> Stanford 1976 [1973]<sup>3</sup> Westermann & Bryan 1952: 120–121<sup>4</sup> Cook 1976 [1972]<sup>5</sup> Preston 1854: 12<sup>6</sup> Iwara 1990<sup>7</sup> Barnwell 1980, Oyama & Barnwell 1985<sup>8</sup> Cook 1976 [1968], Essien 1987<sup>9</sup> Westermann & Bryan 1952: 135–136, Cook 1976 [1968]<sup>10</sup> Faraclas 1984, Rowland Oke 2003<sup>11</sup> Hyman et al. 2002, Hyman & Udoh 2002<sup>12</sup> Connell [ms.]<sup>13</sup> Isukul [ms.]<sup>14</sup> Gardner 1976 [1966/67]<sup>15</sup> Bond 2003, 2006<sup>16</sup> Westermann & Bryan 1952: 137, Ikoru 1996: 114–132, Bond 2006, Williamson et al. 2007<sup>17</sup> Wilson Kpàkpàn Nwí-Bàrì 2002

Table 3.2

	subject	object	possessive	independent / emphatic
Bekwarra	<i>o-</i>	<i>wō</i>	<i>íwòṅ</i>	<i>àwō</i>
Boki	<i>wɔ'ɔ</i>	<i>yɔ</i>		
Kohumono	<i>a- / a-</i>	<i>àgwɔ̀</i>	<i>-ó / -ɔ̀</i>	<i>àgwɔ̀</i>
Kukele	<i>wě</i>	<i>wě, (a)gwě</i>	<i>yǒ</i>	<i>wěni</i>
Lokaa	<i>a-, aa- (neg.)</i>	<i>ɔ́, wɔ́</i>	<i>-wù</i>	<i>àwú</i>
Mbembe	<i>a-</i>	<i>-ɔ</i>	<i>-ɔ</i>	<i>kɔ</i>
Ibibio	<i>ɛ- (pos.), ù- (neg.)</i>	<i>fìèn, -ú-</i>	<i>ṅfò</i>	<i>àfò</i>
Efik	<i>à- / ò- / è-</i>	<i>fì</i>	<i>(ò)fò</i>	<i>àfò</i>
Obolo	<i>o-, ù-</i>	<i>ùy</i>	<i>ùɔ</i>	<i>òwù</i>
Legbo	<i>a-, aà- (neg.)</i>	<i>-ɔ̀</i>		
Proto-Lower Cross				<i>*òfò</i>
Ogbia	<i>(a)nwá</i>	<i>maɛnwá</i>	<i>-íyóm</i>	
Abua	<i>na</i>	<i>ɲí-na</i>	<i>a-na</i>	<i>na</i>
Elemé	<i>ɔ-, àǎ̀, rò- (neg.)</i>	<i>-rũ</i>	<i>-yo</i>	
Kana	<i>ò- / ó-</i>	<i>ā</i>	<i>ō, á-lō (emph.), ā (body parts)</i>	<i>òò / òò<sup>18</sup>, ā-lō</i>
Tee	<i>o</i>	<i>o, òò</i>	<i>o</i>	

Unclear forms include, among others, Abua *-na* and the Elemé object suffix *-rũ*. The Elemé negative prefix *rò-* has the same origins as its 1sg. counterpart, but we have no evidence of its relation to the object marker.

<i>*o- / *a-</i>	subject
<i>*wo / *wu / *we</i>	object
<i>*òwò</i>	independent, possessive
<i>*wo-n(V) / *we-n(V)</i>	object / possessive / emphatic

### 1st person plural (Table 3.3)

In Delta Cross, prefixed predicative person markers are gradually replaced by independent pronouns. Still, they are sometimes found in V-shape as both subject prefixes and object suffixes.

*befě* in Kukele is almost identical to Proto-Bantoid *\*(à)bèce* ‘we’ < *\*(à-)ba-ice* (Babaev 2008: 158). Another retention is *àdò* ‘we, us’ in Kohumono which should go back to *\*àtò* with an intervocal voicing, probably cognate to the Proto-Bantu subject prefix *\*tù-* ‘we’. *tu / ti* are common 1pl. pronouns in Kwa and Gur languages as well (cf. Бабаев 2010), but whether Bekwarra *itēn* ‘our’ is a related form is obscure.

Connell (ms.) reconstructs *\*àjit* ‘you (emph.)’ for Proto-Lower Cross. We would also suggest *\*àjin* here, which may be related to Abua *a-yira* ‘our (incl.)’. Namely, the *-ar / -rV* -like forms which are found in the Central Delta and Ogoni languages are probably cognate with

<sup>18</sup> *òò* occurs in nominal affirmative and focalised negative constructions; *òò* is used in non-focus negative phrases, both verbal and nominal Ikoro 1996: 118–119.

Table 3.3

	subject	object	possessive	independent/emphatic
Bekwarra	<i>àbèrē é-</i>	<i>tī</i>	<i>ítēn</i>	<i>àbèrē</i>
Boki	<i>bεvε</i>	<i>bε</i>		
Kohumono	<i>e- / ε-</i>	<i>àdò</i>		<i>àdò</i>
Kukele	<i>befē ba</i>	<i>befē</i>	<i>-afi</i>	
Lokaa		<i>mǎn</i>	<i>-mì, -mò</i>	<i>àmǎn</i>
Mbembe	<i>mo- / mɔ-</i>	<i>mína</i>	<i>-amina</i>	<i>mína</i>
Ibibio	<i>ì-</i>	<i>ṅṅìn, -ì-</i>	<i>ṅṅìn</i>	<i>ṅṅìn</i>
Efik	<i>ì-</i>	<i>ṅṅìn</i>	<i>ṅṅìn</i>	<i>ṅṅìn</i>
Obolo	<i>é-, mí- (past), ème</i>	<i>èmey</i>	<i>ème</i>	<i>èzi / èjì</i>
Legbo	<i>me-, mè (neg.)</i>	<i>-mán</i>		
Proto-L.Cross				<i>*àjìt</i>
Ogbia	<i>iyər / iyər</i>	<i>məiyər</i>	<i>-íyər</i>	
Abua	<i>yoor (excl.), yira (incl.)</i>	<i>i-yoor (excl.), i-yira (incl.)</i>	<i>a-yoor (excl.), a-yira (incl.)</i>	<i>yoor (excl.), yira (incl.)</i>
Eleme	<i>rε- / nε-</i>	<i>-í</i>		<i>èbai</i>
Kana	<i>ì- / í-</i>	<i>ī</i>	<i>ī, á-lī (emph.)</i>	<i>ì, ā-lī (emph.)</i>
Tee	<i>i</i>	<i>i, ì</i>	<i>plii</i>	<i>iri</i>

them as well. The intervocal consonant *j / y* is seen as *-z-* in Obolo and can be related to *-f-* in Kukele, an Upper Cross tongue. Phonologically, it would be reasonable to reconstruct *\*àci(n)* for Proto-CR, which would be close to the Bantoid possessive forms: cf. Proto-Manenguba *\*àcí* or Kenyang *-esé*. The final Lower Cross consonant acts in a number of Lower Cross dialects which we did not include in the 3.3: Ebughu *ijìn*, Enwang and Uda *àjìn* (Connell, ms.).

In Upper and Lower Delta languages person markers are *me / mo* (subject), *mV(n)* (object, possessive), *\*(a)mi(na)* (non-subject, independent). These forms have parallels in languages as far away as Gur and Mande, but are rarely found in BC.

So, the tentative reconstructions for Proto-CR are these:

- \*e- / \*i-* subject
- \*àci(n)* independent
- \*mV* subject / object / independent

## 2nd person plural

In Eleme, the subject marker is suffixed to the verbal predicate. To distinguish the exact meaning of the phrase, the independent pronoun of the respective person is used together with this prefix:

- òbàù dose dé-í ñja* (1)  
 2.PL must eat-2.PL food  
 ‘you must eat food’ (Bond 2003)

In Ibibio, markers of the 2nd and 3rd persons have been equalised:

Table 3.4

	singular	plural
1st	<i>ɲ-</i>	<i>ì-</i>
2nd	<i>è-</i>	<i>è-</i>
3rd	<i>é-</i>	<i>é-</i>

Table 3.5

	subject	object	possessive	independent / emphatic
Bekwarra	<i>āmīn</i>	<i>āmīn</i>	<i>ínèn</i>	<i>āmīn</i>
Boki	<i>bēva</i>	<i>bē</i>		
Kohumono	<i>a- / ə-</i>	<i>ànò</i>		<i>ànò</i>
Kukele	<i>be ba</i>	<i>bewě, begwě</i>	<i>-ani</i>	<i>begwe</i>
Lokaa		<i>bá:</i>	<i>-wù</i>	<i>ábá:</i>
Mbembe	<i>ma-</i>	<i>bɔŋa</i>	<i>-abɔŋa</i>	<i>bɔŋa</i>
Ibibio	<i>è- (pos.), í- (neg.)</i>	<i>ndùfò, -í- / -é-</i>	<i>ndùfò / mbùfò</i>	<i>ndùfò / mbùfò</i>
Efik	<i>è- (pos.), í- (neg.)</i>	<i>mbùfò</i>	<i>mbùfò</i>	<i>mbùfò</i>
Obolo	<i>é-, mí- (past)</i>	<i>éy</i>	<i>ébe</i>	<i>èɲì</i>
Legbo	<i>ba-, b̀̀ (neg.)</i>	<i>-bɔ</i>		
Ogbia	<i>ɲín</i>	<i>máɲín</i>	<i>-íɲín</i>	
Abua	<i>ɲina</i>	<i>i-ɲina</i>	<i>a-ɲina</i>	<i>ɲina</i>
Eleme	<i>nɛ-, -ì, rò- (neg.)</i>	<i>-i</i>		<i>òbàù</i>
Kana	<i>bì-</i>	<i>ī</i>	<i>bùì, á-bààlò (emph.), ī (body parts)</i>	<i>bùì, ā-bààlò (emph.)</i>
Tee	<i>b̀̀</i>	<i>b̀̀̀̀</i>	<i>ɲlòo</i>	<i>borò</i>

We mostly see a vocalic prefix as subject marker, sometimes (but not everywhere) identical to the one we saw in the 1pl. Here, it is reflected as *e-* or *i-*, but we also find *a-* in Kohumono. These vocalic markers are never found in the possessive series or as independent pronouns.

The non-subject markers with the *-VnI-* shape include Kohumono *ànò* (object / emphatic), Bekwarra *ínèn* and Kukele *-ani* (possessive), Obolo *èɲì* (emphatic), and Ogbia / Abua *-íɲín / ɲina*. Similar forms are found in Bendi, Upper Cross, Lower Cross and Central Delta languages, encompassing most of CR. Eleme (the Ogoni group) has a subject prefix *nɛ-* which may complete the picture. We can only suggest *\*anV / \*inI(n)* with a non-subject meaning based on the comparison with the other person / number forms of the paradigm.

The nasality in Boki *bē* ‘us’ suggests a cognate with *\*be-ne* we reconstructed for Proto-Bantoid (Babaev 2008: 159), with the same root and a prefixed plural class marker *\*ba- / \*be-*. The independent pronoun *bēva* ‘we’ in Boki may suggest *\*be-nue* as well. In Kana, as Ikoro (1996: 115) suggested, emphatic pronoun *ā-bààlò* ‘you pl.’ and *á-bààlò* ‘your pl.’ seem to be derived from the 3pl. *bà* ‘they’. The widespread typological feature of replacing the 2pl. pronoun by the 3pl. one may be suggested for Kukele, Tee and Legbo.

The Lower Cross *\*mbùfò < \*ò-bù-fò* form is a probable pluralisation of the 2sg. *\*òfò* with the same prefix marking the independent pronoun. A similar model can be seen in Kukele *bewě* ‘you pl.’ < *\*be-i-we*, Legbo *b̀̀* vs. 2sg. *-ɔ́*, Tee *b̀̀̀̀* vs. 2sg. *ò̀̀̀*, and Kana *á-bààlò* vs. 2sg. *ā-lò*. Ko-



humono has merged the 2sg. and 2pl. pronouns. The process of eliminating the suppletion between the two 2nd person pronouns must have started no later than in Proto-Delta Cross, and the proto-form might have been *\*be-wo* / *\*ba-wo* / *\*be-we* for object and independent forms. The Eleme negative marker *rò-* lost a distinction in number and marks both the 2nd person singular and plural.

The 2pl. forms for Proto-CR might be given as follows:

<i>*e-</i> / <i>*i-</i>	subject
<i>*ànV(n)</i> / <i>*inI(n)</i>	independent
<i>*ba-wo</i> / <i>*be-wo</i>	subject / non-subject (Proto-Delta Cross)

The summarising table demonstrates the resulting tentative forms for Proto-CR:

Table 3.6

	prefixed (subject)	object	independent (non-subject)
1sg.	<i>*N-</i> , <i>*mi</i> , <i>*ma</i> (perf.)	<i>*m(i)</i>	<i>*àmì</i> , <i>*minV</i>
2sg.	<i>*o-</i> / <i>*a-</i>	<i>*wo</i> / <i>*wu</i> / <i>*we</i> , <i>*won(V)</i> / <i>*wen(V)</i>	<i>*òwò</i> , <i>*wonV</i> / <i>*wenV</i>
1pl.	<i>*e-</i> / <i>*i-</i> , <i>*mV</i>	<i>*mV</i>	<i>*àci(n)</i> , <i>*mV</i>
2pl.	<i>*e-</i> / <i>*i-</i> , <i>*ba-wo</i> / <i>*be-wo</i>		<i>*ànVn</i> / <i>*inI(n)</i> , <i>*ba-wo</i> / <i>*be-wo</i>

The comparison between the selected Proto-CR forms and the Proto-Bantoid person markers (Babaev 2008: 160) is shown below.

Table 3.7

	Proto-Bantoid	Proto-CR
1sg. subject	<i>*ni-</i>	<i>*N-</i>
1sg. independent	<i>*(à)me</i>	<i>*àmì</i>
2sg. subject	<i>*v-</i>	<i>*o-</i>
2sg. independent	<i>*(à)we</i>	<i>*òwò</i>
1pl. non-subject	<i>*(bè-)c(u)e</i>	<i>*àci(n)</i>
2pl. non-subject	<i>*(bè-)n(u)e</i>	<i>*ànVn</i> / <i>*inI(n)</i>

The similarities between the two columns seem evident. The independent pronouns in Proto-Bantoid and Proto-CR are virtually identical. Still, the variability of some vowels does matter and deserves a more detailed analysis following this brief survey. In the singular, *\*mI* / *\*amI* with a final front vowel is common for both families. The CR form *\*mi-nV* finds exact parallels in a number of Bantoid tongues including Bantu Western zones A-C, H and K (Babaev 2008: 162–166). The 2sg. prefix *\*o-* is also a clear cognate, and the independent CR pronoun *\*òwò* directly corresponds to widespread variants *owo* / *àwe* / *ewe* / *wo* / *we* in Bantu and Bantoid.

In the plural, person markers are more diverse. Both the Bantoid languages and some CR languages use pronouns with the prefixed pluraliser *\*be-* or, alternatively, with a prefixed vowel of either front or back row: *\*acV* / *\*icV* for 1pl. and *\*anV* / *\*inV* for 2pl. The original meanings of this vowel prefix are still traceable in Bantu tongues where forms in *\*a-* are possessives, while *\*i-* marks the independent pronoun (Babaev 2008: 139).

While the Bantu prefixed subject markers differ in origin from the non-subject (originally independent) pronouns, this difference can only be seen in the 2sg. in CR. As well as in Ban-



toid (beyond Bantu), CR person markers are mostly syntactically independent. Proto-Bantu subject markers *\*ni-* ‘I’, *\*tù-* ‘we’, and *\*mv-* ‘you pl.’ (Babaev 2008: 148) only have a few cognates in CR, such as Kohumono *àdò* ‘we’.

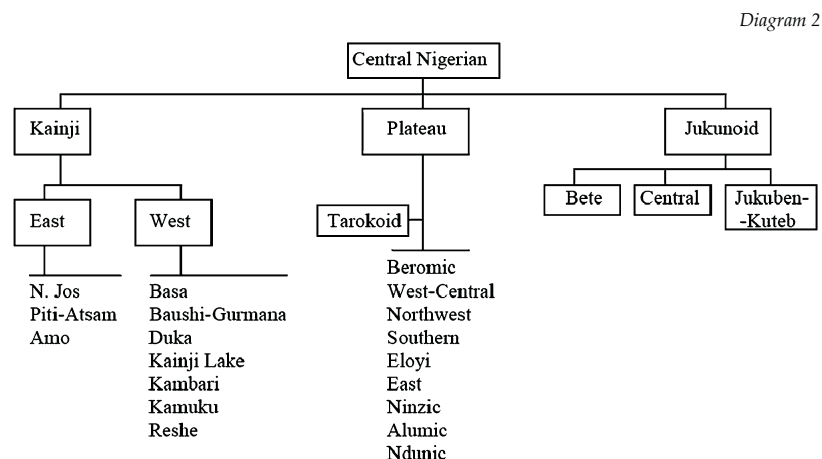
#### 4. Plateau, Kainji and Jukunoid (Central Nigerian)

Since Greenberg (1963) first identified the Plateau and Jukunoid languages stretching as a wide belt from W to E through the center of Nigeria, too many pages have been covered with discussions on both the external and the internal classification of this language stock. This discussion was fuelled by the clear fact that the languages which Greenberg defined with these two terms (and which we now know as Plateau, Kainji and Jukunoid) are indeed very close to each other, and definitely to the other tongues of East BC, namely CR and Bantoid. However, the lack of reliable language data and, more essentially, the lack of comparative research did not allow to make any distinct conclusions on the classification of the language groups of the Nigerian Plateau.

The Plateau and Kainji languages in North Nigeria border an area inhabited by Chadic-speaking people, which makes many minority languages endangered due to the cultural advance of Hausa. In the south, besides Chadic, Nupoid and Idomoid neighbours are located. The Jukunoid area borders Bantoid to the south.

At least six different classifications of these three language families were presented between 1963 and 2000 (Williamson 1971; Hoffmann 1976; Bennett & Sterk 1977; Gerhardt & Jockers 1981; Gerhardt 1989; Crozier & Blench 1992) resulting in the most widely referred internal grouping suggested by Williamson & Blench (2000). The new idea of the latter paper was to suggest the existence of a Central Nigerian (CN) proto-language, from which the languages of all the three groups originated. However, many issues of this subgrouping still remain unsettled. The idea of the Proto-Plateau language is challenged, since there are only a handful of lexical innovations that we can attribute to it. The Jukunoid languages indeed form a very tight group of around 20 languages (Lewis 2009), but whether they once formed a community with Tarokoid, a Plateau group, as suggested by Shimizu (1975), is a question mark. The internal structure of the Plateau family, the most diversified of all the three, must comprise at least ten groups of the same chronological level, which is not fully acceptable. One of the recent analyses of the issue by Roger Blench presented in 2005 bears a telling — and, for the moment, rhetoric — heading: “Is there a boundary between Plateau and Jukunoid”?

The most recent classification of CN, which is a compilation of (Blench 2004), looks as follows:



Blench is undoubtedly the most fruitful modern collector of Plateau language data, having gathered material from a great number of languages of the area and identified at least a dozen new ones. He has presented plenty of field notes with lexical data, which also include some morphological analysis. His online databases include initial steps in reconstructing Proto-Kainji and Proto-Plateau, which would have otherwise remained without any attempts at reconstruction. Beyond this, there has been no attempt to build a reliable system of proto-language phonology or morphology for either of the two. The most recent comparative survey of the Plateau and Kainji groups is contained in (Gerhardt 1989); however, it only briefly describes the morphology, mainly the nominal class systems, to which Gerhardt (1983) dedicated a special paper earlier. Williamson & Blench (2000: 32) give a short summary of morphological features characterising CN: “full or reduced noun class systems, mostly with prefixes; widespread verbal extensions; object pronouns following transitive verbs”, etc. The reference material for comparative lexical data is contained in “Benue-Congo Comparative Wordlists” (Williamson & Shimizu 1968; Williamson 1973).

For Jukunoid, a large study was conducted by Shimizu (1971–80), who has reconstructed the core of Proto-Jukunoid and has also dealt with personal pronouns. It clearly demonstrates that the Jukunoid morphology was more innovative than that of both Plateau and Kainji: this is seen in the noun class system, of which Jukunoid languages have preserved only -V class suffixes where Kainji and Platoid have CV-/V- prefixes corresponding to the Bantu class markers, which are considered more archaic (Gerhardt 1989: 372). However, the pronominal paradigms of some Jukunoid tongues were not yet recorded by that time, and the decades following Shimizu’s fundamental work filled some important gaps.

De Wolf (1968) suggested that the Proto-Kainji noun class system also showed signs of assimilating prefixed consonants. As we will see below, the situation in person marking is nearly the same.

In the present section, we will perform a brief comparison of person markers of all the three CN families, thus trying to draw a route towards the Proto-CN reconstruction, and further comparing the results with the rest of previously reconstructed East Bantoid systems. We must admit, however, that this analysis is still far from being complete.

Subject markers in CN languages precede the predicate. An object marker is usually suffixed to the verb, though in some Central Plateau languages it is placed between the tense/aspect element and the verb root, just like in Bantu. The difference between subject and object markers is often limited to suprasegmental features.

Possessive pronouns are mostly suffixed to the noun, but can also act independently following the noun (often called “possessive adjectives”), in which case they are sometimes grammaticalised compounds of a substantive stem with a possessive suffix, as in Berom or Tarok (Longtau 2008). Possessive pronouns can be linked with the nouns by an associative particle: Wapan *atan bú ù* ‘your house’, lit. ‘house of you’. And, similar to Bantoid, there is a general trend to bind the particle with the possessive: Jukun Takum *bú* ‘your’ < *bú-ú* ‘of-you’.

A phenomenon of “recapitulative”, “copy” or “repeating” pronouns is witnessed in a number of languages of the Plateau. This means that the subject pronoun placed before the predicate is also copied following it, being either suffixed or independent. This construction can have different meanings. Copy pronouns can follow intransitive verbs and mark negation, as in Migili (Lijili, S.Plateau) or Kente (Jukunoid). They are marking the plural in Izere (C.Plateau):

<i>jín</i>	<i>rus</i>	<i>jín</i>	<i>agabu</i>	(2)
2.PL	beat	2.PL	dog	
‘you are beating the dog’				

and the unexpected (“unanticipated”) action, including the negative action, in Wapan (Jukunoid):

*Be ci zhenzhen kù* (3)  
‘they are royalty’ (neutral)

vs.

*Be ci bé zhenzhen kù* (4)  
‘they are royalty’ (though you did not expect them to be).

In Kuteb (Jukunoid) it marks the completive aspect and the negative, and Koops (2007) calls this construction a “possessive verb”, since the postpositive pronoun has the same form in Kuteb as the possessive one. Postpredicative repeating pronouns are an areal feature found in a number of languages of Central Sudan, including Adamaua-Ubangian (Gbaya, Sangi) and Chadic. Whether it originates from NC or beyond, is not clear (Gerhardt 1989: 374).

The number of series of person markers varies in CN languages. In Jukun, at least seven sets of personal pronouns can be distinguished, varying in tone, vowel length and prefixed particles, as shown below (Storch, field notes):

Table 4.1

	subject	emphatic subject	object	possessive	emphatic possessive	recapitulative possessive	recapitulative
1sg.	<i>m</i>	<i>ám</i>	<i>m</i>	<i>ń / búm</i>	<i>ábúm</i>	<i>ń...ń</i>	<i>ń</i>
2sg.	<i>u</i>	<i>áu</i>	<i>u</i>	<i>bú</i>	<i>abu</i>	<i>bú...bú</i>	<i>ú</i>
3sg.	<i>kū</i>	<i>ákū</i>	<i>kū</i>	<i>bá</i>	<i>aku</i>	<i>bá...bá</i>	<i>á</i>
1pl.	<i>ī</i>	<i>ái / ábí</i>	<i>ī</i>	<i>bí</i>	<i>ábí</i>	<i>bí...bí</i>	<i>í</i>
2pl.	<i>ni</i>	<i>ánī</i>	<i>nī</i>	<i>bíní</i>	<i>ábíní</i>	<i>bíní...bíní</i>	<i>ní</i>
3pl.	<i>bi</i>	<i>ábī</i>	<i>bī</i>	<i>bíbí</i>	<i>ábíbí</i>	<i>bíbí...bíbí</i>	<i>bí</i>

Note that tense / aspect / modality markers can also vary in tone, so the perfective paradigm will be suprasegmentally different from that of the progressive, which further increases the number of series. In Tables 4.2–4.6, we combined all subsets of person markers of selected CN languages into four columns: subject, object, possessive and independent (emphatic), marking specific semantics where necessary.

### 1st person singular (Table 4.2)

The subject pronoun for Proto-Plateau and Proto-Kainji may be schematically reconstructed as *\*mI*, with a front vowel. There are a number of languages demonstrating *ma-* / *ma* or *mu*, but the proto-language front vowel may be supported by Tesu, Kwanka, Toro *myε* / *mya*: these probably added the aspect marker *-a-* to the earlier *\*mi-*. In the north, forms derived from *\*mi-n-* demonstrate an extension similar to what we saw in CR in Section 3. The Proto-Jukunoid form is *\*mi* / *\*me* (Shimizu 1971–80). Hone preserved the prefixed subject marker *n-*, comparable to CR and Bantoid *N-*.

The object pronoun for Proto-CN can be reconstructed as *\*me* / *\*mε*. Relic forms in *\*-n* / *\*nI* are also noted, marking the indirect object in Eggon, and the direct object in the present tense in Izere. In Fyam, *\*n-* is found as the independent pronoun. Independent pronouns seem quite uniform and allow to reconstruct *\*me* / *\*ame* identical to the forms in CR.

	subject	object	possessive	independent / emphatic
<b>Plateau</b>				
Tesu (Alumic) <sup>19</sup>	<i>myá</i>			
Ayu <sup>20</sup>	<i>mè</i>			
Berom (Beromic) <sup>21</sup>	<i>má-, mé</i>	<i>me, mé / mé</i>	<i>hɔŋ, -náŋ</i>	<i>mé</i>
Cara (C, NC) <sup>22</sup>	<i>mini</i>			
Izere (C, SC) <sup>23</sup>	<i>mì (pres.), tí (past)</i>	<i>ní (pres.), tìin (fut.)</i>	<i>-naan</i>	
Irigwe (C, SC) <sup>24</sup>	<i>ńjé</i>	<i>ńíŋ, N²-</i>	<i>ńjé</i>	<i>ńjé</i>
Fyam (SE) <sup>25</sup>	<i>náá, ín</i>	<i>-uŋ, -iŋ</i>	<i>-naŋ, náŋ (emph)</i>	<i>mé</i>
Horom (SE) <sup>26</sup>	<i>me</i>			
Hyam (Hyamic) <sup>27</sup>	<i>mi</i>			
Kulu (N) <sup>28</sup>	<i>amin, amii</i>	<i>mi</i>	<i>mii</i>	
Kanufi (Ninzic) <sup>29</sup>	<i>inbɔ</i>			
Bu (Ninzic) <sup>30</sup>				<i>amɛ</i>
Mada (Ninzic) <sup>31</sup>	<i>ŋgə</i>	<i>ŋgə</i>	<i>ŋgə</i>	
Eggon (W, SW) <sup>32</sup>	<i>me</i>	<i>me (dir.), nè (indir.)</i>	<i>mě</i>	
Che (W, SW) <sup>33</sup>		<i>mì</i>		
Kwanka (W, SW) <sup>34</sup>	<i>myɛ</i>			
Ake (W, SW) <sup>35</sup>				<i>àmù</i>
Idun (W, NW) <sup>36</sup>	<i>mi (fut., progr.), múm</i>			
Yeskwa (W, NW) <sup>37</sup>	<i>ɛma / ma, má (past), màá (fut.), máá (progr.), mma (cond.)</i>	<i>ma</i>	<i>-ma / ma</i>	<i>gɛma</i>

<sup>19</sup> Blench & Kato 2007<sup>20</sup> Blench 2006e<sup>21</sup> Bouquiaux 1970<sup>22</sup> Blench 2006h<sup>23</sup> Blench & Kaze 2006<sup>24</sup> Blench & Gya 2008<sup>25</sup> Nettle 1998<sup>26</sup> Blench 2006g<sup>27</sup> Blench 2006b<sup>28</sup> Shimizu 1996<sup>29</sup> Blench 2006d<sup>30</sup> Blench 2006f<sup>31</sup> Blench 2006c<sup>32</sup> Sibomana 1985: 54–61<sup>33</sup> Blench et al. 2006<sup>34</sup> Blench 2007b<sup>35</sup> Blench 2007a<sup>36</sup> Blench 2008a<sup>37</sup> Blench 2008b

	subject	object	possessive	independent / emphatic
Tarok (Tarokoid) <sup>38</sup>	<i>N, mi</i> (fut.)	<i>mi</i>	<i>mi, -jímí /-nyími</i>	<i>̀m̀m̀àmi</i>
Sur (Tarokoid) <sup>39</sup>	<i>mi</i>			
Toro <sup>40</sup>	<i>mya</i>			
Lijili (S) <sup>41</sup>	<i>n</i>		<i>me</i>	<i>̀m̀ε</i>
<b>Kainji</b>				
Hun-Saare (W, Duka) <sup>42</sup>	<i>εm, mε</i> (neg., nominal predicate, fut.)	<i>mε</i>	<i>re, de</i>	<i>mε</i>
C'Lela (W, Duka) <sup>43</sup>	<i>m̀à, m̀î</i>	<i>mε, m̀î</i>	<i>rí</i>	
Ror (W, Duka) <sup>44</sup>	<i>ām / mē</i>			
Laru (W, Kainji Lake) <sup>45</sup>	<i>ma</i>			
Basa (W, Basa) <sup>46</sup>	<i>n</i> (past), <i>ma</i> (pres.)	<i>ma</i>		
Pongu (W, Kamuku) <sup>47</sup>	<i>gēm</i>			
Tsvadi (W, Kambari) <sup>48</sup>				<i>omú</i>
Cicipu (W, Kambari) <sup>49</sup>	<i>m̀ù, m-</i>	<i>-n</i>	<i>vô</i>	<i>̀àm̀ù</i>
<b>Jukunoid</b>				
Jukun Takum (C, Jukun) <sup>50</sup>	<i>m, m</i>	<i>m</i>	<i>ń</i>	<i>̀àm</i>
Jibu (C, Jukun) <sup>51</sup>	<i>m, m̀ì</i>			
Hone (C, Jukun) <sup>52</sup>	<i>n-, ín-</i> (subj.), <i>m̀ù</i>	<i>-ì</i> (dir.), <i>-é</i> (dir.neg.), <i>ỳd̀m / ỳím̀ù</i> (indir.)	<i>-mii, -um</i>	<i>̀àm̀ù</i>
Wapan (C, Kororofa) <sup>53</sup>	<i>m̀, ǹ</i> (before <i>m, n, d</i> )	<i>m̀, m̀ì</i> (indir.)	<i>-m̀</i>	<i>̀àm̀ì</i>

<sup>38</sup> Sibomana 1981/82, Longtau & Blench, forthcoming, Longtau 2008

<sup>39</sup> Blench 2006i

<sup>40</sup> Blench 2006j

<sup>41</sup> Stofberg 1978: 331

<sup>42</sup> Cressmann & Skitch 1980 [1974]

<sup>43</sup> Rikoto et al. 2002

<sup>44</sup> Blench, manuscript a

<sup>45</sup> Blench, manuscript b

<sup>46</sup> Blench & al. 1991

<sup>47</sup> Blench, manuscript a

<sup>48</sup> Blench 2007c

<sup>49</sup> McGill [online]

<sup>50</sup> Welmers 1949, Koops 2007: 258

<sup>51</sup> Priest [ms.]

<sup>52</sup> Storch 2005

<sup>53</sup> Evenhouse [ms.]

	subject	object	possessive	independent / emphatic
Etkywan (C, Kpan-Icen) <sup>54</sup>				<i>āmē</i>
Bete <sup>55</sup>				<i>mba</i>
Yukuben (Yukuben-Kuteb) <sup>56</sup>		<i>m̃</i>		<i>āmà</i>
Kuteb (Yukuben-Kuteb) <sup>57</sup>	<i>m</i>	<i>m, me</i>	<i>-ʼm, anám</i>	<i>ame</i>
Lufu <sup>58</sup>	<i>m</i>			<i>mma</i>

Possessive suffixes in Plateau and Kainji often use the subject marker *\*mI*. The Proto-Plateau form *\*-naŋ* is derived from *\*-nam*, found also in Jukunoid (Kuteb) *-anám*, comparable with Kohumono (CR) *ànù̀m* ‘I, me’ (Section 3 above). The Kuteb construction, according to Koops (2007: 107), is a merger of the nominalizer *-a-*, the possessive particle *-na-* and the person marker. A similar construction underlies Berom *hɔŋ*. The Jukunoid languages have *\*-m* as the possessive suffix.

Proto-CN markers can be suggested as such:

<i>*mI</i>	subject
<i>*me / *mɛ</i>	object
<i>*me / *ame</i>	independent / emphatic

For Proto-Jukunoid we may reconstruct *\*m* (subject), *\*m(e)* (object), *\*-m* (possessive suffix) and *\*ame* (emphatic pronoun).

### 2nd person singular (Table 4.3)

Subject markers of the 2sg. may be preliminarily suggested as *\*o* and *\*wo* in Proto-Plateau. The non-labialised form is seen in Berom, Fyam and Idun, sometimes with a “personal theme prefix” such as Berom *h-* (Bouquiaux 1970). The Jukunoid data and some Kainji tongues (Hun-Saare) support the hypothesis that the subject form in Proto-CN was *\*o / \*u*, while the object / possessive marker was *\*wo / \*wu*, and the independent emphatic pronoun is seen as *\*awo / \*awu*. This distribution finds support in all the three families, but too much diffusion still blurs the picture. *-m-* in Yukuben is a non-personal affix for all the singular paradigm: *āmà* ‘I’, *àmū* ‘you’, *āmí* ‘he’.

Various nominal formants are used for creating the possessive markers: some of these are Berom *m-*, Tarok *-ji-* / *-nyi-*, Kuteb *aná-*, C’Lela *ro-*, etc.

Lijili (Plateau) and Cicipu (Kainji) emphatic pronouns *ìwḍ / ìwɔ* correspond to the independent pronouns of Bantoid (Babaev 2008: 144–145). In general, the paradigms of these two languages are strikingly similar to that of the Bantu 2sg. markers, which is noteworthy for their classification.

<sup>54</sup> Koops 2007: 258

<sup>55</sup> Koops 2007: 258

<sup>56</sup> Rennison, p.c.

<sup>57</sup> Koops 2007

<sup>58</sup> Koops 2007: 258

Table 4.3

	subject	object	possessive	independent / emphatic
<b>Plateau</b>				
Tesu	<i>nɔ̀n</i>			
Ayu	<i>ɲuk</i>			
Berom	<i>hó-, ò</i>	<i>hò</i>	<i>-mò, mo</i>	<i>hó, ó</i>
Cara	<i>wunmi</i>			
Izere	<i>wán</i> (pres.), <i>tá</i> (past)	<i>ká</i> (pres.), <i>tàa</i> (fut.)	<i>-fwan</i>	
Irigwe	<i>ɲwe</i>	<i>ré</i>	<i>ɲwê</i>	<i>ɲwé</i>
Fyam	<i>ti, wú</i>	<i>-o</i>	<i>-mé, náme</i> (emph.)	<i>wéé</i>
Horom	<i>awɔ</i>			<i>awɔ</i>
Hyam	<i>ɲu</i>			
Kulu	<i>aɲgu, un, ɲu</i>	<i>ɲ</i>	<i>ɲ</i>	
Kanufi	<i>wɔ</i>			
Bu				<i>awu</i>
Mada	<i>wə</i>	<i>wə</i>	<i>wə</i>	
Eggon	<i>ɲwo</i>	<i>ɲwo</i> (dir), <i>dɲwò</i> (indir)	<i>-ɲo, bɲo / bmo</i>	
Che	<i>ɲu</i>		<i>ì-mín</i>	
Kwanka	<i>wo</i>			
Ake				<i>àɲɔ̀</i>
Idun	<i>ɲwu</i> (fut., progress.), <i>ɲɔ</i>			
Yeskwa	<i>ɛmu, mú</i> (past), <i>mùú</i> (fut.), <i>múú</i> (progr.), <i>mmu</i> (cond.)	<i>mu</i>	<i>-mu / mu</i>	<i>gɛmu</i>
Tarok	<i>ɓú, u / ùpán</i> (indirect)	<i>ɓú</i>	<i>ɓú, -jibú / -nyibu</i>	<i>ɲmàɓú</i>
Sur	<i>ɓu</i>			
Toro	<i>anɔ</i>			
Lijili	<i>ɔ</i>		<i>wɔ</i>	<i>ìwò</i>
<b>Kainji</b>				
Hun-Saare	<i>o, wɔ</i>	<i>wɔ</i>	<i>ru, du</i>	<i>wɔ</i>
C'Lela	<i>vò</i>	<i>vò</i>	<i>ròvò</i>	
Ror	<i>bɔ̄</i>			
Pongu	<i>gà</i>			
Tsuvadi				<i>‘avó</i>
Cicipu	<i>vù, v-</i>		<i>ávù</i>	<i>ìvɔ</i>

	subject	object	possessive	independent / emphatic
<b>Jukunoid</b>				
Jukun Takum	<i>u</i>	<i>u</i>	<i>ú</i>	<i>áú</i>
Jibu	<i>wù</i>			
Hone	<i>ɔ-, vùù</i>	<i>-ù (dir.), -ó (dir.neg.), yúù (indir.)</i>	<i>-wuu</i>	<i>ávùù</i>
Wapan	<i>ù</i>	<i>ù</i>	<i>-ú</i>	<i>áú</i>
Etkywan				<i>ābō</i>
Bete				<i>owu</i>
Yukuben		<i>mū</i>		<i>āmū / òmū</i>
Kuteb	<i>u</i>	<i>fu</i>	<i>-'fu, anáfu</i>	<i>afu, abi (dial.)</i>
Lufu	<i>u</i>			<i>āu</i>

Proto-CN markers can be reconstructed as follows:

<i>*o / *u</i>	subject
<i>*wo / *wu</i>	object, possessive
<i>*awo / *awu</i>	emphatic

### 1st person plural (Table 4.4)

It is possible to suggest two variants of proto-language person markers for the 1pl.: *\*ti / \*ati*. *\*a-* was added to construct the emphatic pronoun. Proto-Jukunoid had *\*ti* for subject and object, *\*(a)ti* for possessive, and *\*atī* as a stressed pronoun. The dental obstruent has disappeared in much of Central Jukunoid regularly, cf. the Proto-BC plural class suffix *\*ti-*, becoming *\*-i* in Central Jukunoid (Shimizu 1971: 197). Based on the forms above, we can even see the direction of the reduction process: *\*t* (Kuteb) > *d* (Bete) > *j* (Etkywan) > *y* (Hone) >  $\emptyset$  (Jukun).

The vocalism of Kainji and Plateau forms varies greatly, and the amount of data is not sufficient for any representation, so the tentative conclusions on Proto-CN markers can only be made according to the Jukunoid forms and the other person / number forms of the paradigm.

The *\*-Vt* forms, with the back vowel, are suffixed to various nominalisers used to form the possessive markers, but in Plateau they are also widespread in the subject series. Tesu *mbɔɔɔ*, Berom *wot*, Cara *futte* and Kwanka *wòrà* may all descend from Proto-Plateau *\*βot(V)*, where *\*βo-* is a reflex of the NC plural noun class marker *\*bV-*. The possessive form in Proto-Plateau could be *\*mot*, found in Beromic, Southeast and Southwest subgroups. Kanufi and Che have subject forms from *\*tot*.

In Kainji, there is an evident lack of data that we could gather. However, the Cicipu subject marker *tù* correlates with Bantu *tu-/tu-* ‘we/us’ (Babaev 2008: 139), and C’Lela *cín* ‘we two’ directly corresponds to Proto-CR *\*àcl(n)* ‘we’.

Hun-Saare and Ror are among the few Kainji languages with a distinction between exclusive and inclusive pronouns. This feature is a clear innovation.

<i>*ti</i>	subject, object
<i>*at(i)</i>	emphatic, possessive



Table 4.4

	subject	object	possessive	independent / emphatic
<b>Plateau</b>				
Tesu	<i>mbɔɔ</i>			
Ayu	<i>bibîr</i>			
Berom	<i>wot &lt; *hwot<sup>59</sup></i>	<i>wot, hót</i>	<i>-mòt, mot</i>	<i>hót, wót</i>
Cara	<i>futte</i>			
Izere	<i>yír...yír (pres.), tí...yír (past)</i>	<i>ní...yír (pres), tìin...yír (fut)</i>	<i>-feyir</i>	
Irigwe	<i>nji</i>	<i>nij</i>	<i>nji</i>	<i>nji</i>
Fyam	<i>tí, tík</i>	<i>-té</i>	<i>-mót, ná- mun (emph.)</i>	<i>móti</i>
Horom	<i>taŋ</i>			
Hyam	<i>hera</i>			
Kulu	<i>andil</i>	<i>yaa</i>	<i>(ε)ndal</i>	
Kanufi	<i>tot</i>			
Bu				<i>kita</i>
Mada	<i>tə</i>	<i>tə</i>	<i>tə</i>	
Eggon	<i>gi</i>	<i>gi (dir.), dâgi (indir.)</i>	<i>bi</i>	
Che	<i>tút</i>		<i>ì-mot</i>	
Kwanka	<i>wòrà</i>			
Ake				<i>ani</i>
Idun	<i>mén</i>			
Yeskwa	<i>εmbi, mbí (past), mbú (fut.), mbú (progr.), mbi (cond.)</i>	<i>mbi</i>	<i>-mbi / mbi</i>	<i>gεmbi</i>
Tarok	<i>í, í / î (fut.)</i>	<i>yí</i>	<i>yí, -jiiyí / -nyiyí</i>	<i>m̀m̀àyí</i>
Sur	<i>yiyi</i>			<i>yiyi</i>
Toro	<i>anɔɔ</i>			
Lijili	<i>la</i>		<i>la</i>	<i>ila</i>
<b>Kainji</b>				
Hun-Saare	<i>a (incl.), tε (excl.)</i>	<i>na (incl.), tε (excl.)</i>	<i>na (incl.), tε (excl.)</i>	
C'Lela	<i>cín (dual), mà, nà, cwán</i>	<i>cwán</i>	<i>á'cò</i>	
Ror	<i>ín (incl.), ìt (excl.)</i>			
Laru	<i>ti</i>			
Pongu	<i>gáátù</i>			
Cicipu	<i>tù, ti-</i>		<i>ttù</i>	<i>òtù</i>

<sup>59</sup> Bouquiaux 1970: 165

	subject	object	possessive	independent / emphatic
<b>Jukunoid</b>				
JukunTakum	<i>i</i>	<i>i</i>	<i>í</i>	<i>ái</i>
Jibu	<i>yí</i>			
Hone	<i>i-, yì</i>	<i>-yì (dir.), yéyì (indir.)</i>	<i>-yì</i>	<i>áyì</i>
Wapan	<i>i</i>	<i>i</i>	<i>-i</i>	<i>ái</i>
Etkywan				<i>ājē</i>
Bete				<i>dēdē</i>
Yukuben		<i>zī</i>		<i>ādí / ēdí / āzī / ēzī</i>
Kuteb	<i>tī</i>	<i>tī</i>	<i>-tī, anátī</i>	<i>atī</i>
Lufu	<i>ī</i>			<i>ái</i>

## 2nd person plural (Table 4.6)

The Proto-Jukunoid forms can be suggested as *\*ni* (predicative subject / object), *\*-ni* (possessive), and *\*ánī* (emphatic). The Jibu pronoun *nìŋ* might reflect the original form with the final *\*-n* lost in other Jukunoid languages. This pronoun is seen in Plateau, where it might be reconstructed as *\*ni(n)*, corresponding to the Proto-CR *\*inI(n)*. Clear cognates within the Bantoid family include Vute (Mambiloid) *ním* and Manenguba (Bantu A15) *\*jî-* (subject) and *\*-àjî* (possessive) (Babaev 2008: 156).

The tentative Proto-Plateau form *\*bIn(I)* is witnessed in Ayu *bì-bìn*, Bu *biyi*, and probably some other languages as well. It is not present in either Kainji or Jukunoid. The possessives for Plateau languages are often derived from *\*-min*, with the possessive particle *\*m-*. Another form — *\*(a)nu* — should be suggested on the basis of forms in Plateau (Ake, Toro, Kulu, Toro) and West Kainji (Hun-Saare, C'Lela and Ror).

The only two person markers of 1pl. which we can tentatively reconstruct for Proto-CN are *\*ni(n)* and *\*(a)nu*.

The combined chart of Proto-CN pronouns which may be proposed according to these thoughts is given in 4.5:

Table 4.5

	subject	object	possessive	independent / emphatic
1sg.	<i>*mI</i>	<i>*me / *mε</i>	Jukunoid <i>*-m</i>	<i>*me / *ame</i>
2sg.	<i>*o / *u</i>	<i>*wo / *wu</i>	<i>*wo / *wu</i>	<i>*awo / *awu</i>
1pl.	<i>*ti</i>	<i>*ti</i>	<i>*-at(i)</i>	<i>*at(i)</i>
2pl.	<i>*ni(n), *(a)nu</i>	<i>*ni(n), *(a)nu</i>	<i>*ni(n), *(a)nu</i>	<i>*ni(n), *(a)nu</i>

The emphatic series is marked by the *\*a-* prefix, and it would be logical to support the *\*ajni(n)* form for the 2pl., though it is only seen in Jukunoid.

Table 4.6

	subject	object	possessive	independent / emphatic
<b>Plateau</b>				
Tesu	<i>mɔɔ</i>			
Ayu	<i>bìbìn</i>			
Berom	<i>yìn &lt; hyìn</i>	<i>hín, yín, én</i>	<i>-min, min</i>	<i>hín, yín, én</i>
Cara	<i>yimi</i>			
Izere	<i>ɲín...ɲín (pres.), tá...ɲín (past)</i>	<i>ká...ɲín (pres.), tâa...ɲín (past)</i>	<i>-feɲin</i>	
Irigwe	<i>ɲí</i>	<i>ɲíɲ</i>	<i>ɲì</i>	<i>ɲí</i>
Fyam	<i>tí ...-n, wúń</i>	<i>-múń</i>	<i>-múń, námot (emph.)</i>	<i>múńi</i>
Horom	<i>min</i>			
Hyam	<i>ɲì</i>			
Kulu	<i>anuɲ</i>	<i>ma</i>	<i>(ε)nuɲ</i>	<i>anuɲ</i>
Kanufi	<i>ayin</i>			
Bu				<i>biyi</i>
Mada	<i>gyə</i>	<i>gyə</i>	<i>gyə</i>	
Eggon	<i>gimí</i>	<i>gymí / bmí / mo (dir.), dâmo (indir.)</i>	<i>b`mí</i>	
Che	<i>yheɲi</i>		<i>bɔmin</i>	
Kwanka	<i>ɲinà</i>			
Ake				<i>ánù</i>
Idun	<i>mbî</i>			
Yeskwa	<i>εmi, mí (past), mîi (fut.), mîi (progr.), mi (cond.)</i>	<i>mi</i>	<i>-mi / mi</i>	<i>gemi</i>
Tarok	<i>ó / opən (indirect)</i>	<i>wó, opən (indirect)</i>	<i>wó, -jivó / -nyivó</i>	<i>ɲmàwó</i>
Sur	<i>ɲinín</i>			<i>ɲinín</i>
Toro	<i>anɔɔ</i>			
Lijili	<i>yi</i>		<i>yi</i>	<i>iyi</i>
<b>Kainji</b>				
Hun-Saare	<i>nɔ</i>	<i>nɔ</i>	<i>nɔ</i>	
C'Lela	<i>nwá</i>			
Ror	<i>nɔ̄</i>			
Laru	<i>darni</i>			
Pongu	<i>gáhĩ</i>			
Cicipu	<i>dô, i-</i>		<i>đđó</i>	<i>ídó</i>

	subject	object	possessive	independent / emphatic
<b>Jukunoid</b>				
JukunTakum	<i>ni</i>	<i>ni</i>	<i>ní</i>	<i>ánī</i>
Jibu	<i>nìŋ</i>			
Hone	<i>nán-</i> / <i>nán</i>	<i>-nán</i> (dir.), <i>yánàn</i> (indir.)	<i>-nan</i>	<i>ánán</i>
Wapan	<i>ni</i>	<i>ni</i>	<i>-ni</i>	<i>áni</i>
Etkywan				<i>áŋē</i>
Bete				<i>lēlē</i>
Yukuben		<i>ŋī</i>		<i>āŋí / ēŋí</i>
Kuteb	<i>nī</i>	<i>nī</i>	<i>- 'nī, anánī</i>	<i>anī</i>
Lufu	<i>nī</i>			<i>ánī</i>

## 5. Edoid

The relatively small group of 33 Edoid languages is spoken in Southern Nigeria, in three areas separated from each other by either natural or ethnic borders. This division is the basis for the internal classification of Edoid: since Elugbe (1979, 1989) this family is usually divided into four main groups: North Western (NWE) and North Central (NCE) languages are spoken in the northern half of Bendel State; the South Western (SWE) languages border them from the south, and Delta Edoid (DE) is scattered in the language-dense district of the Niger Delta, to the east of the larger bulk of Edoid. The bordering ethnic groups speak West BC (Defoid, Igboid, Nupoid) and Ijoid languages.

Though the Edoid language area is rather compact, and even the earliest scholars noticed their mutual resemblance (cf. Koelle 1854), common morphosyntactic features of Edoid were not recorded until Strub (1915–16) gave a brief grammar description of one of Yekhee dialects, and Westermann (1926) analysed nominal class prefixes of some of the family's languages. By that time, the Edoid languages were seen as members of the Kwa family, and this tradition lasted until the 1970's when a few groups of Nigerian languages were unified under the term 'West BC' and thus joined the (New) BC family (Bennett & Sterk 1977).

The phonetic system of Proto-Edoid must be characterised by an extremely rich consonant inventory, including the fortis/lenis distinction of all obstruents and nasals, and two implosives which are only preserved now in DE (Elugbe 1989: 297). A ten-vowel system is reconstructed, in which *a, e, i, o, u* are joined by their 'non-expanded pharynx', more open pairs (Elugbe 1989: 298). Vowel harmony is one of the distinctive features in Edoid morphemes, including person markers. +ATR and -ATR vowels vary within the same lexeme, influenced by vowel harmony and the consonant environment. These varying vowels will be indicated by the respective capital letters for Proto-Edoid (i.e. \**mI* for \**mi / mi*).

Materials discussing Edoid grammar reconstructions mostly belong to Elugbe (1979, 1980, 1983, 1984, 1989). Neither these nor any other materials available to us cover the reconstruction of person markers or pronouns.

Personal pronouns are mostly independent in Edoid, though the subject forms are sometimes used as prefixed markers. A subject marker serves as a clitic following a nominal group subject, e.g. in Ivbie:

Ogele ɔ ma yεεε (5)  
 Ogele 3SG much know  
 ‘Ogele knows much’ (Emuekpere-Masagbor 1997: 190).

Subject markers are obligatory after emphatic pronouns and in copulative constructions with a nominal predicate. In Degema, both the independent pronoun and the concord person prefix are used in a verb phrase to show the emphasis:

mée mi-món óji (6)  
 I 1SG-see.FACT him  
 ‘I saw him’ (Kari 2007: 92).

Four series of pronouns are rarely observed: the object marker coincides either with the subject or the possessive one. In some NCE tongues, all four series of markers given in 5.1. are identical. Possessive markers are located post-nominally and are syntactically independent.

### 1st person singular

Table 5.1

	subject	object	possessive	independent / emphatic
Yekhee (NCE) <sup>60</sup>	ȳ, ì- / í-	me, mē	me, mhè	me, mē
Edo (NCE) <sup>61</sup>	i	βέ	βέ	(i)mε
Ivbie (NCE) <sup>62</sup>	mhi / i	mhε	mhε	mhεmhε
Okpamheri (NWE) <sup>63</sup>	mi, mε	mε	hmε	mεmε
Epie (DE)		méni, mε	mε	mεmε
Degema (DE) <sup>64</sup>	me- / mε-, mi- / mi- (past, neg.)	mé / mé	mée / méeε	méeē
Engenni (DE) <sup>65</sup>	mi / mi	-me / -mε	me / mε	me / mε
Urhobo (SWE) <sup>66</sup>	mí, mé	βε, mε	mε	oma-mε
Okpe (SWE) <sup>67</sup>	mì / mi			
Uvbie (SWE) <sup>68</sup>	mì, mèmè (progr.)			

There are three candidates for Proto-Edoid: \*mi and \*i as the subject pronoun, and \*me as a non-subject marker. The emphasis is often expressed by reduplication: Okpamheri, Epie mεmε, Ivbie mhεmhε.

<sup>60</sup> Strub 1915–16: 458–459, Elimelech 1976

<sup>61</sup> Amayo 1980 [1975]

<sup>62</sup> Emuekpere-Masagbor 1997

<sup>63</sup> Elugbe 1980

<sup>64</sup> Kari 2002; 2003; 2004, Thomas & Williamson 1967

<sup>65</sup> Thomas 1976 [1968]

<sup>66</sup> Ukere 1986

<sup>67</sup> Omamor 1988

<sup>68</sup> Omamor 1988

The *i*- forms in NCE may be derived from lenition and further elimination of the initial *\*m*- before a front vowel. However, in Ivbie, the lenited form *mhi* is interchanged with *i*, which may demonstrate a different origin of the latter. The independent pronoun (*i*)*mε* is found throughout the family.

We may suggest the following Proto-Edoid forms:

<i>*mI</i>	subject (independent)
<i>*I-</i>	subject (prefixed)?
<i>*mE</i>	object / possessive
<i>*mE / *mEmE</i>	emphatic

## 2nd person singular

Table 5.2

	subject	object	possessive	independent / emphatic
Yekhee	<i>ū</i>	<i>ē, ugwε, ugwεguε</i>	<i>è, itsyε</i>	<i>ē, ugwε, ugwεguε</i>
Edo	<i>u</i>	<i>wé, é, ó, ǒ, á, á, ǎ, ǎ, rùé, rùé</i>	<i>rué / rué</i>	<i>(u)wè</i>
Ivbie	<i>u</i>	<i>giε, ε</i>	<i>giε</i>	<i>giε</i>
Okpamheri	<i>u</i>	<i>wε</i>	<i>bwε</i>	<i>wεwε</i>
Epie				<i>wɔwɔ</i>
Degema	<i>mv- / mu- v- / u- (neg.) e- / ε- (neg. imp.)</i>	<i>wó</i>	<i>wó</i>	<i>wó, uwɔ (Atala)</i>
Engenni	<i>bu / bv</i>	<i>-wo / -wɔ</i>	<i>wo / wɔ</i>	<i>bo / bɔ</i>
Urhobo	<i>wε, wɔ</i>	<i>wε</i>	<i>wén</i>	<i>oma-wen</i>
Okpe	<i>wù / wù</i>			
Uvbie	<i>ù, guè (progr.)</i>			

We might reconstruct *\*u* and *\*wo / \*we* for subject and non-subject markers respectively. Emphatic pronouns which may be suggested are *\*wεwε / \*wɔwɔ* or *\*uwε / \*uwɔ*.

Yekhee and Ivbie demonstrate a development *\*we > \*gwe (> giε)*. Edo's collection of object markers vary depending on the phonetic environment. Degema has *mu*, which consists of the original *u* attached to the affirmative prefix *m-*. In the reconstructions below, we show capital letters to avoid duplicating forms with alternating +ATR or -ATR vowels.

<i>*U</i>	subject
<i>*wO, *wE</i>	non-subject
<i>*wEwE / *wOwO, *uwE / *uwO</i>	emphatic

## 1st person plural (Table 5.3)

Table 5.3

	subject	object	possessive	independent / emphatic
Yekhee	<i>mā</i>	<i>mā</i>	<i>ma</i>	<i>mā</i>
Edo	(i)mà	ímà	ímà, βá	(ì)mà
Ivbie	<i>eye</i>	<i>eye</i>	<i>eye</i>	
Okpamheri	<i>mani</i>	<i>mani</i>	<i>hmani</i>	
Degema	<i>me-</i> / <i>mε-</i> , <i>e-</i> / <i>ε-</i> (neg.)	<i>ení</i>	<i>néni</i>	<i>eni</i>
Engenni	<i>eni</i>	<i>eni</i>	<i>eni</i>	<i>eni</i>
Urhobo	<i>aβáre</i>	<i>aβáren</i>	<i>raβare</i>	<i>oma-ráβáre</i>
Okpe	<i>àmí</i> / <i>àmí</i>			
Uvbie	<i>órìn</i>			

The only form which can be supposed for Proto-Edoid is *\*ma* / *\*mE*, observed in NCE and NWE. No external cognates for this form were found in West BC. However, we find some examples of subject *\*mV-* in Central Nigerian: Idun (NW Plateau) *mεn*, C'Lela (W Kainji) *mà*; and in CR: Mbembe *mo-* / *mɔ-*, Legbo *me* / *mè* (negative). Maybe at least some of them were generated under the influence of the 1sg. marker *\*me* / *\*mi* (which is clear for Okpe). Typologically, this seems reasonable. Building the plural of the pronouns by changing the vocalism of the singular one is a common feature in Indo-European and beyond (Indo-European 1sg. *\*me* ~ 1pl. *\*mes* / *\*mos*, Turkic 1sg. *\*ben* ~ 1pl. *\*bir*, Mongolian 1sg. *\*bi* ~ 1pl. *\*ba*, Hungarian 2sg. *\*te* ~ 2pl. *\*ti*, etc.); this method of building the plural is one the most widespread in the languages of the world (Siewierska 2004). Moreover, a vowel change is the core method of building the plural in the Edoid nominal system, and it is no wonder that it might be used for person markers as well.

Outside the BC family, similar markers of 1pl. are found in Kwa, namely the Gbe group (*\*mí* 'we') and in C and W Tano languages (*\*ame* 'we incl.') (Бабаев 2010). Several Grusi (the Gur family) languages show *ma* as an object marker (Seeger 2002–07).

Other forms in Table 5.3 do not allow to suggest any immediate reconstructions. We should note the resemblance of Urhobo *aβáre* 'we' to Bekwarra (Bendi, CR) *àbèrē* and Tesu (Alumic, Plateau) *mbɔrɔ*. However, in Edoid there are no traces of the NC plural class prefix *\*ba-*, so this form seems to be a frozen relic of this plurality marker.

## 2nd person plural (Table 5.4)

Most languages show various stages of levelling in the four series of pronouns. Yekhee, Edo, Ivbie and Engenni have basically only one pronoun for all meanings, Degema has two.

We tentatively reconstruct the Proto-Edoid pronoun as *\*wa* / *\*βa*. *\*ba*-forms in other BC languages must be cognate with this marker, which obviously originated from the plural noun class prefix.



Table 5.4

Yekhee	<i>βā, iβā</i>	<i>βā, iβā</i>	<i>βa</i>	<i>βā, iβā</i>
Edo	<i>(ù)wà</i>	<i>úwà</i>	<i>úwà</i>	<i>(ù)wà</i>
Ivbie	<i>vhe / vha</i>	<i>vhe</i>	<i>vhe</i>	<i>vhevhe</i>
Okpamheri	<i>we</i>	<i>wa</i>	<i>buba</i>	
Degema	<i>ma- / ma-, a- / ə- (neg.)</i>	<i>máaŋ / máəŋ</i>	<i>máaŋ / máəŋ</i>	<i>máaŋ / máəŋ</i>
Engenni	<i>ba</i>	<i>-ba</i>	<i>ba</i>	<i>ba</i>
Urhobo	<i>wá, βán</i>	<i>wá, βán</i>	<i>rowán, roβán</i>	<i>oma-rowan</i>
Okpe	<i>àrí / àri</i>			
Uvbie	<i>màmá</i>			<i>màmá</i>

Summing up Section 5, the following hypothetical reconstructions for Proto-Edoid are suggested:

Table 5.5

	subject	object / possessive	independent / emphatic
1sg.	<i>*mI, *I-(?)</i>	<i>*mE</i>	<i>*mE, *mEmE</i>
2sg.	<i>*U</i>	<i>*wO / *wE</i>	<i>*wOwO / *wEwE, *uwO / *uwE</i>
1pl.	<i>*ma / *mE</i>	<i>*ma / *mE</i>	<i>*ma / *mE</i>
2pl.	<i>*wa / *βa</i>	<i>*wa / *βa</i>	<i>*wa / *βa</i>

## 6. Defoid and Igboid

Yoruboid and Akokoid languages were unified under the Defoid brand (Crozier & Blench 1992), although no consensus has been reached on this among Africanists (Capo 1989), and Igboid languages are defined as their close relatives in (Williamson & Blench 2000). To date, there is no comparative grammar or complex phonetic research written on any of the families in West BC, except some hard-to-find dissertation manuscripts in Nigerian universities. The situation is complicated further by the lack of data on many languages, beyond 100/200-item wordlists which rarely include personal pronouns and never deal with paradigms.

Around two dozens of Defoid tongues (the exact number is still risky to claim) are spoken in Southeastern Nigeria, the adjacent border areas of Benin and some districts of Central Togo. Yoruba or Yoruboid lects dominate the family, Yoruba itself having the greatest number of speakers in Nigeria, Benin and as far as West Indies, Brazil and USA. Igala, an isolated tongue considered a close relative of Yoruboid, is in use on the Eastern bank of the Lower Niger, and is classified as a separate subgroup. The other branch of the proposed Defoid stock, Akokoid, is a cluster of as many as ten less known dialects spoken to the SW of the Niger/Benue confluence.

The internal subclassification of Yoruboid has been proposed by Akinkugbe (1976, 1978), analysed by Capo (1989) and is based on a number of innovations separating Igala from the main Yoruboid core. There is a draft Proto-Yoruboid Swadesh list put online by Guillaume

Segerer<sup>69</sup> (later referred to as [online]), but the conclusions there are apparently made upon Yoruba and Igala only, and only subject markers are included.

An attempt to classify Akokoid is being made by Blench (in progress). Ohiri-Aniche (1991, 1999) showed that Akokoid is as close to Edoid as it is to Yoruboid. Of all the Akokoid languages, Arigidi has enjoyed the most attention (e.g., Abiodun 1984), but mainly in regard to phonetics, so that personal marking remained almost untouched. For Yoruba, personal pronouns were studied by Bamgbose (1966), Fresco (1968), Omamor (1976) and others. There is a brief comparison of four basic personal pronouns of western Nigerian languages in Ohiri-Aniche (1999: 90). For Akokoid tongues, we only give Arigidi data from his paper.

The Igboid language area lies in the left-bank half of the Lower Niger basin in Nigeria. The total number of languages (some of which should rather be called dialect clusters) does not exceed ten, but the dialects of one idiom may vary greatly. Igbo, the largest language of the family with over 25 million speakers, includes over 30 dialects recognised so far (Lewis 2009). However, they have some slight differences in morphology. Outside Igbo, on which there exists extensive linguistic literature (see Manfredi 1989 for a survey), the only language thought to represent a separate branch of Igboid is Ekpeye, a tiny language spoken between the Orashi and Sambreiro rivers. Ekpeye is surrounded by Delta Cross, Edoid and Ijoid languages and must have undergone some influence from them. Some of the few sources of language data are the two papers by Clark (1969, 1972) and an unpublished wordlist on Roger Blench's website (Blench 2006a).

The Proto-Igboid phonetic system, according to Manfredi (1989: 344), included palatals such as *c, j, ʃ, ɟ*, labiovelars *kw, gw, nw, hw, kp*, and *gb*, and a 9- or 10-item vocalic system which we find in all Igboid lects. The vowel harmony is widespread, and the vowel system with the +/-ATR opposition must go back to the proto-language. Person markers of both Igboid and Defoid have plenty of tense/aspect, polarity and modality variations, mostly for the subject, marked by both vowel alternations and difference in tone. Tones, widely used for building various syntactic meanings all over West BC and Kwa (Creissels 2000: 238), are important since a great lot of person markers in these languages have a V form, as a result of reduction of earlier CV markers. In a number of tongues, subject, object and possessive markers fully coincide on the segmental level, with only occasional tonal distinctions. Direct object markers usually resemble their subject counterparts (sometimes with a different tone as well), or coincide with the possessive pronouns.

### 1st person singular (Table 6.1)

We cannot go further than *\*mI* (with an undefined front vowel) for the reconstruction of all the three right-hand sets: subject, object and possessive pronouns. Proto-Igboid must have had *\*mi*, but the Igbo forms have undergone reduction everywhere, so we can only rely upon the stressed pronouns here. Segerer (online) suggested Proto-Yoruboid *\*mĩ* without any supporting evidence. The independent pronoun is *\*amI* for both Proto-Igboid and Proto-Defoid. Izi *-bedva* is a particle used to indicate emphasis in all person markers.

<sup>69</sup> <http://sumale.vjf.cnrs.fr/NC/docs.php>

Table 6.1

	subject	object	possessive	independent / emphatic
<b>Defoid</b>				
Igala <sup>70</sup>	<i>mi, nń, nná</i> (progr.)	<i>mi</i>	<i>mi</i>	<i>omi</i>
Yoruba <sup>71</sup>	<i>mo / mò / mà,</i> <i>n, ń</i> (neg., past)	<i>mi</i>	<i>mi</i>	<i>èmi, òmí</i> (Oka)
Arigidi <sup>72</sup>				<i>ame</i>
<b>Igboid</b>				
Igbo <sup>73</sup>	<i>m-, m / mv</i>	<i>m</i>	<i>m</i>	<i>àmí</i>
Echie <sup>74</sup>	<i>m(v)</i>	<i>-m</i>	<i>m(v)</i>	
Ogba <sup>75</sup>			<i>m</i>	
Ikwere <sup>76</sup>	<i>m</i>	<i>m̃</i> (before obj.), <i>mé</i>	<i>mé</i>	<i>mé, mé</i> (contrastive)
Izi <sup>77</sup>	<i>mu</i>	<i>mu</i>	<i>mu</i>	<i>mbedva</i>
Ekpeye <sup>78</sup>	<i>me / mε / ma / mɔ /</i> <i>mo-</i> (tense), <i>N-</i>	<i>-m̃</i>	<i>mê</i>	<i>mê</i>

## 2nd person singular (Table 6.2)

Table 6.2

	subject	object	possessive	independent / emphatic
<b>Defoid</b>				
Igala	<i>wε, ε, wεé</i> (progr.)	<i>wε</i>	<i>wε</i>	<i>uwe, εε</i>
Yoruba	<i>o, ò</i> (before particle), <i>wà</i> (before <i>á</i> )	<i>ɔ, ε</i>	<i>ε, rε</i>	<i>ìwɔ, ùwɔ</i> (Oka)
Arigidi				<i>árɔ̃</i>
<b>Igboid</b>				
Igbo	<i>i- / ɪ-, gɪ</i>	<i>gɪ</i>	<i>i</i>	<i>àgí</i>
Echie	<i>ì, gɪ</i>		<i>gɪ</i>	
Ikwere	<i>ɪ</i>	<i>í, ì</i> (before object)	<i>í</i>	<i>jí, jé</i> (contrastive)
Ogba	<i>yυ, ɪyυ</i>			
Izi	<i>i / ɪ</i>	<i>ngu</i>	<i>ngu</i>	<i>gvbedva</i>
Ekpeye	<i>i- / ɪ- / e- / ε-</i> (tense)	<i>yɔ, -ɪ</i>	<i>yò</i>	<i>yò</i>

<sup>70</sup> Philpot 1935<sup>71</sup> Bamgbose 1966, Ohiri-Aniche 1999: 90<sup>72</sup> Ohiri-Aniche 1999: 90<sup>73</sup> Фихман 1975<sup>74</sup> Blench, manuscript c<sup>75</sup> Blench 2005<sup>76</sup> Segerer 2002–07<sup>77</sup> Meier 1976<sup>78</sup> Clark 1976

Yoruboid and Akokoid data are too scarce to make ultimate conclusions: emphatic *\*uwe / \*uwo* might be the only form reconstructible for Proto-Yoruboid. Segerer (online) reconstructs Proto-Yoruboid *\*o*. Arigidi shows an interesting form which is comparable with Eleme (CR, Ogoni) object *\*-rũ* ‘us’ and Edo non-subject *rué / rué* ‘our, us’. Let us also add here Ukaan (Section 8 below) *ihyè-rɔ̀* ‘you (sg.)’ (Ohiri-Aniche 1999: 90).

Proto-Igboid markers can be reconstructed as *\*i-* (subject) and probably *\*gi / \*gu-*.

### 1st person plural (Table 6.3)

Table 6.3

	subject	object	possessive	independent / emphatic
<b>Defoid</b>				
Igala	<i>a, aá</i> (progr.)	<i>wa</i>	<i>wa</i>	<i>a<sup>w</sup>a</i>
Yoruba	<i>a, à</i> (before particle or <i>á</i> )	<i>wa</i>	<i>wa</i>	<i>à<sup>w</sup>a</i>
Ede dialects		<i>wa, γa, a</i>		
Arigidi				<i>áò</i>
<b>Igboid</b>				
Igbo	<i>a<sup>n</sup>i / a<sup>n</sup>u</i>	<i>a<sup>n</sup>u</i>	<i>a<sup>n</sup>u</i>	<i>a<sup>n</sup>í</i>
Echie	<i>ɲu</i>	<i>ɲu</i>	<i>ɲu</i>	
Ikwere	<i>a</i>	<i>áyì</i>	<i>áyì</i>	<i>áyìlé</i>
Ogba	<i>yê</i>			
Izi	<i>a<sup>n</sup>i</i>	<i>a<sup>n</sup>i</i>	<i>a<sup>n</sup>i</i>	<i>a<sup>n</sup>ibedva</i>
Ekpeye	<i>a-</i> (excl.), <i>a-...-nì</i> (incl.)	<i>ye</i>	<i>yè</i>	<i>yê</i>

Yoruboid markers may be derived from *\*a* for the subject marker, *\*wa* for object and possessive, and *\*awa* for the independent emphatic pronoun. Proto-Igboid must have had *\*(a)ni*, though note *\*a* in Ikwere and Ekpeye (plural exclusive). The *\*(a)ni* form has a number of notable cognates across BC: let us name Efik and Ibibio (Lower CR) *ɲìn* ‘we’, Ake (South West Plateau) *ani* ‘we’, Degema and Engenni (Delta Edoid) *eni* ‘we’.

### 2nd person plural (Table 6.4)

The Yoruba object and possessive pronoun *ɲín* ‘you, your’ reminds of *\*ɲin* which we have suggested in CR (see Section 3 above). A slightly different marker is found in Proto-Igboid *\*(V)nu*. This pronoun appears below in Idomoid (Section 7) and Oko (Section 8) as well.

Arigidi *ámô* has much in common beyond Defoid. It must be cognate to the Proto-Bantu subject prefix *\*mv-* (Babaev 2008: 147) and probably Ukaan (Ikakumo dialect) *ihyè-mò*. This correspondence, if verified, would be enough to suggest a Proto-BC retention.

Table 6.4

	subject	object	possessive	independent / emphatic
<b>Defoid</b>				
Igala	<i>amε, mε, mεε</i> (progr.)	<i>mε</i>	<i>mε</i>	<i>amε</i>
Yoruba	<i>ε, ê</i> (before particle or <i>á</i> )	<i>yín</i>	<i>yín</i>	<i>êyín, êηwĩ</i> (Oka)
Arigidi				<i>ámô</i>
<b>Igboid</b>				
Igbo	<i>unù</i>	<i>unù</i>	<i>unù</i>	<i>unû</i>
Echie	<i>nv</i>	<i>nv</i>	<i>nv</i>	
Ikwere	<i>ánv</i>	<i>ánv</i>	<i>ánv</i>	<i>ánvélé</i>
Ogba	<i>vnv</i>			
Izi	<i>unu</i> (variable tone)	<i>unu</i>	<i>unu</i>	<i>unubεdva</i>
Ekpeye	<i>i- / ɪ- .. -nì,</i> <i>e- / ε- .. -nì</i>	<i>yónì</i>	<i>yò</i>	<i>yó / yónì</i>

We will briefly summarise our tentative reconstructions in the following charts:

Table 6.5

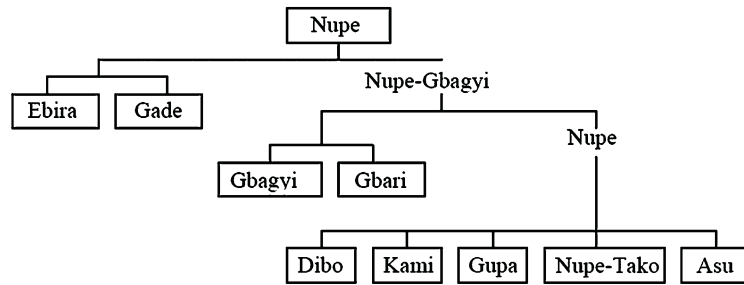
	subject	object / possessive	independent / emphatic
1sg.	<i>*mI</i>		<i>*amI</i>
2sg.	<i>*i</i> (Igboid)	<i>*gi / *gu</i> (Igboid)	<i>*uwe / *uwo</i> (Yoruboid)
1pl.	<i>*a</i> (Yoruboid) <i>*(a)ni</i> (Igboid)	<i>*wa</i> (Yoruboid)	<i>*awa</i> (Yoruboid)
2pl.	<i>*ηin</i> (Yoruboid) <i>*(V)nu</i> (Igboid)		

It is obvious that, although we tend to fit person markers into a traditional “Bantu-style” four-column table, noting subject, object, possessive markers and independent pronouns, this does not seem realistic for West BC. Proto-Igboid must have had only two distinct series of person markers (subject and object), and Yoruboid could have up to three, including independent pronouns with a vocalic prefix.

## 7. Nupoid and Idomoid

The Nupoid family spoken at the Niger-Benue confluence and to the north of it includes around a dozen languages: Blench (1989a) names seventeen, while Ethnologue (Lewis 2009) gives only eleven. They are indeed tightly knit together, which was noticed long ago by Koelle (1854: 8–9), who was the first to describe some of them. The grouping of Nupoid seems to have never been in doubt, as well as its subclassification: the Ebira-Gade group must be quite distant from the Nupe-Gbagyi core, within which Nupe and Gwari (Gbagyi-Gbari) subgroups are usually identified.

Diagram 3



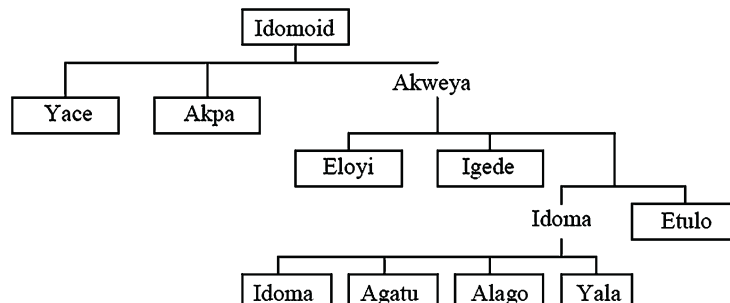
Nupoid comparative research was conducted by Sterk (1977, 1978), Bennett & Sterk (1977), Hyman & Magaji (1970), Madugu (1975, 1979, 1985), to name just a few. The overview is contained in Blench (1989a). It argues for a 10-vowel system for Proto-Nupoid, with a complete series of +ATR and -ATR slots, and a fortis/lenis counter-opposition of consonants. The proto-language consonantal system is given by Blench as follows:

p b	t d	c j	k g	kp gb	
	s z				h
	r				
	l				
m	n				
w	y				

Nupoid languages use at least five tones, with three level tones and two contour tones. Noun classes are mostly marked by innovated suffixes, while some languages only have suffixes marking the plural. The word order of the verb phrase is SVO, the subject and the object of the verb expressed by free-standing pronouns. The object marker may be infixes into the verb root, like in Nupe *kpēyē* ‘to know’ – *mī-kpē-ū-yē* ‘I know it’. Possessive markers are post-posed to nouns.

While the Nupoid language area is mostly surrounded by other West BC tongues, including Igboid and Yoruboid, the smaller Idomoid family of languages lies farther to the east, bordering some of the East BC lects: e.g., Tiv (Bantoid) and Bokyi (CR). Eight or nine Idomoid languages are usually identified, which are classified following Crozier & Blench (1992):

Diagram 4



The idea of close genetic relations between Nupoid and Idomoid was first expressed as early as the mid-nineteenth century, and further supported by Bennett & Sterk (1977) and Blench (1989a). However, Armstrong (1989: 323) considered the Idomoid branch coordinate with all other West BC families. Armstrong had been analysing Idomoid from a comparative

standpoint since 1955, and has described a number of dialects, along with some comparative research papers (1981, 1983). Apart from this, most works are synchronic overviews, some of which are used in the charts below.

Proto-Idomoid must have had a 10-vowel system which is still observed in Igede, but was reduced to seven items in most of the other languages of the family. The consonant system lacks a fortis/lenis opposition characteristic for Nupoid. As well as in the latter, tones play a vital role in morphology, including the pronominal forms which build negative, tense, and object meanings using various tones. Idomoid languages have up to four tones.

In Idomoid, person markers are usually bound to verbs. The only independent pronoun is used in the emphatic context. Possessive suffixes are often attached to the attributive particle which is a former noun, cf. Idoma *kú-m* ‘my’, *kú-nū* ‘his’ etc. This feature is also encountered in East BC languages, including Bantoid and CR languages.

### 1st person singular (Table 7.1)

Table 7.1

	subject	object	possessive	independent / emphatic
<b>Nupoid</b>				
Ebira <sup>79</sup>	<i>ma / me</i>	<i>mi / mu</i>	<i>amu / emi</i>	<i>ɛmu</i>
Gade <sup>80</sup>	<i>ñ</i>			
Nupe-Tako <sup>81</sup>	<i>mi, mii</i> (opt.), <i>mí</i> (neg.)	<i>mî</i>	<i>mî</i>	<i>mî</i>
Gbagyi <sup>82</sup>	<i>mi</i> (tone)	<i>mi</i> (tone)	<i>mi</i> (tone)	<i>òmi</i>
Gbari <sup>83</sup>	<i>myi, mà</i> (fut.), <i>má</i> (past)	<i>myi</i>	<i>myi</i>	
<b>Idomoid</b>				
Idoma <sup>84</sup>	<i>ṁ</i>	<i>-ṁ</i>	<i>-m</i>	<i>àmì</i>
Eloyi <sup>85</sup>	<i>mi-</i> (tone)	<i>mī</i>	<i>-mī</i>	<i>īmī, īmīmī</i>
Igede <sup>86</sup>	<i>ṁ-</i>	<i>-ṁ / -ṁ̄</i>	<i>ṁāṁ</i>	<i>àṁ</i>

For both Nupoid and Idomoid, only two forms may be reconstructed: *\*mi* for the marker expressing the verb subject, object and possession, and *\*Vmi* for the emphatic pronoun.

<sup>79</sup> Scholz 1976 [1973]

<sup>80</sup> Sterk 1977

<sup>81</sup> Smith 1980 [1967]

<sup>82</sup> Hyman & Magaji 1970

<sup>83</sup> Blench & Doma, manuscript

<sup>84</sup> Armstrong 1989, Segerer 2002–07

<sup>85</sup> Mackay 1976 [1968], Blench 2007d

<sup>86</sup> Bergman 1976 [1973]



## 2nd person singular (Table 7.2)

Table 7.2

	subject	object	possessive	independent / emphatic
<b>Nupoid</b>				
Ebira	<i>u / v</i>	<i>wu / wv</i>	<i>awv / ewu</i>	<i>εwv</i>
Nupe-Tako	<i>wo, o / woo (opt.), wo (neg.)</i>	<i>wô / ô</i>	<i>wô / wê / ô</i>	<i>wô</i>
Gbagyi	<i>ho</i>	<i>ho</i>	<i>ho</i>	<i>òho</i>
Gbari	<i>he</i>	<i>he</i>	<i>he</i>	
<b>Idomoid</b>				
Idoma	<i>à</i>	<i>-ò</i>		<i>àwò</i>
Eloyi	<i>ηo- (tone)</i>	<i>ηò, ηa</i>	<i>-ηa</i>	<i>ũη(w)ò</i>
Igede	<i>à-, àhù</i>	<i>-h̄ / -h̄</i>	<i>h̄āh̄</i>	<i>àhù</i>

*u / o / v* markers and *wu / wo* in non-subject series are cognate to 2sg. markers in other BC languages. Pronouns with initial *η-* in Idomoid remind of the *ηo / ηwo* forms in Plateau languages, as well as some Igboid forms.

## 1st person plural (Table 7.3)

Table 7.3

	subject	object	possessive	independent / emphatic
<b>Nupoid</b>				
Ebira	<i>ya / ye, ι / i</i>	<i>yi / yι</i>	<i>ayι / eyi</i>	<i>εyι</i>
Nupe-Tako	<i>yi, yii (opt.), yi (neg.)</i>	<i>yî</i>	<i>yî(-zi)</i>	<i>yî</i>
Gbagyi	<i>yi</i>	<i>yi</i>	<i>yi</i>	<i>òyi</i>
Gbari	<i>yi</i>	<i>yi</i>	<i>yi</i>	
<b>Idomoid</b>				
Idoma	<i>àlò</i>	<i>-(à)lò</i>		<i>àlò</i>
Eloyi	<i>ki-</i>	<i>nzú</i>	<i>-nzú</i>	<i>kínzú</i>
Igede	<i>àhì</i>	<i>-hì</i>	<i>h̄āhì</i>	<i>àhì</i>

The Proto-Nupoid pronouns may be conjectured to have been *\*yi* for subject / object / possession and *\*Vyi* for the emphatic pronoun. The *-zi* suffix in Nupe is a nominal plurality marker.

The Idoma pronoun reminds of *lo-*, which we observed in Bantu zones A-C (Babaev 2008: 171–175), where they are derived from Proto-Bantu *\*tù-*. Igede *àhì* may be a derivation from *\*aci* or *\*ati*, widely witnessed across BC.

## 2nd person plural (Table 7.4)

Table 7.4

	subject	object	possessive	independent / emphatic
<b>Nupoid</b>				
Ebira	<i>wa / we nini</i>	<i>wu / wu nini</i>	<i>awv / ewu nini</i>	<i>εwv nini</i>
Gade	<i>na</i>			
Nupe-Tako	<i>ye, e (before cons.), yee (opt.), yé (neg.)</i>	<i>ê / yê</i>	<i>ê / yê(-zì)</i>	<i>yê</i>
Gbagyi	<i>fye</i>	<i>fye</i>	<i>fye</i>	<i>òfye</i>
Gbari	<i>se</i>	<i>se</i>	<i>se</i>	
<b>Idomoid</b>				
Idoma	<i>àá</i>	<i>àá, -nḍ</i>		<i>àḍ</i>
Eloyi	<i>lu-, lalu- (tone)</i>	<i>yí</i>	<i>-yí</i>	<i>lú(n)yí</i>
Igede	<i>ànḍ</i>	<i>-nḍ</i>	<i>nānḍ</i>	<i>ànḍ</i>

For Idoma and Igede, the proto-form must be *\*-nv / \*-nɔ* and *\*ànḍ / \*ànḍ*. Both *\*nu* and *\*ni*, as we have seen earlier, are frequent all across the BC area. But there is a clear distribution: *\*nu / \*Vnu* is seen in Igbooid and Bantoid, while *\*ni* is in Yoruboid, Central Nigerian and CR (but note *anu(η)* in some Plateau dialects in Section 4). We need a much deeper insight to reconstruct any proto-forms here.

In conclusion, we must say that the pronominal systems of Nupoid and Idomoid do not share enough similarities to justify the hypothesis of a specific mutual genetic relationship. Beyond the 1sg., there is hardly a single pronoun that we can reconstruct for ‘Proto-Nupoid-Idomoid’. The idea of the NOI unity which we mentioned earlier does not seem to find evident support in the systems of person marking.

## 8. Benue-Congo Isolates

**Oko** (Ogori) is an unclassified BC language spoken by some 10,000 people in the Kogi State of Central Nigeria. Together with Ukaan and Akpes, Oko has long been regarded as an isolate due to significant lexical differences with the surrounding BC tongues. However, Williamson & Blench (2000) included it into West BC, placing it into the same stock with Nupoid and Idomoid.

Table 8.1

	subject	object	possessive	independent / emphatic
1sg.	<i>ì- / è- (perf.) mè- / mà- (progr.)</i>	<i>mame, -mu / -mɔ</i>	<i>m-</i>	<i>àmé</i>
2sg.	<i>ù- / ɔ- (perf.) wè- / wà- (progr.)</i>	<i>wawɔ, -wu / -wɔ</i>	<i>w-</i>	<i>àḍ, awɔ</i>
1pl.	<i>tì- / tè- (perf.) tè- / tà- (progr.)</i>	<i>tatɔ, -tu / -tɔ</i>	<i>t- / tì-</i>	<i>àtɔ</i>
2pl.	<i>nḍ- / nè- (perf.) nè- / nà- (progr.)</i>	<i>nancɔ, -nu / -nɔ</i>	<i>n-</i>	<i>ànɔ</i>

There are at least two works on personal pronouns in Oko. We quote their forms below in accordance with works by Elugbe & Ologori (1980) and Atoyebi (2008). The variants divided by a slash mark depend on vowel harmony. See Table 8.1.

There are a few remarkable things about the Oko person marking system. First, its subject markers differentiate between progressive (incomplete) and perfective (complete) aspects. The difference lies mainly in the vocalism of the forms, but in the 1sg. they are built upon suppletive roots. This seems to have been the impact of an old aspect marker which, as it usually happens in BC, used to follow the subject marker, but finally merged with it, changing the vocalism of the latter. However, there is some additional morphological sense here. Complete markers are immediately linked to the verb root, while progressive ones are prefixed to the tense / aspect markers:

*ì-sú-já* (7)  
 1SG.S.COMP-marry-3SG.O  
 ‘I married him’, but

*mà-áka-gām*                      *óbīn*    *usie* (8)  
 1SG.S.PROGR-FUT-greet      king    tomorrow  
 ‘I will greet the king tomorrow’ (Atoyebi 2008: 31–32).

The only predicative marker preceded by complete person markers is that of negation: here the complete marker is used to prevent homophony, because in Oko the negative prefix is *-mè-*. It seems that the complete markers are more ancient, and the progressive ones may have originated from complex auxiliary constructions.

Second, there are bound and independent object forms, and the latter are built by means of reduplication. Finally, the possessive markers are prefixed: a feature rarely met across BC.

The roots used for marking persons are definitely cognate with those found in East BC. See the comparison in 8.1. showing structural, segmental and tonal similarities between Oko and NW Bantu independent pronouns:

Table 8.1

	Oko	Viya (Bantu B301)
1sg.	<i>àmé</i>	<i>mé</i>
2sg.	<i>àv / avɔ</i>	<i>wè</i>
1pl.	<i>átɔ</i>	<i>itú</i>
2pl.	<i>ànɔ</i>	<i>inú</i>

Note that Oko has no Bantu neighbours — it is surrounded by West BC languages, so that the chance of language contact with Bantu as the reason for such a striking similarity is quite low. Since the Bantu system is considered archaic compared to other BC families, Oko pronouns seem to enjoy a common retention of the original pronouns with Bantu.

**Akpes** is another isolate tongue of C Nigeria, spoken to the southwest of the Niger-Benue confluence squeezed between larger Edoid and Yoruboid areas. It consists of about ten dialects, most of which are mutually intelligible. Disputes on the external classification of Akpes have lasted ever since Hansford & al. (1976) first described it as ‘Kwa unclassified’. The view of Akpes as a BC language was expressed in (Williamson 1989) and reproduced in (Williamson & Blench 2000), supported by Elugbe (2001). Some classify Akpes as an Edoid language (Agoyi 1997).

The independent subject pronouns of Akpes are cited by Ohiri-Aniche (1999: 90).

Table 8.3

1sg.	<i>òṛì</i>
2sg.	<i>òsì</i>
1pl.	<i>àbès</i>
2pl.	<i>àbèṛ</i>

The 1sg. form is probably derived from \**òmì*. The two plural forms correspond perfectly to East BC: we saw (*à*)*bèse* < \*(*à*)*bè-c(u)e* in both Bantoid (Babaev 2008: 148) and CR languages (Section 3), there are remnants of this form in the languages of the Plateau (Section 4) and beyond. The same goes for \*(*à*)*bè-n(u)e* as the 2pl. form. The 2sg. form is a mystery, never found among BC languages. It can be a loanword, e.g. from some Gur languages of the area: cf. Ntcham *-sī*, Akasele *isé* or Konkomba *sí*, all denoting 2sg. and all spoken in Togo.

However, the three cognate forms out of four strongly suggest a common retention shared by Akpes and East BC languages.

Another unclassified BC language of C Nigeria is **Ukaan**, known by only a few papers published or placed online so far (Jungraithmayr 1973: 47–48), (Ohiri-Aniche 1999: 90), (Blench 2005b). Four dialects are identified, and each possesses its own grammatical peculiarities making them rather dissimilar to each other. Relying on the abovementioned works, we hereby present the forms from three of them: Ishe, Ikakumo and Ikan (only for the 1sg., taken from [Adekanye & Salffner 2007]).

Table 8.4

	subject	object	possessive	independent / emphatic
1sg.	<i>dò</i> (Ishe) <i>ìhyè-ji</i> (Ikakumo) <i>ja</i> (Ikan)	<i>wity</i> (Ishe)	<i>-j / -ji</i> (Ikan)	<i>joo</i> (Ikan)
2sg.	<i>hò</i> (Ishe) <i>ìhyè-rò</i> (Ikakumo)			
1pl.	<i>bà</i> (Ishe) <i>ìhyè-bò</i> (Ikakumo)	<i>úbò</i> (Ishe)		
2pl.	<i>mà</i> (Ishe) <i>ìhyè-mò</i> (Ikakumo)			

The Ikakumo dialect, whose subject pronouns are actually nouns with possessive suffixes, shares three forms of its pronominal markers with Arigidi (Akokoid) (see Section 6), and the Ishe dialect also corresponds with them. Based on this, we can confirm that the pronominal system of Ukaan is that of a West BC language quite close to Akokoid.

## 9. Conclusive notes

We have compiled data on person markers in some 300 languages of the BC family which might lead us to some preliminary conclusions on the Proto-BC system of person marking. Even

## 1st person singular

Table 9.1

	subject	object	possessive	independent / emphatic
Bantoid	* <i>ŋi-</i>	*( <i>à</i> ) <i>me</i>	*( <i>à</i> ) <i>me</i>	*( <i>à</i> ) <i>me</i>
Bantu	* <i>ŋi-</i>		*( <i>à</i> ) <i>me</i>	*( <i>i</i> ) <i>me</i>
CR	* <i>N-</i> , * <i>mi</i>	* <i>m(i)</i>	<i>àmì</i>	*( <i>à</i> ) <i>mì</i> , * <i>minV</i>
CN	* <i>me</i>			* <i>ame</i>
Edoid	* <i>mi</i> , * <i>i-</i>	* <i>me</i>	* <i>me</i>	* <i>meme</i>
Yoruboid	* <i>mI</i>			* <i>amI</i>
Igboïd	* <i>mI</i>			* <i>amI</i>
Nupoid	* <i>mi</i>	* <i>mi</i>	* <i>mi</i>	* <i>Vmi</i>
Idomoid	* <i>mi</i>	* <i>mi</i>	* <i>mi</i>	* <i>Vmi</i>
Oko	<i>ì-</i> , <i>mè-</i>	<i>mamε</i> , <i>-mu</i>	<i>m-</i>	<i>àmé</i>
Akpes				<i>òjù</i>
Ukaan	<i>dà</i> , <i>ja</i> , <i>-ji</i>		<i>-ji</i>	<i>joo</i>

## 2nd person singular

Table 9.2

	subject	object	possessive	independent / emphatic
Bantoid	* <i>ò-</i>	*( <i>à</i> ) <i>ue</i>	*( <i>à</i> ) <i>ue</i>	*( <i>à</i> ) <i>ue</i>
Bantu	* <i>ò-</i>	* <i>-kv-</i>	*( <i>à</i> ) <i>we</i>	*( <i>i</i> ) <i>we</i>
CR	* <i>o-</i> , * <i>a-</i>	* <i>wo</i> / * <i>wu</i> / * <i>we</i> , * <i>wo-n(V)</i> / * <i>we-n(V)</i>	* <i>òwò</i> , * <i>wo-n(V)</i> / * <i>we-n(V)</i>	* <i>òwò</i>
CN	* <i>o</i> , * <i>u</i>	* <i>wo</i> , * <i>wu</i>	* <i>wo</i> , * <i>wu</i>	* <i>awo</i> , * <i>awu</i>
Edoid	* <i>U</i>	* <i>wO</i> , * <i>wE</i>	* <i>wO</i> , * <i>wE</i>	* <i>wEwE</i> / * <i>wOwO</i> , * <i>uwE</i> / * <i>uwO</i>
Yoruboid				* <i>uwe</i> , * <i>uwo</i>
Igboïd	* <i>i</i>	* <i>gu</i> , * <i>go</i>	* <i>gu</i> , * <i>go</i>	* <i>gu</i> , * <i>go</i>
Nupoid	Nupe <i>wo</i>	Nupe <i>wô</i>	Nupe <i>wô</i>	Nupe <i>wô</i>
Idomoid	Idoma <i>â</i>	Idoma <i>-ò</i>		Idoma <i>âwò</i>
Oko	<i>ù-</i> , <i>wè-</i>	<i>wawɔ</i> , <i>-wɔ</i>	<i>w-</i>	<i>àv</i> , <i>awɔ</i>
Akpes				<i>òsì</i>
Ukaan	<i>hò</i> , <i>-rò</i>			

tentatively, this is no easy task. BC is an old family (over 6,000 years of age, according to [Williamson 1988: 104]), and the diversity of pronominal systems in all of its 900+ languages is much greater than that of Indo-European or Semitic. The data is scarce or incomplete, there is no historical track of the language development, and the lack of regular phonetic correspondences and reconstructed phonetic systems makes any comparative analysis highly vulnerable. Finally, various phonological processes of merging person markers with predicative markers

## 1st person plural

<i>*e-</i> / <i>*i-</i>	subject
<i>*àci(n)</i>	independent
<i>*mV</i>	subject / object / independent

Table 9.3

	subject	object	possessive	independent / emphatic
Bantoid	<i>*tù-</i>	<i>*(bè-)ce</i>	<i>*(bè-)ce</i>	<i>*(bè-)ce</i>
Bantu	<i>*tù-</i>		<i>*(a)cue</i>	<i>*icue, *(bè-)cue</i>
CR	<i>*e- / *i-, *mV</i>	<i>*mV</i>		<i>*àci(n), *mV</i>
CN	<i>*ti</i>	<i>*ti</i>	<i>*-at(i)</i>	<i>*at(i)</i>
Edoid	<i>*ma / *mE</i>	<i>*ma / *mE</i>	<i>*ma / *mE</i>	<i>*ma / *mE</i>
Yoruboid	<i>*a</i>	<i>*wa</i>	<i>*wa</i>	<i>*awa</i>
Igbooid	<i>*(a)ni</i>			
Nupoid	<i>*yi</i>	<i>*yi</i>	<i>*yi</i>	<i>*Vyi</i>
Idomoid	Idoma àlò	Idoma (à)lò		Idoma àlò
Oko	<i>tì- / tè-</i>	<i>tatɔ , -tu</i>	<i>t-</i>	<i>àtɔ</i>
Akpes				<i>àbès</i>
Ukaan	<i>bò / bà</i>			<i>úbɔ</i>

## 2nd person plural

Table 9.4

	subject	object	possessive	independent / emphatic
Bantoid	<i>*mv-</i>	<i>*(bè-)ne</i>	<i>*(bè-)ne</i>	<i>*(bè-)ne</i>
Bantu	<i>*mv-</i>		<i>*(a)nue</i>	<i>*(i)nue</i>
CR	<i>*e-, *i-</i>	<i>*ànI, *inI(n)</i>	<i>*ànI, *inI(n)</i>	<i>*ànI, *inI(n)</i>
CN	<i>*ni(n), *(a)nu</i>	<i>*ni(n), *(a)nu</i>	<i>*ni(n), *(a)nu</i>	<i>*ni(n), *(a)nu</i> Jukunoid <i>*ánī</i>
Edoid	<i>*wa / *βa</i>	<i>*wa / *βa</i>	<i>*wa / *βa</i>	<i>*wa / *βa</i>
Yoruboid	<i>*nin</i>			
Igbooid	<i>*(V)nu</i>			
Idomoid		<i>*nv / *nɔ</i>	<i>*nv / *nɔ</i>	<i>*ànù / *ànɔ</i>
Oko	<i>n̄- / nè-</i>	<i>nanɔ , -nu / -nɔ</i>	<i>n-</i>	<i>ànɔ</i>
Akpes				<i>àbèɲ</i>
Ukaan	<i>mò / mà</i>			
Ukaan	<i>bò / bà</i>			<i>úbɔ</i>

of tense, aspect, modality and polarity have made the situation in many languages obscure. We are confident that much more work is needed to create a reliable reconstruction of the proto-language system of person marking for BC. The following are just some conclusive

charts following the discussion in this paper. Bantu and Bantoid forms are given from (Babaev 2008: 148).

We may conclude that the object and possessive markers are mostly marginal compared to subject markers and emphatic pronouns. The object is marked in all languages of the stock, but in the majority there is no specific object series, and the subject markers or emphatic pronouns are used instead. For a number of languages (Bantoid, CR) we reconstruct a cumulative “non-subject” series that marks all of the remaining three meanings.

Similarly, the possessive meaning is built by means of the emphatic pronoun, usually suffixed or postposed to the possessed noun. Where proto-languages seem to have distinguished between emphatic and possessive pronouns (in Bantu, CN), the distinction was only a matter of a vowel prefix (or particle).

In Bantu, the independent pronouns constitute a separate person marking series which does not seem to be etymologically cognate with subject markers: *\*àme* and *\*ni-* in the 1sg. (Babaev 2008). Bantu is considered to be the most archaic of all BC branches, since it has preserved a number of important morphological retentions lost in the western part of the BC area. One of the examples is certainly the noun class system: Bantu preserved old class prefixes in their original CV-form, while most of its relatives in Nigeria either retained only V- prefixes or dropped every trace of the original system, sometimes (as in Jukunoid) elaborating a new one.

We face a similar situation in person marking. In Proto-BC, there seem to have been prefixed person markers of the verb subject and the independent pronouns, which were based on the same root as the non-subject markers (i. e. object and possessive). Originally this root was probably a nominal one: note that non-subject pronouns in Bantoid and CR languages can use the plural noun class prefix *\*bV-*, which the subject markers do not show.

Later the prefixes were weakened in the majority of BC branches: they were probably reduced to V-shape already in Proto-West BC. Just as it happened with noun classes, the CV-marker was weakened to V- or -V, and then finally was lost. The process was encouraged by the presence of a clear “substitute”, i. e. an independent pronoun with emphatic semantics that could instantly replace the prefixed subject marker. The BC languages of present-day Nigeria mostly use subject pronouns independently, and they are generally identical with or at least related to the emphatic pronouns.

Some traces of the former subject prefixed series are still found. In some languages, they are found in V-shape prefixes (e.g. for the 1sg. *ì-* or *̀-* which is a syllabic nasal bearing its own tone).

It is typologically common that when a language merges two series of pronouns in one, the resulting paradigm is mixed. Turkic languages, for instance, have adopted person markers of the preterite from the ancient possessive paradigm only in the 2sg. and 1pl. forms, using the original subject marker series for all the other slots. Unifications of several pronominal paradigms into one are commonly known in the world’s languages, and non-subject forms adopted to mark the subject are also not unknown: let us recall Indo-European *\*me-* ‘me’, which replaced older *\*eg’Hom* for the nominative case in Irish, Hindi or Farsi, or Indo-European *\*nos* ‘us’ which was generalised for the subject meaning in Latin and all Romance languages.

Syntactic shifts of this and other kinds are quite habitual for BC languages. In a number of languages, the pronominal paradigm is further unified by merging plural pronouns together as it happened in CR (cf. the Old English plural personal affix *-ap*). The plural particle *\*bV-* which in Proto-BC marked the 3pl. spread to the 2pl. and further to 1pl., squeezing out the original person markers; this was noted as early as Delafosse (1904: 29). Another common means of building the plural innovations is to add *\*bV-* to the corresponding singular pronoun, as in Tee (CR) *b̀̀̀̀* ‘you pl.’ vs. *̀̀̀* ‘you sg.’ Proto-Edoid chose to analogically level the paradigm based on the singular pronoun roots: *\*mi* / *\*me* ‘I’ ~ *\*ma* ‘we’ in the 1st person and *\*we* / *\*wo*



‘you sg.’ ~ *\*wa* ‘you pl.’ in the second, and Delta Cross languages use the same affixes for marking singular and plural of the 2nd and 3rd persons.

We may suggest that there were only two series of person markers in Proto-BC: the subject prefixed series and the non-subject independent series:

Table 9.5

	subject	non-subject
1sg.	<i>*N-</i>	<i>*mI</i>
2sg.	<i>*o- / *u-</i>	<i>*wV</i>
1pl.	<i>*tu- / *ti-</i>	<i>*(bV)ce</i>
2pl.	<i>*mV- ?</i>	<i>*(bV)nV</i>

Some possible reflexes of these tentative reconstructed person markers are given below. Bantu and Bantoid forms are taken from (Babaev 2008), all other forms are mentioned in the charts of the present article.

**\*N-**

Bantu: Oroko *N̄-*, Ndumu *n*, Enya *n-*, Shi *ñ-*, Digo *n-*, Tongwe *N-*, Kongo *N-*, Haya *N-*, Pimbwe *n-*, Yao *n-*, Tswana *n-*;

Other Bantoid: Samba-Daka *m̄*, Tiv *m*, Noni *N-*, Mbe *ñ-*, Kenyang *n- / n-*, Yemba *N*, Aghem *N*, Mundani *ñ-*  
CR: Bekwarra *N-*, Kohumono *N-*, Mbembe *N-*, Ibibio *N̄-*, Eleme *N̄-*

CN: Tarok *n / m,*, Lijili *n*, Basa *n* (past), Hone *n-*, Jukun *m*

Defoid: Igala *n̄*, Yoruba *n, η* (neg., past)

Igboid: Igbo *m-*, Ikwere *m*, Ekpeye *N-*

Nupoid: Gade *n̄*

Idomoid: Idoma *m̄*, Igede *m̄-*

**\*mI**

Bantu: Basaa *mε*, Tsogo *mé*, Kele *eme*, Bila *ime*, Shambala *imi*, Kongo *-àmè*, Luvale *àmi*, Herero *àmì*. Zulu *mìná*

Other Bantoid: Samba-Daka *mèè*, Mambila *me*, Esimbi *me-*, Kenyang *mε*, Bamileke *mé*, Jarawa *mì / m̄*, Bwazza *m̄*

CR: Boki *me*, Kukele *m̄*, Bekwarra *àmì*, Lokaa *-mì*, Obolo *mi-*, Efik *m̄*, Ogbia *mí / m̄*, Eleme *-m̄*

CN: Berom *mé*, Horom *me*, Eggon *me*, Tarok *mi*, Lijili *me*, Hun-Saare *mε*, Etkywan *āmē*, Kuteb *me*

Edoid: Yekhee *me / mē*, Edo (*i*)*mε*, Epie *mε*, Engenni *mi / m̄*, Urhobo *mí / mé*

Defoid: Igala *mi*, Yoruba *mi*, Arigidi *ame*

Igboid: Igbo *àmì*, Ikwere *mé*, Ekpeye *mè*

Nupoid: Ebira *m̄ / m̄*, Gbagyi *m̄*

Idomoid: Eloyi *m̄*, Idomo *àmì*

Oko *àmé*

Akpes *òjì*

**\*o- / \*u-**

Bantu: Oroko *ò-*, Bafia *ù-*, Mpongwee *ò-*, Nkengo *u-*, Kumu *v-*, Digo *v-*, Sukuma *v-*, Shambala *u-*, Vili *ù-*, Shi *ù-*, Kwangari *o-*, Songe *o-*, Pimbwe *v-*, Matuumbi *v-*, Herero *ù-*, Venda *ù-*, Tswana *ò-*

Other Bantoid: Tikar *ù*, Tiw *ú*, Mbe *ò-*, Kenyang *ó-*, Limbum *à*, Yemba *ò*, Lamnso' *ã*

CR: Bekwarra *o-*, Kohumono *a-*, Mbembe *a-*, Efik *ò-*, Eleme *ɔ-*

Central Nigerian: Berom *hó-*, Tarok *u*, Lijili *ɔ*, Hun-Saare *o*, Wapan *ù*, Kuteb *u*

Edoid: Edo *u*, Uvbie *ù*

Defoid: Yoruba *o*

Nupoid: Ebira *v*

Idomoid: Igede *à-*, Idoma *à*

Oko *ù-* (perf.)

**\*wV**

Bantu: Oroko *owa*, Tsogo *èwè*, Babole *àwé*, Lega *ugwe*, Rombo *-afo*, Nilamba *ve*, Shambala *iwe*, Ha *we(we)*, Kwangari *ove*, Nyakyusa-Ngonde *ugwe*, Matuumbi *weé-nga*, Herero *òvè*, Venda *ìwè*, Copi *awe*  
 Other Bantoid: Mambila *wɔ*, Tikar *wǔ*, Esimbi *wo-*, Noni *wo*, Mbe *-wê*, Kenyang *wɔ̀*, Limbum *wè*, Yemba *wù*, Mundani *wèè*, Mbula *wé*  
 CR: Bekwarra *wō*, Kohumono *àgwɔ̀*, Kukele *wě*, Lokaa *àwú*, Obolo *òwù*  
 CN: Izere *wán*, Fyam *wéé*, Horom *awɔ̀*, Bu *awu*, Tarok *wa*, Lijili *ìwɔ̀*, Hun-Saare *wɔ̀*, Jibu *wu*, Bete *owu*  
 Edoid: Edo *wé*, Okpamheri *wε*, Degema *wó*, Urhobo *wε*, Okpe *wù / wò*  
 Defoid: Igala *wε*, Yoruba *ìwɔ̀*  
 Nupoid: Ebira *wu / wv*, Nupe-Tako *wó*  
 Idomoid: Idoma *àwɔ̀*  
 Oko *-wu / -wɔ̀*

**\*tu- / \*ti-**

Bantu: Bafia *tì-*, Viya *tù-*, Babole *tò-*, Nande *tu-*, Kuria *tu-/to-*, Sukuma *tv-*, Swahili *tu-*, Vili *tù-*, Rwanda *tu-*, Dciriku *tu-*, Lunda *tu-*, Pimbwe *tv-*, Sena *ti-*, Matuumbi *tv-*, Umbundu *tù-*  
 Other Bantoid: Meta *t̥*, Bamum *-út̥* (incl.)  
 CR: Kohumono *àd̥*  
 CN: Berom *wo-t*, Fyam *tí*, Kanufi *tot*, Mada *tə*, Che *tút*, Hun-Saare *tε* (excl.), Laru *ti*, Pongu *gáátù*, Cicipu *tù*, Kuteb *tī*  
 Oko *tì-* (perf.)

**\*(bV)ce**

Bantu: Oroko *isε*, Manenguba *-àcí*, Basaa *βés*, Pinji *àsé*, Mboshi *bísí*, Lega *bíswé*, Taita *ísì*, Nyamwezi *iswe*, Gogo *ase*, Haya *icwe*, Luyana *aci*, Nyakyusa-Ngonde *vswe*, Ndonga *tse*  
 Other Bantoid: Tiv *sé*, Noni *bese*, Kenyang *bèsé*, Aghem *sè* (incl.), Duguri *sú*  
 CR: Kukele *bešě*, Obolo *èzi*  
 CN: C'Lela *cín* (dual)  
 Akpes *àbès*

**\*mV-**

Bantu: Bulu *mi-*, Mituku *mu-*, Kamba *mv-*, Langi *mv-*, Zalamo *mu-*, Haya *mu-*, Luyana *mu-*, Kwezo *mú-*, Lungu *mu-*, Tumbuka *mu-*, Matuumbi *mv-*, Herero *mù-*, Zulu *mu-*  
 Other Bantoid: Dong *mε-re*, Meta *mbú*  
 CN: Tesu *mɔ̀wɔ̀*, Horom *min*, Yeskwa *mí* (past),  
 Edoid: Degema *ma- / mɔ̀-*. Uvbie *mámá*  
 Defoid: Igala *mε*, Arigidi *ámô*  
 Ukaan *mà / mɔ̀*

**\*(bV)nV**

Bantu: Nen *nú*, Viya *-àno*, Kele *enú*, Lengola *-anú*, Bila *βénú*, Rombo *ɲwé*, Sukuma *βiɲwe*, Gogo *abe*, Vili *bèno*, Haya *ɲwε*, Luvale *ènu*, Luba *nu-* (subject), Herero *ènè*, Copi *anu*  
 Other Bantoid: Mambila *ben*, Mesaka *bèn*, Noni *ben*, Mbe *ènǒ*, Bamum *-ñn*, Bamileke *bíné*, Lamnso' *ven*, Bwazza *wún*  
 CR: Bekwarra *ínèn*, Kohumono *ànɔ̀*, Mbembe *boɲa*, Obolo *èɲì*, Abua *ɲina*  
 CN: Ayu *bìnìn*, Izere *ɲín*, Irigwe *ɲí*, Kulu *anunɲ*, Ake *ánù*, Toro *anɔ̀wɔ̀*, Hun-Saare *nɔ̀*, C'Lela *nwá*, Wapan *ánì*, Kuteb *anī*  
 Edoid: Urhobo *βán*  
 Igbooid: Igbo *unù*, Ikwere *ánv*, Ekpeye *yónì*  
 Idomoid: Idoma *ànɔ̀*, Igede *ánv*  
 Oko *ànɔ̀*  
 Akpes *àbèɲ*

### Abbreviations

C	central	fut.	future	perf.	perfective
comp.	completive	imp.	imperative	pl.	plural
cond.	conditional	incl.	inclusive	pos.	positive
cons.	consonant	indir.	indirect speech	pres.	present
dial.	dialectal	L	Lower	progr.	progressive
dir.	direct speech	N	north(ern)	S	south(ern)
E	east(ern)	neg.	negative	sg.	singular
emph.	emphatic	opt.	optative	U	Upper
excl.	exclusive	p.c.	personal communication	W	west(ern)

### Literature

- ABIODUN, Michael. 1984. *Phonology of Arigidi*. B.A. Essay. Ilorin: University of Ilorin.
- ADEKANYE, Fred & Sophie SALFFNER (eds.). 2007. *Ikann Proverbs, Riddles and a Story*. Ikakumo-Aworo.
- AGOYI, Taiwo Odeyemi. 1997. *The category of number and the genetic classification of Èkìròmì*. Seminar paper, University of Ilorin.
- AKINKUGBE, Femi. 1976. An Internal Classification of the Yoruboid Group (Yoruba, Isekiri, Igala). In: *JWAL*, 11.1–2, pp. 1–19.
- AKINKUGBE, Femi. 1978. *A Comparative Phonology of Yoruba Dialects, Isekiri, and Igala*. Ph.D. Thesis, University of Ibadan.
- ALEX, Inoma. 1989. *A Reconstruction of the Segmental Phonology of Proto-Central Delta*. M.A. Thesis. Port Harcourt: University of Port Harcourt.
- AMAYO, Airen. 1980 [1975]. Edo. In: KROPP DAKUBU, M.E. (ed.) *West African Language Data Sheets*. Vol. 2. West African Linguistic Society.
- ARMSTRONG, Robert. 1981. The Idomoid language sub-family of the Eastern Kwa borderland. In: JUNGRAITHMAYR, H. (ed.), *Berliner Afrikanische Vorträge*. Berlin: Reimer. Pp. 7–33.
- ARMSTRONG, Robert. 1983. The Idomoid languages of the Benue and CR valleys. In: *JWAL*, 13.1, pp. 91–149.
- ARMSTRONG, Robert. 1989. Idomoid. In: *The Niger-Congo languages*, ed. by John BENDOR-SAMUEL, pp. 323–336. Lanham — New York — London: University Press of America.
- ATÓYÈBÍ, Joseph Dele. 2008. Morphophonemic Determination of Person Forms in Òko. In: *Newcastle Working Papers in Linguistics*, University of Newcastle. Vol. 14: Pp. 23–38.
- BABAEV, Kirill. 2008. Reconstructing Niger-Congo Personal Pronouns I: Proto-Bantoid. In: *Journal of West African Linguistics*. Vol. XXXV: 1–2. Pp. 131–183. [<http://www.nostratic.ru/benue-congo1.pdf>]
- BAMBOSE, Ayo. 1966. A Grammar of Yoruba. In: *West African Languages Monographs*, 5. Cambridge University Press.
- BARNWELL, Katharine. 1980. Mbembe. In: KROPP DAKUBU, M.E. (ed.) *West African Language Data Sheets*. Vol. 2. West African Linguistic Society.
- BENNETT, P. & J. STERK. 1977. South-Central Niger-Congo: a Reclassification. In: *Studies in African Linguistics* 8, 1977: pp. 241–273.
- BERGMAN, Richard. 1976 [1973]. Igede. In: KROPP DAKUBU, M.E. (ed.) *West African Language Data Sheets*. Vol. 1. West African Linguistic Society.
- BLENCH, Roger. 1989. Nupoid. In: *The Niger-Congo languages*, ed. by John BENDOR-SAMUEL, pp. 305–322. Lanham — New York — London: University Press of America.
- BLENCH, Roger. 2005a. *A Dictionary of Ogbà, an Igboid Language of Southern Nigeria*. Manuscript [<http://www.rogerblench.info>].
- BLENCH, Roger. 2005b [1984]. The Ukaan Language: Bantu in South Western Nigeria? Revised paper presented at the Hamburg Conference ‘Trends in the Historical Study of African Languages’, Hamburg, 3–7 Sep 1984.
- BLENCH, Roger. 2006a. *A Dictionary of Ekpeye, an Igboid Language of Southern Nigeria*. Manuscript [<http://www.rogerblench.info>].
- BLENCH, Roger. 2006b. *A Wordlist of the Hyam Language of Central Nigeria and Its Affinities*. Manuscript [<http://www.rogerblench.info>].

- BLENCH, Roger. 2006c. *Mada Pronouns*. Manuscript [<http://www.rogerblench.info>].
- BLENCH, Roger. 2006d. *The Anib (Kanufi) Language of Central Nigeria and Its Affinities*. Manuscript [<http://www.rogerblench.info>].
- BLENCH, Roger. 2006e. *The Ayu Language of Central Nigeria and Its Affinities*. Manuscript [<http://www.rogerblench.info>].
- BLENCH, Roger. 2006f. *The Bu Language of Central Nigeria and Its Affinities*. Manuscript [<http://www.rogerblench.info>].
- BLENCH, Roger. 2006g. *The Cara Language of Central Nigeria and Its Affinities*. Manuscript [<http://www.rogerblench.info>].
- BLENCH, Roger. 2006h. *The Horom Language of Central Nigeria and Its Affinities*. Manuscript [<http://www.rogerblench.info>].
- BLENCH, Roger. 2006i. *The Sur (Tapshin) Language of Central Nigeria and Its Affinities*. Manuscript [<http://www.rogerblench.info>].
- BLENCH, Roger. 2006j. *The Toro Language of Central Nigeria and Its Affinities*. Manuscript [<http://www.rogerblench.info>].
- BLENCH, Roger. 2007a. *The Ake Language of Central Nigeria and Its Affinities*. Manuscript [<http://www.rogerblench.info>].
- BLENCH, Roger. 2007b. *The Kway Language of Central Nigeria and Its Affinities*. Manuscript [<http://www.rogerblench.info>].
- BLENCH, Roger. 2007c. *A Sketch of Phonology and an Extended Wordlist of TsuVadi, a Kambari Language of Western Nigeria*. Manuscript [<http://www.rogerblench.info>].
- BLENCH, Roger. 2007d. *The Eloyi Language of Central Nigeria and Its Affinities*. Manuscript [<http://www.rogerblench.info>].
- BLENCH, Roger. 2008a. *The Idū Language of Central Nigeria. Phonology, Wordlist and Suggestions for Orthography Changes*. Manuscript [<http://www.rogerblench.info>].
- BLENCH, Roger. 2008b. *The Nyankpa (Yeskwa) Language of Central Nigeria*. Manuscript [<http://www.rogerblench.info>].
- BLENCH, Roger. (Manuscript a). *The West Kainji Language of Northwestern Nigeria: Comparative Wordlist*. [<http://www.rogerblench.info>].
- BLENCH, Roger. (Manuscript b). *The Laru People and Their Language*. [<http://www.rogerblench.info>].
- BLENCH, Roger. (Manuscript c). *A Dictionary of Echie, an Igboid Language of Southern Nigeria*. [<http://www.rogerblench.info>].
- BLENCH, Roger & Barau KATO. 2007. *The Tasu Language of Central Nigeria and Its Affinities*. Manuscript [<http://www.rogerblench.info>].
- BLENCH, Roger & Bitrus Bulus KAZE. 2006. *A Dictionary of the Izere Language of Fobur*. Manuscript [<http://www.rogerblench.info>].
- BLENCH, Roger & Musa DOMA. *A Dictionary of the Gbari Language*. Manuscript [<http://www.rogerblench.info>].
- BLENCH, Roger & Gya DANIEL. 2008. *Rigwe Pronouns*. Manuscript [<http://www.rogerblench.info>].
- BLENCH, Roger & al. 1991. *Dictionary of the Language of the Bas(s)a People*. Manuscript [<http://www.rogerblench.info>].
- BLENCH, Roger, Ruth ADIWU & Gideon ASUKUTU. 2006. *The Ce (Rukuba) Language of Central Nigeria and Its Affinities*. Manuscript [<http://www.rogerblench.info>].
- BOND, Oliver. 2003. *Personal Inflection in Eleme*. Manuscript.
- BOND, Oliver. 2006. *Selected Proceedings of the 35th Annual Conference on African Linguistics*, ed. John Mugane et al., 234–244. Somerville: Cascadilla Proceedings Project.
- BOND, Oliver & Gregory ANDERSON. 2006. Divergent Structure in Ogonoid Languages. In: COVER, Rebecca T. & KIM, Yuni, (eds.), *Proceedings of 31st Annual Meeting of the Berkeley Linguistic Society: Special Session on Languages of West Africa*. Berkeley: Berkeley Linguistics Society.
- BOUQUIAUX, Luc. 1970. *La langue birom*. Paris: Les belles lettres.
- CAPO, Hounkpati. 1989. Defoid. In: *The Niger-Congo languages*, ed. by John BENDOR-SAMUEL, pp. 275–290. Lanham — New York — London: University Press of America.
- CASSETTA, Pete & Lydia CASSETTA. 1994. *The Ugare Verb System*. Manuscript.
- CLARK, David. 1969. *A grammatical study of Ekpeye*. Ph.D. Department of Linguistics and Phonetics, SOAS.

- CLARK, David. 1972. A four-term person system and its ramifications. *Studies in African Linguistics*, 3: 97–105.
- CLARK, David. Ekpeye. 1976. Izi. In: KROPP DAKUBU, M.E. (ed.) *West African Language Data Sheets*. Vol. 1. West African Linguistic Society.
- CONNELL, Bruce. 1988. *Lower Cross Reconstructions*. Manuscript.
- CONNELL, Bruce. 1991. *Phonetic aspects of the Lower Cross Languages and their implications for sound change*. Ph.D. Linguistics. Edinburgh: University of Edinburgh.
- CONNELL, Bruce. 1995. The Lower Cross languages: a prolegomena to the classification of the CR languages. *Journal of West African Languages*, XXIV 1: pp. 3–46.
- CONNELL, Bruce. *Lower Cross Wordlists*. Manuscript.
- COOK, T.L. 1976 [1968]. Efik. In: KROPP DAKUBU, M.E. (ed.) *West African Language Data Sheets*. Vol. 1. West African Linguistic Society.
- COOK, T.L. 1971. Report of the Benue-Congo Working Group of the West African Linguistic Society. *Actes du 8e Congrès de la Société linguistique de l'Afrique Occidentale*. Annales de l'Université d'Abidjan, Serie H, 2: 559–574, Abidjan: Université d'Abidjan.
- COOK, T.L. 1976 [1972]. Kohumono. In: KROPP DAKUBU, M.E. (ed.) *West African Language Data Sheets*. Vol. 1. West African Linguistic Society.
- CREISSELS, Denis. 2000. Typology. In: HEINE, Bernd and NURSE, Derek (eds). *African Languages – An Introduction*. Cambridge: Cambridge University press, pp. 231–258.
- CRESSMANN, Esther & Donna SKITCH. 1980 [1974]. Duka. In: KROPP DAKUBU, M.E. (ed.) *West African Language Data Sheets*. Vol. 2. West African Linguistic Society.
- CROZIER, D. & Roger BLENCH. 1992. *Index of Nigerian Languages* (edition 2). SIL, Dallas.
- DELAFOSSÉ, Maurice. 1904. *Vocabulaires comparatifs de plus de 60 langues ou dialectes parlés à la Côte d'Ivoire et dans la régions limitrophes*. Paris: E.Leroux.
- DE WOLF, Paul. 1968. Problems in Reconstructing the Noun Class System of the Languages of Plateau 1-a and 1-b. In: *Benue-Congo Working Group of the West African Linguistic Society*. Leiden Conference Draft.
- DIMMENDAAL, Gerrit. 1978. *The Consonant of Proto-Upper Cross and Their Implications for the Classification of Upper Cross languages*. Leiden: Department of African Linguistics.
- ELIMELECH, Baruch. 1976. *A Tonal Grammar of Etsako*. WPP, no. 35. Los Angeles: University of California.
- ELUGBE, Ben Ohi. 1979. Some Tentative Historical Inferences from Comparative Edoid Studies. In: *Kiabàrà*, vol. 2.1, pp. 82–101.
- ELUGBE, Ben Ohi. 1980. Reconstructing the Lenis Feature in Proto-Edoid. In: *JALL*, vol. 2, pp. 39–67.
- ELUGBE, Ben Ohi. 1983. Noun Class Prefixes in Proto-Edoid. In: *PALL*, vol. 1, pp. 59–75.
- ELUGBE, Ben Ohi. 1984. Morphology of the Gerund in Degema and Its Reconstruction in Proto-Edoid. In: *SAL*, vol. 15, pp. 77–90.
- ELUGBE, Ben Ohi. 1989. Edoid. In: *The Niger-Congo languages*, ed. by John BENDOR-SAMUEL, pp. 291–304. Lanham — New York — London: University Press of America.
- ELUGBE, Ben Ohi. 2001. *The Classification of Akpes and Aika*. Delivered at CALL 31, Leiden.
- ELUGBE, Ben Ohi & OLOGORI of Ogori. 1980. Ogori. In: KROPP DAKUBU, M.E. (ed.) *West African Language Data Sheets*. Vol. 2. West African Linguistic Society.
- EMUEKPERE MASAGBOR, G.A. 1997. *Preverbal Subject Markers in Ivoie*. Ph.D. Thesis. University of Sherbrooke.
- ESSIEN O.E. 1987. The Aspectual System of Ibibio. In: ODDEN D. (ed.), *Current Approaches to African Linguistics*. Vol. 4. Pp. 151–165.
- EVENHOUSE, W. *Jukun Wukari: a Brief Description in English*. Manuscript.
- FARACLAS, Nicholas. 1984. *A Grammar of Obolo*. Bloomington: Indiana University Linguistics Club.
- FARACLAS, Nicholas. 1989. CR. In: *The Niger-Congo languages*, ed. by John BENDOR-SAMUEL, pp. 377–400. Lanham — New York — London: University Press of America.
- FRESCO, E. 1968. Two Dialects of Igala and Yoruba: Some Comparisons. In: *Research Notes*, 4, pp. 32–46. Ibadan: University of Ibadan.
- GARDNER, Ian. 1976 [1966/67]. Abua. In: KROPP DAKUBU, M.E. (ed.) *West African Language Data Sheets*. Vol. 1. West African Linguistic Society.
- GERHARDT, Ludwig & Heinz JOCKERS. 1981. Leksikostatistischen Klassifikationen von Plateausprachen. In: *Berliner Afrikanische Vorträge (Akten des 21 DOT)*, pp. 25–54.
- GERHARDT, Ludwig. 1983. *Beträge zur Kenntnis der Sprachen des Nigerianischen Plateaus*. Glückstadt: J. J. Augustin.



- GERHARDT, Ludwig. 1989. Kainji and Platoid. In: *The Niger-Congo languages*, ed. by John BENDOR-SAMUEL, pp. 359–376. Lanham — New York — London: University Press of America.
- GREENBERG, Joseph. 1963. *The Languages of Africa*. The Hague: Mouton; IJAL 29.1, Bloomington: Indiana University Press.
- HANSFORD, Keir, John BENDOR-SAMUEL & Ron STANFORD (eds.). 1976. *An Index of Nigerian Languages*. Ghana: SIL.
- HOFFMANN, Carl. 1976. The Languages of Nigeria by Language Family. In: K. HANSFORD et al. (eds.), *An Index of Nigerian Languages*. Ghana SIL. Pp. 169–190.
- HYMAN, Larry & Daniel MAGAJI. 1970. *Essentials of Gwari Grammar*. Occasional Publication, 27. IAS: University of Ibadan.
- HYMAN, Larry & Imelda UDOH. 2002. Length harmony in Leggbó: A counter-universal? In: *Workshop on the Phonology of African Languages (WOPAL)*, Vienna.
- HYMAN, Larry, Narrog HEIKO, Mary PASTER & Imelda UDOH. 2002. Legbo Verb Inflection: A Semantic and Phonological Particle Analysis. In: *Proceedings of 28th Annual Meeting of the Berkeley Linguistics Society*.
- IKORO, Suana M. 1989. *Segmental phonology and lexicon of Proto-Keggoid*. University of Port Harcourt: M.A. thesis.
- IKORO, Suana. 1996. *The Kana Language*. Leiden: Research School CNWS.
- ISUKUL, Caroline. *A Grammar of Agholo (Ogbia)*. Manuscript [<http://www.rogerblench.info>].
- IWARA, Alexander. 1990. A phonemic analysis of long vocoids and phonetic diphthongs in Lokaa. *African notes: bulletin of the Institute of African Studies*. Ibadan: University of Ibadan, pp. 74–79.
- IWARA, Alexander. 2002. The Tonal Phonology and Phonetics in the Future Negative in Lokaa. In: MUGAN, J. (ed.), *The Linguistic Typology and Representation of African Languages*, vol. 5. Africa World Press. Pp. 103–116.
- JUNGRAITHMAYR, H. 1973. Notes on the Ishe dialect of Ukaan (Akoko, Western State, Nigeria). *Africana Marburgensia*, 6,1:39–57.
- KARI, Ethelbert. 2002. Distinguishing between Clitics and Words in Degema, Nigeria. In: *African Study Monographs*, 23(4), pp. 177–192.
- KARI, Ethelbert. 2003. Clitic Doubling: Insights from Degema. In: *Journal of Asian and African Studies*, no. 66, pp. 235–248.
- KARI, Ethelbert. 2004. *A Reference Grammar of Degema*. Köln: Rüdiger Köppe.
- KARI, Ethelbert. 2007. Vowel Harmony in Degema, Nigeria. In: *African Study Monographs*, 28(2), pp. 87–97.
- KOELLE, Sigismund. 1854. *Polyglotta Africana*. London: Church Missionary House.
- KOOPS, Robert. 2007. *A Grammar of Kuteb, a Jukunoid Language of East-Central Nigeria*. Köln: Rüdiger Köppe. In Press.
- KROPP DAKUBU, Mary Esther. *West African Language Data Sheets*. Vols 1–2. West African Linguistic Society. 1976–1980.
- LEWIS, M. Paul (ed.). 2009. *Ethnologue: Languages of the World*, 16th edition. Dallas: SIL International [<http://www.ethnologue.com>].
- LONGTAU, Selbut. 2008. *The Tarok Language: Its Basic Principles and Grammar*. Dart-Jos: KWEF.
- LONGTAU, Selbut & Roger BLENCH. *Tarok Dictionary and English-Tarok Finderlist*. Forthcoming.
- MACKAY, V. 1976 [1968]. Eloyi. In: KROPP DAKUBU, M.E. (ed.) *West African Language Data Sheets*. Vol. 1. West African Linguistic Society.
- MADUGU, Isaac. 1975. *A grammar of Kwa-type verb serialisation; its nature and significance for current generative theory*. Ph.D. thesis. Los Angeles: UCLA.
- MADUGU, Isaac. 1979. The Nupe verb and diachrony. In: *Ibadan Journal of Humanistic Studies*, 1, pp. 74–97.
- MADUGU, Isaac. 1985. Complex verbs in Nupe and Yoruba. *SAL*, 16:3, pp. 295–321.
- MANFREDI, Victor. 1989. Igboid. In: *The Niger-Congo languages*, ed. by John BENDOR-SAMUEL, pp. 337–358. Lanham — New York — London: University Press of America.
- MCGILL, Stuart. *A Cicipu / English / Hausa Lexicon*. [<http://www.cicipu.org/dictionary.html>]
- MEEUSSEN, Achille. 1965. *Reconstructions grammaticales du bantou*. Tervuren: Musée Royal de l’Afrique Centrale (MRAC).
- MEIER, Paul. 1976. Izi. In: KROPP DAKUBU, M.E. (ed.) *West African Language Data Sheets*. Vol. 1. West African Linguistic Society.
- NETTLE, Daniel. 1998. Materials from the South-Eastern Plateau Languages of Nigeria. In: *Afrika und Übersee*, Band 81. Pp. 253–279.
- OHIRI-ANICHE, Chinyere. 1991. *A reconstruction of Proto-Igboid-Yoruboid-Edoid*. Ph.D. Thesis. Port Harcourt: University of Port Harcourt.

- OHIRI-ANICHE, Chinyere. 1999. Language Diversification in the Akoko Area of Western Nigeria. In: Roger BLENCH, Matthew SPRIGGS (eds.). *Archaeology and Language IV: Language Change and Cultural Transformation*, pp. 79–94. Routledge.
- OMAMOR, Augusta. 1976. The Place of Isekiri in Relation to Yoruba. In: *Research Notes*, vol. 7, nos. 1–2. Ibadan: University of Ibadan.
- OMAMOR, Augusta. 1988. Okpe and Uvwie: a Case of Vowel Harmony Galore. In: *Journal of West African Languages*, 18.1, pp. 47–64.
- OYAMA, Ebinda, & Katharine BARNWELL. 1985/1995. *Mbembe-English Dictionary*. Manuscript.
- PHILPOT, W. T. A. 1935. Notes on the Igala Language. In: *Bulletin of the School of Oriental and African Studies*, 7, pp. 897–912.
- PRESTON, B. 1854. *A Grammar of the Bakële Language, with Vocabularies*. New York: J. P. Prall.
- PRIEST, Perry. *Jibu Dictionary*, ed. by Roger Blench. Manuscript. [<http://www.rogerblench.info>]
- RIKOTO, B., P. SENCHI, Steve DETTWEILER & Sonia DETTWEILER. 2002. *K'Batksa: C'Lela—C'Turanci—C'Gana Dictionary*. Lelna Language Development / Translation Association.
- ROWLAND Oke, Mary. 2003. *Description systématique de la langue obolo-andoni*. Paris: L'Harmattan.
- SCHOLZ, Hans-Jürgen. 1976 [1973]. Igbirra. In: KROPP DAKUBU, M.E. (ed.) *West African Language Data Sheets*. Vol. 1. West African Linguistic Society.
- SEGERER, Guillaume. 2002–2007. *Les marques personnelles dans les langues africaines*, [base de données en ligne]. <http://sumale.vjf.cnrs.fr/pronoms>.
- SEGERER, Guillaume. 2008. Les langues Niger-Congo. In: ЖЕЛТОВ А. Ю. [ZHELTOV, A. Y.] 2008. *Языки нигер-конго: структурно-динамическая типология*. СПб: Издательство Санкт-Петербургского университета. [*The Niger-Congo Languages: Structural and Dynamic Typology*. Saint-Petersburg: Saint-Petersburg University.]
- SHIMIZU, Kiyoshi. 1971. *Comparative Jukunoid: an Introduction*. Ph.D. thesis. Ibadan: University of Ibadan.
- SHIMIZU, Kiyoshi. 1971–1980. *Comparative Jukunoid*, vols 1–3. Vienna: Afro-Pub.
- SHIMIZU, Kiyoshi. 1975. A Lexicostatistical Study of Plateau Languages and Jukun. *Anthropological Linguistics*, 17, pp. 413–418.
- SHIMIZU, Kiyoshi. 1996. *A Kulu vocabulary and fragments of Kulu grammatical structures*. AFLANG [[http://www.aa.tufs.ac.jp/~P\\_aflang/](http://www.aa.tufs.ac.jp/~P_aflang/)].
- SIBOMANA, Leo. 1981/82. Tarok III: Das Verbalsystem und der Satz. In: *Afrika und Übersee*, Band 64, 2, pp. 237–247.
- SIBOMANA, Leo. 1985. A Phonological and Grammatical Outline of Eggon. In: *Afrika und Übersee*, Band 68, pp. 43–68.
- SIEWIERSKA, Anna. 2004. *Person*. Cambridge: Cambridge University Press.
- SMITH, Neil. 1980 [1967]. Nupe. In: KROPP DAKUBU, M.E. (ed.) *West African Language Data Sheets*. Vol. 2. West African Linguistic Society.
- STANFORD, B. 1976 [1973]. Bekwarra. In: KROPP DAKUBU, M.E. (ed.) *West African Language Data Sheets*. Vol. 1. West African Linguistic Society.
- STERK, Jan Pieter. 1977. *Elements of Gade Grammar*. Doctoral dissertation, University of Wisconsin, Madison.
- STERK, Jan Pieter. 1979. Fortis / Lenis Contrast in Upper Cross Consonants: a Survey. In: *Kiabàrà*: 2.2: 57–68.
- STOFBERG, Yvonne R. 1978. Migili grammar. In: *Language Data, African Series*, 12. Dallas: Summer Institute of Linguistics.
- STORCH, Anne. 2005. Jukun Pronominal Systems from an Areal Perspective. Paper presented at the Jukun Conference, Vienna, 18–20 Nov 2005.
- STRUB, E. 1915–1916. Essai d'une grammaire de la langue kukuruku. In: *Anthropos*, vol. 10, pp. 450–465; vol. 11, pp. 888–907.
- THOMAS, Elaine. 1976 [1968]. Engenni. In: KROPP DAKUBU, M.E. (ed.) *West African Language Data Sheets*. Vol. 1. West African Linguistic Society.
- THOMAS, Elaine & Kay WILLIAMSON. Wordlists of Delta Edo: Epie, Engenni, Degema. In: *Occasional Papers* 8, p. 105. Accra: Institute of African Studies, University of Ibadan.
- UKERE, Anthony. 2005 [1986]. *Urhobo-English Dictionary*. Benin City: Elupeju Press.
- WELMERS, William. 1949. *Speaking Jukun (Takum Dialect)*. Lupwe.
- WESTERMANN, Diedrich. 1926. Das Edo in Nigerien. In: *MSOS*, vol. 29.3, pp. 32–58.
- WESTERMANN, Diedrich, & Margaret BRYAN. 1952. *The Languages of West Africa. Handbook of African Languages* 2. London: Oxford University Press.
- WILLIAMSON, Kay. 1971. The Benue-Congo Languages and Ijo. *CTL*: 7, pp. 245–306.





## The Swadesh wordlist. An attempt at semantic specification<sup>1</sup>

The paper is a thematical follow-up to the refinements of the lexicostatistical method suggested in [Starostin G. 2010]. It discusses the issue of synonymy/polysemy, a well-known obstacle in the compilation of Swadesh wordlists for various languages, and presents a list of both syntactic/semantic contexts and explanatory notes that could help reduce the ambiguity issue in the creation and quantitative analysis of such wordlists. The notes and contexts are partially based on linguistic tradition and partially on theoretical and/or pragmatic considerations, some of which are stated explicitly.

*Keywords:* lexicostatistics, glottochronology, comparative method, language relationship, semantic reconstruction, Swadesh wordlist.

In the middle of the 20<sup>th</sup> century the American linguist Morris Swadesh came up with a wordlist of 200 items that, according chiefly to his own linguistic intuition and experience, represented the core basic vocabulary of the human language *par excellence* (Swadesh 1952 w. prev. lit.). Still later, for the sake of convenience, the same author reduced this list to the 100 stablest items (Swadesh 1955).

Despite the fact that, in the rarest of cases, some languages have been noted to lack the equivalents for one or two items even on the shortened list (at least as inherited words) due to the absence of corresponding *realia*,<sup>2</sup> over the years the 100-wordlist has proven its worth as a useful linguistic tool, and is now widely applied in various lexicostatistical and glottochronological studies.<sup>3</sup>

One important feature that has made the Swadesh wordlist so popular among linguists is its relative simplicity and transparency. However, that same simplicity — each item is defined by a single English word — in many particular situations comes across as a significant flaw, since the words have been originally listed without any specific comments or explanations. Because Modern English is a natural language, not a meta-language of semantic primitives, such laconicity may cause plenty of problems during the construction of the lists as well as critical faults in their procedural analysis.

---

<sup>1</sup> This article is, to a large extent, based on the collective results of several sessions of the Nostratic Seminar (Center for Comparative Linguistics of the Institute of Oriental Cultures and Antiquity, Russian State University for the Humanities, 27.04.06–01.06.06). The authors express their gratitude to Eugenia Chekmeneva, Albert Davletshin, Vladimir Dybo, Alexandra Evdokimova, Yuri Koryakov, Tatiana Mikhailova, Oleg Mudrak, Elena Parina, Mikhail Zhivlov and other participants who came up with important suggestions that have been incorporated into the paper.

<sup>2</sup> E.g., ‘horn’ for languages of Oceania or ‘(domesticated) dog’ for the greater part of pre-Neolithic cultures.

<sup>3</sup> Although the amount of such studies, carried out on the basis of lexica from languages all over the world, is immense, to the best of our knowledge, there is no single overview of the various applications of the Swadesh wordlist that have been published since its creation. [Renfrew, McMahon & Trask 2000], with its ample bibliographical references, may be a good starting point for those wishing to estimate the true influence of Swadesh on statistical methods in historical linguistics.

Thus, in compiling a Swadesh wordlist for Latin, we will already have trouble with filling the very first slot: English ‘all’ corresponds both to Latin *omnis* (= ‘each one’, with plural objects) and *totus* (= ‘entire’, with a singular object), and there are no hints on which option to choose. Likewise, the single English word ‘round’ corresponds to no less than two different words in some Turkic languages: ‘2D round’ (as in *a round dish*) and ‘3D round’ (as in *a round stone*). Examples can be multiplied; for a good selection of similar cases, see [Campbell 1998: 181]. The problem is well recognized among linguists and makes compilation and analysis of the wordlists a very confusing occupation, especially for “beginners”.

Sometimes the ambiguous selections that researchers undertake at their personal initiative violate the very purpose of the 100-wordlist. For instance, in a vast majority of particular studies the slot ‘breast’ is filled in with the equivalent for ‘female breast(s)’; however, the meaning ‘female breast(s)’ in the world’s languages is very frequently expressed by various baby-talk formations (*titi, nana, mama*, etc.) that can hardly serve as reliable *comparanda* in one’s investigation of language relationship. (Strictly speaking, relegation of words like ‘father’ and ‘mother’ to the second half of the 200-wordlist would automatically imply the relegation of ‘female breast(s)’ to the same group). In African linguistics, there is a frequent tendency to replace the verb ‘to fly’ with the noun ‘fly’ (insect) — probably because the verb ‘to fly’ is, in that particular area, quite prone to massive areal borrowing. This, however, is very unpractical, since it makes the African wordlists incompatible with wordlists collected by other researchers in other areas (at least, for that particular item).

The purpose of the current paper is, therefore, to attempt to define the semantics that underlies the various items on the Swadesh list with better precision and clarity than it is generally understood by simply glancing at the “bare” wordlist as published in Swadesh’s original works and later studies. We do not pretend to completeness of the offered solutions. A definitive treatment of this problem should eventually include formal interpretations of the items that make proper use of a semantic metalanguage, either an existing one or one specially designed for such a purpose. Another significant issue is that of *justification* of the wordlist: as far as we know, neither Swadesh, nor any other linguist has ever offered an empirically or logically grounded explanation of why we are supposed to take *this* particular group of words and not any other to represent the generally most stable layer of the lexicon in languages worldwide. Such a justification should be conducted on a wide empirical basis and combine results of massive statistical calculations with the analysis of the formal interpretations mentioned above. Our goals are far more modest: to offer easily understandable and applicable specifications of the “Swadesh meanings” that

- reflect and respect as much of the “tradition” of collecting and compiling Swadesh wordlists as possible;
- do not violate either common sense or the criterion of (near-)universality;
- facilitate the job of the compiler working with live informants, massive textual corpora, or detailed dictionaries.

It should be stressed that, in a few cases at least, our choice of specific meaning *may* differ from the original meaning implied by M. Swadesh himself, but such differences would be difficult to prove, and would hardly have any crucial importance altogether.

Some of the decisions presented and described below are entry-specific, but there are also several principal methodological guidelines that we try to apply consistently. These are as follows:

1) In all cases, as is prescribed by the standard procedure, we advocate the choice of the most stylistically *basic*, *neutral*, and *unmarked* word, uncumbered by either pragmatic connotations or strongly emphasized additional, “extra” semantic features. Thus, *belly* is used instead of *abdomen* (“scientifically colored” talk), *paunch* (referring to a specific type of belly, often “emotionally marked”) or *tummy* (colloquial or “baby talk”) for the Modern English 100-wordlist.

2) The word also has to be encountered in *unbound* use, not constrained exclusively to specific idiomatic constructions or compound forms. E.g., for Modern Japanese in the meaning ‘sun’ we have to choose the Chinese borrowing *taiyo*: rather than the original form *hi*, since the latter only truly occurs in bound form in the modern language.

3) A frequent source of synonymity and, therefore, confusion concerning the Swadesh wordlist is the phenomenon of *suppletion* within a paradigm. Because proper grammatical suppletion is sometimes hard to distinguish from lexical variation (e.g., the use of special collective forms of nouns; aspectual forms of verbs, etc.), it is strongly recommended to avoid including suppletive stems as synonyms altogether.<sup>4</sup> One should attempt to determine the most unmarked, morphologically unencumbered stem, which typically corresponds to such morphological values as singular number (for nouns), positive degree of comparison (for adjectives), 3<sup>rd</sup> person / present tense / singular subject or object of action (for verbs); however, these recommendations are not to be taken as categorical, because some languages impose their own standards of markedness (for instance, treat the plural meaning in nouns and the past tense in verbs as the “simplest” form rather than the singular meaning and the present tense).

4) The list is generally anthropocentric, meaning that most of the anatomic terms and action verbs (such as ‘sit’, ‘lie’, ‘eat’, etc.) are to be taken as human body parts and actions; e.g., the wordlist for Modern Russian should contain *pom* ‘mouth (of a human)’, but not *nacmь* ‘mouth (of an animal)’, and the wordlist for German should contain *essen* ‘to eat (of a human)’, but not *fressen* ‘to eat (of an animal)’. Accordingly, the meaning ‘claw’ on the list is replaced by ‘(human) fingernail’ (as is very much the accepted practice among wordlist compilers anyway). Understandable exceptions include the words for ‘feather’, ‘horn’, ‘tail’, ‘to fly’ and a few others. Obviously, in many cases there will be no differentiation whatsoever, simplifying the task of the compiler.

5) Many of the items on the Swadesh list form functional antonymous pairs, e.g., ‘big : small’, ‘black : white’, ‘man : woman’, or groups of two or three elements with tight semantic connections, e.g., ‘sun : moon’, ‘eat : drink’, ‘sit : stand : lie’ etc. It is very useful to seek out syntactic contexts that contrast these elements, since otherwise we increase the risk of a non-basic term creeping in. For instance, if we elicit (either from a live informant or from a large dictionary) several quasi-synonymous words for ‘to eat’, it is highly likely that the most neutral will be the one used in a context like ‘I need to eat and drink to stay alive’.

6) One of the cases where synonymity is unavoidable is when we deal with a *transitional* stage in language history, during which an older word is gradually being “ushered out” by a more recent replacement (e.g., the relation between ‘stone’ and ‘rock’ in American English). If one can put up a good argument for such a case, both words should be accepted for the language in question. This phenomenon is most easily noticeable in languages with a long written history (such as Egyptian or Chinese), but can be observed in living languages as well, where sometimes the difference is generation-related.

---

<sup>4</sup> One possible exception can be made for suppletion in the direct / indirect stems of personal pronouns (e.g., *I / me*, *we / us*). The reason for this is that such suppletion is exceptionally frequent in the world’s languages, which makes this kind of information very important for lexicostatistical purposes.

7) Synonymity must also be considered (but can sometimes be avoided) in dealing with the phenomenon of *compound* forms found on the list, consisting of two or more root morphemes; this is especially typical of “monosyllabic” languages that frequently recur to compounding in order to avoid excessive homonymy (e.g., Sino-Tibetan, Kra-dai, Abkhaz-Adyghe), but in small doses can be found almost anywhere. With binary comparison, this is never a problem, since the rule of thumb here is that the lexemes are to be scored as matches if at least one morpheme in each is cognate (thus, Abaza *lámħa* ‘ear’ ← \**lə+məħa* = Ubykh *La+kʷəma* id., even though the second morphemes in these words are different). Difficulties may, however, arise with groups of more than two languages, where one might encounter more complicated situations, e.g., Old Chinese 齒 \**thəʔ* ‘tooth’ ≠ Modern (Beijing) Chinese 牙 *yá* ‘id.’ (← Old Chinese \**ɦrāʔ* ‘tusk’), but what about Sichuan 牙齒 *ya-tsi* ‘tooth’, a collocation of both? It seems that the only reasonable solution here is to admit synonymity and double the word on the list, scoring it once with Old Chinese and then again with Beijing Chinese.<sup>5</sup>

Most of these criteria — not always, but more frequently than not — help to single out the most *statistically frequent* of the possible equivalents. The criterion of frequency of use cannot be construed as a strong demand, but it may be related to the idea of writing down the equivalent that “first springs to mind” when dealing with an informant, or the one that is listed before all others in a context-less dictionary (unless, of course, it lists all the synonyms in alphabetic order). Thus, if a language offers the choice of several different quasi-synonyms for the verb ‘to eat’, and it is not easy to place formal borders between the various shades of meanings, it always makes sense to look for the most frequently encountered item; this, among other things, should be the default method of working with material from ancient languages, represented only by written texts.

As a scientifically inferior, but pragmatically more helpful, alternative to formal semantic definitions we currently suggest the method of *diagnostic contexts* for the Swadesh items. Such contexts can be of great help during fieldwork with informants, since they help to verify the correctness of the lexical equivalents that are provided. We list two or three syntactic contexts that, in most cases, should guarantee the same equivalent; if the elicited equivalents differ from context to context (based on our own experience, such a situation is not likely to occur frequently), an additional check on the issue of semantic markedness is necessary.

If the wordlist is compiled based on secondary literature, such as a detailed dictionary or a corpus of texts, the diagnostic contexts that we offer cannot be used directly (unless, by pure chance, they happen to be reproduced in the dictionary or the corpus), but can still be made indirect use of, since they give a general idea of the kind of speech situations in which we expect the word in question to show up. (Obviously, if the wordlists are extracted from a relatively small dictionary that does not provide any syntactic contexts for its entries, the method cannot be used, and the results will, almost inevitably, be slightly distorted.)

The selected contexts, along with accompanying notes, are presented below in the form of a table, based on the standard 100-wordlist (as it is reproduced, e.g., in various publications by Sergei Starostin, see [Старостин 2007] both in Russian and English) and also featuring 10 additional words from Sergei Yakhontov’s 100-wordlist that are carried over from the second part of the 200-item Swadesh wordlist (see [Бурлак & Старостин 2005: 12–13] for details).

<sup>5</sup> Of course, if we are willing to take into consideration the factor of language history, the Sichuan binome 牙齒 *ya-tsi* is almost certainly a homonymy-reducing extension of the original 齒 \**thəʔ*, whereas the Beijing monosyllabic form 牙 *yá* is a proper lexical replacement, meaning that, strictly speaking, the Sichuan word should be scored together with Old Chinese, not with the Beijing form. In most cases, however, we are not that well familiar with the history of compared languages/dialects to be able to easily reach such decisions.



Yakhontov's additional entries are marked by the "+" sign and placed at the end of the list (No.101–110).

The table's structure is as follows. The two columns marked "Eng." and "Rus." contain the traditional one-word-definitions in English and Russian. The next column contains the proposed syntactic contexts both in English and Russian. In the last column we give additional comments on the entry.

In a few cases, the authors were unable to reach a single decision on which particular semantic equivalent should be taken as the default value, so several rows have been divided into two separate entries, e.g., No.107a 'thin (of flat objects)' vs. No.107b 'thin (of elongated objects)'; two distinct sub-entries have also been introduced for suppletive personal pronouns, e.g., No.45a 'I (direct stem)' vs. No.45b 'me, mine (indirect stem)'. We suggest that such quasi-duplicates should be treated as synonyms at the subsequent mathematical stage of lexicostatistic procedure.

### Literature

- Campbell 1998 — Lyle CAMPBELL. *Historical Linguistics. An introduction*. Edinburgh University Press.
- Renfrew, McMahon & Trask 2000 — C. RENFREW, A. MCMAHON & L. TRASK (eds.). *Time Depth in Historical Linguistics*. Cambridge: The McDonald Institute for Archaeological Research, 2000.
- Starostin G. 2010 — G. STAROSTIN. Preliminary lexicostatistics as a basis for language classification: A new approach // *Journal of Language Relationship*, No. 3. P. 79–116.
- Swadesh 1952 — M. SWADESH. Lexicostatistic dating of prehistoric ethnic contacts // *Proceedings of the American Philosophical Society* 96. P. 452–463.
- Swadesh 1955 — M. SWADESH. Towards greater accuracy in lexicostatistic dating // *International Journal of American Linguistics* 21. P. 121–137.
- Бурлак & Старостин 2005 — С. А. БУРЛАК, С. А. СТАРОСТИН. *Сравнительно-историческое языкознание [Comparative Linguistics]*. 2-е изд. М., 2005.
- Старостин 2007 — С. А. СТАРОСТИН. *Труды по языкознанию [Works in Linguistics]*. М., 2007.

Статья продолжает серию публикаций (см. Starostin G. 2010), посвященных уточнению лексикостатистического метода классификации языков. В статье обсуждается проблема синонимии/полисемии, что, как известно, на практике является главной сложностью при составлении списков Сводеша. Для каждого элемента 110-словного списка приводится набор синтаксических и семантических контекстов, помогающих при работе в поле, а также комментарии, объясняющие, что именно подразумевается под данным сводешевским словом. При составлении контекстов и комментариев авторы частично основывались на лингвистической традиции, а частично — на теоретических и прагматических соображениях, которые в статье эксплицитно указываются.

Number	ENG	RU	Contexts	Notes
1.	all ( <i>omnis</i> )	все ( <i>omnis</i> )	<ol style="list-style-type: none"> <li>1. <i>He has cut down all the trees</i></li> <li>2. <i>He has killed all animals</i></li> <li>3. <i>All these men are brothers</i></li> </ol> <ol style="list-style-type: none"> <li>1. Он срубил все деревья</li> <li>2. Он убил всех животных</li> <li>3. Все эти мужчины — братья</li> </ol>	<p>The “plural” ‘all’ (= Latin <i>omnis</i>). Not to be confused with ‘each, every’ (cf. context 3). If possible, should be separated from the “singular” ‘all’ (Latin <i>totus</i>, in contexts like: <i>All of the water in the lake/pool became frozen/evaporated; He has eaten all the meat</i>), which we recommend to exclude from the list.</p> <p>‘Все’ применительно к совокупности объектов (латин. <i>omnis</i>). Отличать от ‘каждый’ (см. контекст 3). По возможности отличать от «сингулярного» ‘весь’ (латин. <i>totus</i>; контексты вроде: <i>Вся вода в озере/луже высохла/заморозилась; Он съел всё мясо</i>), которое мы рекомендуем исключить из списка.</p>
2.	ashes	зола	<ol style="list-style-type: none"> <li>1. <i>The campfire has left only ashes.</i></li> <li>2. <i>The wind scattered the ashes.</i></li> <li>3. <i>He scooped up a handful of ashes.</i></li> </ol> <ol style="list-style-type: none"> <li>1. От костра осталась одна зола и угли</li> <li>2. Ветер развеял золу</li> <li>3. Он взял в горсть золу</li> </ol>	<p>Ashes as the basic result of combustion of wood, grass, dry dung (not after cremation, as in Russian <i>прах</i>). Not to be confused with special words for “fly ash”, “tobacco ashes”, “volcanic ashes”, or “embers” (hot ashes/coals, cf. context 3).</p> <p>Зола как продукт сгорания дерева, травы, кизяка (но не после кремации, как в русском <i>прах</i>). Отличать от маркированных «табачного пепла», «вулканического пепла», а также «жара костра» (неостывшая зола, см. контекст 3).</p>
3.	bark	кора	<ol style="list-style-type: none"> <li>1. <i>He tore some bark off the tree.</i></li> <li>2. <i>A piece of bark came loose from the tree.</i></li> </ol> <ol style="list-style-type: none"> <li>1. Он содрал с дерева кору</li> <li>2. От дерева отвалился кусок коры</li> </ol>	<p>The neutral word for ‘bark’, applicable to as many different tree species as possible.</p> <p>Нейтральная ‘кора’, сочетающаяся с максимальным количеством видов деревьев.</p>



Number	ENG	RU	Contexts	Notes
4.	belly	живот	<ol style="list-style-type: none"> <li>1. <i>While fighting, he punched him in the belly.</i></li> <li>2. <i>Wrap this belt around your belly.</i></li> <li>1. Он в драке ударил его в живот.</li> <li>2. Завяжи пояс на животе.</li> </ol>	<p>Part of the human body located directly above the pelvis. Not to be confused with various terms that denote internal organs (<i>stomach, intestines</i>) or semantically/stylistically marked words (<i>paunch</i>).</p> <p>Часть туловища между грудью и пахом. Отличать от многочисленных внутренних органов (<i>желудок, кишки</i> и т. п.) и семантически/стилистически маркированных слов (<i>пузо</i> и т. п.)</p>
5.	big, large	большой	<ol style="list-style-type: none"> <li>1. <i>This person has a big nose, that one has a small nose.</i></li> <li>2. <i>A big stone [leaf, tree, pool, etc...]</i></li> <li>1. У этого человека большой нос, а у того — маленький.</li> <li>2. Большой камень [лист, дерево, водоем...]</li> </ol>	<p>An adjective or verb that may be applied to different objects. The contexts are the same as in No.78 'small'.</p> <p>Прилагательное/глагол, сочетающееся с максимальным количеством объектов. Контексты те же, что и для №78 'маленький'.</p>
6.	bird	птица	<ol style="list-style-type: none"> <li>1. <i>Something is moving in the bushes, I cannot tell if it is a bird or an animal.</i></li> <li>2. <i>Birds lay eggs, animals and people bear children.</i></li> <li>3. <i>There is a bird flying on high, I cannot tell what kind.</i></li> <li>1. В кустах что-то шевелится, я не понимаю, это птица или зверь.</li> <li>2. Птицы несут яйца, а звери и люди рожают.</li> <li>3. Вон там далеко летит птица. Слишком далеко, я не могу разобрать какая именно.</li> </ol>	<p>Some languages make a contrast between 'small bird' / 'large bird' or have a special word for 'large (predatory) bird'. The contexts have been specially chosen so as to assist in revealing the more neutral term.</p> <p>В некоторых языках могут противопоставляться 'большая (хищная) птица' vs. 'маленькая птица' ит. п. Контексты подобраны так, чтобы попытаться выявить наименее маркированную лексему.</p>

Number	ENG	RU	Contexts	Notes
7.	to bite	кусать	<ol style="list-style-type: none"> <li>1. <i>The dog bit him.</i></li> <li>2. <i>He bit his opponent in a fight.</i></li> <li>1. Собака укусила его.</li> <li>2. Он в драке укусил противника.</li> </ol>	<p>Not to be confused with 'to bite off' (while eating) or 'to gnaw'. Said of animals or people (not a separate word like 'to sting' used for insects, snakes, stinging plants etc.).</p> <p>Отделять от 'откусывать' (кусание при принятии пищи), 'грызть'. Отделять от 'кусать' ('жалить') применительно к насекомому, змее, жгучему растению.</p>
8.	black	черный	<p>[As on the RGB 0:0:0 card; the natural color of coal]</p> <p>[Карточка RGB 0:0:0; естественный цвет угля]</p>	<p>Not to be confused with 'dark', 'dirty' etc.</p> <p>Отличать от 'тёмный', 'грязный' и т. п.</p>
9.	blood	кровь	<ol style="list-style-type: none"> <li>1. <i>If you cut yourself, blood will flow.</i></li> <li>2. <i>There's blood flowing from his wound.</i></li> <li>3. <i>Blood is red.</i></li> <li>1. Если порезаться, потечет кровь.</li> <li>2. У него из раны течет кровь.</li> <li>3. Кровь красная.</li> </ol>	<p>Of human blood. To be distinguished from 'coagulated blood', 'ichor', 'menstrual blood' etc. We know of no languages that distinguish lexically between 'arterial blood' (bright red) and 'venous blood' (dark red, black).</p> <p>О человеке. Отличать от 'запекшейся крови', 'сукровицы', 'менструальных выделений' и т. п. Нам неизвестны языки, лексически различающие артериальную кровь (<i>светло-красная</i>) и венозную кровь (<i>темно-красная, черная</i>).</p>

Number	ENG	RU	Contexts	Notes
10.	bone	кость	<ol style="list-style-type: none"> <li>1. <i>His bones are frail.</i></li> <li>2. <i>He was sucking the marrow out of a bone.</i></li> <li>3. <i>The body consists of flesh, fat and bones.</i></li> <li>4. <i>He choked on a bone.</i></li> </ol> <ol style="list-style-type: none"> <li>1. <i>У него хрустят кости.</i></li> <li>2. <i>Он высасывал мозг из кости.</i></li> <li>3. <i>Тело состоит из мяса, жира и костей.</i></li> <li>4. <i>Он подавился костью.</i></li> </ol>	<p>General term for human and animal bones (anatomical, not as part of food). Not to be confused with specific designations of various bones (like ‘rib’, ‘vertebra’) and a possible designation of ‘tubular bone in general’.</p> <p>Общее слово для кости как у человека, так и у животного (анатомический термин, не ‘кость как продукт, еда’). Отличать от конкретизированных обозначений различных костей (вроде ‘ребро’, ‘позвонок’) и от возможного специального обозначения трубчатой кости вообще.</p>
11.	breast	грудь	<ol style="list-style-type: none"> <li>1. <i>He hit him in the breast (chest) while fighting.</i></li> <li>2. <i>His breast (chest) was decorated with ornaments.</i></li> </ol> <ol style="list-style-type: none"> <li>1. <i>Он в драке ударил его в грудь.</i></li> <li>2. <i>На груди у него были украшения.</i></li> </ol>	<p>Frontal part of the human body, located between the neck and the upper belly (region of the thorax). If there is a common term, applicable to both the male and female breasts, we recommend choosing this common term (e.g., English <i>breast</i>). If no common term is present, the word for <i>male breast</i> (= <i>chest</i>) must be chosen. Not to be confused with specific designations of the female mammary gland, nipples, internal organs (e.g., lungs) or thoracic cage.</p> <p>Человеческая грудь как передняя часть туловища между шеей и животом. Если в языке есть общий термин, обозначающий как мужскую, так и женскую грудь, рекомендуется выбирать именно его. В противном случае следует брать обозначение мужской груди. Отличать от конкретизированного обозначения женских молочных желез или сосков; от внутренних органов (легкие и т. п.) и от грудной (реберной) клетки.</p>

Number	ENG	RU	Contexts	Notes
12.	to burn (trans.)	жечь, сжечь	<ol style="list-style-type: none"> <li>1. <i>We have to burn fuel (wood) all night long, so as to have light and warmth.</i></li> <li>2. <i>Why are you burning your clothes?</i></li> </ol> <ol style="list-style-type: none"> <li>1. <i>Надо жечь костер/поленья/траву всю ночь, чтобы мы не замерзли и чтобы было светло.</i></li> <li>2. <i>Зачем ты жжешь свою одежду?</i></li> </ol>	<p>A transitive verb. To be distinguished from both inchoatives like 'to set on fire' and "emphatic perfectives" like 'to burn down'.</p> <p>Переходный глагол (не 'гореть'). Отличать от 'зажечь', 'поджечь' и от маркированного 'сжечь дотла'.</p>
13.	cloud	облако	<ol style="list-style-type: none"> <li>1. <i>There is a white cloud in the sky.</i></li> <li>2. <i>Rain does not come from (white) clouds, only from rain clouds.</i></li> </ol> <ol style="list-style-type: none"> <li>1. <i>На небе белое облако.</i></li> <li>2. <i>Облако не дает дождя, а туча дает.</i></li> </ol>	<p>White cumulus clouds, as distinguished from rain clouds (black). The original wordlist never specified the exact type of cloud, but for languages that distinguish lexically between the two, the immediate association seems to be with white rather than black cloud (e.g., Russian-based wordlists always use <i>облако</i> 'white cloud' rather than <i>туча</i> 'black cloud, rain cloud').</p> <p>Белое кучевое облако; отличать от темной 'тучи' (дождевого облака).</p>
14.	cold	холодный	<ol style="list-style-type: none"> <li>1. <i>The water is too cold.</i></li> <li>2. <i>I do not know if this stone is still cold or whether it already got hot in the sun.</i></li> </ol> <ol style="list-style-type: none"> <li>1. <i>Эта вода слишком холодная.</i></li> <li>2. <i>Я не знаю, этот камень холодный или он успел нагреться на солнце.</i></li> </ol>	<p>The regular antonym of 'hot'. Should be separated from more specific temperature characteristics ('ice-cold', 'cool', etc.).</p> <p>Стандартный антоним слова 'горячий'. Отличать от оттенков холодности: 'ледяной', 'прохладный' и т. п.</p>

Number	ENG	RU	Contexts	Notes
15.	to come	приходить	<ol style="list-style-type: none"> <li>1. <i>He came to me / my place.</i></li> <li>1. <i>Он пришел ко мне (в мой дом).</i></li> </ol>	<p>Of humans: to reach (preferably on foot) a destination. Should be distinguished from 'enter' ('penetrate some confined space'), and, if possible, from 'arrive' and 'reach' (with strong emphasis either on the circumstances of one's arrival or on the very facts).</p> <p>О человеке: достигать (желательно, пешком) точки-ориентира. Следует отличать как от 'входить' ('внутри ограниченного пространства'), так и от 'прибывать' или 'достигать' (упор на обстоятельность достижения цели или на сам факт того, что она достигнута).</p>
16.	to die	умирать	<ol style="list-style-type: none"> <li>1. <i>He was very old; he died a year ago.</i></li> <li>2. <i>All who live, die sooner or later.</i></li> <li>1. <i>Он был очень старым. Он умер год назад.</i></li> <li>2. <i>Все живущие рано или поздно умирают.</i></li> </ol>	<p>Non-violent human death (but to be separated from possible marked verb 'to die of old age'). Neutral stylistics, to be distinguished from various polite or euphemistic forms.</p> <p>Ненасильственная смерть человека (однако отличать от возможного маркированного глагола 'умереть от старости'). Нейтральный термин; отличать от специальных вежливых или эвфемистических форм.</p>
17.	dog	собака	<ol style="list-style-type: none"> <li>1. <i>A dog is a man's best friend.</i></li> <li>2. <i>The dog dwells together with people, wild animals do not.</i></li> <li>1. <i>Собака — друг человека</i></li> <li>2. <i>Собака живет с человеком, а дикие звери — нет.</i></li> </ol>	<p>Domesticated dog. Not to be confused with various breeds or specific words for dogs of different ages or sexes ('puppy', 'bitch' etc.).</p> <p>Одомашненная собака. Отличать от названий разных пород или половых возрастных категорий ('пёс', 'щенок', 'сука' и т. п.).</p>
18.	to drink	пить	<ol style="list-style-type: none"> <li>1. <i>He is drinking water.</i></li> <li>2. <i>I need to eat and drink to stay alive.</i></li> <li>1. <i>Он пьет воду.</i></li> <li>2. <i>Чтобы жить, надо есть и пить.</i></li> </ol>	<p>To be distinguished from specialized terms ('to gulp', 'to drink to emptiness', 'to drink an alcoholic beverage', etc.).</p> <p>Отличать от специализированных понятий вроде 'хлебать', 'выпить залпом', 'выпить до дна', 'пить алкогольный напиток' и т. п.</p>

Number	ENG	RU	Contexts	Notes
19.	dry	сухой	<ol style="list-style-type: none"> <li>1. <i>Dry earth</i></li> <li>2. <i>My clothes are already dry, but his are still wet.</i></li> <li>1. <i>Сухая земля</i></li> <li>2. <i>Моя одежда уже сухая, а его пока еще мокрая.</i></li> </ol>	<p>Antonym of 'wet', applicable to as many different types of objects as possible (to be separated from specialized terms like <i>desiccated</i>, <i>dehydrated</i> etc.).</p> <p>Антоним 'мокрого', сочетающийся с максимальным количеством различных видов объектов. Отличать от специализированных терминов (<i>сушеный</i>, <i>иссушенный</i>, <i>обезвоженный</i> и т. п.).</p>
20.	ear	ухо	<ol style="list-style-type: none"> <li>1. <i>He punched him in the ear while fighting.</i></li> <li>2. <i>To cut off someone's ear.</i></li> <li>3. <i>That man has large ears, this one has small ones.</i></li> <li>1. <i>Он его в драке ударил в ухо.</i></li> <li>2. <i>Отрезать человеку ухо.</i></li> <li>3. <i>У того человека большие уши, а у этого — маленькие.</i></li> </ol>	<p>Separate from 'internal ear' as well as the abstract 'hearing' (as in, <i>have a good ear for music</i>).</p> <p>Отличать от 'внутреннего уха' (т. е. внутренних органов) и 'слуха'.</p>
21.	earth	земля	<ol style="list-style-type: none"> <li>1. <i>While digging the hole, he was throwing out the earth.</i></li> <li>2. <i>He took a handful of earth.</i></li> <li>3. <i>I don't know whether that site contains sand or earth.</i></li> <li>1. <i>Он рыл яму и выбрасывал наружу землю.</i></li> <li>2. <i>Он взял в горсть земли.</i></li> <li>3. <i>Я не знаю, что на том участке: земля или песок.</i></li> </ol>	<p>'Earth' as 'soil' (but not as a specialized professional term, e.g., Eng. 'soil' itself). To be separated from 'ground' (= 'surface'), 'land' (= 'dry land' or 'property'), etc. Also to be separated from specific kinds of soil ('red earth', 'black earth', 'clayey soil' etc.). Contexts are the same as in No.71 'sand'.</p> <p>Земля как почва (собственно слово 'почва' не подходит как профессионально маркированный или «научный» термин). Отличать от поверхности земли ('ground'), от суши ('land') и т. п. Отличать от конкретизированных типов почвы: 'красная земля', 'черная земля', 'глинистая земля' и т. п. Те же контексты, что и в №71 'песок'.</p>

Number	ENG	RU	Contexts	Notes
22.	to eat	есть	<ol style="list-style-type: none"> <li>1. <i>I need to eat and drink to stay alive.</i></li> <li>2. <i>He is eating [insert primary food product(s) of given culture].</i></li> <li>3. <i>When he lived there, he ate and drank plenty.</i></li> </ol> <ol style="list-style-type: none"> <li>1. <i>Чтобы жить, надо есть и пить.</i></li> <li>2. <i>Он ест [базовый продукт для данной культуры].</i></li> <li>3. <i>Когда он там жил, он ел и пил вволю.</i></li> </ol>	<p>In some languages different verbs can be used for different food products (“eat meat”, “eat cereals”, “eat fruits / nuts”, etc.) or food types (“eat hard food”, “eat soft food”, “eat liquid food”, etc.). In such cases, one should probably choose the verb that agrees with the most basic food product for the civilization in question (= statistically frequent). Separate from marked verbs like ‘to eat up’ etc. and various nursery or polite words.</p> <p>Для языков, противопоставляющих разные глаголы ‘есть’ по продуктам (мясо, злаки, фрукты, суп и т. п.) или типам пищи (твердая, мягкая) предлагается брать глагол, сочетающийся с базовым для данной культуры продуктом, т. к. он, скорее всего, будет наиболее частотным. Отличать от маркированных глаголов вроде ‘съесть всё до конца’ и т. п. и различных «детских» или вежливых слов вроде <i>кушать</i>.</p>
23.	egg	яйцо	<ol style="list-style-type: none"> <li>1. <i>Birds lay eggs, animals do not.</i></li> <li>2. <i>He had an egg in his hand.</i></li> <li>3. <i>Do you eat eggs?</i></li> </ol> <ol style="list-style-type: none"> <li>1. <i>Птицы несут яйца, а звери нет</i></li> <li>2. <i>Он взял в руку яйцо.</i></li> <li>3. <i>Ты ешь яйца?</i></li> </ol>	<p>Note that the notion ‘to lay eggs’ can frequently be expressed <i>syncretically</i> by a separate verb; this should be rejected.</p> <p>‘Снести яйцо’ может быть синкретически выражено отдельным глаголом, что не подходит как элемент списка.</p>
24.	eye	глаз	<ol style="list-style-type: none"> <li>1. <i>He plucked his eye out.</i></li> <li>2. <i>Most people have got two eyes, but this man only has one eye.</i></li> </ol> <ol style="list-style-type: none"> <li>1. Он выколол ему глаз.</li> <li>2. У большинства людей два глаза, а у этого человека только один глаз.</li> </ol>	<p>Singular “eye” (to avoid suppletive stems in the plural). Not to be confused with ‘eyeball’ or ‘eyesight’.</p> <p>«Глаз» (ед. ч., чтобы избежать возможного супплетивизма с формой мн. ч.). Следует отличать от ‘зрачок’, ‘зрение’.</p>



Number	ENG	RU	Contexts	Notes
25.	fat	жир	<ol style="list-style-type: none"> <li>1. <i>The body consists of flesh, fat and bones.</i></li> <li>2. <i>One can get fat from many different animals.</i></li> <li>3. <i>Meat and fat can be eaten.</i></li> <li>4. <i>Here are some pieces of fat.</i></li> </ol> <ol style="list-style-type: none"> <li>1. <i>Тело состоит из мяса, жира и костей.</i></li> <li>2. <i>Жир можно добывать из разных животных.</i></li> <li>3. <i>Мясо и жир можно есть.</i></li> <li>4. <i>Вот куски жира.</i></li> </ol>	<p>Animal fat as a general substance used for eating or various economic needs. Should be distinguished from “vegetable fat” or “butter”, (liquid) “oil”, although many languages use the same word for these notions, as well as specific words denoting different kinds of fat contained in particular body parts (e.g., ‘suet’) or specially treated (e.g., boiled) fat.</p> <p>Общее слово для животного жира как продукта, используемого в хозяйстве. Отличать от ‘растительного жира’, ‘масла’ (в т. ч. жидкого) там, где в языке есть соотв. дифференциация, а также от внутреннего жира и специально обработанного или приготовленного жира.</p>
26.	feather	перо	<ol style="list-style-type: none"> <li>1. <i>Adult birds are covered with feathers.</i></li> <li>2. <i>Birds have feathers, beasts have fur.</i></li> <li>3. <i>He plucked one feather from the bird’s wing.</i></li> <li>4. <i>He stuck a feather in his hair as an ornament.</i></li> </ol> <ol style="list-style-type: none"> <li>1. <i>Взрослые птицы покрыты перьями.</i></li> <li>2. <i>У птиц есть перья, а у зверей — мех.</i></li> <li>3. <i>Он вырвал у птицы из крыла одно перо.</i></li> <li>4. <i>Он воткнул в волосы перо как украшение.</i></li> </ol>	<p>The most general term should be chosen. If no single term for all kinds of feathers exists, the reasonable choice is ‘quill, flight feather of a bird’. Not to be confused with ‘fluff’, ‘down’, etc. The word should be applicable to different bird species.</p> <p>Должен быть выбран наиболее нейтральный термин. Если такого термина нет, то уместно выбрать ‘маховое перо птицы’. Отличать от ‘луха’ и т. п. Обязательно проверить, чтобы слово можно было употреблять для разных видов птиц.</p>

Number	ENG	RU	Contexts	Notes
27.	fire	огонь	<ol style="list-style-type: none"> <li>1. <i>I see a fire in the distance.</i></li> <li>2. <i>He put the fire out with water.</i></li> <li>3. <i>Fire can burn you.</i></li> </ol> <ol style="list-style-type: none"> <li>1. Вдалеке я вижу огонь</li> <li>2. Он потушил огонь водой</li> <li>3. Огонь может обжечь</li> </ol>	Should be kept separate from 'flame' (= 'tongues of fire', stylistically marked), 'bonfire', 'sacred fire', etc. Отличать от стилистически маркированного 'пламени', 'костра', 'священного огня' и т. п.
28.	fish	рыба	<ol style="list-style-type: none"> <li>1. <i>All fish live in water.</i></li> <li>2. <i>Birds fly, fish swim.</i></li> <li>3. <i>I want to catch this fish.</i></li> </ol> <ol style="list-style-type: none"> <li>1. Все рыбы живут в воде.</li> <li>2. Птицы летают, а рыбы плавают.</li> <li>3. Я хочу поймать эту рыбу.</li> </ol>	Live fish, to be distinguished from special terms for already caught or prepared fish. Живая рыба; следует отличать от выловленной или разделанной.
29.	to fly	лететь, летать	<ol style="list-style-type: none"> <li>1. <i>Birds can fly, beasts and people cannot.</i></li> <li>2. <i>Little chicks do not yet know how to fly.</i></li> <li>3. <i>There's a bird in the sky. Where is it flying?</i></li> </ol> <ol style="list-style-type: none"> <li>1. Птицы летают, а люди и животные — нет.</li> <li>2. Маленькие птенцы еще не умеют летать.</li> <li>3. Вот в небе птица. Куда она летит?</li> </ol>	Of birds, denoting neutral movement through the sky (but not 'to fly off', 'to fly away' etc., frequently denoted by different verbs). Should be kept apart from specific types of flying ('to soar', 'to flit', etc.). On rare occasions a neutral iterative verb 'to fly' ( <i>Where is the bird flying?</i> ) and a more marked habitual verb 'to fly' ( <i>Little chicks do not yet know how to fly</i> ) can contain different roots that is a symmetric correspondence to the frequent root opposition 'to go' ~ 'to walk' (see No.33). In such a case both roots can be taken as synonyms. NB: this entry does <i>not</i> refer to the noun 'fly' (insect)! О птицах; имеется в виду нейтральный тип перемещения по воздуху (но не 'взлетать', 'улетать' и т. п.; эти значения часто передаются особыми корнями). Следует отличать от конкретных типов полета ('парить', 'порхать'). Изредка язык может различать корни глаголов 'лететь' (нейтральный итератив) и 'летать' (более маркированный хабиуталис), что является симметричной ситуацией к частой корневой оппозиции 'идти' ~ 'ходить' (см. №33). В таких случаях оба корня можно брать как синонимы.

Number	ENG	RU	Contexts	Notes
30.	foot	нога	<ol style="list-style-type: none"> <li>1. <i>He put the shoe on his foot.</i></li> <li>2. <i>He kicked him with his foot.</i></li> <li>3. <i>Feet are for walking.</i></li> </ol> <ol style="list-style-type: none"> <li>1. Он надел обувь на ногу.</li> <li>2. Он ударил его ногой.</li> <li>3. Ноги нужны для того, чтобы ходить.</li> </ol>	Denoting the part of the leg below the ankle joint. In many languages, “leg” and “foot” are either not distinguished at all, and the particular meaning “foot” is translated as “smth. of the leg”, e.g., “head of leg”, “end of leg”, etc. In these cases, the “leg” root is still carrying the bulk of the meaning.  Часть ноги ниже лодыжки (но не специализированный анатомический термин ‘ступня’). В языках, не различающих обозначения ‘ноги выше ступни’ и ‘ступни’, последнее значение часто передается как композит: «голова ноги», «конец ноги» и т. п. Во всех таких случаях основным корнем следует именно ‘ногу’.
31.	full		<ol style="list-style-type: none"> <li>1. <i>This jar is full, that one is empty.</i></li> <li>2. <i>The jar (bag, pit, etc.) is already full, we cannot put in anything else.</i></li> </ol> <ol style="list-style-type: none"> <li>1. Этот сосуд пустой, а это полный.</li> <li>2. Сосуд/мешок/яма полон(а), больше ничего не поместится.</li> </ol>	Adjective or verb, referring to the full state of a vessel. Can be used without dependents, thus, should not be confused with ‘filled with smth.’. Also to be distinguished from specialized words like ‘completely full’, ‘full to the brim’, ‘full with small/large objects’, etc.  Прилагательное или глагол (‘быть полным’), применительно к наполненному суду. Слово должно иметь возможность независимого употребления, так что следует отличать от ‘наполненный чем-либо’. Также отличать от маркированного ‘полностью заполненный’, ‘полный до краев’ или от конкретизированного ‘полный мелких предметов’ и т. п.
32.	to give	давать	<ol style="list-style-type: none"> <li>1. <i>I am giving my father a stone.</i></li> <li>2. <i>My father is giving me a stone.</i></li> <li>3. <i>He gave him a stick, but then snatched it back.</i></li> </ol> <ol style="list-style-type: none"> <li>1. Я даю отцу камень.</li> <li>2. Отец дает мне камень.</li> <li>3. Он дал ему палку, но потом выхватил ее обратно.</li> </ol>	To be distinguished from ‘to hand’, ‘to offer’ etc., as well as special polite forms if encountered. Some languages make a clear and regular distinction between ‘to give (to 1 <sup>st</sup> /2 <sup>nd</sup> person)’ and ‘to give (to 3 <sup>rd</sup> person)’ (cf. the difference between contexts 1 and 2); it is permissible to use these as synonyms if no third ‘recipient-neutral’ party is found.  Отличать от ‘протягивать’, ‘вкладывать в руку’ и т. п., а также от специальных «вежливых» форм. В некоторых языках отчетливо различаются понятия ‘давать (1-му/2-му л.)’ и ‘давать (3-му л.)’ (ср. разницу между контекстами 1 и 2); оба корня позволительно брать как синонимичные в том случае, если невозможно обнаружить третий предикат, нейтральный по отношению к лицу реципиента.

Number	ENG	RU	Contexts	Notes
33.	to go	идти	<ol style="list-style-type: none"> <li>1. <i>He was standing here, but now he is going that way</i> [indicate direction w. hand].</li> <li>2. <i>Today I work at home, tomorrow I'm going to the next village.</i></li> <li>3. <i>Where are you going?</i></li> <li>1. <i>Он стоял тут, а теперь он идет туда</i> [показать рукой направление].</li> <li>2. <i>Сегодня я буду работать дома, а завтра пойду в соседнее селение.</i></li> <li>3. <i>Куда ты идешь?</i></li> </ol>	<p>Indicates movement from point A to point B; on foot by default, but with more emphasis on the fact of movement than on the mode of movement. A continuative/durative/imperfective verb. The original Swadesh item 'to walk', referring either to movement with no specific direction (<i>walking in the park</i>) or emphasizing the idea of using one's feet rather than any means of transport (<i>I'm going to walk to the station</i>), was rather ambiguous and tended to be replaced by 'to go' already in the early wordlists; we recommend preserving this tradition. NB: in many languages this may be the same root as 'to go away' (the antonym of 'to come'), but the correlation is not necessary (cf. Russian <i>идти, not уходить</i>).</p> <p>Нейтральное передвижение из точки А в точку В; по умолчанию ногами, но с большим упором на факт передвижения, нежели на способ передвижения. Длительное действие. Исходный элемент в списке Сводеша — 'to walk', т. е. 'гулять, расхаживать' (<i>walking in the park</i>) или специфически 'идти (именно) пешком' (<i>I'm going to walk to the station</i>); двусмысленность и расплывчатость термина заставляла многих составителей списков уже на раннем этапе заменить его на 'идти' (<i>go</i>); мы рекомендуем придерживаться этой традиции. NB: во многих языках лексически совпадает с 'уходить' (= антоним <i>to come</i> 'приходить'), но это совпадение не является правилом (ср. русское <i>идти vs. уходить</i>).</p>
34.	good	хороший	<ol style="list-style-type: none"> <li>1. <i>You are a good man, he is a bad man.</i></li> <li>2. <i>A good knife.</i></li> <li>3. <i>A good house.</i></li> <li>4. <i>Good food.</i></li> <li>1. <i>Ты — хороший человек, он — плохой.</i></li> <li>2. <i>Хороший нож.</i></li> <li>3. <i>Хороший дом.</i></li> <li>4. <i>Хорошая еда.</i></li> </ol>	<p>Should be combined with as many different objects as possible. It also makes sense to look for a basic antonymous pair ('good' / 'bad').</p> <p>Проверить на максимальную сочетаемость с объектами. Стоит искать также как один из членов базисной антонимической оппозиции 'хороший' : 'плохой'.</p>

Number	ENG	RU	Contexts	Notes
35.	green	зеленый	[As on the RGB 0:255:0 and similar cards. The natural color of fresh, newly sprung grass.] [Карточка RGB 0:255:0 и близкие к ней. Цвет молодой свежей травы.]	
36.	hair	волосы	1. <i>He pulled him by his hair.</i> 2. <i>My hair grows better than his.</i> 1. Он дернул его за волосы. 2. У меня волосы лучше растут, чем у него.	Collective or plural noun, applicable to head hair of human. Should be strictly separated from words that specifically designate 'a single hair', 'body hair (fur)' or 'facial hair'. Also to be distinguished from numerous specific kinds of hair ('long hair', 'blonde hair', 'curly hair', various types of coiffures etc.). Волосы на голове человека. Следует строго отличать от (отдельного) 'волоска', 'волос на теле' (шерсть, мех) и 'волос на лице' (борода, усы). Не путать с конкретизированными типами волос (напр., 'длинные волосы', 'светлые волосы', 'курчавые волосы', разные типы причесок и т. п.).
37.	hand	рука	1. <i>Man has two hands.</i> 2. <i>They cut off his hands.</i> 3. <i>People have fingers on each hand.</i> 1. У человека две руки. 2. Ему отрубили обе руки. 3. У человека на каждой руке пальцы.	Specifically 'hand', i. e. the body part at the end of the arm below the wrist. Not to be confused with specialized terms like 'palm', 'back side of the hand', 'right hand', 'left hand' etc. Lack of the 'hand' : 'arm' opposition should be addressed the same way as with 'foot' : 'leg', see No.30. Часть руки от запястья до пальцев, включая пальцы. Отличать от 'ладони', 'тыльной стороны руки', 'правой руки', 'левой руки', 'левой руки', анатомического термина 'кисть' и т. п. Отсутствие оппозиции 'рука (кисть)' : 'рука (выше кисти)' следует трактовать так же, как в случае с 'ногой', см. №30.
38.	head	голова	1. <i>While fighting, he hit him on the head.</i> 2. <i>This stone is the same size as a human head.</i> 1. Он в драке ударил его по голове. 2. Вот камень размером с человеческую голову.	Not to be confused with: (a) narrower terms like 'occiput', 'crown (of the head)', etc.; (b) 'brain', 'mind', etc.; (c) metaphorical 'head' > 'top', 'chief', etc. Отличать от: (а) более узких терминов — 'затылок', 'макушка' и т. п.; (б) 'мозг', 'ум' и т. п.; (в) метафорического переноса 'голова' > 'верхушка', 'начальник' и т. п.

Number	ENG	RU	Contexts	Notes
39.	to hear	слышать	<ol style="list-style-type: none"> <li>1. <i>You speak too quietly, I cannot hear you.</i></li> <li>2. <i>Suddenly I heard a shot.</i></li> <li>1. Ты тихо говоришь, я тебя не слышу.</li> <li>2. Я неожиданно услышала выстрел.</li> </ol>	<p>Not to be confused with 'to listen' (sometimes, but far from always, expressed by the same root). To be separated from 'to hear' in the meaning 'to know', i.e. 'to hear that X' or 'to hear of X' (<i>I heard that he died; I heard of it</i>).</p> <p>Отличать от 'слушать' (иногда, но далеко не всегда, то же корень) и 'прислушиваться'. Отличать от 'слышать' в значении 'знать', т.е. 'слышать, что X' или 'слышать о X-е' (<i>Я слышала, что он умер; Я слышала об этом</i>).</p>
40.	heart	сердце	<ol style="list-style-type: none"> <li>1. <i>He cut the heart out of the carcass.</i></li> <li>2. <i>A human heart is larger than a dog's heart.</i></li> <li>1. Он вырезал сердце из туши.</li> <li>2. Человечье сердце больше собачьего.</li> </ol>	<p>Both of humans and animals. Understood as a body part rather than in any abstract sense ('heart' as a vessel for one's mind or soul).</p> <p>О человеке и животном. Часть тела; отличать отместилица души, ума и т. п.</p>
41.	horn	рог	<ol style="list-style-type: none"> <li>1. <i>Some animals have horns, some don't.</i></li> <li>2. <i>People don't have horns, some animals do.</i></li> <li>3. <i>He grabbed [the beast] by the horn.</i></li> <li>4. <i>He cut off [the beast's] horn.</i></li> <li>1. Бывают звери с рогами, а бывают без.</li> <li>2. У человека нет рогов, а у некоторых зверей есть.</li> <li>3. Он схватил [зверя] за рог.</li> <li>4. Он отрезал рог у [зверя].</li> </ol>	<p>Applied to the prevailing/endemic species of the given region. If there are different words for 'hollow horn' and 'antler' and both kinds of species are endemic, the preferred choice is 'hollow horn' ('antler' is generally less stable and more likely to have a "descriptive" origin). However, the word for 'antler' can be used if there are no prevailing cavicorn endemics in the region (e.g., in North Siberia). If there are no cornute species in the region altogether, the slot remains empty (e.g., Oceania). To be strictly separated from <i>termini technici</i> like 'hunting horn' or 'drinking horn' which may be formed from different roots.</p> <p>Обозначение 'рога' у широко распространенных и исконных в данном регионе видов. Если язык лексически различает 'полый рог' и 'ветвистые рога', причём оба типа животных относятся к числу исконных, то предпочтительно слово для 'полорога' (для лексически менее устойчивого понятия 'ветвистые рога' характерны различные описательные обозначения). Обозначение 'ветвистых рогов' следует брать, только если исконного широко распространённого полорогого вида в регионе нет (случай Северной Сибири). Если исконных рогаемых видов в регионе нет вообще, строка не заполняется (случай Океании). Строго отличать от технических терминов, таких как 'охотничий рог', 'рог-сосуд'.</p>



Number	ENG	RU	Contexts	Notes
42a.	I	я		Stylistically neutral 1 <sup>st</sup> sg. person; various honorific forms must be excluded. Direct stem. Стилистически нейтральное личное местоимение 1 л. ед. ч. Различные формы вежливости должны быть исключены. Прямая основа.
42b.	me	меня		Indirect stem of the same pronoun (for languages which distinguish between the two on a paradigmatic level). Это же, но косвенная основа (для языков, которые имеют соответствующее различие на парадигматическом уровне).
43.	to kill	убивать	<ol style="list-style-type: none"> <li><i>The man killed a beast / another man.</i></li> <li><i>He is so strong / agile that he can kill any beast / any man.</i></li> </ol> <ol style="list-style-type: none"> <li>Человек убил зверя / другого человека.</li> <li>Он такой сильный / ловкий, что может убить любого человека / любое животное.</li> </ol>	By an animate agent (i.e. excluding diseases, colliding stones etc.). This verb should preferably be able to take both humans and common animals as its object, so as to rule out stylistic and archaic variants (e.g., 'to slay'). Singular object use (i. e. not 'to slaughter' or various suppletive roots in specific paradigms, see the introduction). С одушевленным агентом (т. е. исключая случаи болезни, сорвавшегося вниз камня и т. п.). Желательно, чтобы позиция объекта для данного глагола могла быть заполнена названиями как людей, так и обычных животных, в целях фильтрации стилистически маркированных и архаичных вариантов (вроде 'умерщвлять'). Отличать от 'убивать' с множественным объектом ('вырезать' и т. п.; см. также введение относительно стратегии, выбираемой для супплетивных парадигм).
44.	knee	колени	<ol style="list-style-type: none"> <li><i>He punched the man with his knee while fighting.</i></li> <li><i>My knee hurts.</i></li> <li><i>Bend your knees.</i></li> </ol> <ol style="list-style-type: none"> <li>Он ударил его в драке коленом.</li> <li>У меня болит колено.</li> <li>Согни колени.</li> </ol>	The knee joint (outside part; to be distinguished from 'knee cap' as a bone). Коленный сустав (внешняя часть ноги). Следует отличать от внутренней стороны колена и 'коленной чашечки' (кость).



Number	ENG	RU	Contexts	Notes
45.	to know	знать	<ol style="list-style-type: none"> <li>1. <i>What's happened? I don't know, maybe he knows.</i></li> <li>2. <i>Everybody knows that clouds bring rain.</i></li> <li>1. <i>Что случилось? — Я не знаю, может быть, он знает.</i></li> <li>2. <i>Все знают, что тучи приносят дождь.</i></li> </ol>	<p>Expresses previous possession of knowledge of a situation; thus, to be separated from (a) <i>acquisition</i> of knowledge — 'to recognize', 'to learn', 'to get to know', etc.; (b) <i>acquaintance</i> with an object/person — 'to know the way', 'to know one's name' etc. (as in French <i>connaitre</i>).</p> <p>Глагол, обозначающий уже имеющее место знание относительно какой-л. ситуации; следует, таким образом, отличать от (a) глаголов узнавания ('узнавать', 'поинимать' и т. п.), (б) ситуаций знакомства с каким-л. объектом/человеком (<i>знать до-рогу, знать имя</i> и т. п., как франц. <i>connaitre</i>).</p>
46.	leaf	лист	<ol style="list-style-type: none"> <li>1. <i>He plucked one leaf from the tree/bush.</i></li> <li>2. <i>Trees / bushes bear leaves, but dead trees / bushes have no leaves.</i></li> <li>1. <i>Он сорвал с дерева/куста один лист.</i></li> <li>2. <i>На деревьях/кустарнике растут листья, а у мертвого дерева/кустарника листьев нет.</i></li> </ol>	<p>Should be applicable to as many different plant species as possible. Should also be separated from special collective terms, such as 'foliage' or 'greenery' (hence the singular use in ex. 1). Not 'leaf of grass' (for languages that distinguish between 'tree leaf' and 'grass leaf').</p> <p>Желательна сочетаемость с максимальным разнообразием видов растений. Отличать от коллективных 'листва' или 'зелень' (ср. пример 1) и от обозначения листа у травянистых растений.</p>
47.	to lie	лежать	<ol style="list-style-type: none"> <li>1. <i>Man can stand, sit, or lie.</i></li> <li>2. <i>Why are you lying in bed?</i></li> <li>1. <i>Человек может стоять, сидеть и лежать.</i></li> <li>2. <i>Почему ты лежишь в постели?</i></li> </ol>	<p>A verb of continuative/durative action. It is generally not recommended to fill in the slot with the inchoative verb 'to lie down', although such practice is rather frequent (and, in truth, both notions are quite commonly expressed with the same root). Should be strictly separated from lying in particular postures ('lie on one's side', 'lie face down' etc.). Cf. similar requirements for No.75 'to sit', No.80 'to stand'. NB: this entry does not mean 'to speak untruthfully'!</p> <p>Глагол состояния; не рекомендуется подменять инхоативным глаголом 'ложиться', хотя подобного рода практика существует (на самом деле, оба понятия часто выражаются одним и тем же корнем). Строго отличать от лежания в конкретных позах ('лежать на боку', 'лежать ничком' и т. п.). Ср. аналогичные требования для No.75 'сидеть', No.80 'стоять'.</p>

Number	ENG	RU	Contexts	Notes
48.	liver	печень	<ol style="list-style-type: none"> <li>1. <i>He cut the liver from the carcass.</i></li> <li>2. <i>Human liver is larger than a dog's liver.</i></li> </ol> <ol style="list-style-type: none"> <li>1. Он вырезал печень из туши.</li> <li>2. Человечья печень больше собачей.</li> </ol>	<p>Applicable to both humans and animals. 'Liver' as a body organ rather than an abstract vessel for the mind, soul, etc., see 'heart'.</p> <p>О человеке и животном. Отличать отместилища души, ума и т. п. (см. 'сердце').</p>
49.	long	длинный	<ol style="list-style-type: none"> <li>1. <i>Here is a long stick, here is a short one.</i></li> <li>2. <i>This rope is long, I need a short one.</i></li> </ol> <ol style="list-style-type: none"> <li>1. Вот длинная палка, а вот короткая.</li> <li>2. Эта веревка длинная, мне нужна короткая.</li> </ol>	<p>Of physical objects. Not to be confused with 'high', 'tall' (vertical, not horizontal) in languages that distinguish between the two. Antonymous to No.105 'short'.</p> <p>О физических объектах. Не смешивать с 'высокий' (вертикальное положение) в тех языках, которые различают эти два слова. Антоним — №105 'короткий'.</p>
50.	louse	вошь	<ol style="list-style-type: none"> <li>1. <i>There is a large louse creeping across the top of his head.</i></li> <li>2. <i>He has no lice.</i></li> </ol> <ol style="list-style-type: none"> <li>1. У него по макушке ползет большая вошь.</li> <li>2. У него нет вшей.</li> </ol>	<p>Head louse as a basic human parasite. Should be kept separate from body lice, pubic lice, animal/plant species of lice, as well as nits.</p> <p>Головная вошь как обычный паразит человека. Отличать от платяной вши, лобковой вши, травяной вши и т. п., а также от гниды.</p>

Number	ENG	RU	Contexts	Notes
51.	man (male)	мужчина	<p>1. <i>Men are doing [X], women are making [Y].</i></p> <p>2. <i>Women bear children, men do not.</i></p> <p>1. Мужчины делают [X], а женщины делают [Y].</p> <p>2. Женщины рожают детей, а мужчины — нет.</p>	<p>Adult human being of the male sex, antonymous to 'woman'. Should be kept apart from 'male' (of any species) if the latter is distinct from 'man' (male human), as well as from 'person / human being' in general. Context No. 1 is culturally conditioned, in that X and Y should be understood as designations of typical male / female occupations for the given society (e.g., 'hunting', 'cooking', etc.). If, instead of a general term for 'man', one finds differentiation by age groups, the respective lexemes for the age of ca. 20–45 must be treated as synonymous. To be separated from specific designations of 'bachelor', 'widower' etc. Same contexts as in No.98 'woman'.</p> <p>Взрослый человек мужского пола. Отличать от понятия 'самец' ('male'), если для последнего существует отдельное слово, равно как и от обозначения человека вообще ('person' / 'human being'). Контекст 1 является культурно обусловленным; здесь следует подставлять наименования типичных для данного социума мужского и женского занятий (охота, готовка и т. п.). Если в языке нет общего слова для 'мужчины', а имеется спецификация по возрастным группам, то соответствующие лексемы для возраста примерно 20—45 лет надо брать как синонимы. Отличать от специальных наименований холостяка, вдовца и т. п. Те же контексты, что и для No.98 'женщина'.</p>
52.	man (person)	человек	<p>1. <i>The dog is a man's (best) friend.</i></p> <p>2. <i>There is something moving in the bushes, I do not know if it is a man or a beast.</i></p> <p>3. <i>Man can talk, an animal cannot.</i></p> <p>1. Собака — друг человека</p> <p>2. В кустах что-то шевелится, я не понимаю, это человек или зверь.</p> <p>3. Человек умеет говорить, а животное — нет.</p>	<p>The basic term for a human being regardless of sex. Singular (plural is frequently suppletive, e.g., 'people'). NB: The original Swadesh list included this word as 'person' in order to avoid misunderstandings, even though the basic English equivalent (in non-PC speech) for the term is also 'man'.</p> <p>Базисный термин, обозначающий человека как вид, безотносительно к половой принадлежности. Ед. ч. (мн. ч. часто образуется супплетивно, ср. <i>люди</i>). NB: в исходном списке Сводеша слово обозначено как 'person' во избежание неоднозначности понимания, несмотря на то, что в стандартном английском языке (без налета поправок) правильный эквивалент — <i>man</i> (= 'мужчина').</p>

Number	ENG	RU	Contexts	Notes
53.	many, a lot of	много	<ol style="list-style-type: none"> <li>1. <i>There are many / there is a lot of rocks on the ground.</i></li> <li>2. <i>I have many / a lot of friends.</i> <ol style="list-style-type: none"> <li>1. <i>На земле лежит много камней.</i></li> <li>2. <i>У меня много друзей.</i></li> </ol> </li> </ol>	<p>Adjective or adverb; applicable to countable objects.</p> <p>Прилагательное или наречие, употребляется со счетными объектами.</p>
54.	meat	мясо	<ol style="list-style-type: none"> <li>1. <i>Which animal's meat is this?</i></li> <li>2. <i>He got / brought plenty of meat.</i></li> <li>3. <i>He ate all the meat and all the fat.</i> <ol style="list-style-type: none"> <li>1. <i>Это мясо какого животного?</i></li> <li>2. <i>Он достал/принес много мяса.</i></li> <li>3. <i>Он съел все мясо и весь жир.</i></li> </ol> </li> </ol>	<p>Meat as food. To be distinguished from specific kinds of meat (from various body parts, various animals etc.) and from marked terms like 'flesh'.</p> <p>Мясо как пищевой продукт. Отличать от наименований конкретных видов мяса (из разных частей туши, от разных животных и т. п.) и от маркированных слов вроде 'плоть'.</p>
55.	moon	луна	<ol style="list-style-type: none"> <li>1. <i>The sun shines during the day, the moon shines at night.</i></li> <li>2. <i>There is a full moon in the sky.</i> <ol style="list-style-type: none"> <li>1. <i>Днем светит солнце, а ночью — луна.</i></li> <li>2. <i>На небе круглая луна.</i></li> </ol> </li> </ol>	<p>To be distinguished from 'crescent moon' if a separate root exists, as well as from 'month'.</p> <p>Отличать от 'месяца' (= неполная луна) и календарного месяца.</p>
56.	mountain	гора	<ol style="list-style-type: none"> <li>1. <i>A mountain is higher than a hill.</i></li> <li>2. <i>It takes a whole day to climb this mountain.</i> <ol style="list-style-type: none"> <li>1. <i>Гора выше холма.</i></li> <li>2. <i>На эту гору подниматься целый день.</i></li> </ol> </li> </ol>	<p>The concept is relative, depending on the landscape of the region (for some languages, the term 'hill' itself functions as a proper equivalent).</p> <p>Относительное понятие, зависящее от ландшафта данного ареала (для некоторых языков в качестве эквивалента приходится брать понятие 'холм').</p>

Number	ENG	RU	Contexts	Notes
57.	mouth	рот	<ol style="list-style-type: none"> <li>1. <i>He took a piece (of food) in his mouth.</i></li> <li>2. <i>He rinsed his mouth with water.</i></li> <li>3. <i>What a big mouth you have!</i></li> <li>1. Он положил в рот кусок [пищи]</li> <li>2. Он прополоскал рот водой</li> <li>3. Какой у тебя большой рот!</li> </ol>	<p>'Mouth' as part of face ('lips' + 'oral cavity'), but should be kept separate from 'lips' <i>per se</i>, as well as 'cheek pouches' etc. It seems that some languages may possess two separate terms — one for 'oral cavity' (<i>His mouth is full of water</i>) and another for 'external parts of the mouth' (<i>He has a very wide mouth</i>) (plus a special designation for 'lips' <i>per se</i>); in such a case both terms can be accepted as synonyms.</p> <p>'Рот' как часть лица ('губы' + 'ротовая полость'); следует, однако, отличать от собственно 'губ', равно как 'защечных пауз' и т. п. Видимо, некоторые языки могут обладать двумя отдельными словами — одно для 'ротовой полости' (<i>Ego рот молон воды</i>), а второе для 'внешних частей рта' (<i>У него очень широкий рот</i>) (при этом есть и особое выражение собственно для 'губ'); в таких случаях оба слова можно брать как синонимы.</p>
58.	nail	ноготь	<ol style="list-style-type: none"> <li>1. <i>Man has five fingers on each hand, and each finger has a nail.</i></li> <li>2. <i>He broke one of his (finger)nails.</i></li> <li>1. У человека на руках 10 пальцев, на каждом пальце по ногтю.</li> <li>2. Он сломал ноготь на большом пальце.</li> </ol>	<p>Common term for fingernails (<i>not</i> toenails). The original Swadesh wordlist had 'claw' (= of animals), but this has on practice been replaced by 'nail' quite a long time ago.</p> <p>Общее слово для ногтей на пальцах рук (не ног). В исходном списке Сводеша на этом месте было 'claw' = 'коготь' (животного), но на практике уже давно используется 'ноготь' (человека).</p>
59.	name	имя	<ol style="list-style-type: none"> <li>1. <i>Tell me your mother's / father's name.</i></li> <li>2. <i>A new-born baby does not yet have a name.</i></li> <li>1. Скажи имя твоей матери / твоего отца.</li> <li>2. У новорожденных еще нет имени.</li> </ol>	

Number	ENG	RU	Contexts	Notes
60.	neck	шея	<ol style="list-style-type: none"> <li>1. <i>He hit him in the neck while fighting.</i></li> <li>2. <i>He tightened the rope around his neck.</i></li> <li>1. <i>Он в драке ударил его по шее.</i></li> <li>2. <i>Он затянул веревку вокруг шеи.</i></li> </ol>	<p>To be distinguished from 'throat' (internal) or specialized terms like 'front / back of the neck'.</p> <p>Отличать от 'глотки' ('внутренняя часть шеи'), а также от конкретизированных понятий 'передняя часть шеи', 'задняя часть шеи' и т. п.</p>
61.	new	новый	<ol style="list-style-type: none"> <li>1. <i>He got himself some new clothes.</i></li> <li>2. <i>He made himself a new tool.</i></li> <li>1. <i>Он приобрел себе новую одежду.</i></li> <li>2. <i>Он сделал себе новый инструмент.</i></li> </ol>	<p>A common term for 'just made' and 'just acquired', applicable to objects; to be distinguished from 'young'.</p> <p>Общее слово для 'только что сделанный' и 'только что приобретенный' по отношению к вещам; отличать от 'молодого'.</p>
62.	night	ночь	<ol style="list-style-type: none"> <li>1. <i>Night comes after the day.</i></li> <li>2. <i>The night is bright when the moon is shining.</i></li> <li>1. <i>Ночь сменяет день.</i></li> <li>2. <i>Ночь светла, когда светит луна.</i></li> </ol>	<p>'Night' as the time period devoid of sunshine, opposed to 'day'. Not to be confused with adverbial temporal usage ('by night', 'nightly'), if such forms are represented by a different root.</p> <p>Ночь как время суток без солнечного света, противопоставленное 'дню'. Отличать от адвербиального употребления (ночью) в том случае, если соответствующие формы представлены другим корнем.</p>
63.	nose	нос	<ol style="list-style-type: none"> <li>1. <i>While fighting, he hit him in the nose.</i></li> <li>2. <i>We breathe and smell with the nose.</i></li> <li>1. <i>Он в драке ударил его в нос.</i></li> <li>2. <i>Мы дышим и нюхаем носом.</i></li> </ol>	<p>Should be distinguished from 'nostrils', "internal nose" (i. e. internal organs), as well as 'flair' or 'scent'.</p> <p>Отличать от 'ноздря', «внутреннего носа» (т. е. внутренних органов) и 'нюха'.</p>

Number	ENG	RU	Contexts	Notes
64.	not	не	<ol style="list-style-type: none"> <li>1. <i>I do not understand</i> [in response to a question].</li> <li>2. <i>You do not / did not see</i> [X].</li> <li>3. <i>He does not / did not love</i> [X].</li> </ol> <ol style="list-style-type: none"> <li>1. <i>Я не понимаю</i> [ответ на непонятно заданный вопрос].</li> <li>2. <i>Ты не видишь</i> [X].</li> <li>3. <i>Он не любит</i> [X].</li> </ol>	<p>The basic negation of an assertion (indicative, <i>not</i> prohibitive). Used with the most simple or statistically frequent verbal forms (e.g., Present Indefinite for Modern English). If different tense / aspect etc. forms use different negative markers and no preferences can be easily formulated, the suppletive variants should be treated as synonyms. In languages that use a special negative verbal stem, this particular stem should be used. Should be strictly separated from the negative answer “No”.</p> <p>Базовое отрицание при утверждении (индикатив, не прохибитив), сочетающееся с самыми морфологически простыми и/или частотными временными формами (напр., настоящее время для русского языка). Если для разных видовременных форм используются разные маркеры отрицания и никаких предпочтений установить не удается, допустимо брать супплетивные варианты как синонимичные. В языках, выражающих отрицание с помощью специальной негативной глагольной основы, должна браться именно эта основа. Следует строго отличать от негативного ответа «Нет».</p>
65.	one	один	<ol style="list-style-type: none"> <li>1. <i>There is one leaf on the ground, but many leaves on the tree.</i></li> <li>2. <i>Here is one stone, here are two stones, here are three stones.</i></li> </ol> <ol style="list-style-type: none"> <li>1. <i>На земле лежит один лист, а на дереве висит много листьев.</i></li> <li>2. <i>Вот один камень, вот два камня, вот три камня.</i></li> </ol>	<p>Cardinal numeral. It is recommendable to keep this apart from the respective element of the numeric row <i>one, two, three, four...</i> used in counting. Same contexts as in No.91 ‘two’.</p> <p>Количественное числительное. Желательно отличать от соответствующего элемента при счете <i>один (раз), два, три, четыре...</i> Контексты те же, что и при №91 ‘два’.</p>
66.	rain	дождь	<ol style="list-style-type: none"> <li>1. <i>Rain is water from the sky.</i></li> <li>2. <i>Look at the rain, it is not very heavy yet.</i></li> </ol> <ol style="list-style-type: none"> <li>1. <i>Дождь — это небесная вода.</i></li> <li>2. <i>Посмотри на дождь, он пока еще не очень сильный.</i></li> </ol>	<p>Should be separated from specific kinds of rain (‘heavy’, ‘light’, ‘shower’, etc.). It is not recommended to fill the slot with the basic term for ‘rainy season’ (frequent practice, e.g., in Africanist tradition) or the verb ‘to rain’.</p> <p>Отличать от конкретных видов дождя (‘ливень’ и т. п.). Не рекомендуется заполнять эту позицию словом для ‘дождевого сезона’ (типичная замена для африканистики), а также глагольной формой ‘идти (о дожде)’.</p>



Number	ENG	RU	Contexts	Notes
67.	red	красный	[As on the RGB 255:255:0 and similar cards. The natural color of arterial blood from a fresh wound.] [Цвет артериальной крови, текущей из раны. Карточка RGB 255:0:0 и несколько оттенков].	
68.	road	дорога	1. <i>There is a road from my village to the next one, but no road to the far away village.</i> 2. <i>He is walking on the road.</i> 1. <i>Из моего селения в соседнее есть дорога, а в то дальнее селение дороги нет.</i> 2. <i>Он идет по дороге.</i>	The most typical kind of walkway for the given region (path, trail, serpentine etc.); in many cultures, 'path' is the correct equivalent (since 'road' may be taken to represent a cultural term imposed by a technologically superior civilization). Наиболее обычный тип дороги для данного ареала (тропа [но не звериная], серпантин и т. д.). Во многих случаях 'тропа' оказывается предпочтительнее 'дороги', т. к. противопоставлена последней как культурному термину, заимствованному от более технологически развитой цивилизации.
69.	root	корень	1. <i>This tree/bush has very deep roots.</i> 2. <i>He stumbled against the root of this tree/bush.</i> 3. <i>Look how long this root is.</i> 1. <i>У этого дерева/кустарника глубокие корни.</i> 2. <i>Он споткнулся о корень этого дерева/кустарника.</i> 3. <i>Посмотри, какой длинный корень.</i>	Generic term referring to roots of trees/shrubs. Not to be confused with 'tubers', 'edible roots' of specific plants, or grass roots. Общее название корня дерева (или кустарника). Отличать от корней травянистых растений; не путать также с клубнями и съедобными корнями отдельных растений.

Number	ENG	RU	Contexts	Notes
70a.	round (3D)	круглый	<ol style="list-style-type: none"> <li>1. <i>A round stone.</i></li> <li>2. <i>A round fruit.</i></li> </ol> <ol style="list-style-type: none"> <li>1. <i>Вот круглый камень.</i></li> <li>2. <i>Вот круглый плод.</i></li> </ol>	<p>3D round, i. e. of spherical objects. 3D круглый, т. е. о сферических объектах.</p>
70b.	round (2D)	круглый	<ol style="list-style-type: none"> <li>1. <i>A round dish.</i></li> <li>2. <i>A round leaf.</i></li> </ol> <ol style="list-style-type: none"> <li>1. <i>Вот круглое блюдо.</i></li> <li>2. <i>Вот круглый лист.</i></li> </ol>	<p>2D round, i. e. of flat objects. Some languages lexically distinguish a third kind of roundness (terete, round in section, e.g., a stick). This term is not included in the list. 2D круглый, т. е. о плоских и т. п. объектах. В некоторых языках лексически выделяется третий тип круглых объектов: круглый в сечении (напр., палка). Подобный термин не включается в список.</p>
71.	sand	песок	<ol style="list-style-type: none"> <li>1. <i>While digging the hole, he was throwing out the sand.</i></li> <li>2. <i>He took a handful of sand.</i></li> <li>3. <i>I don't know whether that site contains sand or earth.</i></li> </ol> <ol style="list-style-type: none"> <li>1. <i>Он рыл яму и выбрасывал наружу песок.</i></li> <li>2. <i>Он взял в горсть песка.</i></li> <li>3. <i>Я не знаю, что на том участке: земля или песок.</i></li> </ol>	<p>Same contexts as in No.21 'earth'. Not to be confused with 'dust', 'gravel' (= 'small stones'), 'coarse sand' (= 'with small stones in it'). Те же контексты, что и в №21 'земля'. Отличать от 'пыли', 'гальки' (= 'камешки') и т. п.</p>

Number	ENG	RU	Contexts	Notes
72.	to say	сказать	<ol style="list-style-type: none"> <li>1. <i>What did you say?</i> [= a request to repeat something that was misheard]</li> <li>2. He said that he would come tomorrow.</li> <li>1. <i>Что ты сказала?</i> [Переспрос после непонятной реплики]</li> <li>2. Он сказал, что завтра придет.</li> </ol>	<p>A neutral designation of a single speech act. Should be distinguished from durative action ('to talk'), but also not confused with certain special markers of direct speech, e.g., Old Indian <i>iti</i>.</p> <p>Нейтральное обозначение однократного речевого акта. Отличать от длительного действия ('говорить'), а также от специальных показателей прямой речи, таких, как санскр. <i>iti</i>.</p>
73.	to see	видеть	<ol style="list-style-type: none"> <li>1. <i>It is dark. I do not see you.</i></li> <li>2. <i>I see the animal.</i></li> <li>1. Темно, я тебя не вижу</li> <li>2. Я вижу зверя.</li> </ol>	<p>Durative process of perception. Should not be confused with intentional action ('to look') or various inchoative or punctive verbs ('glance', 'notice' etc.).</p> <p>Основное обозначение длительного процесса восприятия. Не следует смешивать с намеренным действием ('смотреть'), а также с различными инхоативами или пунктивными ('взглянуть', 'замечить' и т. п.).</p>
74.	seed	семя	<ol style="list-style-type: none"> <li>1. <i>When the seed falls on the ground, it grows into a plant.</i></li> <li>2. <i>Seeds were scattered on the ground.</i></li> <li>1. Если семя попадет в землю, оно прорастет.</li> <li>2. Семена рассыпались по земле.</li> </ol>	<p>Generic term for 'seed' of any plant; the word that can be applied to as many different plant species as possible. Not specifically 'fruit-stone', although the same root may be used for both small and large seeds. To be kept separate from more specialized terms ('seed for sowing', 'moist seed', 'seed of a particular group of plants, e.g., cereals').</p> <p>Общий термин для семени любого растения (проверить на сочетаемость с максимальным разнообразием видов). Отличать от 'косточки', хотя во многих языках 'большое' и 'маленькое' семя могут обозначаться одинаково. Также отличать от более узкоспециальных терминов ('посевные семена', 'влажное семя', 'семя опр. группы растений, напр. злаковых').</p>

Number	ENG	RU	Contexts	Notes
75.	to sit	сидеть	<ol style="list-style-type: none"> <li>1. <i>Man can stand, sit, or lie.</i></li> <li>2. <i>It is comfortable to sit on this stone/log.</i></li> <li>1. <i>Человек может стоять, сидеть и лежать.</i></li> <li>2. <i>На этом камне/бревне удобно сидеть.</i></li> </ol>	<p>Static verb. It is not recommended to fill in the slot with an inchoative form ('to sit down'). Should be separated from verbs depicting specific, "non-generic" ways of sitting, such as 'squatting', 'sitting in a lotus position' etc. See No.47 'to lie', No.80 'to stand'.</p> <p>Длительное состояние. Не рекомендуется подставлять инхоативный глагол 'сесть'. Отличать от «нестандартного» сидения в специфических позах (на корточках, в позе лотоса и т. п.). См. также №47 'лежать', №80 'стоять'.</p>
76.	skin	кожа	<ol style="list-style-type: none"> <li>1. <i>To paint on one's skin.</i></li> <li>2. <i>I have cut my skin.</i></li> <li>1. <i>Рисовать на коже.</i></li> <li>2. <i>Я порезал свою кожу.</i></li> </ol>	<p>Human skin. Кожа человека.</p>
77.	to sleep	спать	<ol style="list-style-type: none"> <li>1. <i>People usually sleep in the night and keep awake by day.</i></li> <li>2. <i>Do not wake him up, he is sleeping.</i></li> <li>1. <i>Обычно ночью люди спят, а днем бодрствуют.</i></li> <li>2. <i>Не буди его, он спит.</i></li> </ol>	<p>Static verb. Should be kept apart from inchoatives like 'to fall asleep'.</p> <p>Длительное состояние. Отличать от инхоативов вроде 'заснуть'.</p>
78.	small, little	маленький	<ol style="list-style-type: none"> <li>1. <i>This person has a big nose, that one has a small nose.</i></li> <li>2. <i>A small stone [leaf, tree, pool, etc...].</i></li> <li>1. <i>У этого человека большой нос, а у того — маленький.</i></li> <li>2. <i>Маленький камень [лист, дерево, водоем...].</i></li> </ol>	<p>An adjective or verb that may be applied to different objects. The contexts are the same as in No.5 'big'.</p> <p>Прилагательное/глагол, сочетающееся с максимальным количеством объектов. Контексты те же, что и для №5 'большой'.</p>

Number	ENG	RU	Contexts	Notes
79.	smoke	дым	<ol style="list-style-type: none"> <li>1. <i>With no wind, smoke from the bonfire rises straight up.</i></li> <li>2. <i>Smoke gets in my eyes.</i></li> <li>1. <i>Когда нет ветра, дым от костра поднимается прямо вверх.</i></li> <li>2. <i>Дым ест мне глаза.</i></li> </ol>	<p>Smoke as the result of combustion of wood, grass, dry dung (not after cremation). Should be distinguished from 'vapour' or specific kinds of smoke (e.g., '(thick) black smoke').</p> <p>Дым как результат горения дерева, травы, кизяка (но не после кремации). Следует отличать от 'пара', а также специальных видов дыма (напр., '(густой) черный дым').</p>
80.	to stand	стоять	<ol style="list-style-type: none"> <li>1. <i>Man can stand, sit, or lie.</i></li> <li>2. <i>He is standing on his own two feet.</i></li> <li>3. <i>I am tired of standing. I want to lie down.</i></li> <li>1. <i>Человек может стоять, сидеть и лежать.</i></li> <li>2. <i>Он стоит на двух ногах.</i></li> <li>3. <i>Я устал стоять, я хочу лечь.</i></li> </ol>	<p>Static verb. It is not recommended to fill the slot by an inchoative form like 'to stand up'. See No.47 'to lie', No.75 'to sit'.</p> <p>Длительное состояние. Не рекомендуется подставлять инхоативный глагол 'встать'. См. №47 'лежать', №75 'сидеть'.</p>
81.	star	звезда	<ol style="list-style-type: none"> <li>1. <i>At night we see the stars and the moon, during the day we see the sun.</i></li> <li>2. <i>There is a bright star over there.</i></li> <li>1. <i>Ночью на небе видны звезды и луна, а днем солнце.</i></li> <li>2. <i>Вон яркая звезда.</i></li> </ol>	<p>Should be kept apart from 'constellation' (a group of stars) as well as names of specific stars.</p> <p>Отличать от 'созвездия' (группы звезд) и наименований конкретных звезд.</p>

Number	ENG	RU	Contexts	Notes
82.	stone	камень	<ol style="list-style-type: none"> <li>1. He took a stone in his hand.</li> <li>2. He threw a stone at the bird and killed her.</li> <li>3. He used a stone to beat a peg into the ground.</li> <li>4. Stones can be of different colors.</li> </ol> <ol style="list-style-type: none"> <li>1. Он взял в руку камень.</li> <li>2. Он бросил камень в птицу и убил ее.</li> <li>3. Он забил в землю колышек камня.</li> <li>4. Камни бывают разных цветов.</li> </ol>	<p>An average-size stone that can be taken into one's hand (and used as a tool). It is preferable to separate this term from words for notably smaller stones ('pebble', 'grave', etc.) and notably bigger stones ('boulder', 'block' etc.).</p> <p>Камень среднего размера, который можно взять в руку (и использовать как оружие). Отличать от специализированных названий мелких камней ('галька', 'щебень' и т. п.) и крупных камней ('глыба', 'скала', 'валун' и т. п.).</p>
83.	sun	солнце	<ol style="list-style-type: none"> <li>1. At night we see the stars and the moon, during the day we see the sun.</li> <li>2. Look at the sun.</li> </ol> <ol style="list-style-type: none"> <li>1. Ночью на небе видны звезды и луна, а днем солнце.</li> <li>2. Посмотри на солнце.</li> </ol>	<p>Not to be confused with the abstract terms 'sunshine', 'heat of the sun' etc.</p> <p>Отличать от абстрактных терминов — 'солнечный свет', 'жара' и т. п.</p>

Number	ENG	RU	Contexts	Notes
84.	to swim	плыть, плавать	<ol style="list-style-type: none"> <li>1. <i>Is there anyone among your friends who cannot swim?</i></li> <li>2. <i>If you do not learn to swim, you will drown.</i></li> <li>3. <i>There is someone swimming out there.</i></li> </ol> <ol style="list-style-type: none"> <li>1. <i>Кто-нибудь из твоих друзей не умеет плавать?</i></li> <li>2. <i>Если ты не научишься плавать, ты утонешь.</i></li> <li>3. <i>Вон плавают какой-то человек.</i></li> </ol>	<p>Of human, not of fish, other animals, or inanimate objects. Should be kept apart from more specific terms like ‘to cross, to swim across (a water body)’, ‘to bathe / wash oneself’, or designations of specific techniques of swimming.</p> <p>О человеке, не о рыбах, других животных или неодушевленных предметах. Следует отличать от более конкретного ‘переплывать’ (т. е. пересекать водоем), а также от обозначений конкретных техник плавания и от ‘купаться’, ‘мыться’.</p>
85.	tail	хвост	<ol style="list-style-type: none"> <li>1. <i>Animals have tails, man does not have a tail.</i></li> <li>2. <i>[The beast] is swishing its tail to drive away the flies.</i></li> <li>3. <i>He cut off [the beast's] tail.</i></li> </ol> <ol style="list-style-type: none"> <li>1. <i>У зверей есть хвост, а у человека нету.</i></li> <li>2. <i>[Зверь] машет хвостом, отгоняя насекомых.</i></li> <li>3. <i>Он отрезал у [зверя] хвост.</i></li> </ol>	<p>Of mammals (not birds, fish, snakes, etc.). Generic term, applied to as many different species as possible. Not to be confused with possible narrow terms, such as ‘small tail as of a goatling’ etc.</p> <p>О млекопитающем (не о рыбах, птицах, змеях и т. п.). Общий термин, сочетающийся с максимальным разнообразием видов. Отличать от возможных специализированных обозначений, напр., ‘хвостик козленка’ и т. п.</p>



Number	ENG	RU	Contexts	Notes
86.	that	тот	<p>1. <i>I am well acquainted with this tale, but do not remember that one.</i></p> <p>2. <i>I know this man, but I do not know that one.</i></p> <p>1. Я хорошо знаю это предание/сказку и совсем плохо помню то предание/сказку.</p> <p>2. Я знаком с этим человеком и не знаком с тем.</p>	<p>Demonstrative pronoun with adjectival functions. Not to be confused with 'another'. A complex case, since in a great part of the world's languages the pronominal systems of spatial deixis are much more complex than the binary opposition <i>this</i> ~ <i>that</i> in Modern English. Such systems can include several degrees of proximity (both on the horizontal and vertical axis), references to the location of the speaker and the addressee, or indication of some special features of the object (e.g., animate/inanimate or visible/invisible). It is recommended to try to reveal the most basic binary opposition between 'this' (close) and 'that' (far away). If this is completely impossible, several (standardly three) pronouns should be listed as synonyms.</p> <p>Указательное местоимение с адъективными функциями. Отличать от 'другой'. Сложный случай, т. к. в большом числе языков местоименные системы пространственного дейксиса намного более развиты, нежели русская бинарная оппозиция <i>этот</i> ~ <i>тот</i>. Подобные системы могут включать в себя несколько степеней близости (причем как по горизонтальной, так и вертикальной осям), указание на место расположения говорящего и слушающего, спецификацию определенных свойств объекта (одушевленный ~ неодушевленный, видимый ~ невидимый). Рекомендуется попробовать выделить наиболее базовую бинарную оппозицию 'этот' — 'тот'. Если это невозможно, то несколько (стандартно три) местоимений должны быть перечислены как синонимы.</p>
87.	this	этот	<p>1. <i>I am well acquainted with this tale, but do not remember that one.</i></p> <p>2. <i>I know this man, but I do not know that one.</i></p> <p>1. Я хорошо знаю это предание/сказку и совсем плохо помню то предание/сказку.</p> <p>2. Я знаком с этим человеком и не знаком с тем.</p>	<p>See No.86 'that'. См. №86 'тот'.</p>

Number	ENG	RU	Contexts	Notes
88.	tongue	язык	<ol style="list-style-type: none"> <li>1. <i>He showed him his tongue.</i></li> <li>2. <i>He bit his tongue while falling, and now it hurts.</i></li> </ol> <ol style="list-style-type: none"> <li>1. Он показал ему язык.</li> <li>2. Когда он упал, он прикусил язык. Ему больно.</li> </ol>	Часть тела, не 'речь'.
89.	tooth	зуб	<ol style="list-style-type: none"> <li>1. <i>The grown-up person has a lot of teeth, the newborn baby has no teeth.</i></li> <li>2. <i>Old men usually have few teeth left.</i></li> <li>3. <i>You can break a tooth while eating hard food.</i></li> </ol> <ol style="list-style-type: none"> <li>1. У взрослого человека много зубов, а у новорожденных зубов нету.</li> <li>2. У стариков обычно остается мало зубов.</li> <li>3. О твердую пищу можно сломать зуб.</li> </ol>	<p>Generic term; to be strictly separated from specific kinds of teeth (molars, cuspids, wisdom teeth etc.). If no single term can be determined (e.g., in Basque, which lexically separates 'back tooth, molar' from 'front tooth, incisor' and uses the compound of both to denote 'teeth' in general), two roots can be taken as synonyms.</p> <p>Общий термин; строго отличать от наименований конкретных групп зубов (клыки, коренные зубы, зубы мудрости и т. п.). Если общий термин неопределим (ср. ситуацию в баскском, где 'задний (коренной) зуб' лексически противопоставлен 'переднему зубу, резцу', а для выражения идеи 'зубов' вообще используется композит из этих двух корней), разрешается брать два синонима.</p>

Number	ENG	RU	Contexts	Notes
90.	tree	дерево	<ol style="list-style-type: none"> <li>All trees <i>have</i> roots.</li> <li>The tree <i>has</i> a thick trunk, the bush <i>has</i> a thin one.</li> <li>There is a tall tree <i>growing over</i> there.</li> <li>He cut <i>down</i> the tree.</li> </ol> <ol style="list-style-type: none"> <li>Все деревья <i>имеют</i> корни.</li> <li>У дерева <i>толстый ствол, а у кустика тонкий</i>.</li> <li>Вон <i>растет</i> высокое дерево.</li> <li>Он <i>срубил</i> дерево.</li> </ol>	<p>Not to be confused with designations of particular tree species, as well as the collective 'forest' or 'wood (material)'.</p> <p>Не путать с наименованиями конкретных видов деревьев; отличать также от собирательного 'лес' и 'дерево как материала (= древесина)'.</p>
91.	two	два	<ol style="list-style-type: none"> <li>There are two leaves on the ground, but many leaves on the tree.</li> <li>Here is one stone, here are two stones, here are three stones.</li> </ol> <ol style="list-style-type: none"> <li>На земле <i>лежат</i> два листа, а на дереве <i>висит</i> много листьев.</li> <li>Вот один камень, вот два камня, вот три камня.</li> </ol>	<p>Cardinal numeral. It is recommendable to keep this apart from the respective element of the numeric row <i>one, two, three, four...</i> used in counting. Same contexts as in No.65 'one'.</p> <p>Количественное числительное. Желательно отличать от соответствующего элемента при счете <i>один (раз), два, три, четыре...</i> Контексты те же, что и при №65 'один'.</p>

Number	ENG	RU	Contexts	Notes
92.	warm	теплый	<ol style="list-style-type: none"> <li>1. <i>The water is warm, not too hot and not too cold.</i></li> <li>2. <i>The stone is warm under the sun.</i></li> <li>3. <i>A warm wind is blowing.</i></li> </ol> <ol style="list-style-type: none"> <li>1. <i>Эта вода не горячая и не холодная, она теплая.</i></li> <li>2. <i>Теплый на солнце камень.</i></li> <li>3. <i>Дует теплый ветер.</i></li> </ol>	<p>'Warm' can be understood as 'having a temperature slightly above the norm', with a generally positive connotation. This is a dubious and often hard to elicit item on the Swadesh list; while it should be strictly distinguished from 'hot', it may be useful to add 'hot' as a synonym (92b?) or, in the future, even to replace 'warm' with 'hot' completely.</p> <p>'Теплый' следует понимать как 'температура слегка выше нормы', с позитивными коннотациями. Несколько сомнительный элемент списка, для которого не всегда удается подобрать адекватный эквивалент; следует строго отличать от 'горячего', но, возможно, имеет смысл добавить и 'горячий' в качестве синонима (92b?), а в будущем — вообще заменить 'теплый' на 'горячий'.</p>
93.	water	вода	<ol style="list-style-type: none"> <li>1. <i>People cannot live without water.</i></li> <li>2. <i>He is drinking water.</i></li> <li>3. <i>She dipped some water in her jug.</i></li> </ol> <ol style="list-style-type: none"> <li>1. <i>Человек не может жить без воды.</i></li> <li>2. <i>Он пьет воду.</i></li> <li>3. <i>Она набрала в сосуд воды.</i></li> </ol>	<p>Sweet water, good for drinking. Should be kept apart from more specialized terms: 'salt water' (= 'sea (water)'), 'foul water', 'running / flowing water' (= 'stream', 'river'), etc.</p> <p>Пресная вода, пригодная для питья. Отличать от специализированных терминов: 'морская вода', 'загнившая вода', 'текущая вода' (= струя, родник, ручей) и т. п.</p>
94a.	we (incl.)	мы (incl.)	<ol style="list-style-type: none"> <li>1. <i>We are friends (with you), but he is not our friend.</i></li> <li>1. <i>Мы (с тобой) друзья, а он — не наш друг.</i></li> </ol>	<p>Stylistically neutral 1<sup>st</sup> pl. person; various honorific forms must be excluded. Dual forms, if different, should not be taken into account (for simplicity reasons, since special dual stems for non-singular pronouns are not very widespread). Direct stem.</p> <p>Стилистически нейтральное личное местоимение 1 л. мн. ч. Различные формы вежливости должны быть исключены. Формы двойственного числа, если таковые существуют, не учитываются в целях экономии (специальные двойственные основы местоимений в языках мира встречаются не очень часто). Прямая основа.</p>
94b.	we (incl.)	мы (incl.)		<p>Indirect stem of the same pronoun (for languages which distinguish between the two on a paradigmatic level).</p> <p>Это же, но косвенная основа (для языков, которые имеют соответствующее различие на парадигматическом уровне).</p>

Number	ENG	RU	Contexts	Notes
94c-d	we (excl.)			If the languages in question make a systematic differentiation between exclusive and inclusive forms, both should be taken into account as synonyms. Direct and indirect stems. Если языки систематически проводят различие между эксклюзивом и инклюзивом, к сравнению в качестве синонимов привлекаются обе формы. Прямая и косвенная основы.
95.	what	что	<ol style="list-style-type: none"> <li>1. <i>Who came and what did he bring?</i></li> <li>2. <i>What are you making out of this piece of wood?</i></li> </ol> <ol style="list-style-type: none"> <li>1. Кто пришел и что он принес?</li> <li>2. Что ты изготавливаешь из этой древесины?</li> </ol>	<p>Interrogative (not relative) pronoun, referring to an inanimate object; syntactically should be able to occupy the position of direct object.</p> <p>Вопросительное (не относительное) местоимение, относящееся к неодушевленному объекту; в синтаксическом отношении должно встречаться в позиции прямого дополнения.</p>
96.	white	белый	<p>[As on the RGB 255:255:255 card. The natural color of snow, paper, babies' teeth etc.]</p> <p>[Карточка RGB 255:255:255. Естественный цвет снега, бумаги, зубов у маленьких детей и т. п.].</p>	
97.	who	кто	<ol style="list-style-type: none"> <li>1. <i>Who came and what did he bring?</i></li> <li>2. <i>Who killed this dog?</i></li> </ol> <ol style="list-style-type: none"> <li>1. Кто пришел и что он принес?</li> <li>2. Кто убил эту собаку?</li> </ol>	<p>Interrogative (not relative) pronoun, referring to an animate subject/object.</p> <p>Вопросительное (не относительное) местоимение, относящееся к одушевленному субъекту/объекту.</p>

Number	ENG	RU	Contexts	Notes
98.	woman	женщина	<p>1. <i>Men are doing [X], women are making [Y].</i></p> <p>2. <i>Women bear children, men do not.</i></p> <p>1. <i>Мужчины делают [X], а женщины делают [Y]</i></p> <p>2. <i>Женщины рожают детей, а мужчины — нет.</i></p>	<p>Adult human being of the female sex, antonymous to 'man'. Should be kept apart from 'female' (of any species) if the latter is distinct from 'woman' (female human). Context No. 1 is culturally conditioned, in that X and Y should be understood as designations of typical male / female occupations for the given society (e.g., 'hunting', 'cooking', etc.). If, instead of a general term for 'woman', one finds differentiation by age groups, the respective lexemes for the age of <i>ca.</i> 20–45 must be treated as synonymous. If, instead of a general term for 'woman', there is an opposition by marital status ('unmarried' ~ 'married' ~ 'widowed'), it is advisable to use the word for 'married woman' as the most prototypical concept. Contexts are the same as in No.51 'man (male)'.</p> <p>Взрослый человек женского пола. Отличать от понятия 'самка' ('female'), если для последнего существует отдельное слово. Контекст 1 является культурно обусловленным; здесь следует подставлять наименования типичных для данного социума мужского и женского занятий (охота, готовка и т. п.). Если в языке нет общего слова для 'женщины', а имеется спецификация по возрастным группам, то соответствующие лексемы для возраста примерно 20—45 лет надо брать как синонимы. Если в языке нет общего слова 'женщина', а имеется противопоставление по семейному положению (незамужняя ~ замужняя ~ овдовевшая), рекомендуется выбирать слово 'замужняя женщина' как наиболее прототипическое понятие. Те же контексты, что и для №51 'мужчина'.</p>
99.	yellow	желтый	<p>[As on the RGB 255:255:0 card and similar ones. The natural color of egg yolk.]</p> <p>[Естественный цвет желтка в яйце. Карточка RGB 255:255:0 и несколько оттенков.]</p>	
100a.	you (thou)	ты		<p>Stylistically neutral 2<sup>nd</sup> sg. person; various honorific forms must be excluded. Direct stem.</p> <p>Стилистически нейтральное личное местоимение 2 л. ед. ч. Различные формы вежливости должны быть исключены. Прямая основа.</p>

Number	ENG	RU	Contexts	Notes
100b.	you (thou)	тебя		Indirect stem of the same pronoun (for languages which distinguish between the two on a paradigmatic level). Это же, но косвенная основа (для языков, которые имеют соответствующее различие на парадигматическом уровне).
101.	far +	далеко +	<p>1. <i>That tree is far away, I cannot see it; this one is quite near, I can see it clearly.</i></p> <p>1. Это дерево далеко, я его плохо вижу, а это дерево близко, я его хорошо вижу.</p>	For languages which distinguish between adjectives and adverbs, adverbial use of the word is preferred. Cf. No.103 'near'. Имеется в виду именно адverbиальное употребление, а не адъективное (для языков, различающих эти части речи). Ср. №103 'близко'.
102.	heavy +	тяжелый +	<p>1. <i>He is carrying a heavy load, but my load is light.</i></p> <p>2. <i>This stone is too heavy.</i></p> <p>1. Он несет тяжелый груз, а у меня легкий груз.</p> <p>2. Этот камень слишком тяжелый.</p>	Said of an object with a large mass. О предмете, обладающем большим весом.
103.	near +	близко +	<p>1. <i>That tree is far away, I cannot see it; this one is quite near, I can see it clearly.</i></p> <p>1. Это дерево далеко, я его плохо вижу, а это дерево близко, я его хорошо вижу.</p>	For languages which distinguish between adjectives and adverbs, adverbial use of the word is preferred. Cf. No.101 'far'. Имеется в виду именно адverbиальное употребление, а не адъективное (для языков, различающих эти части речи). Ср. №101 'далеко'.



Number	ENG	RU	Contexts	Notes
104.	salt +	соль +	<ol style="list-style-type: none"> <li>1. <i>People have to eat salt.</i></li> <li>2. <i>This kid wants some salt.</i></li> <li>3. <i>Here is some white salt.</i></li> </ol> <ol style="list-style-type: none"> <li>1. Людям нужно есть соль.</li> <li>2. Этот ребенок хочет соли.</li> <li>3. Вот белая соль.</li> </ol>	<p>Salt as natural product. Should be distinguished from specific designation of 'packaged salt', bought in stores; also from 'saltlick', 'saltpan', etc.</p> <p>Натуральный продукт. Следует отличать от покупной соли, а также от названий мест, содержащих соль ('соляное озеро', 'лизунец' и т. п.).</p>
105.	short +	короткий +	<ol style="list-style-type: none"> <li>1. <i>Here is a long stick, here is a short one.</i></li> <li>2. <i>This rope is long, I need a short one.</i></li> </ol> <ol style="list-style-type: none"> <li>1. Вот длинная палка, а вот короткая.</li> <li>2. Эта веревка длинная, мне нужна короткая.</li> </ol>	<p>Of physical objects. Not to be confused with 'low', 'short (of stature)' (vertical, not horizontal) in languages that distinguish between the two. Antonymous to No.49 'long'.</p> <p>О физических объектах. Не смешивать с 'низкий' (вертикальное положение) в тех языках, которые различают эти два слова. Антоним — №49 'длинный'.</p>
106.	snake +	змея +	<ol style="list-style-type: none"> <li>1. <i>A snake has eyes, a worm has no eyes.</i></li> <li>2. <i>Birds fly in the sky, snakes crawl on the ground.</i></li> <li>3. <i>The snake is basking in the sun.</i></li> </ol> <ol style="list-style-type: none"> <li>1. У змеи есть глаза, а у червя нету.</li> <li>2. Птицы летают в небе, а змеи ползают по земле.</li> <li>3. Змея греется на солнце.</li> </ol>	<p>Generic term for 'snake' or 'small snake' (see context 2). Should be kept apart from No.109 'worm' (context 1).</p> <p>Общий термин для 'змеи' или 'маленькой змеи' (см. контекст 2). Следует отличать от №109 'червь' (см. контекст 1).</p>

Number	ENG	RU	Contexts	Notes
107a.	thin (2D) +	тонкий +	<ol style="list-style-type: none"> <li>1. <i>This is a thin skin.</i></li> <li>2. <i>This is a thin leaf.</i></li> </ol> <ol style="list-style-type: none"> <li>1. <i>Вот тонкая шкура.</i></li> <li>2. <i>Вот тонкий лист.</i></li> </ol>	<p>2D thin, i. e. of flat objects. 2D тонкий, т. е. о плоских объектах.</p>
107b.	thin (1D) +	тонкий +	<ol style="list-style-type: none"> <li>1. <i>This is a thin rope/sinew.</i></li> <li>2. <i>This is a thin branch.</i></li> </ol> <ol style="list-style-type: none"> <li>1. <i>Вот тонкая веревка/жила.</i></li> <li>2. <i>Вот тонкая ветка.</i></li> </ol>	<p>1D thin, i. e. of elongated objects. 1D тонкий, т. е. о вытянутых объектах.</p>
108.	wind +	ветер +	<ol style="list-style-type: none"> <li>1. <i>Clouds are driven by the wind.</i></li> <li>2. <i>He wet his finger to learn from where the wind was blowing.</i></li> <li>3. <i>What (kind of) wind blows in summer/winter?</i></li> </ol> <ol style="list-style-type: none"> <li>1. <i>Облака гонит ветер.</i></li> <li>2. <i>Он намочил палец, чтобы узнать, откуда ветер.</i></li> <li>3. <i>Какой ветер обычно бывает летом/зимой?</i></li> </ol>	<p>Generic term for 'wind' as movement of air; should be separated from more narrow designations of particular winds (characterized, e.g., by a certain direction or a certain level of intensity), as well as special adverbial forms such as 'alee', etc.</p> <p>Общий термин: 'ветер' как движение воздуха. Следует отличать от конкретных наименований ветров (различающихся, напр., по сторонам света или по степени силы). Также отличать от наречий 'с подветренной стороны' и 'с заветренной стороны'.</p>

Number	ENG	RU	Contexts	Notes
109.	worm +	червь +	<ol style="list-style-type: none"> <li>1. <i>A snake has eyes, a worm has no eyes.</i></li> <li>2. <i>A worm has crawled out of the ground over there.</i></li> </ol> <ol style="list-style-type: none"> <li>1. <i>У змеи есть глаза, а у червя нету.</i></li> <li>2. <i>Вон червь выполз из земли.</i></li> </ol>	<p>Earthworm. Not to be confused with №106 'snake' or with 'intestinal worm' (e.g., 'tapeworm') and 'larvae' (incl. 'caterpillars').</p> <p>Земляной червь. Отличать от №106 'змея', а также от 'глиста' и 'личинки' (вкл. 'гусеница').</p>
110.	year +	год +	<ol style="list-style-type: none"> <li>1. <i>The first month of the year is called [X].</i></li> <li>2. <i>How many months are there in a year?</i></li> </ol> <ol style="list-style-type: none"> <li>1. <i>Первый месяц в году называется [X].</i></li> <li>2. <i>Сколько месяцев в году?</i></li> </ol>	<p>Should be distinguished from adverbial forms ('annually', 'all the year round', etc.); from specific nominal or adverbial designations of particular time periods ('this year', 'past year', etc.); from special terms for half-years or seasons; from 'year' as an age marker ('he is five years old', etc.). If the language has no common term for 'year' whatsoever, the slot should remain empty.</p> <p>Отличать от наречных образований 'ежегодно', 'круглогодично' и т. п.; от именных/наречных обозначений конкретных временных отрезков ('этот год', 'прошлый год' и т. п.); от названий отдельных полугодий; от года как единицы возраста ('ему пять лет' и т. п.). Если общий термин 'год' в языке отсутствует, строка не заполняется.</p>

## Uto-Aztecan comparative linguistics and etymological databases

The paper presents a general survey of Uto-Aztecan historical studies. It discusses the genetic classification of the family, phonological reconstructions of individual daughter languages and of Proto-Uto-Aztecan itself. Lexical material used in the paper is taken from electronic databases created for the “Evolution of Human Languages” Project (Santa Fe Institute).

*Keywords:* Uto-Aztecan languages, lexicostatistics, comparative linguistics, etymological databases.

The project “Evolution of Human Languages” was launched in 2002 by Murray Gell-Mann and Sergei Starostin. One of its main aims is to build a network of etymological databases for all language families of the world, in which, for each known family, we plan to have:

- a system of phonological correspondences connecting many, if not all, languages within this family;
- an extensive etymological corpus supporting these correspondences;
- a lexicostatistical classification based on reliable etymologies.

After this type of work was performed for some better known families (results are available at [starling.rinet.ru](http://starling.rinet.ru)), it became clear that in many cases, instead of a full-scale study of every language family, we can only conduct a preliminary survey based on existing material.

The current paper briefly presents the results of such a survey on Uto-Aztecan (UA), one the better known language families of Northern America. Due to the nature of the paper, no detailed discussion of literature and alternative views is included.

The family consists of about 60 languages. Its generally assumed classification is given in [Campbell 1997: 134]:

### Uto-Aztecan

#### *Northern Uto-Aztecan*

##### Numic

##### Western

Northern Paiute (Oregon, Idaho, Nevada)

Monache (California) [obsolescent]

##### Central

Shoshone, Panamint [obsolescent], Comanche [obsolescent] (Nevada, Utah, Wyoming, Oklahoma)

##### Southern

Southern Paiute (Utah, Nevada, California, Arizona)

Ute, Chemehuevi [obsolescent] (Utah, Colorado, California, Arizona)

Kawaiisu [obsolescent] (California)

Tübatulabal [moribund] (California)

## Takic

Serran: Serrano [moribund], Kitanemuk [extinct] (California)  
 Cahuilla [moribund?], Cupeño [moribund] (California)  
 Luiseño-Juaneño [obsolescent] (California)  
 Gabrielino [extinct], Fernandeseño [extinct] (California)

Hopi (Arizona)

*Southern Uto-Aztecan*

## Pimic (Tepiman)

Pima-Papago (Upper Piman) (Arizona, Sonora)  
 Pima Bajo (Lower Piman, Nevome) (Sonora)  
 Northern Tepehuan, Southern Tepehuan (Sonora, Durango, Jalisco)  
 Tepecano [extinct] (Jalisco)

## Taracahitic

## Tarahumaran

Tarahumara (Chihuahua)  
 Guarijio (Chihuahua, Sonora)  
 Tubar [extinct?] (Chihuahua)  
 Cahitan (Yaqui-Mayo-Cahita) (Arizona, Sonora)

## Opatan

Opata [extinct] (Sonora)  
 Eudeve [extinct] (Sonora)

## Corachol-Aztecan

## Cora-Huichol

Cora (Nayarit)  
 Huichol (Nayarit, Jalisco)

## Nahuan (Aztecan)

Pochutec [extinct] (Oaxaca)  
 Core Nahua

Pipil (El Salvador, extinct in Guatemala and Nicaragua)  
 Nahuatl: many dialects (Mexico).

“[Languages] known to have fewer than 10 speakers are specified as ‘moribund’; languages with more than 10 but fewer than 100 speakers are labeled ‘obsolescent’” (p. 133).

An additional list of extinct languages which may have belonged to UA is given on pp. 133–134 of Campbell’s book as well as in other sources, like.

In a somewhat impressionistic way, typical of the rest of the book as well, Campbell does not justify his classification. The only modern UA classification that is based on complete, explicit presentation of the evidence is W. R. Miller’s paper from 1984.<sup>1</sup> Using lexicostatistics, Miller (p. 13) comes to the following conclusion: “thus there seem to be five groups that are more or less coordinate with each other: Numic, Tübatulabal, Takic, Hopi and SUA” (South Uto-Aztecan); this is supported by the matrix of ‘cognate density’. Further investigation of this matrix (see the table on p. 114 below) suggests, however, that lexicostatistics allows to generate a different classification.

Let us reduce Miller’s matrix to a compressed version, where each transparently obvious group is represented by the lowest percentage that one of its members shares with other groups. This version looks as follows:

<sup>1</sup>Hale’s earlier lexicostatistical classification [Hale 1958–59] also presents the data.

	*Numic	*Takic	Hopi	*Sonoran	*Aztecan
*Numic	×	22	22	23	14
*Takic	22	×	26	24	18
Hopi	22	26	×	26	24
*Sonoran	23	24	26	×	29
*Aztecan	14	18	24	29	×

It can be seen that \*Numic and \*Aztecan form two ends of the spectrum, while the other groups are somewhere in between. Normally, \*Sonoran should be grouped with \*Aztecan, but it shows much higher figures with \*Numic. Such discrepancies are a clear indication that we are dealing with the effects of undetected borrowings within related language groups.

Restricting ourselves to reconstructed intermediate 50-item proto-lists<sup>2</sup>, we can overcome this obstacle and obtain the following Uto-Aztecan classification (utza.dbf<sup>3</sup>):

### Uto-Aztecan

1. \*Numic
2. Main group
  - a. Central:
    - (i) Tubatulabal, Hopi  
Kitanemuk
    - (ii) \*Cupan
  - b. Southern:
    - (i) \*Sonoran
    - (ii) \*Aztec

The weakest point of this classification is that it places Hopi close to Tübatulabal; this may also reflect later contacts between the two, not detected so far.

According to glottochronological calculations<sup>4</sup> based on the complete 100-item lists for attested languages, the Proto-UA language began to disintegrate approximately 4,500 – 4,200 years ago. It should be mentioned, however, that calculations based on Sergei Yakhontov/Sergei Starostin's 35-item list reveal an entirely different age – only 3,500 years. This is surprising, since usually calculations based on the shorter list generate slightly older dates than those based on the longer list. The reasons for such a discrepancy remain unknown.

### Numic languages

The Numic group is usually divided into three branches, with two languages each – Southern: Kawaiisu and Ute (Chemehuevi, Southern Paiute and Ute); Central: Comanche and Shoshoni (Panamint and Shoshoni dialects); Western: Mono and Northern Paiute.

This classification is accepted by Miller (1984), but the matrix of similarities given in his paper suggests a different tree structure:

<sup>2</sup>See [G. Starostin: 2009].

<sup>3</sup>All of the databases mentioned in this paper will be available on EHL's website (<http://starling.rinet.ru>) in the near future and can also be obtained upon request from the author.

<sup>4</sup>If not indicated otherwise, all the lexicostatistical and glottochronological results have been generated with the aid of STARLING, a computerized linguistic environment developed by S. Starostin.

	Mon	NPa	Pan	Sho	Com	Kaw	Che	Spa	Ute
Mono	×	77	59	58	57	52	50	53	52
N.Paiute	77	×	58	58	58	56	55	58	57
Panamint	59	58	×	87	79	54	61	62	59
Shoshoni	58	58	87	×	88	55	58	62	61
Comanche	57	58	79	88	×	49	54	59	59
Kawaiisu	52	56	54	55	49	×	75	79	76
Chemehuevi	50	55	61	58	54	75	×	86	78
S.Paiute	53	58	62	62	59	79	86	×	87
Ute	52	57	59	61	59	76	78	87	×

Numic (more than 49%)

A. Central-Northern Numic (more than 57%)

1. Western: Mono and Northern Paiute (77%)
2. Central: Panamint, Shoshoni, and Comanche (more than 79%)

B. Southern (more than 75%)

3. Ute: Chemehuevi, Southern Paiute, and Ute (more than 78%)
4. Kawaiisu

The same classification is obtained using S. Starostin's 35-item list (numi.dbf).

### Southern Numic languages

Phonologically, the Ute languages are very similar to each other. Southern Paiute, studied in [Sapir 1930], has the following system:

Initial consonants:

p-	m-	w-	
t-	n-		
c-		y-	s-
k-			
k <sup>w</sup> -			
ʔ-			h-

Word-medial position (single consonants and clusters):

-v-	-pp-	-mp-	-mm-		-w-	
-r-	-tt-	-nt-, -nč-	-nn-	-n-		
-c-	-cc-	-nc-			-y-	-ss-
-ʏ-	-kk-	-ŋk-		(-ŋ-)		
-ʏ <sup>w</sup> -	-kk <sup>w</sup> -	-ŋk <sup>w</sup> -		-ŋ <sup>w</sup> -		
-ʔ-						-h-

Final consonants: none



Vowels (short and long):

i, i:	ɨ, ɨ:	u, u:
e, e:		o, o:
	a, a:	

The language is well known for its “consonant gradation”. Sapir (1930, 62) talks about spirantization (S), gemination (G), and nasalization (N):

S: *uɣ<sup>w</sup>i-vi* ‘grass’+Abs

G: *kučča-ppi* ‘ashes’+Abs

N: *miyi-mpici* ‘gopher’+Abs

Morphophonemic [Sapir 1931], accentual [Miller 1980], and other explanations were suggested for these phenomena, but the real reason for “gradations” seems to be purely historical. A simple reconstruction may be based on the following considerations:

– the stem in Southern Paiute is usually formed by two syllables and thus may have the form of (a) CVPPV, where PP is any geminated stop, affricate or ?; (b) CVPV, where P is any other consonant or a cluster of nasal + stop;

– in most cases Gemination is associated with the first type of stems, whereas Spirantization is mainly found within the second: (a) CVPPV + G, (b) CVPV + S;

– therefore, Spirantization and Gemination, predetermined by the consonants in the second syllable, must have developed later in the history of the languages;

There are, however, also other cases, each with a significantly smaller number of examples: (c) CVPPV + S: *kicci-vi* ‘saliva’ Abs, (d) CVPV + G: *movi-ppi* ‘nose’ Abs. Historically these can be regarded as showing traces of final consonants: \*h (or \*s) and \*ʔ. Nasalization can be explained as a trace of a lost final nasal: (e) \*CVPVN, \*CVPPVN.

As a result, we can identify five types of stems in Southern Paiute:

(a) \*CVPPV#: \**kučča*# > *kučča-ppi* ‘ashes’

(b) \*CVPV#: \**uɣ<sup>w</sup>i*# > *uɣ<sup>w</sup>i-vi* ‘grass’

(c) \*CVPPVh: \**kiccih* > *kicci-vi* ‘saliva’

(d) \*CVPVʔ: \**moviʔ* > *movi-ppi* ‘nose’

(e) \*CVPVN, \*CVPPVN: *miyiN* > *miyiN-mpici* ‘gopher’, \**wikkuN* > *wikku-mpicc* ‘buzzard’

The same three types of “consonant gradation” are found in the other Ute languages as well as Kawaiisu.

The Kawaiisu phonological system, defined on the original stems of the language, is found in [Zigmond et al. 1990, pp. 5–6]:

p	v	b	m	w	
t	r	d	n		
c	ʃ				s
č	ʒ				š
k	ɣ				
k <sup>w</sup>	ɣ <sup>w</sup>				
ʔ					h

i, i:	ɨ, ɨ:	u, u:
e, e:		o, o:
	a, a:	

The Proto-Southern Numic system of consonants is reconstructed as follows:

Finals:

*#	*-ʔ	*-h	*-N
----	-----	-----	-----

Initial position:

*p	*m	*w	
*t	*n		*s
*c		*y	
*k	*ŋ		
*k <sup>w</sup>			
*ʔ			*h

Medial position:

*-b-	*-pp-	*-m-	*-mm	*-w-	
*-d-	*-tt-	*-n-			*-s-
*-ʒ-	*-cc-			*-y-	
*-g-	*-kk-	*-ŋ-			
*-g <sup>w</sup> -	*-kk <sup>w</sup> -				
	*ʔ				*-h-

The supporting correspondences are given in the following table:<sup>5</sup>

Kawaiisu	Proto-S. Numic	S. Paiute	Chemihuevi	Ute	Note
p	*p-	p	p	p	
v	*-b-	v	v	v	
p	*-pp-	pp	p	p	
b	*-mb-	mp	mp	p	< *N-pi
m	*m-	m	m	m	
m	*-m-	mm	m	m	
w	*w-	w	w	w	
w	*-w-	#	#	#	
w		ŋ <sup>w</sup>	w	w	-VN, iwV
t	*t-	t	t	t	
r	*-d-	r	r	r	

<sup>5</sup> The sources are as follows: Kawaiisu — [Zigmond, et al. 1990]; Southern Paiute — [Sapir 1931]; Chemihuevi — [Press 1979]; Ute — [Givon 1979].

Kawaiisu	Proto-S. Numic	S. Paiute	Chemihuevi	Ute	Note
t	*-tt-	tt	t	t	
n	*n-	n	n	n	
n	*-n-	n	n	n	
c	*c-	c	c	č	
ʒ, ʒ̣	*-ʒ-	c	c	č	
c	*-cc-	cc	c	č	
s, š	*s-	s	s	s	
s, š	*-s-	ss	s	s	
y	*y- ,	y	y	y	
y	*-y-	y	y	y	
k	*k-	k	k	k	
g	*-g-	g	g	g	
k	*-kk-	kk	k	k	
ŋ	*-nk-	ŋk	ŋk	k	
n	*-ŋ-	ŋ	ŋ, (u+m+u)	#	
k <sup>w</sup>	*k <sup>w</sup> -	k <sup>w</sup>	k <sup>w</sup>	k <sup>w</sup>	
ɣ <sup>w</sup>	*-g <sup>w</sup> -	w	g <sup>w</sup>	ɣ <sup>w</sup> -, ɣ(+i)-	
k <sup>w</sup>	*-kk <sup>w</sup> -	kk <sup>w</sup>	K <sup>w</sup>	k <sup>w</sup>	
ʔ	*ʔ-	#, ʔ	#, ʔ	#	
ʔ	*-ʔ-	#, ʔ	ʔ	ʔ	
h	*h-	ʔ, #	h	ʔ, #	
h	*-h-	#	h	#	

The etymological data supporting these correspondences can be found in the database numiet.dbf. The vocalic system of Proto-Southern Numic requires further investigation.

### Central Numic languages

All Central Numic languages have six vowels (i, e, i, a, u, and o) that can be short or long. The systems of initial consonants of modern Central Numic languages are identical to those of Southern Numic, while in the medial position there occur many more phonemes (single consonants or clusters).

Medial position consonants of Shoshoni [SOD]:

b	pp	mb	ʔb	m	(mm)	hm		w	hw			
d	tt	nd	ʔd	n	nn	hn	ʔn				s	ss
ʒ	cc	nʒ						y	hy	ʔy		
g	kk	ŋg		ŋ		(hŋ)						
g <sup>w</sup>	kk <sup>w</sup>	ŋg <sup>w</sup>		ŋ <sup>w</sup>								
	ʔ										h	

It is possible that all clusters containing  $\text{ʔ}$ ,  $h$  and nasal consonants are not actually found within roots, but are rather the results of various morphophonemic processes.

Another noticeable Central Numic feature is the existence of an additional type of “consonantal gradation”. Thus, in Panamint one finds the following types of articulation:

- spirantized: [naʔavaʔa] ‘on the bighorn’ (naka ‘bighorn’, paʔa ‘on’)
  - geminated: [tuappaʔa] ‘on the son’ (\*tuaʔ ‘son’, paʔa ‘on’)
  - nasalized: [piyimbaʔa] ‘on the duck’ (piyin ‘duck’, paʔa ‘on’)
  - aspirated: [hainčifaʔa] ‘on the friend’ (hainčih ‘friend’, paʔa ‘on’)
- (Miller et al. 2005: 415)

Apparently the “aspirated gradation” corresponds to the final  $*-h$  that is reconstructed for Southern Numic.

The following reconstruction of Proto-Central Numic is based only on my analysis of four languages / dialects: Panamint [Dayley 1989], Shoshoni and Western Shoshoni [SOD], Comanche [Robinson and Armagost 1990].

The sets of initial and final consonants are identical to those of Proto-Southern Numic, with the exception of  $*\eta-$ , not found in PCN.

The set of PCN medial position consonants is as follows:

*-b-	*-ʔb-	*-pp	*-mb-		*-mm-	*-hm-	*-ʔm-	*-w-	*-hw-	
*-d-	*-ʔd-	*-tt-	*-nd-	*-n-	*-nn-	*-hn-	*-ʔn-		*-ss	
*-ʒ-	*-ʔʒ-	*-cc	*-nʒ-					*-y-	*-hy-	*-ʔy-
*-g-	*-ʔg-	*-kk	*-ŋg-	*-ŋ-		(*-hŋ-)	*ʔŋ			
*-g <sup>w</sup> -		*-kk <sup>w</sup> -	*-ŋg <sup>w</sup> -	(*-ŋ <sup>w</sup> -)						
	*ʔ								*-h-	

It should be noted that some of the reconstructed medial consonants may in fact be clusters found at morpheme boundaries.

The following correspondences support this reconstruction:

Proto-C. Numic	Panamint	Shoshoni	W. Shoshoni	Comanche	Notes
*p	p	p	p	p	
*-b-	p	b	p	b / hp, p	
*-ʔb-	p	:ʔb, (:b)	p	b $\checkmark$ , :b $\checkmark$	
*-pp-	pp	p	pp	hp	
*-mb-	mp	mb	mp	:p, p	
*m-	m	m	m	m	
*-ʔm-	m	:m, m	m	m	
*-m-	mm	mm	m	m	
*-hm-	hm	hw	hm	Hm	
(*-)	m	ʔw, :ʔw	:w	ʔm, :m $\checkmark$	2 examples
*w-	w	w	W	w	
*-w-	w	#	#	w	
*-hw-	hw		hw, h	h	

Proto-C. Numic	Panamint	Shoshoni	W. Shoshoni	Comanche	Notes
*t-	t	t	t	t	
*-d-	t	t	t	r	
-ʔd-	t	:ʔd, ʔd	t	:r, (e:t)	
*-tt-	tt	t	t	ht	
*-nd-	nt	nd	nt	t	
*n-	n	n	n	n	
*-nn-	nn	nn	n	n	
*-n-	n	n	n	n	
*-hn-	hn	hn	hn	hn	
*-ʔn-	nn	:ʔn		ʔn, :n	
*c	c	c	t	c	
*-ʒ-	c	ʒ, ʒ̃	c	c	
*ʔʒ	c	:ʔʒ, :ʒ	c	c	
*-cc-	cc	c, č	cc	hc, :c	
*-nʒ-	nc	nʒ, nč		c	
*s-	s	s	s	s	
*-ss-	s	s, š	s	s	
*y-	y	y	y	y	
*-y-	y	y	y	y	
*-hy-	hy	hy	hy	hy	
*-ʔy-	y	ʔy	y	:y, y	
*k	k	k	k	k	
*-g-	k	g	k	k	
*-kk-	kk	k	kk	hk, (:k)	
*-ŋg-	ŋk	ng	nk	k, :k	
-ʔg-	kVʔ	:ʔg	k	:k	
*-ŋ-	ŋ	n	n	n	
*-ʔŋ-	ŋ	nn		n	
(*-hŋ-)	hm	hn	hm	Hm	1 example
*k <sup>w</sup>	k <sup>w</sup> / k(+u)	k <sup>w</sup> / k(+u)	k <sup>w</sup> / k(+u)	k <sup>w</sup> / k(+u)	
*-g <sup>w</sup> -	k <sup>w</sup>	g <sup>w</sup> / (u+)g	ku / (u+)k	k <sup>w</sup> / (u+)k	
*-kk <sup>w</sup> -	kk <sup>w</sup>	/ (u+)k	k <sup>w</sup> /	k <sup>w</sup> / (u+)hk	
*-ŋg <sup>w</sup> -	ŋk <sup>w</sup>	ng <sup>w</sup>	nk <sup>w</sup>	k <sup>w</sup>	
(*-ŋ <sup>w</sup> -)	ŋ <sup>w</sup>			w	1 example
*ʔ-	#	#	#	#	
*-ʔ-	ʔ	ʔ	ʔ	ʔ	
*h-	h	h	h	h	
*-h-	#, h	h	h	h	

The following reconstruction of Proto-Numic is based on a comparison of Proto-Central and Proto-Southern systems. A significant gap, however, is the absence of data on Western Numic languages.

A preliminary reconstruction of the vowel system seems to be quite simple (it is not quite clear if we need to reconstruct long vowels):

*i	*i	*u
		*o
	*a	

The consonantal system of Proto-Numic is as follows:

Initial consonants:

*p	*m	*w	
*t	*n		*s
*c		*y	
*k			
*k <sup>w</sup>			
*ʔ			*h

Final consonants:

*#	*-ʔ	*-h	*-N
----	-----	-----	-----

Medial consonants:

*-b-	*-pp	(*-mb-)	*-m-	(*-hm-)	*-w-		
*-d-	*-tt-		*-n-	*-hn-			*-s-
	*-cc				*-y-	*-hy-	
*-g-	*-kk	*-ŋg-	*-ŋ-	*-hŋ-			
*-g <sup>w</sup> -	*-kk <sup>w</sup> -		*-ŋ <sup>w</sup> -				
	*ʔ						*-h-

Some of the medial combinations seem to have been clusters, found on morpheme boundaries. It is possible that further investigation will fill in a few gaps in the reconstruction, such as an unexpected absence of \*-ʒ-.

Correspondences supporting the reconstruction are:

*Proto-Numic	PCN	PSN	Notes
*p-	*p-	*p-	
*-b-	*-b-	*-b-	
*-pp-	*-pp-	*-pp-	few examples
(*-mb-)	*-mb-	*-mb-	alternatively: *-N-pa
*t-	*t-	*t-	
*-d-	*-d-	*-d-/-c-	c — next to *i

*Proto-Numic	PCN	PSN	Notes
*-tt-	*-tt-	*-tt-	few examples
*c-	*c-	*c	
*-cc-	*-cc-	*-cc-	
*k-	*k-	*k-	
*-g-	*-g-	*-g-	
*-kk-	*-kk-	*-kk-	
*-ŋg-	*-ŋg-	*-ŋg-	
*k <sup>w</sup> -	*k <sup>w</sup> -	*k <sup>w</sup>	
*-g <sup>w</sup> -	*-g <sup>w</sup> -	*-g <sup>w</sup> -	
*-kk <sup>w</sup> -	*-kk <sup>w</sup> -	*-kk <sup>w</sup> -	
*m-	*m-	*m-	
*-m-	*-m-, *-mm-	*-m- / *-w-	
(*-hm-)	*-hm-	*-m-	1 example
*n-	*n-	*n	
*-n-	*-n-, *-nn-	*-n-	
*-hn-	*-hn-	*-n-	
*-ŋ-	*-ŋ-	*-ŋ-	
*-hŋ-	*-hŋ-	*-ŋ-	
*-ŋ <sup>w</sup> -	*-ŋ <sup>w</sup> -, *-m-	*-w- / #	# after *o
*w-	*w-	*w-	
*-w-	*-w-	*-w-	few examples
*y-	*y-	*y-	
*-y-	*-y- / #	*-y- / #	# after *i
*-hy-	*-hy-	*-y-	
*s-	*s-	*s-	
*-s-	*-s-, *-ss-	*-s-	
*h-	*h-	*h-	
*-h-	*-h-, #	*-h-, #	
*ʔ-	*-ʔ-	*ʔ-	
*-ʔ-	*-ʔ-	*-ʔ-	

Examples supporting these correspondences are collected in numiet.dbf. Further investigation is needed to finalize the set of PN medial consonants and firmly distinguish them from other possible morpheme-boundary clusters.

The reconstructed system allows for further ‘fine tuning’. E. g., one can suggest that the two types of stops and affricates actually represent an opposition between plain and glottalized consonants: \*p ~ \*pʔ, \*t ~ \*tʔ, etc., which is neutralized in the initial position.

The EHL Numic etymological database (numiet.dbf) consists of more than 600 entries. Most of these represent modified versions of etymologies found on such lists as [Iannucci 1972] and [Miller 2003].



Glottochronological calculations suggest that the disintegration of Proto-Numic may have begun about 2,500 years ago.

### Takic languages

Miller (1984) distinguishes two branches of Takic languages:

- a. Serrano-Gabrielino
  - (1) Serranan: Serrano, Kitanemuk
  - (2) Gabrielino (Gabrielino, Fernandino)
- b. Cupan
  - (1) Cupeño, Cahuilla
  - (2) Luiseño

His lexicostatistical matrix, however, does not support any reasonable classification of Serrano and Gabrielino, probably due to the quality of the wordlists:

	Gabrielino	Serrano	Cahuilla	Cupeño	Luiseño
Gabrielino	×	45	42	34	38
Serrano	45	×	50	38	35
Cahuilla	42	50	×	65	50
Cupeño	34	38	65	×	48
Luiseño	38	35	50	48	×

It is possible to improve the results for Serrano by including forms only from Kroeber (1907) and by adding Kitanemuk words collected by Harrington [Anderton 1988]. A new matrix (based on taki.dbf) is:

	Kitanemuk	Serrano	Cahuilla	Cupeño	Luiseño
Kitanemuk	×	75	50	50	50
Serrano	75	×	61	58	52
Cahuilla	50	61	×	80	61
Cupeño	50	58	80	×	61
Luiseño	50	52	61	61	×

Some differences in the percentages of lexical similarities between the languages are mainly due to differences in the wordlists. Nevertheless, the overall pattern remains the same, suggesting that Miller's classification (without Gabrielino) may be accepted:

- Takic** (more than 52%)
  - a. Kitanemuk + Serrano (75%)
  - b. Cupan (more than 61%)
    - Cahuilla + Cupeño (80%)
    - Luiseño

The three Cupan languages are quite well described. The system of Luiseño can be presented as follows:

Vowels (spelled as i, u, e, a, o):

i		u
ε		ɔ
	a	

According to Bright (1965), all vowels may be short / long and stressed / unstressed, but certain structures predetermine the choice of length and / or strength. For example, ‘in final accented position, only long vowels occur’ (p. 344); ‘when an initial syllable has a long vowel, it is always accented’ (p. 345).

Consonants:<sup>6</sup>

p	v	m	w	
t		n	l	s
			r	ʃ
č			y	
k		ŋ		x
k <sup>w</sup>				k <sup>w</sup>
q				
q <sup>w</sup>				
ʔ				h

Elliott (1999: 14–15) writes that (all?) consonants between two short vowels are either long or accompanied by ʔ if the first vowel is stressed.

My analysis of Tactic comparisons suggests that the vocalic length distinction in Luiseño is of later origin<sup>7</sup>. The rules to eliminate it are as follows:

– the root may consist of one or two syllables: CV, CVC(C), CVC(C)V, CVC(C)VC(C);

– in **nouns**, monosyllabic roots always have long vowels unless the final consonant is ʔ: *ku:ŋ* ‘husband’, *wiʔ* ‘fat, grease’;

– in disyllabic words stress can be placed on the first or the second syllable. If a stressed syllable ends in a consonant<sup>8</sup>, the vowel is long, otherwise the vowel is short: *ʔáana* ‘ant’, *ʔamúu* ‘mescal plant, Agave deserti’, *ʔéčvu-š* ‘left hand’, *qawiʔ* ‘forehead’;

– **verbs** are predominantly monosyllabic with short vowels. Disyllabic verbs follow the same rules as nouns: *ʔuláʔq=i-* ‘to sew (vt.)’.

When modern Luiseño forms do not follow the suggested pattern, we assume that they may have recently lost a consonant: *čáay-i-* ‘to sift, winnow’ < \*\**čáʔay-i-* (cf. Cahuilla *čáʔa-l* ‘sieve, trap’), *qíin-i* ‘to plow’ < \*\**qíyín-i* (cf. Cahuilla *qíyine-* id.).

<sup>6</sup> I was not able to complete a list of consonantal clusters acceptable in Luiseño roots.

<sup>7</sup> For other views on stress in Cupan languages see Hill and Hill 1968, Munro 1990, etc.

<sup>8</sup> One should mention, however, that in some cases it is not clear how to divide a word into syllables, cf. *kúumal* ‘poker’ < \*\* *kíuu=mal* (*ku-* ‘fire’).

The Proto-Cupan system is reconstructed as followed:

Consonants:

*p	*v	*m	*w	
*t		*n	*l	*s
*č			*y	
*k		*ŋ		*x
*k <sup>w</sup>				
*q				
*q <sup>w</sup>				
*ʔ				*h

Vowels:

*i	*i	*u
*e (= *ε)		*o (= *ɔ)
	*a	

The correspondences behind these proto-phonemes are mostly trivial.

The EHL etymological database (takiet.dbf) currently contains 562 entries that reproduce, with some modifications, Miller's list of Proto-UA etymologies [Miller 2003], with additional comparisons from [Munro 1990] and some other sources. No direct search for new etymologies has been conducted.

Glottochronology suggests that the Proto-Takic language may have begun to disintegrate about 3,200 years ago.

### Sonoran languages

A list of Sonoran languages can be found in [Miller 1984]:

1. Tapiam: Upper Piman (Papago, Pima, Nevome), Lower Piman, Northern Tepehuan, Southern Tepehuan (Southern Tepehuan, Tepecano)
2. Taracahitian
  - (a) Tarahumara (several dialects)
    - Guarijio (Highland, Lowland)
  - (b) Opatian: Opata, Eudeve, Jova?
  - (c) Cahita (Mayo, Yaqui)
3. Tubar
4. Corachol: Cora, Huichol.

The Opatian languages are extinct and I was not able to obtain wordlists for any of them.

**Tapiam.** A brief description of Papago phonology can be found in [Mathiot 1973: 113–114]:

	Bilabial	Dental	Alveolar	Palatal	Velar	Glottal
Tense stop	p	t		č	k	
Lax stop	b	d	ɖ	č̥	g	ʔ
Fricative			s	š		h
Nasal	m	n				
Lateral			l			
Glide	w					

Vowels (short and long):

i	i	u
		o
	a	

Stress is predictable.

The systems of other Tepiman languages are listed in [Bascom 1965].

A reliable reconstruction of Proto-Tepiman was conducted by Bascom in 1965 and is the main source of the Tepiam database (tepiet.dbf, with more than 300 etymologies). Additional lexical data are available only for Papago (Saxton & Saxton 1983; Mathiot 1973) and Nevome (Pennington 1979). Papago is now sufficiently well represented in the database. Nevome, known only from a dictionary composed in the XVII<sup>th</sup> or XVIII<sup>th</sup> centuries, has not been studied.

The following system is proposed for Proto-Tepiman:

Root structure: CVCV with no consonantal clusters.

Consonants (initial and medial):

*p	*b	*m	*v	
*t	*d	*n	*r	*s
*k	*g			
*ʔ	*h			

Vowels (may be short or long):

*i	*i	*u
		*o
	*a	

This reconstruction is based on the following set of phonological correspondences:

Tepiman	N. Tepehuan	S. Tepehuan	U. Pima	L. Pima
*p	p	p	p	p
*b	b	b	b	b

Tepiman	N. Tepehuan	S. Tepehuan	U. Pima	L. Pima
*m	m	m	m	m
*v	v, #	v	v	v
*t	t / tʰ (i+, +i)	t / tʰ (i+, +i)	č / t (+i,i,u)	t / č (i+, +i)
*d	d / dʰ (i+, +i)	d / dʰ (i+, +i)	ž / d (+i,i,u)	d / ž (i+, +i)
*n	n / n̄ (i+, +i)	n / n̄ (i+, +i)	n̄ / n (+i,i,u)	n̄, n (i+, +i)
*r	r / l (i+, +i)	r / lʰ (i+, +i)	ɖ / l (+i)	r / l (+i)
*s	s / š (i+, +i)	s / š (i+, +i)	š / s (+i)	š / s (i+, +i)
*k	k, h	k	k, h	k
*g	g, #	g, ʔ	g	g
*ʔ	#	ʔ	ʔ, #	ʔ, #
*h	h, #	h, #	h, #	h, #
*i	i	i	i	i
*ī	ī	ī	ī	ī
*u	u	u	u	u
*o	o	o	o	o
*a	a	a	a	a

These correspondences are supported by reconstructions, collected in tepiet.dbf.

**Tarahumaran.** The two main sources for Tarahumaran languages are a dictionary of Central Tarahumaran [Hilton et al. 1993], and a grammar and dictionary of Highland Guarijio [Miller 1996].

The Guarijio system of phonemes is as follows:

Consonants (initial and medial):

p	m	w	
t	n	l	s
č		y	
k	g		
ʔ			h

Two types of medial clusters are found in the roots: -ʔ + consonant and -h + consonant.

Vowels:

i		u
e		o
	a	

Stress is usually placed on the second (last) syllable of the root, but there are a significant number of exceptions.

The system of Tarahumara is quite similar, as is the reconstructed Proto-Tarahumaran. Minor issues of the reconstruction are not discussed here.

The Tarahumaran database (taraet.dbf, more than 350 etymologies) is mainly based on Miller's materials. It should be mentioned, however, that the database includes only a small portion of possible etymologies for this group.

**Mayo (Cahita).** Lexical corpora for several closely related Mayo languages / dialects are found in three dictionaries: Mayo [Collard and Collard 1962], Yaqui [Estrada et al 2004] and Yoeme [Molina et al. 1999]. Only the Yoeme dictionary has been used for the time being. The phonological system of this variety is as follows:

Consonants (initial and medial):

b <sup>w</sup>	p	m	v	w	
	t	n	R	l	s
				y	
	k	g			
	ʔ				h

No consonant clusters are found within root morphemes.

Vowels:

i		u
e		o
	a	

All vowels can be short and long; vowel length can sometimes fluctuate under certain conditions, probably caused by morphological factors that are not yet fully understood.

A small Mayo database was created with the sole aim of checking the validity of Yoeme forms used in comparisons with other Sonoran languages.

**Corachol.** For my study, two sources of Huichol data were available: a small dictionary of the language published as [McIntosh & Grimes 1954] and a manuscript of a much bigger dictionary prepared in 1982 [Grimes et al. 1982], which had never been published. It is this manuscript that is used as my primary source of information. The phonology of the language may be presented as follows:

Consonants (initial and medial):

p	m	W		
t	n		r	
c		Y		s
k				
k <sup>w</sup>				
ʔ				h

No consonantal clusters are found.

Vowels (all vowels can be short or long):

i	i	u
e		
	a	

Tones: H(igh) (V́), L(ow) (unmarked). The origin of tones is not yet fully understood.

The Cora data is taken from [McMahon & McMahon 1959], but has not yet been properly analyzed.

A few additional extinct Sonoran languages are known with various degree of detalization. None of them have been used in the reconstruction, but some of their words mentioned in various etymological sources are given in the database.

The Proto-Sonoran system and a list of etymologies (about 350 entries) were produced by [Lionnet 1985]. The phonological system of the proto-language is as follows:

Consonants (initial and medial):

*p	*m	*w	
*t	*n	*l	*s
*č		*y	
*k			
*k <sup>w</sup>			
*ʔ			*h

Proto-Sonoran \*l is not found in the initial position.

The suggested reconstruction is based on the following set of correspondences:

Proto-Sonoran	*Tepiman	Yoeme	Huichol	*Tarahumara
*p-	*p, *v	v, p	h	#
*k <sup>w</sup>	*b	b <sup>w</sup>	k <sup>w</sup>	*w, *k <sup>w</sup> (+u)
*w	*g, v	v, w	w	*w
*t	*t	t	t, č	*t
*č	*s	č	č	*č
*s	*h	s	s	*s
*h	*ʔ	h	#	*ʔ
*ʔ	*ʔ	#	#/ Vʔ-	#
*k	*k	k	k	*k
*m	*m	m	m	*m



Proto-Sonoran	*Tepiman	Yoeme	Huichol	*Tarahumara
*n	*n	n	n	*n
*l	*r, *d	l	l	*l
*y	*d	y	y	*y

Vowels:

*i	*ī	*u
		*o
	*a	

The vocalic correspondences are straightforward.

Although, as has already been mentioned, some Sonoran languages distinguish between long and short vowels, this distinction is not reconstructed for the proto-language.

Proto-Tepiman roots with long vowels normally correspond to Proto-Tarahumara roots with consonantal clusters. This observation permits to reconstruct Proto-Sonoran clusters as \*-ʔC- and \*-hC-. The root structure in the proto-language, best preserved in Guarijio, is \*CV(H)CV. Further study is needed to clarify the origin of long vowels in Yoeme and Huichol.

The Proto-Sonoran etymological database (sonoet.dfb, about 400 entries) consists of all the acceptable comparisons from Lionnet's and Miller's sources. Additional research has also been conducted to find missing reflexes in various languages of the family.

Miller (1984) has identified four equal branches (groups) of the Sonoran languages. An alternative interpretation of his matrix, however, gives us the following tree:

**Sonoran** (34% and more):

A. Papago, Nevome, N. Tepehuan (79% and more)

B. 39% and more:

1. 51% and more:

(a) Guarijio and Tarahumara (83%);

(b) Opta and Eudeve (73%);

(c) Mayo and Yaqui (93%);

2. Tubar

3. Corachol (58%)

This classification is confirmed by my own lexicostatistical study based on 35 and 50 item-lists (sono.dbf). The 100-list gives, however, a different picture:

	Nevo	Papa	Guar	Tara	Yoem	Cora	Huic
Nevome	×	93	56	52	59	46	50
Papago	93	×	53	49	57	46	48
Guarijio	56	53	×	81	66	51	54
Tarahumara	52	49	81	×	61	47	51
Yeome	59	57	66	61	×	53	59
Cora	46	46	51	47	53	×	74
Huichol	50	46	54	51	59	74	×

The differences in classifications are presumably caused by etymological misinterpretations of the (so far) poorly understood data from Cora. Without it, the classifications are identical:

- A. \*Tepiam  
 B 1a. Tarahumaran  
     1b. Yeome  
     2. Huichol

Glottochronology suggests that Proto-Sonoran began to disintegrate about 3,200 years ago.

### Aztecan languages

Miller classifies the Aztecan group into two branches: a. Pochutec, b. General Aztecan: Pipil, Aztec (Classical Aztec. Tetelcingo, Zacapoaxtla, and others). The extinct language Pochutec, known from a short article by Boas (1917), was not included in Miller's matrix. Nevertheless, lexicostatistical calculations support his view:

35-item matrix:

	N.Pu	Tet	Pip	Zac	Mec	Poc
N.Puebla	×	94	94	85	81	75
Tetelcingo	94	×	97	88	84	78
Pipil	94	97	×	91	88	79
Zacapoaxtla	85	88	91	×	81	76
Mecayapan	81	84	88	81	×	78
Pochutec	75	78	79	76	78	×

A detailed discussion of modern Aztec dialects can be found in [Canger 1988]. The comparative phonology of the family is well-known [Dakin 1982]. In 1978 Campbell and Langacker published the following set of Aztecan vocalic correspondences (p. 94):

	C&L	Pochutec	Class.Azt.	Tetelcino	Zacapoaxtla	Pipil
1	*a	e	a	a	a	a
2	*a:	a	a:	ɔ	a:	a:
3	*o	o	o	o	o	o
4	*o:	u	o:	u	o:	u:
5	*i	i	i	ɪ	i	i
6	*i:	i	i:	i	i:	i:
7	*i	o	i	ɪ	i	i
8	*ə	o,e,#	ə	e	e	ee
9	*e:	e	e:	ie	e:	e:

Dakin (1983) has shown that correspondence No. 7 is only found before consonants and thus does not reflect an additional Proto-Aztec vowel. A restricted distribution, how-

ever, is not a sufficient argument against a reconstruction, so the proto-system can be interpreted as:

*i (5)	*i (7)	
*e (8)		*o (3)
	*a (1)	

*i: (6)		
*e: (9)		*o: (4)
	*a: (2)	

The Proto-Aztec consonantal system, suggested by Dakin [1979: 50], is formed by 15 consonants:

*p	*m	*w	
*t	*n		
*λ		*l	
*c			*s
*č		*y	*š
*k			
*k <sup>w</sup>			
*ʔ			

This reconstruction is based on the following set of correspondences:

Proto	Pochutec	Class.Azt.	Tetelcino	Zacapoaxtla	Pipil
*p	p, (b)	p	p	p	p
*m	m	m	m	m	m
*w	w	w	w	w	w
*t	t, (d)	t	t	t	t
*n	n, ní	n	n	n	n
*λ	t, (d)	λ	λ	t	t
*l	l	l	l	l	l
*c	c	c	c	c	c
*s	s	s	s	s	s
*č	č	č	č	č	č
*š	š	š	š	š	š
*y	y	y	y	y	y
*k	g, (k)	k	k	k	k
*k <sup>w</sup>	g, (k)	k <sup>w</sup>	k <sup>w</sup>	k <sup>w</sup>	k <sup>w</sup>
*ʔ	h	ʔ	h	h	h

The phonemes \*č and \*š are often found in proximity with \*i and \*i:; this indicates their secondary origin.

The Aztec database (azttet.dbf) contains about 350 entries, which obviously represents only a small portion of the proto-language lexicon. Glottochronological calculations suggest that the disintegration of Proto-Aztec began about 2,000 years ago; Kaufman [2001] talks about 500 CE.

## Hopi

The dialects of Hopi form a separate branch of UA. Whorf [1946] identifies four of them (a claim that I am not able to verify through lexicostatistics): First Mesa (or Whorf's Polacca), Mishongnovi (or Whorf's Toreva), Shipaulovi (or Whorf's Sipaulovi), Third Mesa (or Whorf's Oraibi). The best known is the dialect of the Third Mesa [Hopi 1998]. Its phonological system is as follows:

Consonants:

p	v	m	w		
t		n	l		s
c				ʒ / r	
k <sup>v</sup>		ń	y		
k		ŋ			
k <sup>w</sup>		ŋ <sup>w</sup>			
q	q <sup>w</sup>				
ʔ					h

The opposition of velars and uvulars is lost before front vowels.  
Vowels (long and short) (the vowels are spelled as i, e, u, ö, a, o):

i	ĩ	
ø		o
ε	a	

Stress is predictable.

Tones: the Third Mesa dialect has developed tonal distinctions on long vowels, diphthongs, and vowel + sonorant sequences (*l, m, n, ŋ, ŋ<sup>w</sup>*). The tones are either falling or level. The falling tone corresponds to one of several sequences in the Mishongnovi dialect recorded by Whorf [Manaster-Ramer 1986]: (a) vowel + preaspirated consonant; (b) vowel + voiceless sonorant; (c) vowel + *h*.

## Proto-Uto-Aztecan

The resulting Proto-UA consonantal system can be reconstructed as follows:

Initial consonants:

*p	*m	*w	
*t	*n		*s
*č		*y	
*k	*ŋ		
*k <sup>w</sup>		*ŷ <sup>w</sup>	
*q			
*q <sup>w</sup>			
*ʔ			*h

## Medial consonants:

*p	*b	(*mb)	*m	(*w)	
*t	*d		*n	*l	*s
*č	*ž			*y	
*k	*g	*ŋg	(*ŋ)		
*k <sup>w</sup>	*g <sup>w</sup>		*ŋ <sup>w</sup>	(*y <sup>w</sup> )	
*q	*G	(*ŋG)			
*q			(*N <sup>w</sup> )		
*ʔ					

Proto-consonants in brackets are not supported by sufficient number of examples.

The system of final consonants is probably the same as that of Proto-Numeric.

The consonantal correspondences are supported by etymologies in utatet.dbf. This database consists primarily of comparisons from Miller's list of 2003. The list, published on the Internet, is Hill's revised and expanded version of Miller's computerized database for Uto-Aztecan cognate sets of 1988. The list was rekeyed and made available for the purposes of the EHL project by Nikolayev; I have converted the list to the current database format, double-checked the actual language forms and often expanded the etymologies, using recently published dictionaries and other data sources.

The set of Proto-UA consonantal correspondences is as follows:<sup>9</sup>

*Proto UA	*Sonor	*Aztec	Tubatulaba	*Takic	Hopi	*Numeric	Notes
*p-	*p	*p, #V:	p	*p	P	*p	
*-b-	*p	*p	p, b	*v	v, p	*b	
(*-p-)	*p			*p		*pp	no good exx.
(*-mb-)	*hp		mb				1 example
*m-, *-m-	*m	*m	*m	*m	m	*m	
(*w-)	*w	#	w	*w	#	*w	1 example
(*-w-)	*w	*w	w	*w	w	*w, #	
*t-	*t	*λ, *t	t	*t	t	*t	
*-d-	*t, ht	*t, *ł	d, l	*l	t	*d	
*-t-	*ht	t	l	l	t	*tt	
*n-, *-n-	*n	*n	n	*n	n	*n	
*-l-	*l	*l	n	*l	n	*hn	
*s-, *-s-	*s	*š, *s	s	*s	s	*s	
*č-	*č	*č, *c		*c	c	*c	
*-ž-	*č, *ž	*c	y, d <sup>z</sup>	*y	*y	*y, *ž	
*-č-	*č, *hč	*č	c	*c		*y	

<sup>9</sup> The set of correspondences is significantly different from the one given in [Voeglin, et al. 1962, Miller 1967].

*Proto UA	*Sonor	*Aztec	Tubatulaba	*Takic	Hopi	*Numic	Notes
*y-, *-y-	*y	*y	y	*y	y	*y	
*k-	*k	*k	k	*k	k, kʷ	*k	
*-g-	*k	*k	g	*k	*k	*g	
*-k-	*k	*k	*k	*k	k, kʷ	*kk	
*-ŋg-	*hk		*ŋg	*k	k	*ŋg	
(*ŋ-)	*n	*n	n	*ŋ	ŋ	*n, *ŋ <sup>(w)</sup>	no good exx.
(*-ŋ-)	*n		n	*n, ŋ	n, ŋ	*ŋ	
*-ŋ <sup>w</sup> -	*m	n	m	m	m	*ŋ <sup>w</sup>	
(*-N <sup>w</sup> -)	*hw		ŋ <sup>w</sup> , k <sup>w</sup>	*w, *ŋ <sup>w</sup>	ŋ <sup>w</sup>		no good exx.
*k <sup>w</sup> -	*k <sup>w</sup>	*k <sup>w</sup>	w	*k <sup>w</sup>	k <sup>w</sup>	*k <sup>w</sup>	
(*-g <sup>w</sup> -)	*k <sup>w</sup>	*w			k <sup>w</sup>	*g <sup>w</sup> -	1 example
(*-k <sup>w</sup> -)		*k <sup>w</sup>			k <sup>w</sup>	kk <sup>w</sup>	1 example
(*ŷ <sup>w</sup> -)	*w	#	w	*w	l	*w	1 example
(*-ŷ <sup>w</sup> -)	*w	*kw	w	w	l	#	
*q	*k	*k	h	*q	*q	*k	
(*-G-)	*hk	*k	h	*q	q	*g	1 example
*-q-	*k	*k	h	*x	q	*g	
(*-ŋG-)	hk	k	ŋ	*x	q	*ŋg	1 example
*q <sup>w</sup> -	*k <sup>w</sup>	*k <sup>w</sup>	w	*q <sup>w</sup>	*k <sup>w</sup>	*k <sup>w</sup>	
(*-q <sup>w</sup> -)	(*p)	*k <sup>w</sup>		*q <sup>w</sup>			1 example
*ʔ	*ʔ	#Vʔ, #	ʔ-, #	*ʔ	#	#	
*-ʔ-	#, *ʔ	#, *ʔ	ʔ-, #	#, *ʔ	#	#	
*h	*h	#Vʔ	h	*h	h	*h	

This reconstruction is, naturally, far from being complete: some proto-phonemes are not supported by a sufficient number of examples, while others demonstrate unexplained double reflexes in daughter-languages. Much more work is needed to obtain a detailed proto-UA reconstruction and a better etymological dictionary. Some etymologies in utatet.dbf may not belong to the Proto-UA level, but rather represent other chronological levels, areal roots, or even borrowings. These problems require further investigation.

It is suggested that the homeland of the Proto-UA family was located somewhere in Arizona and neighboring areas of New Mexico and Mexico (data in [Flower 1983]), based on an analysis of plant and animal names). Jane Hill [2001] made an attempt to demonstrate that the speakers of Proto-UA were agriculturalists, but her etymological conclusions are not well supported, since none of the proposed “agricultural” etymologies can be attributed to the level of the proto-language.

	Mn	NP	Pn	Sh	Cm	Ka	Ch	SP	Ut	Tb	Gb	Sr	Ca	Cu	Lu	Ho	Pp	Nv	NT	Gu	Ta	Op	Bu	Ma	Yq	Tu	Hu	Co	CA	Te	Za	Pl
Mono	x	77	59	58	57	52	50	53	52	39	26	26	29	28	26	33	23	24	25	26	23	26	28	27	29	28	25	25	18	19	17	16
N.Paiute	77	x	58	58	58	56	55	58	52	42	26	24	27	27	27	32	26	26	28	29	27	29	27	28	30	27	24	22	18	18	17	15
Panamint	59	58	x	87	79	54	61	62	61	37	26	24	27	24	25	27	25	24	26	23	21	21	23	25	26	27	23	22	16	16	15	14
Shoshoni	58	58	87	x	88	55	58	62	61	38	26	24	27	24	24	23	25	24	26	23	21	20	22	24	26	28	23	23	16	16	15	14
Comanchi	57	58	79	88	x	49	54	59	59	35	23	21	24	23	22	22	23	23	23	24	21	20	26	27	24	27	21	21	14	14	13	12
Kawaiisu	52	56	54	55	49	x	75	79	76	39	24	26	27	26	24	31	26	24	27	24	22	20	24	25	29	28	23	22	16	17	16	16
Chemehuevi	50	55	61	58	54	75	x	86	78	42	27	28	31	30	27	33	28	26	28	24	22	26	27	26	27	23	21	15	15	16	15	
S.Paiute	53	58	62	62	59	79	86	x	87	39	26	27	31	31	27	31	28	26	30	25	23	24	24	25	29	30	24	22	16	16	17	16
Ute	52	52	59	61	59	76	78	87	x	40	27	27	29	28	26	32	30	27	29	28	26	23	26	27	30	31	25	23	16	17	18	17
Tübatulabal	39	42	37	38	35	39	42	39	40	x	40	35	38	37	34	38	35	35	37	36	32	33	27	28	35	33	32	30	24	25	26	24
Gabrielino	26	26	26	26	23	24	27	26	27	40	x	45	42	34	38	29	35	24	26	29	28	26	35	35	28	24	24	19	20	20	21	21
Serrano	26	24	24	24	21	26	28	27	27	35	45	x	50	38	35	29	27	27	30	34	34	31	26	29	32	28	28	21	22	22	24	23
Cahuilla	29	27	27	27	24	27	31	31	31	38	42	50	x	65	50	31	31	30	32	34	33	33	30	33	35	29	34	24	23	24	20	20
Cupenõ	28	27	24	24	23	26	30	31	28	37	34	38	65	x	48	31	28	27	29	29	26	29	29	26	26	26	26	23	19	20	19	18
Luisenõ	26	27	25	24	22	24	27	27	26	34	38	35	50	48	x	26	25	24	26	28	28	24	25	28	28	23	27	22	19	19	24	24
Hopi	33	32	27	23	22	31	33	31	32	38	29	29	31	31	26	x	32	33	33	32	28	33	35	34	36	30	28	26	24	24	31	30
Papago	23	26	25	25	23	26	28	28	30	35	25	27	31	28	25	32	x	85	79	44	41	40	45	43	45	40	41	34	29	30	31	30
Ne'lvome	24	26	24	24	23	24	26	26	27	35	24	27	30	27	24	33	85	x	79	47	42	44	47	45	47	41	43	34	29	30	32	29
N.Tepahuan	25	28	26	26	23	27	28	30	29	35	26	30	32	29	26	33	79	79	x	47	42	40	45	49	49	46	42	35	30	30	29	33
Guarijio	26	29	23	23	24	24	24	25	28	36	29	34	34	29	28	32	44	47	47	x	83	55	59	58	58	48	51	42	32	32	33	34
Tarahunara	23	27	21	21	21	22	22	23	26	32	28	34	33	26	28	28	41	42	42	83	x	54	52	51	51	44	48	38	33	34	35	38
Opata	26	29	21	20	20	20	26	24	23	35	26	31	33	29	24	33	40	44	40	55	54	x	73	53	55	42	48	35	39	38	39	40
Eudeve	28	27	23	23	22	26	24	26	27	35	26	30	34	29	26	35	45	47	45	59	52	73	x	61	62	51	49	42	40	40	37	39
Mayo	27	28	25	26	24	27	25	27	28	35	29	33	36	26	28	34	43	45	49	58	51	53	61	x	93	51	48	45	38	38	39	39
Yaqui	29	30	26	26	24	29	26	29	30	35	28	32	35	26	28	36	45	47	49	58	51	55	62	93	x	53	51	46	39	39	35	37
Tubari	28	27	27	28	27	28	27	30	31	33	24	28	29	26	23	30	40	41	46	48	44	42	51	51	53	x	41	39	36	35	37	
Huichol	25	24	23	23	21	23	23	24	25	32	24	28	34	26	27	28	41	43	42	51	48	48	49	48	51	41	x	58	39	37	33	35
Cora	25	22	22	23	21	22	21	22	23	30	19	21	24	23	22	26	34	34	35	42	36	35	42	45	46	39	58	x	37	35	80	85
ClAztec	18	18	16	16	14	16	15	16	16	24	20	22	23	19	19	24	29	29	30	32	33	39	40	38	39	36	39	37	x	85	81	77
Telecüingo	19	18	16	16	14	17	15	16	17	25	20	22	24	20	19	24	30	30	30	32	34	38	40	38	39	35	37	35	85	x	85	81
Zacapoaxtla	17	17	15	15	13	16	16	17	18	26	21	20	24	20	19	24	31	31	32	29	33	35	39	37	38	35	35	33	80	85	x	77
Pipil	16	15	14	14	12	16	15	16	17	24	21	19	23	20	18	24	30	30	29	33	34	38	40	39	39	37	37	35	79	81	77	x

Matrix from Miller 1984



## Literature

- ANDERTON, Alice J. (1988). *The Language of the Kitanemuks of California*. PhD. Dissertation. University of California, Los Angeles.
- BASCOM, Burton William. 1965. *Proto-Tepiam (Tepehuan-Piman)*. Ph.D. dissertation, University of Washington.
- BOAS, F. 1917. El dialecto mexicano de Pochutle, Oaxaca. *IJAL* 1:9–44.
- BRIGHT, William, and Jane H. HILL. 1967. The linguistic history of the Cupeno. *Studies in Southwestern Ethnolinguistics*, ed. By Dell Hymes, pp.351–371
- BRIGHT, William. 1965. Luiseño Phonemics. *IJAL* 31:342–345.
- BRIGHT, William. 1965. Luiseño Phonemics. *IJAL* 31:342–345.
- CAMPBELL, Lyle, and Ronald W. LANGACKER. 1978. Proto-Aztecan vowels. *IJAL* 44:85–102, 197–210, 262–279.
- CAMPBELL, Lyle. 1997. *American Indian Languages*. Oxford — New York: Oxford University Press.
- CANGER, Una. 1988. Nahuatl dialectology: A survey and some suggestions. *IJAL* 54:28–72.
- COLLARD, H. y E. COLLARD. 1962. *Castellano-mayo, mayo-castellano*. Instituto Linguístico de Verano.
- DAKIN, Karen. 1979. Phonological Changes in Nahuatl: the Tense / Aspect / Number Systems. *IJAL* 45:48–71.
- DAKIN, Karen. 1982. *La Evolución Fonológica del Protonáhuatl*. Colección Lingüística indígena. Mexico: Universidad Nacional Autónoma de México.
- DAKIN, Karen. 1983. Proto-Aztecan vowels and Pochutec: An alternative analysis. *IJAL* 49:196–203.
- DAYLEY, Jon P. 1989. *Tumpisa (Pamamint) Shoshone Dictionary*. UCPL 115. Berkeley and Los Angeles: University of California Press. Electronic edition 1998.
- ELLIOTT, Eric B. 1999. *Dictionary of Rincon Luiseño*. Ph. D. dissertation.
- ESTRADA FERNÁNDEZ, Zarina et al. 2004. *Diccionario yaqui-español y textos: obra de preservación lingüística*. Hermosillo and México, D.F.: Universidad de Sonora and Plaza y Valdés Editores.
- FELIPE S. Molina, Herminia VALENZUELA, David Leedom SHAUL. 1999. *Yoeme-English English-Yoeme : with a comprehensive grammar of Yoeme language*. New York: Hippocrene Books.
- FOWLER, Catherine S. 1983. Some lexical clues to Uto-Aztecan prehistory. *IJAL* 49:224–57.
- GIVON, T., ed. 1979. *Ute Dictionary*. Preliminary ed. Ignacio, Col.: Ute Press.
- GRIMES, Jose. 1981. El Huichol. MS
- HALE, Kenneth. 1958–59. Internal diversity in Uto-Aztecan: I. *IJAL* 24:101–7, II. *IJAL* 25:114–21.
- HILL, Jane H. 2001. Proto-Uto-Aztecan: a community of cultivators in Central Mexico? *American Anthropologist* 103:913–934.
- HILL, Jane H, and Kenneth C. HILL. 1968. Stress in Cupan (Uto-Aztecan languages). *IJAL* 34:233–241).
- HILL, Jane H., and Rocinda NOLASQUEZ. 1973. *Mulu'wetam: The Fist People*. Banning, Calif.: Malki Museum Press.
- HILTON, Simón K. *Diccionario Tarahumara*. Instituto Linguístico de Verano.
- IANNUCCI, David. 1972. *Numic historical phonology*. Ph.D. dissertation, Cornell University.
- Hopi-English Dictionary of the Third Mesa Dialect*. 1998. Tucson: The University of Arizona Press.
- KAUFMAN, Terrence. 2001. *The History of the Nawa language group from the earliest times to the sixteenth century: some initial results*. MS
- KROEBER, A.L. 1907. Shoshonean dialects of California. *UCPAAE* 4:65–165.
- LAMB, Sydney M. 1964. The classification of the Uto-Aztecan languages: a historical survey. *UCPL* 34:106–25.
- LIONNET, André. 1985 Relaciones internas de la rama sonoreña. *Amerindia*, 10.
- MANASTER-RAMER, Alexis. 1993. On Lenition in some Northern Uto-Aztecan Languages. *IJAL* 59:334–341.
- MANASTER-RAMER, Alexis. 1986. Genesis of Hopi Tones. *IJAL* 52:154–160.
- MATHIOT, Madeleine. 1973. *A Dictionary of Papago Usage*. CV.1. Bloomington: Indiana University.
- MCINTOSH, Juan B. y Jose GRIMES. 1954. *Vocabulario Huichol-castellano. Castellano-huichol*. Instituto Linguístico de Verano.
- MCLAUGHLIN, John E. 1992. Counterintuitive Solution in Central Numic Phonology. *IJAM* 58:158–181.
- MCMAHON, Ambrosio y Miaria Aiton de MCMAHON. 1959. *Cora y Espanol*. Instituto Linguístico de Verano.
- MILLER, Wick R. 1967. *Uto-Aztecan Cognate Sets*. UCPL 48. Berkeley and Los Angeles: University of California.
- MILLER, Wick R. 1982. Southern Paiute and Numic Final Feature. *IJAL* 48:444–449.
- MILLER, Wick R. 1984. The classification of the Uto-Aztecan languages based on lexical evidence. *IJAL* 50:1–24.
- MILLER, Wick R. 1996. *Guarijío: gramática, textos y vocabulario*. Mexico: Universidad nacional autónoma de México.
- MILLER, Wick R. 2003. *Uto-Aztecan Cognate Set*. Revised and expanded by Kenneth C. HILL (ms.).

- MILLER, Wick R., Dirk ELZINGA, and John E. MCLAUGHLIN. 2005. Preaspiration and Gemination in Central Numic. *IJAL* 71:413–44.
- MILLER, Wick R., James L. TANNER, and Lawrence P. FOLEY. 1971. A Lexicostatistic Study of Shoshoni Dialects. *Anthropological Linguistics* 13/4:142–164.
- MUNRO, Pamela. 1983. Selected studies in Uto-Aztecan phonology. *IJAL* 49:277–99.
- MUNRO, Pamela. 1990. Stress and Vowel Length in Cupan Absolute Nouns. *IJAL* 56:217–250.
- PENNINGTON, Campbell W., ed. 1979 *Vocabulario en La Lengua Nevome*, The Pima Bajo of Central Sonora, Mexico. Vol. 2. Salt Lake City: University of Utah Press.
- PRESS, Margaret. 1979. *Chemehuevi: Grammar and Lexicon*. UCPL. 92. Berkeley and Los Angeles: University of California Press.
- ROBINSON, Lila W., and James Armagost. 1990. *Comanche Dictionary and Grammar*.
- SAPIR, Edward. 1931. *Southern Paiute Dictionary*. Publications of the American Academy of Arts and Sciences 65, no. 3.
- SAPIR, Edward. 1931. *The Southern Paiute language*, Parts I, II, and III. AAASP 65.1–730.
- SAXTON, Dean, and Lucille SAXTON. 1983. Dictionary: *Papago and Pima to English, English to Papago and Pima*. Tucson: University of Arizona Press.
- SEILER, Hansjakob, and Kojiro HIOKI. 1979. *Cahuilla Dictionary*. Banning, Calif.: Malki Museum Press.
- SOD: *Shoshoni Online Dictionary*, at: [www.shoshonidictionary.com](http://www.shoshonidictionary.com).
- STAROSTIN, George. 2010. Preliminary lexicostatistics as a basis for language classification: a new approach. *Journal of Language Relationship* 3:79–116.
- VOEGELIN, C. F. 1958. Working dictionary of Tubatulabal. *IJAL* 24:221–28.
- VOEGELIN, C. F., F.M. Voeglin, and Kenneth Hale. 1962. Typological and Comparative Grammar of Uto-Aztecan: I (Phonology). *IJAL Memoir* 17.
- WHORF, Benjamin. 1946. The Hopi language, Toreva dialect. *Linguistic Structures of Native America*, ed. by H. Hoijer, 158–183. New York : The Viking Fund, Inc.
- ZIGMOND, Maurice L., Curtis G. BOOTH, and Pamela MUNRO. 1990. *Kawaiisu. A Grammar and Dictionary with Texts*. UCPL 119. . Berkeley and Los Angeles: University of California Press.

Статья представляет собой краткий обзор наработок по исторической фонологии юто-ацтекской языковой семьи. Основными предметами обсуждения являются генетическая классификация языков этой семьи, реконструкции фонологических систем праязыков ее отдельных ветвей и собственно праюто-ацтекская реконструкция. Лексический материал, на основании которого разработаны соответствующие реконструкции, взят из электронных баз данных, созданных в рамках проекта «Эволюция языка» (Институт Санта-Фе).

## The Austroasiatic central riverine hypothesis<sup>1</sup>

The paper considers the vexing issues of the homeland and dispersal of the Austroasiatic languages. A critical analysis finds little firm support for nested sub-groupings among a dozen recognised branches, while lexical analyses suggest a long-term pattern of contact and convergence within mainland Southeast Asia. These facts are interpreted as consistent with a stable long-term presence in Indo-China, probably centred on the Mekong River. The most geographically distant branch — the Munda of India — is treated as a highly innovative outlier, and the evolution of Munda root structure is reconstructed, consistent with this theory.

*Keywords:* Austroasiatic, Munda, comparative method, language relationship, lexicostatistics.

### Introduction

The question of localizing the Austroasiatic (AA) homeland is a crucial and increasingly topical one for scholars interested in the prehistory of Mainland Southeast Asia. A geneological language classification can inform discussion of this point, since the correlation of geography and phylogenetic distribution may permit inferences concerning migration routes, contacts, and time depths. In this context, it is both intriguing and frustrating that, after more than a century of comparative AA studies, scholars have yet to present an explicitly justified and comprehensive internal genetic classification of the phylum (see Sidwell 2009 for a general discussion).

For sure, there are various proposals in print, and in unpublished sources such as dissertations, conference presentations, and manuscripts circulating informally. But when these disparate sources are tracked down, compared, and analysed, it becomes abundantly clear that there is no scholarly consensus on:

- the relations between AA branches,
- the age or diversity of AA,
- an appropriate program for addressing these issues

Consequently, the field is yet to significantly benefit from multidisciplinary research. Scholars eager to pursue the synthesis of archaeology, genetics and linguistics are exasperated, such as recently expressed by Roger Blench:

Austroasiatic languages are the most poorly researched of all those under discussion. Many are not documented at all and some recently discovered in China are effectively not classified. The genetics of Austroasiatic speakers are almost unresearched. Austroasiatic is conventionally divided into two families, Mon-Khmer (in SE Asia) and Munnda (in India). Diffloth (2005, 79) now considers Austroasiatic to have three primary branches but no evidence for these realignments has been published. Indeed Austroasiatic classification has

---

<sup>1</sup> The present paper includes material first presented in a plenary address delivered to the Southeast Asian Linguistics Society meeting held in HoChiMinh City on May 28 2009. Research that made this paper possible was supported by generous assistance of the National Endowment for the Humanities (Washington). Any views, findings, conclusions or recommendations expressed in this publication do not necessarily represent those of the National Endowment for the Humanities.



been dogged by a failure to publish data, making any evaluation of competing hypotheses by outsiders a merely speculative exercise. (Blench 2008, 117–118)

In this paper it is argued that the thirteen<sup>2</sup> AA branches radiated more or less equidistantly from proto-AA. Consequently, since the greatest number of AA branches is spoken along an axis that runs roughly southeast to northwest along the middle course of the Mekong river, it is reasonable to suggest that the AA languages dispersed along and out from that axis. This theory is provisionally called the *Austroasiatic Central Riverine Hypothesis*.

### Recent thinking – India and China

Since the second half of the 1990s there has been a resurgence of interest in AA linguistics and the homeland question. Among the various suggestions that have been offered over the years, there are two broad lines of inquiry that have been especially emphasized:

1. a western origin, in eastern India or about the Bay of Bengal, or
2. a northern origin, in central or southern China

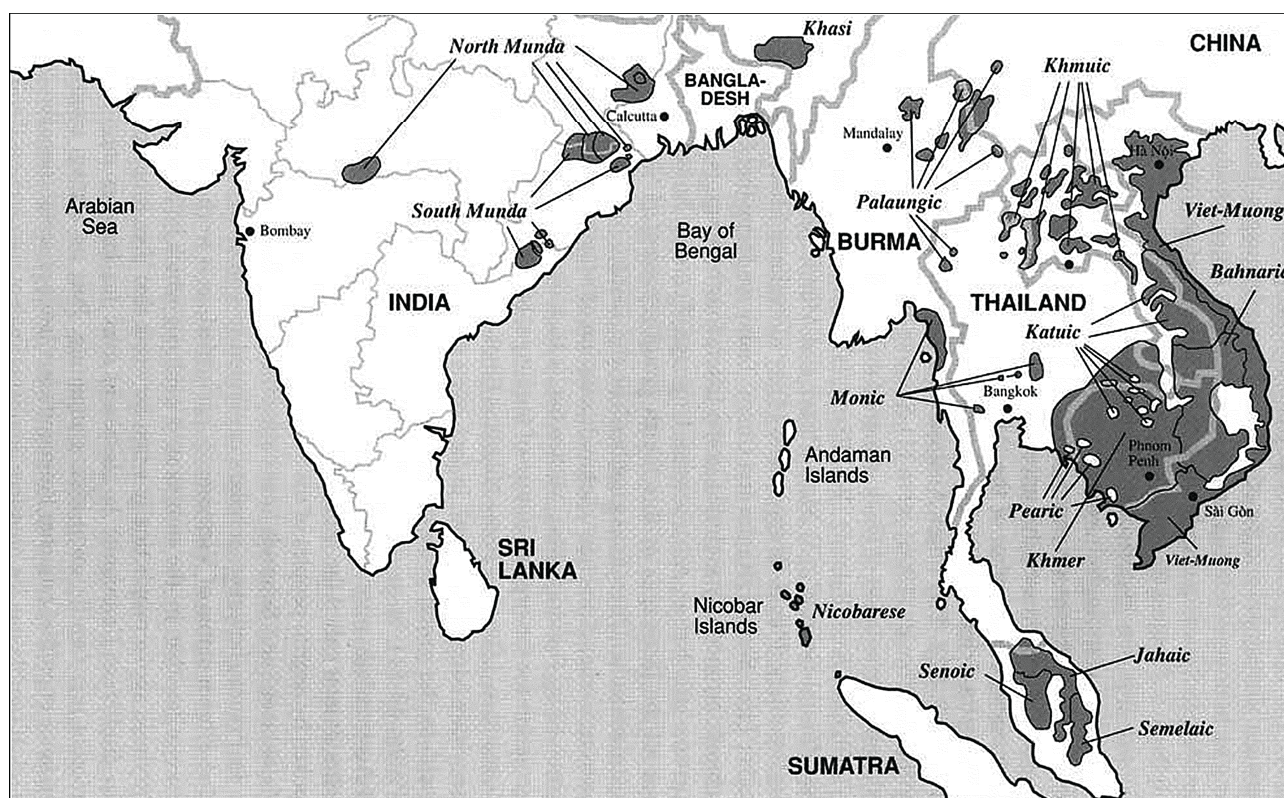


Figure 1: Map of Austroasiatic languages (van Driem 2001)

Gerard Diffloth (2005 and elsewhere) has argued that, on the basis of floral and faunal terms, the AA lexicon rules out a temperate zone (i.e. China) in favour of a tropical homeland. Assuming a primary split between Munda and Mon-Khmer within AA, he proposes a location proximal to the Bay of Bengal (having earlier proposed the Burma-Yunnan border zone).

<sup>2</sup> The figure thirteen assumes the following branches: Munda, Khasi, Palaungic, Khmuic, Mangic, Vietic, Katuic, Bahnaric, Pearic, Khmer, Monic, Aslian, Nicobarese.

George van Driem has been actively promoting this view in print (especially his 2001 handbook). Van Driem makes much of the literature which purports to identify an old Munda substrate in Vedic etc. (especially Kuiper 1955, 1967, 1991, and recently Witzel 1999).

By way of contrast, Peiros (1989, 1998) and Peiros & Shnirelman (1998), have asserted that the AA lexicon indicates a non-tropical, non-coastal location. Combining this idea with data about the mid-Holocene climate, and supposed lexical isoglosses with Hmong-Mien languages (some two dozen or more supposed resemblances), they propose an origin on the mid-Yangtse river. This view dovetails with philologists' suggestions of various AA etymologies for words of uncertain origins in classical languages of China. Mei & Norman (1976) especially are widely cited as proposing AA loans into Chinese, including even the name of the Yangtse itself. Schuessler (2007) more recently proposes hundreds of Mon-Khmer-Old Chinese comparisons, so many that he writes (p.4): "When pursuing OC and TB/ST etyma down to their roots, one often seems to hit Austroasiatic bedrock, that is, a root shared with Austroasiatic."

Generally these philological arguments have the following character: there are a handful of close phonetic and semantic matches which are intriguing, plus there are scores or more of vague resemblances which cannot be organised into any system of regular correspondences. Unfortunately this is what one expects to find when comparing unrelated but typologically similar languages. Presently we have no clear reasons to favour either the sinophiles or the indophiles, indeed shouldn't we begin from the premise embedded in the traditional philologist's wisdom that the search for Latin etymologies should begin on the Tiber? After all, the various philological arguments suppose localising proto-AA in places where little or none of the diversity of the AA phylum is presently found. There are no AA speaking communities along the middle Yangtse or in China's south-eastern provinces. There are no AA speaking communities around the Bay of Bengal, and the closest reasonably diverse group — the Munda — may not be so especially diverse as to justify giving their position any particular weight (see discussion below).

For about a century now, especially since Sapir (1916) it has generally been recognised that, since diversity within a language group tends to increase over time, a region of higher diversity is likely to have been settled longer, and therefore is more likely to be, or be proximal to, the homeland. Dyen (1956) and Diebold (1960) specifically formalized this (and related ideas) under the heading of 'Migration Theory', and this approach is applied here. The efficacy of the theory is readily confirmed in the real world; for example the south of Vietnam was settled by Vietnamese speakers more or less since the formal incorporation of Champa into Vietnam from 1693, and the diversity of Vietnamese dialects in the south is low, but markedly higher in the north, especially for example around Vinh.

The present task is to look at the facts of the AA family and decide whether we can fairly characterize its phylogenetic diversity. While it is uncontroversial that the great bulk of AA languages are spoken in what I am calling the central zone, it is another matter whether they represent the greater proportion of coordinate branches. Consequently it is the issue of the higher branching structure of AA that one must focus..

### **AA Classifications since Pinnow (1959)**

Presently the view is widely received that the AA phylum is composed of two coordinate families; e.g.:

The primary split in the family is between the Munda languages in central and eastern India and the rest of the family. (Anderson 2006, 598)

Although a somewhat different conception is also reported, e.g:

The Austroasiatic language family is conventionally divided into three branches or sub-families, viz. the Munda, the Nicobarese and the Mon-Khmer languages. (van Driem 2001, 262)

Both of these propositions give a special status to Munda, an idea that can be traced directly to the works of Pinnow (1959, 1960, 1963 etc.). In Pinnow's 1960 paper Munda was characterized as preserving ancient morphological complexities that find only murky traces in other AA groups, such as apparent alternations of final consonants in Khmer. Shortly after in 1963 he presented the classification reproduced here at Figure 2, which effectively divides the AA into Western (Munda) and Eastern (Khmer-Nicobar) families, with Mon-Khmer further divided into Nicobarese and 'Palaung-Khmer' sub-families. Thus we see the origins of both views quoted above. Earlier studies, such as Schindt (1906) actually gave Munda a much lower status, sub-grouping it with languages of Indo-China.

- Western group (Nahali-Munda)
- (A) West : Nahali (?)
  - (B) East : Munda
    - (a) North
      - Kherwari (Santali, Mundari, Korwa, etc.)
      - Kurku
    - (b) South
      1. Central : Kharia, Juang
      2. South-East : Sora, Pareng, Gutob, Remo
- Eastern group (Khmer-Nicobar)
- (A) West : Nicobarese (Nancowry, Car, etc.)
  - (B) East : Palaung-Khmer
    - (a) West : Khasi
    - (b) North : Palaung-Wa (Palaung, Wa, Riang, Lawa, etc.)
    - (c) East : Mon-Khmer (Mon, Khmer, Bahnar, Sre, etc.)
    - (d) South : Malacca
      1. Sakai
      2. Jakud
      3. Semang

Figure 2: *Austroasiatic classification by Pinnow (1963)*

However, Pinnow's approach was more typological than historical, and he made an explicit statement to the effect that his Eastern (Khmer-Nicobar) group might not be a single phylogenetic division in opposition to Munda; but this caution appears to have been almost entirely ignored in subsequent scholarship.<sup>3</sup> Subsequently the structure so clearly communicated by Pinnow's figure came to be widely received.

In the following decade Diffloth's (1974) Encyclopedia article became especially influential. There Diffloth presented the modified Pinnow model with three coordinate families (Munda, Nicobarese, Mon-Khmer). His Mon-Khmer family was further sub-divided according to the lexicostatistical studies of Thomas and Headley (1970 especially). Later (e.g. 1979, 1999) Diffloth merged Nicobarese into Mon-Khmer as a sister of Aslian, and more recently (2005) promoted Khasi-Palaungic-Khmuic to form a 'new' third coordinate family.

Such classifications characterize Munda as comparable in terms of historical diversity to the rest of AA, and thus suggest that the geographical centre of that diversity is unlikely to have been in Southeast Asia but further to the west.

Presently it is not possible to assess Diffloth's classifications, since he has not made his data and arguments available. Still, we are able to address the most salient aspect — whether

<sup>3</sup> I have found no explicit recognition of this point in my reading of relevant literature.

it is justified to treat Munda as a primary branch which coordinates with just one or two other primary branches.

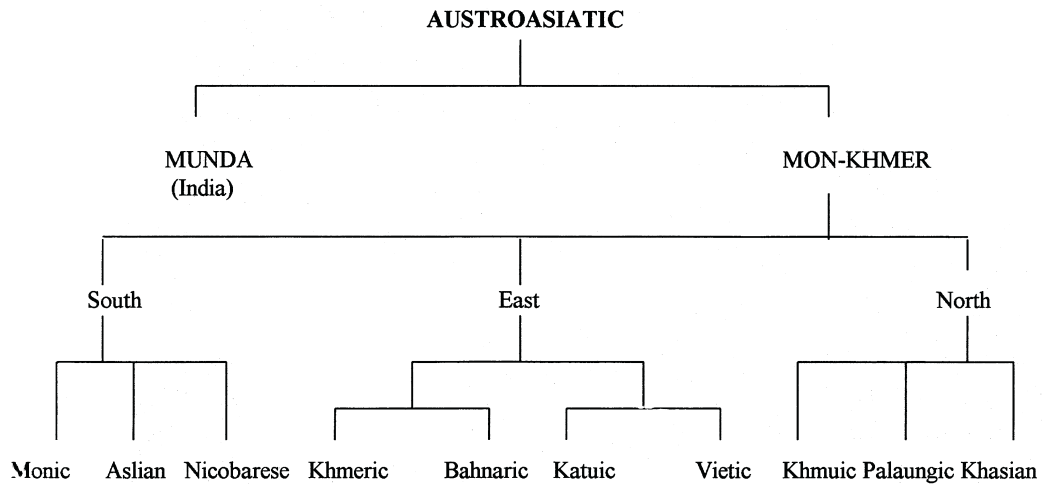


Figure 3: Austroasiatic classification by Diffloth (reproduced from Chazée 1999)

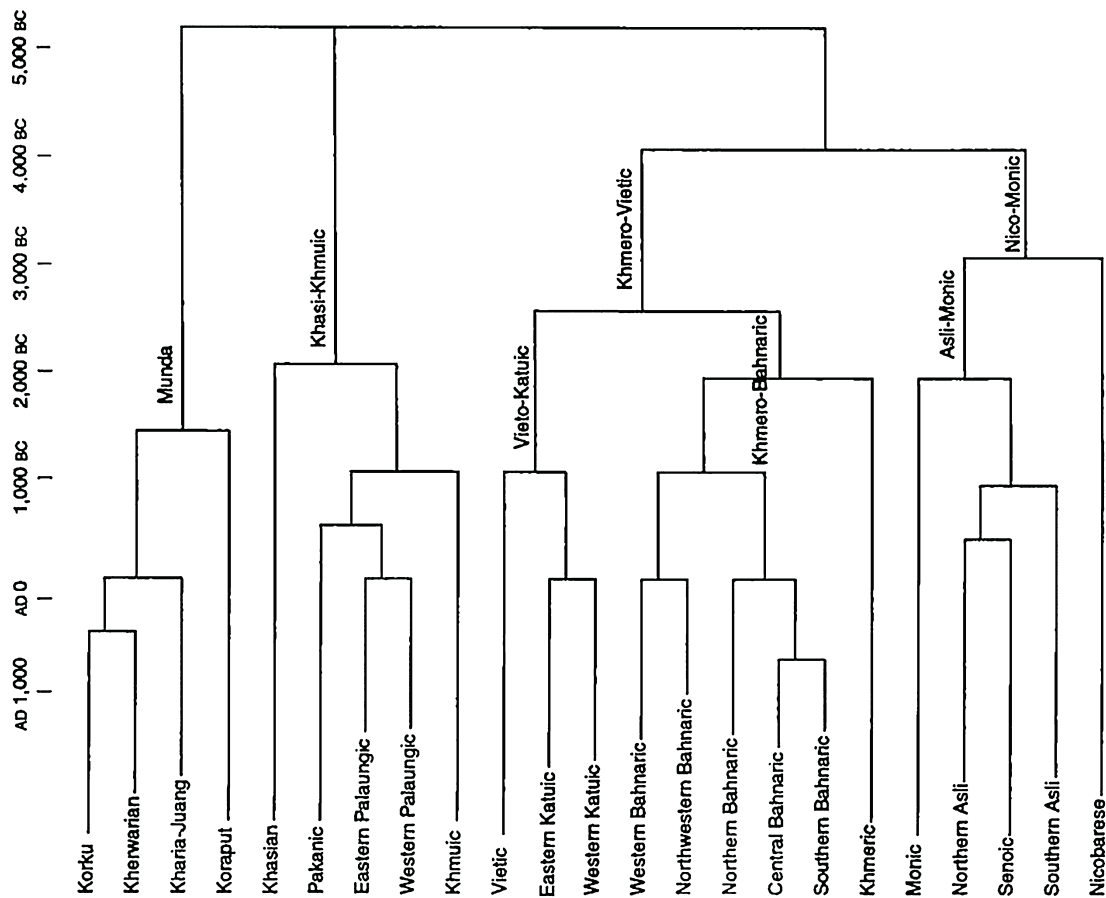


Figure 4: Austroasiatic classification by Diffloth (2005)



## Munda as morphologically archaic?

Uniquely within AA, the Munda languages are highly agglutinative, making extensive use of suffixation. Characterizing the features of Munda as ancient, Pinnow asserted that:

... the Munda languages are undoubtedly more similar to Proto-Austroasiatic than the other members of the family. From a morphological viewpoint they are far more conservative than Nicobarese and Khasi, and from the standpoint of vocabulary they surpass the Mon-Khmer languages in their preservation of ancient word stems and word forms. (Pinnow 1963: 150)

Pinnow's assessment has precise and profound implications for the homeland problem. In asserting that Munda is more conservative, it strongly suggests that the Mon-Khmer languages innovated a typological restructuring away from suffixing etc. a change so profound that it probably did not happen more than once, and hence the Mon-Khmer languages are sprung from a common source, a single branching node in the AA tree. Such a model suggests that the ancestral Mon-Khmers migrating eastward, morphologically simplifying their languages the farther they went, until at the eastern extreme the Vietnamese (for example) eliminated any trace of historical affixes, recycling everything into compact monosyllables.

Some find this view persuasive. So confident is van Driem (2001:299) that he castigates Reid (1994), who attempted to resurrect the 'Austriac' hypothesis by comparing Nicobarese and Austronesian, for not comparing with "the grammatically more conservative Munda languages". Van Driem went on to suggest that the intrusion of Tibeto-Burmans into Bengal stimulated the initial dispersal of AA to the west and east respectively, leaving the gap in the distribution of AA we see there today.

Yet it does not necessarily follow that the history of Mon-Khmer has been one of inexorable morphological loss — since the 1980s, Patricia Donegan and David Stampe have been variously arguing the case for a restructuring from isolating to synthetic typology, such that Pinnow's formulation ought to be reversed, and Mon-Khmer characterised as structurally closer to proto-AA. The argument turns on the issue of phrasal accent. Munda accent is falling while in Mon-Khmer it is rising. This is crucial because the patterns of stress or accent have significant consequences over time, and tend to have strong structural correlates, particular in the distribution of clitics and affixes.

In falling accent languages morphemes that follow the head will tend to be phonetically reduced, and may grammaticalize into suffixes. On the other hand, in rising accent languages, morphemes that precede the head will tend to coalesce into prefixes. This suggests that one may reconstruct the direction of change based on typological observations. This is what Donegan (1993) had in mind when she pointed out that:

Mon-Khmer has neither inflection nor suffixes, and neither did proto-Austroasiatic, but Munda have scores of them. [...]

Second, note the complex derivational and inflectional verb morphology, again largely suffixal. Pinnow, [...], provided etymologies of many of these suffixes. Few of these reconstruct as suffixes to proto-Munda, i.e. they have been developed in the individual languages. (Donegan 1993:341)

And Stampe explained at the 2004 SEALS meeting in Bangkok:

Only Pinnow (1960) and Zide & Anderson (2001) seem to have taken this view. Pinnow's argument was based on evidence for fossil suffixes in Khmer, none of which were supported by the exhaustive study of Khmer morphology by Jenner and Pou (1980–1981), and Pinnow's evidence was explained away by Jacob (1992),<sup>4</sup>

<sup>4</sup> The correct reference appears to be Jacob (1989–1990).

who showed that the variation such as Pinnow cited was expressive, not grammatical. Z & A's argument has two parts: (1) the occurrence of enclitic object pronouns and rarely nouns in Nicobarese and a few other Mon-Khmer languages, and (2) their reconstructions of the proto-Munda verb, which they with no specific evidence see as typical of Austroasiatic as well, with the implication that every trace was lost in Mon-Khmer. (Stampe 2004, 4)

Donegan & Stampe (2004:6) also compare Munda verb morphology to word-order equivalents in Mon-Khmer clauses, e.g:

Sora:	(anin)	əḍ-	məl-	tiy	-dar	-ij	-da	-e
	he	not	want	give	rice	me	-AUX	-3PR
Khmer:	kəət	ʔətcaŋ	ʔaoy	baay	kɲom			

Here we see a complete functional morpheme to morpheme correspondence, even though in this case only one morpheme is actually cognate (Sora **əḍ-** : Khmer **ʔət**). Unambiguously, Mon-Khmer VO word-order is preserved in Sora, grammaticalised as bound morphology, while at the clause level OV word-order prevails.

I would go even further, and propose that AA had had rising phrasal accent for a long time before the proto-language stage and the break-up of the family. Crucial is the observation that while only Munda has extensive suffixing, both Munda and Mon-Khmer have prefixes and infixes, and as Anderson (2004) and others have shown, this prefixing and infixing is cognate and reconstructable to the earliest times.

On typological grounds we might also assume to derive AA infixes from metathesized prefixes. In the present case, we can even suggest the specific source: infixes **-m-** and **-n-** derived from the implosives **ɓ**, **ɗ** which are otherwise rare or missing as prefixes, but securely reconstructable as frequent initials in the proto-language (see for example Shorto 2006). Proto-AA must have had rising accent long enough not only to develop prefixes, but infixes as well. The shift to falling accent in pre-Munda triggered the rise of suffixation, where AA prefixes and infixes remain recognizably intact. Clearly this implies the historical primacy of rising phrasal accent.

The question of accent also offers an explanation for Munda root structure. The problem is to reconcile the monosyllables and trochaic disyllables of the Munda root canon, with the monosyllables and characteristically iambic sesquisyllables of Mon-Khmer languages.

In the first place we can assume that a proportion of roots have always been monosyllables and these require no special explanation. Mon-Khmer languages like Vietnamese (e.g. Ferlus 1998), Nyaheun (e.g. Sidwell & Jacq 2003), U (Svantesson 1988) reduced their sesquisyllables to monosyllables. The question is the approximately half of the lexicon which, in conservative Mon-Khmer languages, is sesquisyllabic. Shorto (2006) reconstructed a sesquisyllabic root canon \*(C)CV(:)C, and we might also apply this formula to proto-AA.

According to Donegan and Stampe's typology the rightmost edge of the phonological word will show reduced phonological and morphological structure. Significantly, the coda position universally in AA languages lacks any contrast of VOT, i.e. in Mon-Khmer we find unreleased voiceless stops with weak glottal coarticulation, and in Munda we find voiced stops with strong glottal coarticulation — in both cases it is a single series without voicing contrast, consistent with the theory.

Additionally, within Mon-Khmer there is frequently contrastive vowel length, while it is uncommon in Munda. In Mon-Khmer sesquisyllabic words mainsyllables are assigned 2 mora, and presyllables have no effective rhythmic weight and their vowels no phonemic value. Consistent with mainsyllables carrying 2 beats, we have the strong tendency for diphthongization

and register splits in Mon-Khmer. By contrast, Munda vowel inventories tend to be small, perhaps only a couple (if any) preserve old length distinctions.

We may suppose that in pre-Munda the phrasal accent fronted, and speakers began to assign a rhythmic beat to the presyllable (and reduplicating or otherwise augmenting monosyllables), taking a mora from the mainsyllable and giving it to the new initial syllable. Reduced to one mora, there would be little structural motivation for length or diphthongization contrasts. We can still see the traces of this process by looking at etymologies that suggest proto-long vowels. Consider the following examples (data and reconstructions from Shorto 2006):

PAA long vowels echoed:

<b>*bluu?</b> ‘thigh’		<b>*jliiŋ</b> ‘long’	
Khmer:	<b>phlɔ̃u</b>	Old Mon:	<i>jlīñ</i>
Bahnar:	<b>blu:</b>	Lawa:	<b>ɰleiŋ</b>
Palaung:	<b>blu</b>	Nicobarese:	<b>cəliŋ</b>
Temiar:	<b>bəlo?</b>	Mundari:	<b>jiliŋ</b>
Nicobarese:	<b>pulɔ:</b>	Bhirhor:	<b>jiliŋ</b>
Sora:	<b>bulu:-n</b>		
Kharia:	<b>bhulu</b>	<b>*kluu?</b> ‘tortoise’	
Mundari:	<b>bulu</b>	Mon:	<b>klao</b>
Ho:	<b>bulu</b>	Stieng:	<b>blɔ:u</b>
Kurku:	<b>bulu</b>	Sora:	<b>ku(:)lu:</b>
		Kharia:	<b>kulu</b>
<b>*brii?</b> ‘forest’		<b>*rk[aw]?</b> ‘husked rice’	
Bahnar:	<b>bri:</b>	Khmer:	<b>ɰŋkɔ:</b>
Praok:	<b>praj</b>	Palaung:	<b>rəko</b>
Jah Hut:	<b>bəri?</b>	Vietnamese:	<b>gəo</b>
Old Khmer:	<b>vrai</b>	Nyahkur:	<b>ŋkɔ́:</b>
Kharia:	<b>biru</b>	Kharia:	<b>'rŋku?b</b>
Mundari:	<b>bir</b>	Sora:	<b>'ruŋku:</b>
Bhirhor:	<b>bir</b>	Juang:	<b>runku:</b>
Ho:	<b>bir</b>	Gutob:	<b>runku:</b>
<b>*jhaam</b> ‘blood’		<b>*kjaal</b> ‘air, wind’	
Khmer:	<b>chì:əm</b>	Old Mon:	<i>kyāl</i>
Bahnar:	<b>pha:m</b>	Bahnar:	<b>kja:l</b>
Nicobarese:	<b>maha:m</b>	Khmer:	<b>khjɔl</b>
Kammu-Yuan:	<b>mà:m</b>	Kuy:	<b>kja:l</b>
Sora:	<b>mijə:m</b>	Kharia:	<b>kɔjɔ</b>
Kharia:	<b>ijām</b>	Juang:	<b>kojo</b>
Santali:	<b>mājām</b>	Santali:	<b>hɔe</b>
Mundari:	<b>mājōm</b>	Mundari:	<b>hɔjɔ</b>

**\*kmuu?** ‘dirty’

Khmer:	<b>khmau</b>
Kuy:	<b>kmau</b>
Palaung:	<b>kəmu</b>
Bhirhor:	<b>humu</b>
Mundari:	<b>humu</b>
Ho:	<b>homu</b>
Kurku:	<b>kumu</b>

**\*smuul** ‘shadow, soul’

Khmer:	<b>srəmaol</b>
Mon:	<b>həmao</b>
Kuy:	<b>smə:l</b>
Bahnar (infix):	<b>pəhŋə:l</b>
Sora:	<b>um'mul</b>
Santali:	<b>umul</b>
Mundari:	<b>umbul</b>

## PAA short vowels – variable/conditioned new initial syllable vowels:

**\*kra?** ‘road, way’

Kuy:	<b>kna:</b>
Praok:	<b>kra</b>
Muong:	<b>khá</b>
Mundari:	<b>hora</b>
Bhirhor:	<b>hora</b>
Kurku:	<b>kora</b>

**\*kla?** ‘tiger’

Bahnar:	<b>kla:</b>
Old Mon:	<b>kla(')</b>
Khasi:	<b>khla:</b>
Khmer:	<b>khla:</b>
Santali:	<b>kul</b>
Mundari:	<b>kul ~ kula</b>
Korwa:	<b>kul ~ kula</b>

**\*[hj]mu?** ‘name’

Old Mon:	<b>jamo', himo'</b>
Praok:	<b>mə</b>
Semai:	<b>muh</b>
Old Khmer:	<b>jmah</b>
Sora:	<b>ə'jam-ən</b>
Kharia:	<b>'jimi, 'jimi</b>
Mundari:	<b>num ~ nutum</b>
Ho:	<b>numu</b>
Kurku:	<b>jumu ~ jimu</b>

**\*tmi?** ‘new’

Mon:	<b>kəmɔɛ?</b>
Khmer:	<b>thmɔj</b>
Thin:	<b>hme</b>
Vietnamese:	<b>mói</b>
Sre:	<b>təme</b>
Kharia:	<b>'tɔnmɛ</b>
Sora:	<b>'tamme:-</b>
Remo:	<b>tamme</b>

**\*kn[i]?** ‘rat,  
mouse’

Old Mon:	<b>kni(')</b>
Bahnar:	<b>kəne:</b>
Kammu Yuan:	<b>kəné?</b>
Khasi:	<b>khnai</b>
Kharia:	<b>kəne</b>
Mundari:	<b>huni</b>
Kurku:	<b>ho'ne'</b>

Occasional reduplication of all or part of monosyllable:

<b>*baʔ</b> 'paddy' (* <b>(kə)ba:ʔ</b> Diffloth 2005)		<b>*[b]uuk</b> 'head'	
Bahnar:	<b>ba:</b>	Stieng:	<b>bu:k</b>
Temiar:	<b>ba:ʔ</b>	Sre:	<b>bou</b>
Danaw:	<b>bā</b>	Kharia:	<b>bəʔ / bəkəʔb</b>
Khasi (prefixed):	<b>kba</b>	Mundari:	<b>bəhəʔ ~ bəʔʔ</b>
Kharia:	<b>baʔ</b>	Sora:	<b>bəʔəb</b>
Mundari:	<b>baba</b>	Juang:	<b>bəkəʔ</b>
Sora:	<b>ba:ba:</b>		
Kurku:	<b>baba</b>		

Simple monosyllabic roots with short vowels have straightforward reflexes:

<b>*kap</b> 'bite'		<b>*mat</b> 'eye'	
Bahnar:	<b>kap</b>	Bahnar:	<b>mat</b>
Kensiu:	<b>kap</b>	Kensiu:	<b>məd</b>
Katu:	<b>kap</b>	Katu:	<b>mat</b>
Car Nicobar:	<b>kap</b>	Car Nicobar:	<b>mət</b>
Sora:	<b>kab</b>	Sora:	<b>mʔo:d</b>
Santali:	<b>haʔb</b>	Kharia:	<b>məʔd</b>
Mundari:	<b>haʔb</b>	Mundari:	<b>məʔd</b>
Kurku:	<b>kap</b>	Kurku:	<b>met</b>
<b>*təh</b> 'breast'			
Bahnar:	<b>təh</b>		
Semelai:	<b>tuh</b>		
Katu:	<b>təh</b>		
Car Nicobar:	<b>təh</b>		
Santali:	<b>toa</b>		
Bhirhor:	<b>toa</b>		
Mundari:	<b>toa</b>		

The following etymon has Eastern reflexes indicating both short and long variants, and Munda reflexes are similarly variable:

<b>*jə[ə]ŋ</b> 'foot/leg'			
Old Mon:	<i>juŋi</i>	Kharia:	<b>gujuŋ</b>
Palaung:	<b>juŋ</b>	Sora:	<b>ju:ŋ</b>
Bahnar:	<b>juŋ</b>	Mundari:	<b>janŋa</b>
Khmer:	<b>cə:ŋ</b>	Kurku:	<b>janŋa</b>

I suggest that such examples indicate that Munda roots are readily derived from sesquisyllabic proto-AA roots, more or less consistently with the root-canon reconstruction of Shorto (2006). Consequently, we cannot characterize the so-called Mon-Khmer languages as forming a sub-group in opposition to Munda. Rather they are relics that maintain (subject to various local changes) AA rising phrasal rhythm and its typological correlates. Thus there is no structural basis for thinking that Munda is “more similar to proto-Austroasiatic”, and we need to turn to other methodologies to investigate this problem.

### Linguistic Phylogeny: phonological reconstruction

For classification of a language family, especially of the structurally isolating type, the gold standard is a phonological reconstruction that identifies innovations with branching nodes. In the case of AA a comprehensive reconstruction has not been presented, rather we have only various branch level studies, and the attempt by Shorto (2006) which does not analyse Munda correspondences (Sidwell in press offers a comprehensive review of such studies). The reconstruction of the proto-AA vowels is very problematic, and results are accruing only gradually. However, the broad outlines of AA consonantism have been fairly well understood for a century already, being largely sorted out by Wilhelm Schmidt (1905, 1906 etc.). The basic consonant system preserved in conservative Mon-Khmer languages (such as Katu) and in the epigraphic record of Mon and Khmer before they underwent devoicing and restructuring, appears to adequately account for the proto-AA consonants without needing to posit additional segments or articulation types.

I reproduce here at Table 1 AA consonant correspondences for initial/prevocalic stops and fricatives — it is my working model at the time of writing. The table compares the oral stops reconstructed for each branch, plus the proto-MK (effectively proto-AA) reconstructions from Shorto (2006). Consonant mergers are shaded; splits are also indicated although without conditioning factors given (these are not yet thoroughly worked out). The table can be compared with the correspondences tabled by Shorto (2006: 52–54).

The most common change we can see is a merger of voiced and implosive stops, which occurred in six out of 12 branches.<sup>5</sup> This is a well distributed and rather trivial change which does not suggest any sub-grouping, not even among neighbouring groups. For example, it occurs in both Pearic and Khmer, although with the former there is a (partial only?) shift to aspiration among the voiceless stops.<sup>6</sup>

Perhaps the most important type of change, the devoicing on initials that is closely associated with restructuring of vowel systems, is not evident at all. It is clear that a voicing contrast is reconstructable for all branches, and such restructuring must have occurred independently after individual branches began diversifying internally.

Thus there is no patterning among the evident splits and mergers which suggest nested sub-branching among branches — their phonological histories appear to be quite independent, so far as I can tell. Consequently we have one other place to look for sub-grouping indications — the lexicon.

---

<sup>5</sup> Mangic is not included.

<sup>6</sup> There was discussion about this apparent Pearic split at the ICAAL4 meeting in Bangkok (October 28 2009); Michel Ferlus proposed that there was no conditioned split, rather there was borrowing of words with plain voiceless stops after the original voiceless series had become regularly aspirated.

PMK	Munda	Khasi	Palaungic	Khmuic	Vietic	Katuic	Bahnaric	Pearic	Pre-Khmer	Old Mon	Aslian	Nicobaric
*p-	*p-	*p <sup>h</sup> -	*p-	*p-	*p-	*p-	*p-	*p <sup>h</sup> - *p-	*p-	*p-	*p-	*f-
*b-	*b-	*p-	*b-	*b-	*b-	*b-	*b-	*b-	*b-	*b-	*b-	*p-
*ɓ-	*b-	*b-	*ɓ-	*b-	*ɓ-	*ɓ-	*ɓ-	*b-	*b-	*ɓ-	*b-	*p- (~*ʔ)
*t-	*t-	*t <sup>h</sup> - *t-	*t-	*t-	*t-	*t-	*t-	*t <sup>h</sup> - *t-	*t-	*t-	*t-	*t-
*d-	*d-	*t- *d-	*d-	*d-	*d-	*d-	*d-	*d-	*d-	*d-	*d-	*t-
*d-	*d-	*d-	*d-	*d-	*d-	*d-	*d-	*d-	*d-	*d-	*d-	*d-
*h-	*G- (Ø- ~ h-)	*h-	*h-	*h- -Ø-	*h-	*h-	*h-	*h-	*h-	*h-	*h-	*h-
*s-	*s-	*f- *s-	*s-	*s-	*s-	*s-	*s-	*s-	*s-	*s-	*s-	*h-
*c-	*s-	*s-	*c-	*c-	*c-	*c-	*c-	*c <sup>h</sup> - *c-	*c-	*c-	*c-	*s-
*j-	*j-	*j- *j-	*j-	*j-	*j-	*j- *f-	*j-	*j-	*j-	*j-	*j-	*c-
*k-	*q- *k-	*k <sup>h</sup> -	*k-	*k-	*k-	*k-	*k-	*k <sup>h</sup> - *k-	*k-	*k-	*k-	*k-
*g-	*g-	*k-	*g-	*g-	*g-	*g-	*g-	*g-	*g-	*g-	*g-	*k-

Table 1: Phonological correspondences for initial/pre-vocalic oral stops between Austroasiatic branches (PMK from Shorto (2006), proto-Munda follows Pinnow, other proto-branch reconstructions by Sidwell);  
(working model at 1/2/10)



## Lexicostatistics

Southeast Asian linguistics has frequently embraced the lexicostatistical method. Typically it is employed as a useful discovery tool, preliminary to a comparative reconstruction, and AA has been the subject of numerous lexicostatistical analyses. The most recent and extensive of these studies is Peiros (2004, in Russian), which utilizes data from over 100 languages. Peiros applies Starostin's glottochronological method which is controversial.<sup>7</sup> His AA classification is reproduced here at Fig. 5;<sup>8</sup> he proposes an elaborate nested branching structure for the phylum; interestingly similar to Pinnow (1960) and Diffloth (1974) in that the highest branches involve Nicobarese, Munda and Mon-Khmer (although in a somewhat different relation).

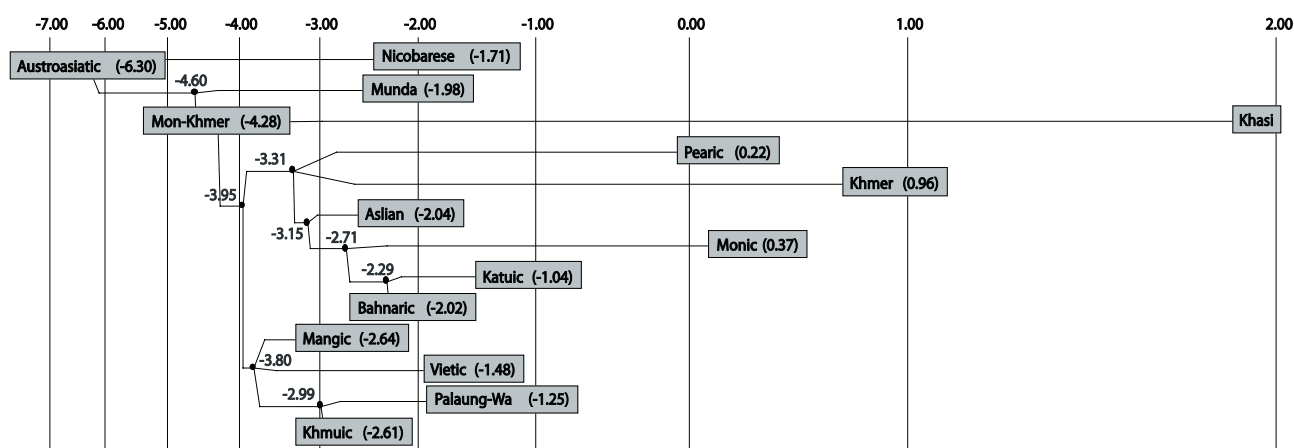


Figure 5: Peiros (2004) Lexicostatistical classification

I decided to replicate the study, but with a smaller, more manageable set of data, in this case 36 languages. A major concern being to use only languages where both the sound correspondences and the major contact languages are reasonably well understood/documented. My data set, with cognate assignments and commentary is publicly available via my website at <http://people.anu.edu.au/~u9907217>. The matrix at Fig. 7 was generated using Glotpc.exe (devised by Jacques Guy, and freely downloaded via link at <http://sil.org>). The results came out quite different to those of Peiros — it is my analysis that he underestimated both borrowings between branches (especially from Mon and Khmer into neighbouring languages) and placed too much importance on some very low cognate percentages (such as Nicobarese) where isolation and social factors must affect the rate of lexical change.<sup>9</sup> Thus I contend that a somewhat flatter tree is indicated, in fact, lexical signals that might support sub-branching are so confused that it is difficult to argue for something other than a simple rake-like tree.

What especially struck me were remarkably parallels with my matrix and the one offered more than 30 years ago by Frank Huffman (1978), reproduced here at Fig. 6.<sup>10</sup> Although my calculations yielded lower overall cognate densities between branches, there is a structural

<sup>7</sup> The software, and Peiros' data and reconstructions, are accessible via: <http://starling.rinet.ru/>.

<sup>8</sup> Peiros (2004) does not provide a matrix of percentages, only various dendrograms and an appendix of data and numbered cognate assignments.

<sup>9</sup> Peiros mentions the possibility of faster rate of change in Nicobarese, but did not take it into account. Readers can compare Peiros' data with my own cognate assignments and analyses at: <http://people.anu.edu.au/~u9907217/lexico/AAclassification.html>.

<sup>10</sup> Huffman's data is available online via: <http://sealang.net/archives/huffman/>

similarity. Both Huffman and I found the highest interbranch percentage is between Katuic and Bahnaric, plus there is a strong tendency for other branches to show a higher percentage between themselves and Katuic-Bahnaric than with other branches, with the effect declining as one moves geographically further way from Katuic-Bahnaric. Logically it does not seem possible to represent these relations as nested branching relationships.

Chart 1: Interbranch Cognate Percentages and Averages

	K-B	Kme	Mon	Pal	V-M	Pea	Asl	Kmu	Kha	Mun	Nic.
Katuic-Bahnaric		47	42	34	33	35	31	32	24	27	20
Khmer	47		33	31	30	39	28	25	24	24	23
Mon	42	33		33	30	22	24	28	22	20	19
Palaungic	34	31	33		26	25	22	26	26	18	23
Viet-Muong	33	30	30	26		24	22	24	23	20	18
Pearic	35	39	22	25	24		24	22	17	19	16
Aslian	30	28	24	22	22	24		24	20	17	19
Khmuic	32	25	28	26	24	22	24		24	20	17
Khasi	24	24	22	26	23	17	20	24		18	14
Munda	27	24	20	18	20	19	17	20	18		17
Nicobarese	20	23	19	23	18	16	19	17	14	17	
Interbranch totals	324	304	273	264	250	243	231	222	212	200	186
Interbranch averages	32.4	30.4	27.3	26.4	25.0	24.3	23.1	22.2	21.2	20.0	18.6

Figure 6: Austroasiatic lexicostatistical matrix by Huffman (1978)

In Huffman's study, inter-branch averages are shown, which makes the effect very clear. In my matrix individual languages are distinguished, and another potentially interesting pattern is noted: the higher percentage scores with Katuic-Bahnaric actually fall away within, rather than just between, branches. For example: within Palaungic Danaw shows markedly lower scores; within Aslian Jahai scores lower; within Khmuic Mal scores very high, Khmu' moderately, and Mlabri very low. Even within Khasian Wa scores significantly higher against Katuic-Bahnaric than do Standard Khasi or Pnar. Such variations appear to be marginal or absent in respect of Munda and Nicobarese.

How can one explain these patterns? It is possible that to some extent the data of individual languages includes some cases of more rapid lexical replacement that are complicating the picture. This said, we still have a stark geographic correlation: the closer any language or branch is to Katuic or Bahnaric, the higher percentage of cognates will be found with Katuic and/or Bahnaric. At the same time, I have not identified shared innovations that would compel one to sub-group Katuic-Bahnaric. Thus I suggest that we are seeing the effects of prolonged contact, centred about the middle Mekong. Branches that moved further away from this contact area were less affected (such as Palaungic or Aslian), and some (such as Mangic and Nicobarese) came into quite different language contact areas or other social conditions, such that their lexicons were likely subject to even more change than might have otherwise occurred.

Huffman suggested that:

This would seem to argue for an eastern (Central Vietnam) center of dispersal and a separate westward migration for each branch of Austroasiatic. (Huffman 1978:5)

The idea would be that various branches, when they were still basically unitary languages, wandered out of the central zone at different times, hence there are no indications of

Munda			Khasi			Palauang			Mangic			Khmuic			Vietic			Katuic			Bahnar			Pearaic			Khmer			Monic			Aslian			Nico.		
Ko	Mu	So	Kh	Fn	Wa	Da	Wa	U	Bu	Ma	Pa	Ml	Km	Ma	Mu	Vi	Ru	Ku	Pa	Ka	Ba	Ba	Jr	St	Ch	Ka	Su	Kh	Mn	Ny	Se	Sl	Ja	Ca	Na			
45	29		9	9	12	9	12	11	10	14	13	14	17	18	10	13	13	22	19	18	15	19	22	12	12	15	12	13	21	19	14	14	8	9	11			
45	36		10	10	14	13	16	15	12	15	15	12	15	20	14	17	16	21	17	15	15	19	20	15	16	18	15	16	26	24	13	15	8	8	11			
29	36		11	9	11	8	11	14	15	11	15	13	15	15	16	17	17	18	15	14	12	14	20	10	10	11	9	10	17	17	13	10	8	4	8			
9	10	11	84	52	16	18	19	14	15	16	12	14	16	13	14	15	14	11	12	10	13	10	13	10	12	13	13	12	13	13	9	8	8	5	10			
9	10	9	84	55	18	19	20	14	16	18	11	15	15	13	14	15	10	13	9	11	10	10	12	12	13	13	14	13	13	11	10	8	5	10				
12	14	11	52	55	15	20	22	12	15	16	12	16	20	13	13	13	19	16	16	12	15	16	15	16	13	15	14	16	16	13	16	9	7	11				
9	13	8	16	18	15	25	28	28	13	14	14	10	17	15	8	10	10	16	14	16	13	13	13	12	11	13	10	10	17	19	12	12	10	4	10			
12	16	11	18	19	20	25	53	53	18	21	21	14	24	20	17	15	19	20	15	18	17	19	17	16	17	18	17	17	23	25	19	16	14	11	16			
11	15	14	19	20	22	28	53	53	18	20	20	18	30	23	11	14	19	19	17	20	15	19	19	17	17	17	17	24	26	17	15	17	15	7	13			
10	12	15	14	14	12	13	18	18	24	46	31	13	19	14	14	15	18	23	20	21	14	19	20	12	16	15	10	11	17	20	14	14	13	8	10			
14	15	11	15	16	15	14	21	20	24	31	31	12	20	15	18	17	20	16	15	13	17	17	17	15	15	13	21	23	17	14	12	10	15	10	15			
13	15	15	16	18	16	14	21	20	46	31	31	13	17	13	17	15	19	18	17	19	15	18	19	15	18	14	15	19	22	14	15	10	7	11	13			
14	12	13	12	11	12	10	14	18	13	12	13	29	31	9	12	12	18	18	17	15	18	18	18	10	11	14	13	13	16	13	12	11	6	10	10			
17	15	15	14	15	16	17	24	30	19	20	17	29	31	14	16	18	24	22	22	22	17	22	21	20	20	22	17	17	21	23	21	16	13	9	13			
18	20	15	16	15	20	15	20	23	14	15	13	31	31	16	16	19	30	27	26	22	22	23	22	21	22	24	19	20	24	22	22	22	15	11	13			
10	14	16	13	13	13	8	17	11	14	18	17	9	14	16	65	51	20	18	19	17	17	16	14	12	13	8	8	17	17	14	13	11	7	10				
13	17	17	14	14	13	10	15	14	15	17	15	12	16	16	65	53	21	15	19	17	15	16	15	16	13	13	9	9	16	16	14	13	7	12	12			
13	16	17	15	15	13	10	19	19	18	20	19	12	18	19	51	53	22	21	21	18	19	21	18	16	16	11	11	20	20	16	15	13	11	15				
22	21	18	14	15	19	16	20	19	23	16	18	18	24	30	20	21	58	53	31	31	32	32	31	32	25	18	18	27	28	26	24	16	14	16				
19	17	15	11	10	16	14	15	17	20	15	17	18	22	27	18	15	58	62	29	26	29	29	22	24	27	18	18	23	26	22	24	15	14	15				
18	15	14	12	13	16	16	18	20	21	15	19	17	22	26	19	19	53	62	33	31	32	32	26	26	28	20	20	26	27	24	26	14	13	15				
15	15	12	10	9	12	13	17	15	14	13	15	15	17	22	17	17	31	29	50	47	47	47	24	22	24	20	23	22	25	29	26	12	13	16				
19	19	14	13	11	15	13	19	19	19	17	18	18	22	23	16	15	31	26	50	41	41	41	22	24	23	20	22	29	30	23	24	14	12	14				
22	20	20	10	10	16	13	17	19	20	17	19	18	21	22	14	16	32	29	47	41	41	41	20	20	23	23	23	31	30	25	20	12	13	15				
12	15	10	12	12	13	12	16	17	12	15	15	10	20	21	12	13	25	22	24	22	20	20	22	26	27	17	19	21	21	17	18	10	13	15				
12	16	11	12	12	15	11	17	17	16	15	18	11	20	22	13	13	26	24	22	24	20	20	26	24	26	76	67	19	21	25	19	20	11	13				
15	18	13	13	13	15	13	18	16	15	15	18	14	22	24	13	13	27	27	28	24	23	23	23	25	28	59	67	23	25	24	23	13	14	15				
12	15	9	13	13	15	10	17	17	10	13	14	13	17	19	8	9	18	18	20	20	20	23	17	19	23	87	87	22	22	19	19	13	10	14				
13	16	10	12	14	14	10	17	17	11	13	15	13	17	20	8	9	18	18	20	23	22	23	19	21	25	87	87	23	23	21	21	12	10	13				
21	26	17	13	13	16	17	23	24	17	21	19	13	21	20	17	16	27	23	26	22	29	31	21	25	24	22	23	80	80	22	22	15	15	15				
19	24	17	13	13	16	19	25	26	20	23	22	16	23	24	17	16	28	26	27	25	30	30	21	24	23	22	23	80	80	22	23	15	14	14				
14	13	13	9	11	13	12	19	17	14	17	14	13	21	22	14	14	26	22	24	29	23	25	17	19	22	19	21	22	22	40	30	15	17	17				
14	15	10	8	10	16	12	16	17	14	14	15	12	16	22	13	13	26	22	24	29	23	20	18	20	23	19	21	22	23	40	24	12	15	15				
8	8	8	8	8	9	10	14	15	13	12	10	11	13	15	11	13	16	15	14	12	14	12	10	11	13	13	12	15	15	30	24	9	12	12				
9	8	4	5	5	7	4	11	7	8	10	7	6	9	11	7	7	14	14	13	13	12	13	10	11	13	14	10	15	14	15	12	9	46	46				
11	11	8	10	10	11	10	16	13	10	15	11	10	13	13	10	12	15	16	15	16	14	15	15	15	15	15	14	13	15	14	17	15	12	12	12			

Figure 7: Austroasiatic lexicostatistical matrix compiled by Sidwell 1/2/10. Branches are boxed, anomalous high percentages are shaded.

sub-grouping between these geographically peripheral groups.<sup>11</sup> Within the central zone there was ongoing differentiation tempered by contact. The ‘least moves’ needed to account for the distribution of the phylum is a simple radiation, rather than a nested tree of multiple nodes. This would seem to be consistent with the Migration Theory approach.

In the absence of a detailed analysis of AA lexical innovations, this is where the available evidence has taken us. The lexicostatistics provide broad indications of the lexical diversity of the phylum and its branches. At least in respect of the basic vocabulary, we cannot say, for example, that Munda is more diverse than, say, Khmuic or Aslian or Mangic. While some branches, such as Khmer and Monic, are quite small, with the data at hand we cannot say that any of the larger branches looks especially older than the others. Thus we have only a weakly branching tree or rake-like radiation.

### Conclusion:

Three independent lines of inquiry – morphological, phonological, and lexical – have failed to provide decisive indications of nested sub-groupings among AA branches, while the lexical data strongly suggests that there is a contact area centred on Katuic and Bahnaric. Until other indications are forthcoming, the most reasonable hypothesis is a simple radiation out of the Mekong valley.

Harry Shorto, who was inclined to accept a Yangtze origin for AA, struggled to reconcile this with what he knew of Mon and Khmer, and reached a similar conclusion. Writing 30 years ago he remarked:

The Northern Mon-Khmers and Khasis are likely to have followed what became a Chinese trade route to India, as the Mundas may well have done before them. But there seems no over-riding reason to trace routes for the Mons and Khmers, and other groups who occupied the river-plains, down the rivers from the hinterland rather than up them from the coast. (Shorto 1979:278)

### Literature

- ANDERSON, G.D.S. 2006. Advances in Proto-Munda reconstruction. *Mon-Khmer Studies Journal*, 34:159–184.
- BLENCH, Roger. 2008. Roger Blench Stratification in the peopling of China: how far does the linguistic evidence match genetics and archaeology? In: Alicia Sanchez-Mazas, Blench, R.M., Ross, M.D., I. Peiros & Marie Lin eds. *Human migrations in continental East Asia and Taiwan. Matching archaeology, linguistics and genetics*. 105–132. London: Routledge.
- CHAZÉE, Laurent. 1999. *The Peoples of Laos: Rural and Ethnic Diversities*: Bangkok White Lotus.
- DIEBOLD, A.R. 1960. Determining the centers of dispersal of language groups. *International Journal of American Linguistics* 26: 1–10.
- DIFFLOTH, Gérard. 1974. Austro-Asiatic Languages. *Encyclopaedia Britannica*. Chicago/London/Toronto/Geneva, Encyclopaedia Britannica Inc. Pp. 480–484.
- DIFFLOTH, Gérard. 1979. Aslian languages and Southeast Asian prehistory. *Federation Museums Journal*. 24ns:3–16.
- DIFFLOTH, Gérard. 1999. Austroasiatic Classification appearing in Chazée, Laurent (1999) *The Peoples of Laos: Rural and Ethnic Diversities*: Bangkok White Lotus.
- DIFFLOTH, Gérard. 2005. The contribution of linguistic palaeontology to the homeland of Austro-asiatic. In: Sagart, Laurent, Roger Blench and Alicia Sanchez-Mazas (eds.). *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*. Routledge/Curzon. Pp. 79–82.

<sup>11</sup> This may turn out to be too strong a claim. There are weak lexical indications of a connection between Khasi and Palaungic, I expect to explore these in the near future.

- DONEGAN, Patricia and David STAMPE. 1983. Rhythm and the Holistic Organization of Language Structure. In RICHARDSON, MARKS and CHUKERMAN (eds.), *Chicago Linguistic Society Papers from the Parasession on the Interplay of Phonology Morphology and Syntax*. Pp. 335–353.
- DONEGAN, Patricia and David STAMPE. 2004. Rhythm and the Synthetic Drift of Munda, *The Yearbook of South Asian Languages and Linguistics*. Berlin and New York, De Gruyter. Pp. 3–36.
- DONEGAN, Patricia. 1993. “Rhythm and vocalic drift in Munda and Mon-Khmer.” *Linguistics of the Tibeto-Burman Area* 16.1:1–43
- DYEN, Isidore. 1956. Language distribution and migration theory. *Language* 32/4: 611–26. (Reprinted in *Linguistic Subgrouping and Lexicostatistics*. The Hague: Mouton. Pp. 50–74.)
- FERLUS, Michel. 1998. Les systèmes de tons dans les langues viet-muong. *Diachronica* 15:1.1–27.
- GRIERSON, George Abraham. 1904. *Mon-Khmer and Siamese-Chinese Families*. Vol. II, *Linguistic Survey of India*: Delhi.
- HUFFMAN, Franklin E. 1978. *On the centrality of Katic-Bahnaric to Austroasiatic*. Unpublished handout distributed at the 3<sup>rd</sup> ICAAL meeting in Mysore.
- JACOB, Judith. 1989–1990. Some comments on the relationship between Khmer words having identical vowel nuclei and final consonants. *The Mon-Khmer Studies Journal*, 18–19:67–76.
- JENNER, Philip & Saveros POU. 1980–81. A lexicon of Khmer morphology. *Mon-Khmer Studies* 9–10.
- KUIPER, F. B. J. 1955. Rigvedic loanwords, in: *Studia Indologica*, ed. Spies, Bonn
- KUIPER, F. B. J. 1967. The genesis of a linguistic area. *Indo-Iranian Journal* 10: 81–102
- KUIPER, F. B. J. 1991. Aryans in the Rigveda, Rodopi (1991).
- MEI TSU-LIN and NORMAN, Jerry. 1976. The Austroasiatics in Ancient South China: some lexical evidence. *Monumenta Serica* 22:274–301
- PEIROS, I. and V. SHNIRELMAN. 1998. ‘Rice in Southeast Asia: a regional interdisciplinary approach’, in R. BLENCH and M. Spriggs (eds) *Archaeology and Language, archaeological data and Linguistic hypotheses* Vol. II: 379–389, London: Routledge
- PEIROS, Ilija J. 2004. *Genetičeskaja klassifikacija avstroaziatskix jazykov*. Moskva: Rossijskij gosudarstvennyj gumanitarnyj universitet (dissertacija).
- PEIROS, Ilija. 1989. Dopolnenie k gipoteze S. A. Starostina o rodstve nostraticheskix i sinokavkazskix jazykov. *Lingvističeskaja rekonstrukcija i drevnejšaja istorija vostoka: Materialy k diskussijam na mezhdunarodnoj konferencii* (Moskva, 29 Maja – 2 Ijunja 1989g.) Chast’ 1. Moskva, Nauka.
- PEIROS, Ilija. 1998. *Comparative Linguistics in Southeast Asia*. Pacific Linguistics Series C-142. Canberra.
- PINNOW, Heinz-Jürgen. 1960. Über der Ursprung der voneinander abweichenden strukturen der Munda- und Khmer-Nikobar-Sprachen. *Indo-Iranian Journal* 4:81–103.
- PINNOW, Heinz-Jürgen. 1963. The position of the Munda languages within the Austroasiatic language family. In *Linguistic Comparison in Southeast Asia and the Pacific*, edited by H. L. Shorto: London: SOAS.
- PINNOW, Heinz-Jürgen. 1963. The position of the Munda languages within the Austroasiatic language family. In: H.L. Shorto (ed.). *Linguistic Comparison in Southeast Asia and the Pacific*. London: SOAS. 140–52
- REID, Lawrence A. 1994. Morphological evidence for Austric. *Oceanic Linguistics* 33(2):323–344.
- SAPIR, Edward (1916). *Time Perspective in Aboriginal American Culture, A Study in Method*. Ottawa: Government Printing Bureau.
- SCHMIDT, Pater Wilhelm. 1905. Grundzüge einer Lautlehre Khasi-Sprache in ihren Beziehungen zu derjenigen der Mon-Khmer-Sprachen. Mit einem Anhang: die Palaung-Wa-, und Riang-Sprachen des mittleren Salwin. *Abh. Bayerischen Akademie der Wissenschaft* 1 (22.3):677–810.
- SCHMIDT, Pater Wilhelm. 1906. Die Mon-Khmer-Völker, ein Bindeglied zwischen Völkern Zentralasiens und Australasiens. *Archiv für Anthropologie* 5:59–109.
- SCHUESSLER, Axel. 2007. *ABC Etymological Dictionary of Old Chinese*. University of Hawaii Press.
- SHORTO, Harry L. 1979. The linguistic proto-history of mainland South East Asia. in: R. B. Smith & W. Watson (eds.), *Early South East Asia*. New York, Kuala Lumpur, Oxford University Press. Pp. 273–278.
- SHORTO, Harry L. 2006. *A Mon-Khmer Comparative Dictionary*: Canberra Pacific Linguistics 579.
- SIDWELL, Paul & Pascale JACQ. 2003. *A Handbook of Comparative Bahnaric: Volume 1, West Bahnaric*. Canberra, Pacific Linguistics 551.
- SIDWELL, Paul. 2009. *Classifying the Austroasiatic languages: history and state of the art*. Munich, Lincom Europa.
- STAMPE, David. 2004. Was Proto-Austroasiatic like Munda or like Mon-Khmer. Conference handout, 14<sup>th</sup> Southeast Asian Linguistics Society meeting, Bangkok, May 2004.



SVANTESSON, Jan-Olof. 1988. U. *Linguistics of the Tibeto-Burman Area*, 11.1:64–133.

VAN DRIEM, George. 2001. *Languages of the Himalayas Volume One*: Leiden Brill.

WITZEL, Michael (1999), “Substrate Languages in Old Indo-Aryan (Rgvedic, Middle and Late Vedic)”, *Electronic Journal of Vedic Studies* 5 (1), <http://www.ejvs.laurasianacademy.com/ejvs0501/ejvs0501article.pdf>

ZIDE, Norman H. & G. D. S. ANDERSON. 2001. “Recent Advances in the Reconstruction of the Proto-Munda (Austroasiatic) Verb” L. Brinton (ed.). *Historical Linguistics 1999*. Amsterdam: John Benjamins. 13–30.

Статья посвящена вопросам прародины и последующего расселения австроазиатских языков. Традиционный «интуитивистский» подход пока не позволяет убедительно классифицировать более мелкие языковые подгруппы внутри уже установленных ветвей этой лингвистической семьи. При этом лексический анализ словарного состава анализируемых языков указывает на наличие длительных контактов и, возможно, случаев языковой конвергенции на территории континентальной Юго-Восточной Азии. По мнению автора, это указывает если не на исконность, то, по крайней мере, на существенную древность присутствия австроазиатских языков в Индокитае (с возможным центром расселения, расположенным в бассейне р. Меконг). Наиболее географически удаленная ветвь австроазиатской семьи — языки мунда в Индии — демонстрирует исключительно высокое число инноваций. В соответствии с этой теорией в статье реконструируется эволюция корневой структуры языков мунда.

## The Afro-Asiatic background of West Rift \*ḥ, \*ḥ, \*ʕ, \*h, and \*ʔ

The main goal of the paper is to assert that the rich laryngeal inventory, reconstructed for the West Rift subgroup of Southern Cushitic languages, must be seriously archaic in origin; with direct parallels to each of its constituents in Semitic and Egyptian languages, it can be deemed as retaining the original phonology of Afro-Asiatic, thus underscoring the significant importance of West Rift languages in the task of reconstructing this macrofamily. The hypothesis is supported by numerous etymologies.

*Keywords:* Afro-Asiatic languages, Southern Cushitic languages, comparative phonology, historical reconstruction

### Introduction

As I have demonstrated elsewhere (Takács 1999; 2001, 55ff.; 2003, 143ff.), the consonantal system of South Cushitic has apparently retained a number of archaic features that make the importance of this remote subbranch (comprising the West Rift, the East Rift languages, Ma'a, and Dahalo in the border area of Kenya and Tanzania) almost unique (along with the equally extraordinary North Bauchi group of West Chadic) in the whole Afro-Asiatic macrofamily from the standpoint of consonantal reconstruction. This is why the pioneer monograph from 1980 by C. Ehret on the historical reconstruction of South Cushitic phonology and lexicon (abbreviated as Ehret) represents up to now — in spite of all its shortcomings — an equally extraordinary treasury of information. The new comparative lexicon of the West Rift languages (Iraqw, Gorowa, Alagwa, Burunge) by R. Kießling & M. Mous from 2004 (referred to below as KM without year) provides us with a fairly solid inner reconstruction of the best known and most reliable group of South Cushitic.

The publication of this magnificent work allows us to tread on a safer path than any time before towards understanding the Afro-Asiatic background of an exciting segment of South Cushitic consonantism (\*ḥ, \*ḥ, \*ʕ, \*h, and \*ʔ) that has been best preserved just in West Rift and Dahalo. Henceforth, in the light of my permanent research on this noteworthy subbranch, I have re-examined at the Starostin Memorial Conference (Moscow, 25–31 March 2008) how this set of the South Cushitic phoneme inventory relates to the corresponding segments in Semitic and Egyptian, where Afro-Asiatic \*ḥ, \*ḥ, \*ʕ, \*h, and \*ʔ have been best retained. Since there is a regular correspondence between the WRift, ERift, Ma'a, and the Dhl. “laryngeals”, some extra-WRift parallels have also been used and discussed in this paper.

### Methodological problems of South Cushitic etymology

Before going into details about the regular correspondences, I have to briefly address a fundamental weakness of the inner and the (luckily not too frequent) extra-SCu. etymologies in all three recent comparative compendia dealing with West Rift (Ehret 1980, Elderkin and



Maghway 1992, Kießling and Mous 2004). Unfortunately, the authors of these works — albeit in different degrees — have all failed to avoid the very same methodological pitfalls of comparison both in terms of phonology and semantics. It has been especially striking to observe in the latter two works the strong inclination towards far-fetched inner SCu. pseudo-etymologies, which has made Southern Cushitic until now a strangely introverted domain uncomparable with the much higher niveau of comparative Agaw (cf. Appleyard 2006) and East Cushitic (cf. Sasse 1979 and 1982).

The general questions around Ehret's pioneering reconstruction have already been dealt with by reviewers (cf. Hetzron & Tálós 1982; Voigt 1983; Zaborski 1984), although a detailed review of this precious and so far unparalleled accumulation of data would be desirable before one gets to completing a new South Cushitic comparative (and etymological) dictionary (for a number of corrected SCu. lexical entries cf. Takács 2000). The lengthy study by Elderkin and Maghway (referred to below as Eld.-Mgw. 1992) has, unfortunately, on each page of it several false omnicomparativistic West Rift etymologies<sup>1</sup> without any reference to the external parallels. The problems are unfortunately rather similar (albeit luckily in a relatively lower degree) also with the most recent lexical comparisons by Kießling and Mous, which can be illustrated with the following instances:

- **WRift \*ṫānc̣-ō** “(molar) teeth” [KM 57]: semantically, this is, in fact, not an exact correspondence to “*Dullay* (sic, for Gollango) /ad'o ‘cheek’” as KM 36 claim (following a bad tradition of Cushitic linguistics). Cf. rather LECu.: Saho aḏ-a “back teeth” [Lmb. 1987, 533] ||| Om. \*aḥ- “tooth” [Bnd. 1988, 145]. *Dullay* \*ṫaḏ(ḏ)- “Backe” [AMS 1980, 232], in turn, fits WRift \*ṫūnc̣-a “cheek” [KM 2004, 309], which KM falsely regard as a WRift innovation from \*ṫuḥ “to wrap, twist, wring”.

- **WRift \*ṫīs** “to do” [KM 166]: its inner derivation proposed by KM was clearly made due to the ignorance of the external cognates: “*may go back to a Pre-WR form \*/ihis, the status of the final consonant as a lexicalized causative extension is confirmed by the intrusion of the durative suffix in \*/iimis*”. It is a pity that KM failed to consider the widely-known cognates of this root, cf. below in the section for SCu. \*ṫ.

- **WRift \*ṫōc** “to shrivel, shrink, dry up” [KM 222] can hardly be cognate with ECU. ṫaw/yš- “grass” [Sasse 1979, 60] as KM think (aside from the semantical difficulties, WRift \*c ≠

<sup>1</sup> E.g., (1) Irq. & Alg. sāp- “to approach” ≠ Irq. céw “near”, Brg. čia “near” (contra Eld.-Mgw. 1992, 53); (2) Irq. daṫāṫ-m- “to be hot” > dàṫrà “ashes” can have nothing to do with Irq. čarāṫ “dust of ashes”, Grw. čaraṫ “dust” (as suggested in Eld.-Mgw. 1992, 53); (3) Irq. ḥorōḥ- “to snore” has nothing in common with Irq. ḥāḥaṫ-ūs- “to yawn”, Alg. ḥāṫḥāṫ-ūs- “to yawn” (contra Eld.-Mgw. 1992, 55); (4) WRift \*gu?- “to sleep” is unrelated to WRift \*qāt- “to lie down, sleep” (contra Eld.-Mgw. 1992, 55); (5) Irq. qaw “to give, bestow” cannot be related to Irq. qēru “intellect” or WRift \*ḥu?- “to know” (contra Eld.-Mgw. 1992, 56 arguing that “*intelligence is something given*”); (6) WRift \*qaf- “bark, skin” ≠ Brg. kaṫafu “gateway” (contra Eld.-Mgw. 1992, 56); (7) WRift \*sag- “head” is not comparable with Irq. seṫēmi “hair” or Dhl. saḥ- “to plait hair” (contra Eld.-Mgw. 1992, 56); (8) WRift \*ḥirf- “to praise” vs. WRift \*ṫuf- “lip” vs. \*ṫūf- “to swell” are also to be separated (contra Eld.-Mgw. 1992, 57); (9) WRift \*ḥāf- “to scrape” ≠ WRift \*ḥīs- “to rub, scrape” (contra Eld.-Mgw. 1992, 57); (10) WRift \*ḥīmp- “to breathe” has nothing to do with Irq. ḥī(n)ṫ- “to breathe” (as in Eld.-Mgw. 1992, 58); (11) Irq. paḥál “bow-legs” ≠ WRift \*pās- “to deviate” (contra Eld.-Mgw. 1992, 58); (12) Irq. ḥorōrmi “shell” is unrelated to WRift \*ṫsuma “fence” (as suggested in Eld.-Mgw. 1992, 58); (13) WRift \*cōṫ- “to strain” ≠ WRift \*corḥ- “to trickle” (as in Eld.-Mgw. 1992, 58); (14) WRift \*dawr- “sky” ≠ WRift \*rāw- “sky” (contra Eld.-Mgw. 1992, 60); (15) WRift \*du?-ūm- “to suckle” is unrelated to Irq. nūnū?- “to suck breast” (contra Eld.-Mgw. 1992, 60); (16) Irq. gumamā?- “to close” has nothing to do with Alg. ḡibis “to stop up” (as maintained in Eld.-Mgw. 1992, 60); (17) Irq. tumbarara?-at- “to kneel” ≠ Irq. sak<sup>w</sup>nene?-it- “to squat” (as in Eld.-Mgw. 1992, 61) etc. The aforementioned lexemes massively confused in this work by Elderkin and Maghway, in fact, represent distinct SCu. (and, eventually, AA) etyma.

ECu. \*š). Cf. rather Ar. biconsonantal \*ʕs “to get dry” [GT] (details below in the section for SCu. \*ʕ). KM confused again two distinct etyma. By the way, A. B. Dolgopolsky (1983, 138, #10.10) affiliated the ECu. stem (probably incorrectly) with Sem. \*ʕiʕ(V)b- “herb(age), grass” [Dlg.], whereby we may assume in the ECu. cluster \*-wš- a lenition from \*-bš-. For ECu. ʕaw/yš- cp. better WCh.: Nbauchi \*(a)was- < \*\*ʕawas- (?) “grass” [GT]: Warji wasə-na, Mburku wāsə, Kariya wasisi, Miya awasu, Jimbin awaši (NBch.: Skn. 1977, 24).

- **WRift \*ʕūnĉ-a** “cheek” [KM 309] cannot be a WRift innovation from \*ʕuĉ “to wrap, twist, wring” as KM proposed, since it was inherited from the common Cushitic anatomic terminology, cf. its cognates without the epenthetic \*-n-, namely SCu. \*ʕoĉ- “cheekbone” [Ehret 1980, 278] || ECu. \*ʕaĉ- “cheek” [GT]: e.g. Gollango ʕaĉ- “Backe” [AMS 1980, 195] | Konso aĉ-á, Dirayta aĉĉ-á (pl.), Mosiya aĉĉ-etá “cheek” (Konsoid: Lmb. 1987, 533, #8.b) ||| Sem. \*ʕaš(aš)- “bone” [Lsl. 1945, 233]. For Sem.-Cu. see Dlg. 1987, 209, #103; 1989, 100, #43.

- **WRift \*ĉāʕ-a** “rock” [KM 36]: KM’s astonishing equation with Somali ɗaʕ-a “sound of a falling stone” is simply a rude error, since the latter evidently derives from Somali ɗaʕ- “to fall” [Rn. 1902, 124]. Instead cf. rather Saho-Afar ɗāy ~ ɗā “Stein, Fels” [Rn. 1890, 127] = ɗā “pebble” [Sasse], in which H.-J. Sasse (1982, 61) rightly surmises an irreg. reflex of his ECu. \*ɗagh- “stone”. Future research has to clear up how the cluster \*-gh- was resolved in Saho-Afar \*-V̄y- and SCu. \*-V̄ʕ-. Cf. perhaps also NEg. dh.wt (fem.pl.) “Steinblöcke” (NE, Wb V 484, 10) = “Geröllblöcke” (GHWb 986), whose -h- is equally irreg.

- **WRift \*ĉub** “rain” [KM 36] = SCu. \*ĉub- “to rain” [Ehret 217, #24] is probably unrelated to ECu. \*ɗub- “to dip in” [Sasse 1979, 30, 61]. A.B. Dolgopolsky (1987, 208, #101) has, besides, long pointed out the correct cognates of the SCu. root: Sem. \*hḏb (i.e., \*hšb) [prefixed \*h-] “heavy shower” [GT]: OSA (Sab.) hḏb “showers of rain” [Biella Kendrick], Ar. haḏaba, impf. -hḏib- “faire tomber de l’eau, pleuvoir (ciel), mouiller” [BK II 1425, cf. DRS 443] = “to rain continuously” [Dlg.] < AA \*ĉ-b “to rain” [GT].

- **WRift \*gad-a** “forest” [KM 37] ≠ ECu. \*garb- “swamp, thicket” [Sasse 1982, 80]. How could one equate SCu. \*-d- with ECu. \*-rb-? Let alone the semantical difficulties.

- **WRift \*gawa** “peak” [KM 37] = PRift \*gawa “above” [GT]: the same is the case with this stem, erroneously equated by KM with ECu. \*gub- “mountain” [Sasse 1979, 15]. (1) For the former root cf. AA \*g-w (act. \*gaw-?) “hill” [GT] > Geez nagəgaw “hill” [Lsl. 1987, 391] ||| Eg. ḏw [reg. < \*gw] “mountain” (PT, Wb V 541–5) > (S) ἄοοϣ (indicating OEg. \*ḏāw < pre-Eg. \*gaw) ||| SBrb.: ETawllemmet & Ayr a-ggu “être au-dessus de, être sur le dessus de”, ETawllemmet i-ggi & Ayr ə-ggi “lieu élevé, hauteur, élévation” [PAM 2003, 192] ||| Ch. \*g-w-y “mountain” [GT]. (2) For ECu. \*gub-, in turn, cp. AA \*gub- “hill” [GT] discussed in EDE I 400.

- **WRift \*hawʔ-ut** (med.) “to go away” [KM 134] can hardly be cognate with Oromo baʔ “to go out” as KM suggest without any reasoning. Naturally, WRift \*h- (= LECu. \*h-) has nothing to do with LECu. \*b- (= WRift \*b-).

- **WRift \*hāmpu** “wing” [KM 148]: its comparison with ECu. \*hūbn- “muscle (?)” [Sasse 1979, 15, 58] is semantically unconvincing and, in addition, would be also phonologically irregular (WRift \*-p- ≠ ECu. \*-b- burdened by a metathesis).

- **WRift \*hitim** (dur.) “to destroy” [KM 154]: I am afraid that it has nothing in common with Somali ḥad- “to steal” as KM envisage, but should be re-analyzed as a haplology for \*hit(i)m-im, cognate with Eg. ḥtm “to destroy” (discussed below).

- **WRift \*hūĉ** “to plait, braid, weave, knit, twist” [KM 156] cannot be compared to both ECu. \*hā/undūr- “navel” [Sasse 1979, 24, 62; 1982, 91] and Somali ḥid- “to tie” [Abr. 1964, 120] < ECu. \*hīĉ- [Sasse 1979, 62]. Only the latter cognate is correct. The WRift-Somali parallel speaks for AA \*h-ĉ, whereas the Eg. cognate (ḥnt3 < \*hntṛ) to ECu. \*hā/undūr- suggests an underlying AA \*h-n-ṭ-r (with \*-ṭ-!).

• **WRift \*ḥūmp-á** “lungs” [KM 155] cannot be identified with LECu.: Oromo *humb-ī* “trunk of elephant, snout of pig” [Gragg 1982, 217: hence Amh. *kun/mbi* “trunk”], for which cf. rather perhaps PAgaw \**gʷəmb-* “nose” [Apl. 1984, 38] = \*[q/ḥ]<sup>wəmb-</sup> [GT, cf. EDE II 577]. Besides, WRift \**p* ≠ ECu. \**b*.

• **WRift \*ḥūnc** “to wash (clothes)” [KM 155] for semantic reasons can hardly have anything to do with LECu.: Somali *ḥundúr* “dysentery” [Abr. 1964, 124].

• **WRift \*ḥar(V)m-** “horn” [GT] = \**ḥadāḥ<sup>w</sup>* ~ \**ḥadamu* “1. horn, tusk, 2. sg. protruding” [KM 318–9] it has, of course, no connection with ECu. \**gās-* “horn” [Ss.] as KM maintain. Cf. SCu. \**ḥarim-* “horn” [GT] below.

• **WRift \*ḥās-īt** (med.) “to be silent” [KM 320] can hardly be regarded as an inner development from WRift \**ḥās-* ~ \**ḥīš-* “to smoothen a surface” (reflected by Irq. *ḥīš-* “to scrape” and Brg. *ḥīla-ḥās-* “to clean a field, rake”!) as KM suggest. Instead, here too, we may safely assume a distinct AA root, cf. Sem. \**ḥšf* “to be calm, still” [GT] (discussed below).

• **WRift \*rawa** “sky” [KM 37] ≠ ECu. \**rōb-* “rain” [Sasse 1979, 22] as KM suggest. For the former cf. LEg. *rj.t* “Himmel” (GR, Wb II 400, 3) ||| Ch. \**r-w* “sky” [JS 1981, 234E] (for Eg.-WCh.-Brg.: HSED #2101). The latter root is (as demonstrated long ago by W. W. Müller 1975, 67, #39) cognate with Sem. \**rbb* “to rain”. KM confused two different AA etyma.

### The Afro-Asiatic correspondences to West Rift and South Cushitic \*ḥ, \*ḥ, \*ʕ, \*h, and \*ʔ

Below those West Rift (and sometimes also some other South Cushitic) lexical items will be examined — *ohne Anspruch auf Vollständigkeit* — that contain these phonemes, primarily in the Anlaut, and whose Semitic and/or Egyptian parallels have been identified (mostly by myself in the course of my current research on South Cushitic historical phonology, ongoing since spring 1998). In rare cases, lone data from some East Cushitic daughter languages (usually equally conservative about \*ḥ, \*ʕ, \*h, and \*ʔ, but not \*ḥ) have also been considered. Since Berber, Omotic, and Chadic languages are very innovative about these old phonemes (with erosions and mergers), reflexes from these Afro-Asiatic branches will only be quoted briefly, provided the Semitic and/or Egyptian cognates are available.

#### SCu. \*ḥ

SCu.	WRift	Qwadza	Asa	Ma'a	Dahalo	Sem.	Eg.	AA
*ḥ	*ḥ-, *-ḥ-	ḥ-, -ḥ-	?	ḥ ~ k	k, h- (?)	*ḥ	ḥ	*ḥ

• **SCu. \*ʔaḥ<sup>w</sup>-** “to listen” [Ehr.] < an incompatible \*\**ʔaḥ<sup>w</sup>-* (note: SCu. \*ʕ-ḥ is unattested!) “to hear” [GT]: WRift \**ʔaḥ-as* and \**ʔaḥ-am-is* “to hear”, \**ʔaḥw-es* (caus.) “to talk” (GT: lit. \*ʔ “to make s’one listen”) [KM 64] | Asa *h-as-* “to hear” [Ehr.] | Dhl. *ʔágaʒʒo* [affix -ʒʒo] “ear” [Tosco 1991, 127] (SCu.: Ehret 1980, 288, #47) || Bed. *ʔángwīl* [affix -īl] “Ohr” [Rn.], Bisharin *ʔanḵwīl* “ear” [Almkvist] || Agaw \**ʔəḥq<sup>w</sup>-* “ear” [Apl. 2006, 59]<sup>2</sup> || LECu.: Saho *oḵká*, pl. *ókā* “Ohr” [Rn. 1890, 23] = *ʔokka* [Bnd.] = *ʔokka* “ear” [Vergari 2003, 66], Saho-Assaorta *ʔoqqá*, pl. *ʔoqqáq* “orecchio” [CR] = *ʔoḵḵwā* [Dlg.: ʕ- “записано ошибочно или вторично”], cf. Afar *ʔokka* “ear-wax (cerumen)” [PH 1985, 61] (Cu.: Crl. 1938 II, 213; Dlg. 1973, 183 with false re-

<sup>2</sup> D. Appleyard’s (l.c.) reluctance to accept its old equation *wih* — as “*probably not related*” — is baseless.

construction; Zbr. 1989, 580, #21; Ehr. 1995, 521, #728) ||| Eg. ṣnh.wj “die zwei Ohren (als Körperteil des Menschen)” (MK, Wb I 204–5)<sup>3</sup> ||| CCh.: Bura ngga “hören”, ngga-ta “hören, fühlen, empfinden” [Hfm. in RK 1973, 93], Chibak ngá-tì “hören” [Hfm. 1955, 135], Margi-Wamdiu nga-ri “to hear” [Krf.], WMargi ngə-dì “to hear”, ngà-dì “1. to hear, 2. feel” [Krf.] | Higi nga-rdì “entendre” [Krf.] (CCh.: Krf. quoted by Jng.-Br. 1990, 77) || ECh.: Mokilko ʔànnìgá “(se) taire” [Jng. 1990, 58] < AA \*ʕaQ<sup>w</sup> ~ \*ʕanQ<sup>w</sup> “to listen, hear” [GT]. For Eg.-Cu. (*sine* SCu.) see Zhl. 1932–3, 166.

• **WRift \*ʔānh** “to be satiated” [KM 58] ||| Ar. nḥw: naḥā I “1. se donner des airs, être fier, bouffi d’orgueil”, X “enfler, gonfler” [BK II 1223] || MSA \*nḥw > Hrs. nḥō “to be glad”, Mhr. hənḥū “to be happy” (MSA: Jns. 1977, 99; 1987, 309) ||| Eg. nḥnh “sich freuen (vom Herzen)” (PT, Wb II 312, 11).

• **WRift \*ʔānh-ari** (affix \*-ar-) “phlegm” [KM 58] ||| Sem. \*naḥ- “mucus, phlegm” [SED I, no. 197: no AA cognates] ||| Eg. nḥ “Speichel” (PT, Wb II 318, 14–15) ||| CCh.: Tera naḥa “saliva” [Nwm.] | Mandara neḥe [Mch.] etc. (Ch.: JI 1994 II 279) < AA \*n-ḥ “mucus” [GT]. For Jbl. nḥḥ “phlegm” [Jns. 1981, 199] note also Mhr. nəḥyōk “phlegm” [Jns. 1987, 308], whose \*-k is not yet explained. Note that a var. root (with an original \*-k) has apparently been retained in WCh.: Nbauchi \*nəḥ- “saliva” [GT]: Warji nənīki-na ~ nənāki, Kariya nənəkí, Miya nínəkí, Siri nīkí, Mburku nənəkə etc. (NBCh.: Skn. in JI 1994 II 278) < AA \*n-k “saliva” [GT].

• **PRift \*caḥ<sup>w</sup>-** “to trap” [GT]: PIRQ. \*cāḥwēli “caltrop, rope trap” [KM] > Irq. caḥweli “trap” | Qwd. caʔ-uko [-ʔ- < \*-ḥ-] “bird trap” (SCu.: Ehret 1980, 355) ||| Eg. šḥt “(Vögel) mit dem Netz fangen” (OK, Wb IV 262–263) = “to trap, snare” (FD 243). The third Eg. root consonant -t is still unexplained. Interestingly, the same problem appears in Eg. šḥt “flechten, weben” (PT, Wb IV 263) ||| ECu. \*sōh- “flechten” [Sasse 1976, 126; 1979, 41].<sup>4</sup>

• **SCu. \*dōḥ-** “to take away” [Ehret]: WRift \*dūḥ “1. to take out, take away, 2. marry” > \*dūḥō “marriage” [KM 100] | Asa duḥ-um- “to marry” [Ehret] | Dhl. dōk- [k reg. < \*ḥ] “to take away” [Ehret] = dōk- “to seize, take away from” [EEN 1989, 28] = dōk- “to take away, kidnap (woman)” [Tosco 1991, 132] (SCu.: Ehret 1980, 166, #34): although SCu. \*d vs. Sem. \*ḍ do not regularly correspond<sup>5</sup>, this root may be an irregular metathetic match to Sem. \*ʔḥḍ “prendre, saisir” [DRS 15] > esp. Akk. aḥāzu “1. nehmen, 2. heiraten” [AHW 18].

• **WRift \*dūḥa**, pl. \*dūḥahē “ditch” [KM 100] ||| Sem.: Geez daḥəḥa “to excavate, hollow out”, dəḥuḥ “hollow, excavated, concave, chasm, cave”, daḥaya “to excavate, dig up, hollow out, dig deep, penetrate” [Lsl. 1987, 128, 130] (listed in DRS 245 with -ḥ-!). Leslau’s (l.c.) proposal on the relationship of this Geez root to Geez daḥaya “grind grain” > madḥe “upper millstone” (listed in DRS 245 with -ḥ-!) as well as to Sem. \*dḥy “to push” has to be declined.

• **SCu. \*ḥaʔ-** “to burn” [GT]: Qwd. ḥaʔo “fire” | Ma’a -ká ~ -ḥa “to be lit, burning”, -kakó ~ -ḥaḥó “heat” | Dhl. kok- (partial redupl.) “to be lit, be burning” (Ehret 1980, 259, #29–#30) || ECu.: (?) Dullay \*ḥa(t)te [suffix -tte?] “fire” [GT after AMS 1980, 241] (SCu.-Dullay: Flm. 1969, 22; Dlg. 1973, 263) ||| Eg. ḥ.t “fire” (PT, Wb III 217–218).

• **SCu. \*ḥab-i** “firewood” [Ehr.] > WRift \*ḥābiya, pl. \*ḥāb-ō “charcoal” [KM 318] | Ma’a lu-hābi “firewood” [Ehr.] (SCu.: Ehr. 1980, 256, #4) ||| Ar. ḥbw > ḥabā I “2. brûler (se dit du feu)” [BK I 537] ||| (???) Eg. ḥb.t “Feuer” (NK, Wb III 252, 16)<sup>6</sup>.

<sup>3</sup> For this Agaw-Eg. etymology see Zhl. 1932–33, 166.

<sup>4</sup> Albeit for other reasons, but Kießling and Mous (2004, 294) have also suggested a biconsonantal origin for the WRift term, namely a semantically far-fetched derivation from a hypothetical WRift \*caḥ<sup>w</sup>- (ideophon.) “sudden hopping or snapping movement” based on the comparison of such unrelated terms as PIRQ. \*caḥ-ūt “to hop”, \*caḥway “grasshopper sp.”, and even \*caḥ-ad “to aim at, shoot and hit”.

<sup>5</sup> Whereas normally SCu. \*ʕ<sub>2</sub> > WRift \*d, Qwd. d, Asa ʕ, Ma’a z, Dhl. ḍ ~ Sem. \*ḍ.

<sup>6</sup> Unless < ḥb.t “Richtstätte (an der die Bösen verbrannt werden)” (since MK, Wb III 252).



• **SCu.** \***hāl-** “to be dry” vs. \***hāla?** “dry” [Ehr.] > Alg. *hala?* “dry”, *hala?ame* “barren (of land)” | Ma’a -*hála* ~ -*hála* “to be dry”, -*halá-ti* “to dry (sg.)”, -*halá?e* “dry” (SCu.: Ehret 1980, 257, #9-#10) ||| Sem. \**hlw* > Ar. *ḥalā* “être vide, vacant” [BK I 626]. For the semantic shift cf. Eg. *šw* “dry” ~ *šw* “empty”.

• **SCu.** \***ḥarim-** “horn” [GT] > WRift \***ḥar(V)m-** “horn” [GT] = \**ḥadān<sup>w</sup>* ~ \**ḥadamu* “1. horn, tusk, 2. sg. protruding” [KM]: Irq. *ḥarmo*, Grw. *ḥarmó*, Alg. *ḥarimo*, Brg. *ḥadaṇ* | Qwd. *ḥalin-ko* [GT: < \**ḥarim-*], Asa *hadon-k* | Ma’a *lu-ḥaremú* ~ *lu-haremú* (SCu.: Ehret 1980, 256, #5; WRift: also KM 318–9) ||| Sem.: Ar. *ḥarm-* ~ *ḥurm-* “pic d’une montagne”, ma-*ḥrim-* “1. sommet saillant d’une colline, 2. pointe d’une arme” [BK I 565]. For the SCu. vs. Ar. meaning cf., e.g., Armenian *sar* “Höhe, Gipfel, Abhang”, Homeric Greek *κάρηνα* (pl.nom.) “Köpfe, Bergesgipfel” < IE \**ker-* “das Oberste am Körper: Kopf, Horn, Gipfel” (IEW 574ff.).

• **SCu.** \***ḥas-** “gully, streambed, dry watercourse” [Ehr.] > Irq. *ḥasa* “valley or ravine without running water” | Dhl. *kaso* “furrow (for irrigation or drainage)” (SCu.: Ehr. 1980, 256, #6) ||| Eg. *ḥz* (old \**-s/\*-z?*) “Kanal, Brunnen” (XVIII., Wb III 332, 4) = “Sumpfloch, Rinnsal, Bach”, cf. *ḥz n mw* “Brunnen (am Grab)” (GHWb 619) < AA \**ḥ-ṣ* “to bore a hole, dig out” [GT]<sup>7</sup> with irreg. SCu. \**-s-* < AA \**-ṣ-*.

• **WRift** \***ḥaš-īṭ** (med.) “to be silent” [KM 320] can hardly be regarded as an inner development from WRift \**ḥāš-* ~ \**ḥīš-* “to smoothen a surface” (reflected by Irq. *ḥīš-* “to scrape” and Brg. *ḥila-ḥāš-* “to clean a field, rake”!) as KM suggest, cf. Sem. \**ḥšṯ* “calm” [GT]: Ar. *ḥašafa* “to be humble, submissive, faint (voice)” || Geez *ḥašʾa* ~ *ḥašʾa* “to be calm, still etc.” (Sem.: Lsl. 1987, 266) ||| CCh.: Glavda *ghačəga-ghačəga* [gh- < \*Q- reg.] “still, calm” [RB 1968, 42] < AA \**Q-š-ṯ* “calm” [GT].

• **WRift** \***ḥāš-** ~ \***ḥīš-** “to smoothen a surface” [KM 320] = WRift \**ḥīš-* “to scrape” [Eld.-Mgw. 1992, 57] = SCu. \**ḥīš-* “to scrape” [GT]: cognate with Bed. *se-has* (caus.) “(ab)reiben, abputzen” [Rn. 1895, 127] ||| Geez *ḥ<sup>w</sup>ašaša* ~ *ḥošaša* “to sweep” and ES \**ḥss* “to sweep” [Lsl. 1987, 266]. It is dubious whether Saho *hašay* “reiben, frottieren”, *hešiy* “(aus)reiben (z.B. Getreidekörner aus den Ähren, zwei Hölzer um so Feuer machen zu können)” [Rn. 1890, 196] | Dobase *ḥaš-* “schaben, Leder bearbeiten” [AMS 1980, 191] | Burji *hāsiy-* “to rub, stroke, massage” [Sasse 1982, 88].

• **WRift** \***ḥaw** “to come” [KM 320] < (?) SCu. \**ḥā-* “to come to” [Ehr. 1980, 256, #2] = \**ḥaw-* / \**ḥ<sup>w</sup>a?* [GT]: cp. perhaps Ar. *ḥwy:* *ḥawā* I “6. emporter, enlever”, II “tirer vers le coucher” [BK I 651].

• **SCu.** \***ḥo?** [\**x-*] “to bend”: Irq. *ḥu?-us-* “to turn one’s back on” | (?) Qwd. *ḥuʾumbayo* “leather wrapping securing bowstring to bow” | Dhl. *koʾʾ-ēd-* “to fold” (SCu.: Ehret 1980, 259) ||| Eg. *ḥ3* [if < \**ḥ?*]<sup>8</sup> “sich beugen” (GR, Wb III 223, 1).

• **WRift** \***ḥoh-** > sg. \***ḥohōḥ<sup>w</sup>** “hollow form”, pl. \***ḥohēri** “hollow forms” [KM 323] || ECu.: perhaps Dullay \**ḥoh-* “hole” [GT] > Dobase *ḥohr-o* & Gollango *hóhn-o* “Loch” (Dullay: AMS 1980, 192) ||| Sem. \**ḥawḥ-* “hollow” [TG]: Hbr. *ḥāwāḥīm* ~ *ḥohīm* (pl.) [< \**ḥawḥ-*] “hole, crevice” [KB 296] = “das Felsspalten” [GB 217] = “hollows, recesses” [Lsl.] || Ar. *ḥawḥ-at-* “aperture in a wall, small doorway between houses” [Lsl.] || Geez *ḥohət* “door(way)” [Lsl. 1987, 260] = *hōht* “hole in the wall” [KB] ||| WCh.: Angas-Sura \**kuk* ~ \**kok* “(i.a.) be empty” [GT 2004, 207]: cf. esp. Angas *kok* “empty” [Ormsby 1914, 209], Mnt. *ku* (so, without -k) “empty” [Ftp. 1911, 216]. The non-reduplicated root appears in Sem. \**ḥwy* “empty” [GT].

<sup>7</sup> Cf. Sem.: Ar. \**ḥz* “to sting, pierce” [Zbr.] = “durchbohren” [Eilers]: Ar. *ḥazza* ~ *naḥaza* ~ *waḥaza* (for the Ar. biconsonantal root see Zbr. 1971, 71, #113; Eilers 1978, 128; Blv. 1993, 34, #22) ||| Brb. \**γ-z* “creuser” [GT after Prs. 1969, 84, #565].

<sup>8</sup> Is GR *ḥ3* to be separated from older Eg. *ḥ3b?*

• **SCu. \*ħor-** “to be in trouble” [Ehr.] > Irq. ħore “trouble, calamity” | Dhl. kor- “to be in trouble” (SCu.: Ehr. 1980, 259, #26) || ECu. \*har- “to become tired or weak” [Sasse 1979, 40] ||| Sem. \*ħwr: Ar. ħwr: ħāra “être faible, débile (homme), faiblir, perdre de son intensité (la chaleur)” [BK I 645] || (?) Mehri ħaur “peu” [Jahn] (Sem.: Lsl. 1938, 192) ||| Eg. ħ3j.t [< \*ħry-t] “Krankheit, Leiden” (MK, Wb III 224), cf. ħ3w.tj “Personifikation des Todes (?)” (PT, Wb III 227, 2). For ECu.-Ar. cf. Dlg. 1987 MS, #44.

• **PIrq. \*ħorōramay** “snail shells” [KM 322] = PIrq. \*ħorōr- [GT]: Irq. ħororami “cowry shell” [Ehret 1980, 259] = ħorōrmi “shell” [Mgw. 1989, 102]:<sup>9</sup> < SCu. \*ħora- “large hard shell of sg.” [Ehret 259, #27] ||| Eg. \*ħ3 [< \*ħr] “Muschel (nur als altes Schriftzeichen für ħ3 belegt)” (Wb III 218, 17) = “bivalve shell” (a lost word, preserved only by the hrgl. representing a bivalve shell, EG 1927, 468, L6) ||| WCh.: Hausa k̄waryár “shell (of tortoise)” [Abr. 1962, 594] | (?) Pero kókólò [irreg. -l- < \*-r-?] “shell” [Frj. 1985, 36]. For Irq.-Hausa see HSED #2071.

• **WRift \*ħoû-** [Ehr.] = \*ħōû- “to grind” [Eld.-Mgw. 1992, 57; KM 322]: Irq. & Grw. ħoû-im- “to grind” [Ehret 1980, 259] = Irq. ħós- “to grind” [Mgw. 1989, 102] = -ħōû-îḡw-, a ħōû-îm “to pound” [Wtl. 1953] = \*ħōû- “to grind”: ħōû-îm- “to be grinding”, ħōû-îḡw- “grinding” [Eld.-Mgw. 1992, 57]<sup>10</sup> ||| Sem.: Akk. ħašû “zerkleinern, verstümmeln” [AHW 335] ||| Eg. ħšħš “Schutt” (NE, Wb III 339, 7) = “rubble” (DLE II 195) ||| WCh.: Angas-Sura \*ğy<sub>1</sub>ēs “to grind” [GT 2004, 147]<sup>11</sup> | Daffo-Butura ġiğyàš “zerbrechen” [Jng. 1970, 86] | Ngizim kàšû “to break into pieces, shatter” [Schuh 1981, 89] || CCh.: Daba ħiz “to grind” [LG], Kola -hèz- [-z- < \*-ž- < \*-š-?] “to grind” [Schubert] (CCh.: JI 1994 II, 171) < AA \*ħ-š “to grind into fine particles” [GT]. For Akk.-WRift cf. Mlt.-Stl. 1990, 62.

• **WRift \*ħu?** “to know”, **\*ħu?-um** (dur.) “to watch, observe” [KM 323–5]: cognate with Dullay \*ħi?- “sehen” [GT]: Harso, Dobase, Gollango ħi?-, Gawwada ħi?- (Dullay: AMS 1980, 168).

• **WRift \*ħuf** “to drink alcohol” [KM 323]: cp. Eg. ħfħf.t “Ausströmen, Erguß” (PT, Wb III 273, 15). For the semantic change cf. Sem. \*šty “to drink” [GT after Rabin 1975, 87, #19] ||| Eg. stj “(eine Flüssigkeit) ausgießen” (PT, Wb IV 328–9).

• **SCu. \*ħul-** “to be sated, full” [GT]: Ma’a -ħú [regular < \*-ħul] “to be sated (of food), full” [Ehret] | Dhl. kul- “to gorge oneself” [Ehret & EEN 1989, 12] (SCu.: Ehr. 1980, 260) ||| NOM.: Gamu kall- “to be satiated” [Sottile 1999, 440] ||| Eg. nħ3ħ3 [\*n-ħl-ħl] “strotzen (von den Brüsten säugender Frauen)” (PT, Wb II 306, 10).

• **WRift \*ħūmpa** “fever illness” [KM 323] ||| Eg. ħnfj (fem. < \*ħnf.t) “1. Feuer, 2. (den Bösen) verbrennen” (LP, Wb III 291, 15–16). Irregular WRift \*-p- vs. Eg. -f-.

• **WRift \*ħup** “to turn oneself face down” [KM 324]: cp. perhaps (with nasal root extension) Ar. ħanafa I “1. (dé)tourner, 2. détourner le museau de la muselière (chameau), 3. tourner la tête et le cou du côté du cavalier” [BK I 641].

<sup>9</sup> The Irq. word has nothing to do either with SCu. \*ħora/\*ħoda “large” (as supposed by Ehret 1980) or with WRift \*šum- “protective enclosure” (as in Eld.-Mgw. 1992, 58).

<sup>10</sup> It seems to me more than dubious whether there was any connection between WRift \*ħoû- “to grind” and SCu. \*ħiš- “to scrape” [GT] as supposed by Elderkin & Maghway (1992, 57), cf. WRift \*ħiš- “to scrape” [Eld.-Mgw.]: Irq. ħiš-it- [x-] “to peel” [Ehret] = ħiš- “to scrape” [Eld.-Mgw.], Alg. ħiš- “to rub” [Ehret] = ħišà “to rub” [Eld.-Mgw.] (WRift: Eld.-Mgw. 1992, 57) | Ma’a -ħi [regular loss of \*-š-] “to scrape off” [Ehret] | Dhl. ħiš- [unexplained h-] “to scrape scales off fish” [Ehret] (SCu.: Ehr. 1980, 258).

<sup>11</sup> Attested in Sura ðiyees “(Korn) mahlen” [Jng. 1963, 64], Mpn. ðiés “to grind” [Frj. 1991, 15], Kofyar eés “to grind” [Ntg. 1967, 12], Chip ?ées “mahlen” [Jng. 1965, 166] = ?es (so!) [Stl. 1972 < ?], Mushere ess (so, -ss) “to grind” [Dkl. 1997 MS], Montol ees “to grind” [Ftp. 1911, 217], Goemay ees “to grind” [Ftp. 1911, 217] = ès, pl. ûas “to grind” [Srl. 1937, 47] = ?ees “mahlen” [Jng. 1962 MS, 7] = ees “to grind” [Hlw. 2000 MS, 8] (AS: Stl. 1972, 181; 1977, 154, #54; 1987, 230, #785).

• **WRift \*hūru?** “to roar, blow, rumble” [KM 324] ||| Sem.: Ar. ḥrr: ḥarra “produire un bruit avec ses ailes en volant, 2. murmurer”, ḥrḥr: ḥarḥara “ronfler” [BK I 551, 556] ||| Eg. ḥrw “Stimme, Geräusch” (PT, Wb III 324).

• **WRift \*hūs** “scrape, gnaw” [KM 324]: akin to HECu.: Gedeo kušš-, Kambata kuss- “to rub (off)” (HECu.: Hds. 1989, 126). Remotely (as a reflex of a PAA var. root?) related to WRift \*hā/iš- “to scrape” [KM, Eld.-Mgw.] (discussed above)?

• **WRift \*hwanšāʔó** “garbage, dirt, refuse” [KM 325] ||| Eg. ḥnš stinken” (Med., Wb III 301).

• **SCu. \*hwar-** “to stir, mix” [GT] = \*hwar- “to stir” [Ehret 1980, 270, #5]: WRift \*hwarānhwar “to stir, jumble, mould” [KM 326]: Irq. ḥwaranḥwar- “to stir” [Ehret] = ḥwara:ḥwār- “to jumble” [Mgw. 1989, 102], Brg. ḥwaraḥwar-es- “to stir” [Ehret] | Ma’a -kwári “to mix” [Ehret] (Ehret 1980, 270) ||| Eg. ḥ3w ~ ḥ3w [< \*ḥrw] “zerstoßen und in eine Flüssigkeit mischen” (Med., Wb III 361, 8) ||| WCh.: Pero gbúru [\*gʷuru] “to mix (with oil or water)” [Frj. 1985, 31] || CCh.: Bura gwar “mischen”, gwar-ta “vermischen” [Hfm. in RK 1973, 88] < AA \*hʷ-r “to mix” [GT].<sup>12</sup>

• **WRift \*hwayaʔi** “evening” [KM 326]<sup>13</sup> ||| WCh.: Angas-Sura \*kʷō > \*kʷo “darkness” [GT 2004, 192] < AA \*h-w, whose metathetic var. (AA \*w-h) is attested in Eg. wḥ “dunkel sein” (Med., Wb I 352, 3–4), wḥ “das Dunkel” (MK, Wb I 352, 6–10), wḥ.t “Dunkelheit” (MK, Wb I 352, 11) ||| CCh.: Bura-Margi \*-wagu “evening” [GT] > Margi wagù, Wamdiu uwàgu, Hildi ʔuwāgu, Kilba ákù, Gava ḥawugà (CCh.: Krf. 1981, #122).

• **WRift \*hwaycā**, pl. \*hwaycō “switch, flexible stick” > \*hwayc-is “whip with a flexible stick” [KM 326–7] ||| Ar. ḥwṭ > ḥūṭ- “1. rameau, branche jeune et tendre, 2. cep de vigne”, hence ḥūṭ-ān-at- “grand, d’une belle taille, et souple comme un rameau (homme, femme)”, mu-ḥāṭ-at- “racines de souchet” [BK I 647].

• **WRift \*hwayla** “child, giving birth” [KM 326] ||| Eg. ḥnw [reg. < \*ḥlw] “Kind” (MK, GR, Wb III 286, 28).

• **WRift \*hwereʔ-es** (caus.) “to fry” [KM 327] ||| Eg. ḥr.t “Flamme” (NK, Wb III 323, 20; GHwB 613) ||| WCh.: Sha hwoḥ [reg. < \*hwor] “brennen (intr.)” [Jng. 1970, 285] < AA \*hʷ-r “to burn” [GT].

• **WRift \*sōḥ** “to urinate” [KM 252] ||| Sem. \*šḥḥ: Ar. šaḥḥa “1. uriner, 2. produire un bruit en coulant, 3. lancer l’urine en long filet”, cf. šaḥšaḥa “lancer l’urine, etc., en filet” [BK I 1200, 1202]. Irregular WRift \*s- vs. Sem. \*š- > Ar. š-.

• **PIrq. \*sāḥi** “gall, bile” [KM]: Irq. saḥi “gall-bladder, gall, bile” [Ehret 1980, 179; 1974, 82] = sāḥi “bile” [Mgw. 1989, 100] = sāḥi [KM], Grw. sāḥi “gall, bile” [KM] (WRift: KM 247) < SCu. \*saḥ- [Ehret] ||| N.Eg. šḥ “gall” (XIX. Med., Wb IV 228, 9–11) = šḥ.w “Galle” (Deines-Grapow 1959, 460–461) = šḥwj “bitter gall” (DLE III 84) ||| CCh.: Logone cekī [c-/ts- < \*s-?] “Galle” [Lks. 1936, 123] < AA \*s-h “gall” [GT]. The Eg.-Irr.-Logone etymology was first suggested by Militarev (1987 MS, #117) and was first published in Orel 1995, 125; HSED #2171. But it has been observed independently also by G. Takács (1996, 145, #49). Logone cekī comes by partial assim. from \*šḥV, being etymologically connected to PKotoko \*šḥV “bitter”: Kotoko šḥáne, Logone šḥa [Lks. 1936, 119], Buduma čay, čei [Lks. 1939, 93], Afade szâsza [sasa] (CCh.: Prh.

<sup>12</sup> To be distinguished from Dhl. ʔōr- “to stir” [EEN 1989, 14], which is perhaps akin to WCh.: Sha yèr “mischen” [Jng. 1970, 289].

<sup>13</sup> Its extended version is WRift \*hwayrā “evening” [KM 326] > i.a. Irq. ḥwera “night, especially the earlier part of it” [Ehr. 1980, 270, #7 with false comparanda], which I was earlier disposed to affiliate with Eg. ḥ3wj [-3- < \*r-] “Abend” (PT, Wb III 225, 17). Or should we assume the same affix \*-r- in Eg. (NB: the Wb III 225 has a hint on the surmised connection of wḥ vs. ḥ3wj)?



1972, 62, #35.2; JI 1994 II, 27). SCu. \*s-, the alternation of s- ~ č- in the CCh. reflexes, and esp. Buduma č-, all point to AA \*s-.

• **WRift \*wah-ā** “thirst” vs. **\*wah-īt** “to be thirsty” [KM 317]: presumably akin to Sem.: Ar. wḥy: wahā “tendre vers, se proposer de, (re)chercher”, cf. wahḥ- “1. peine, fatigue, 2. intention” [DRS 524–5; Dozy II 790; BK II 1503] with a semantic shift “to search” ~ “to desire” ~ “to thirst”<sup>14</sup>.

### SCu. \*ʕ

SCu.	WRift	Qwadza	Asa	Ma'a	Dahalo	Sem.	Eg.	AA
*ʕ	*ʕ	ʕ-, -ʕ-	ʔ-, -ʔ-	ʕ-, -ʔ-, -ʕ	ʕ-, -ʕ-	*ʕ, *ʕ	ʕ	*ʕ, *ʕ

• **WRift \*ʕab** “new” [KM 46, so also Ehret 376] and **\*ʕabāk<sup>w</sup>** “white” [KM 47: fossilized adj. pl. suffix \*-k<sup>w</sup>] = \*ʕabak<sup>w</sup>- “white” [Ehret 376: rare adj. derivative suffix \*-k<sup>w</sup>] — as rightly assumed both by Kießling & Mous (l.c.) and Ehret (l.c.) — may indeed derive from a common etymon, whose basic meaning may be close to “white” and which is preserved in Ar. ʕbw: ʕabā I “1. brillant, avoir de l'éclat (se dit du teint, du visage)” [BK II 160] ||| Eg. ʕbʕb “(von den Strahlen der Sonne, die auf das Gesicht scheinen)” (XVIII., Wb I 178, 4) = “to appear, shine” (FD 41) ||| Om. \*ab “sun” [Bnd. 1994, 1157, #80]. For the shift of meaning Ehret (l.c.) quotes PKalenjin \*lēl “1. white, 2. new” in EAfrica.

• **SCu. \*ʕač-** [\*-tl-] “fat (adj.)” [Ehret 1980, 275] > **WRift \*ʕač-ar** (affix \*-ar) “fat, thick-set, fully-grown, chubby” [KM 62]: Irq. ʕačar “fett”, Brg. ʕačari “fett” | Ma'a eš-amúye [-hl-] “fett” (SCu.: Ehret 1980, 275)<sup>15</sup> || **LECu.:** Afar ʕēd-o “suet, animal fat rendered down and used for water-proofing (graisse de rognon, graisse de boeuf, graisse animale servante à imperméabiliser)” [PH 1985, 59] ||| **Sem.** \*ʕšš > Ar. ʕiḏḏ- “4. ardent et vigoureux à faire qqch., 5. fort, robuste” [BK II 276] | Jbl. ʕəžž “to be fat” [Jns. 1981, 21] ||| **Eg.** ʕḏ [reg. < AA \*ʕč] “Fett” (OK, Wb I 239)<sup>16</sup> ||| perhaps **WCh.** \*[ʕ]aḏa(-r) “Fett” [Stl. 1987, 222; 1986, 94] < AA \*ʕ-č “to be fat” [GT].

• **WRift \*ʕadaʔ** “to be grey” [KM 47] ||| **LECu.** \*ʕad- “white” [Black 1974, 203]: Saho ʕad-o “white” [Welmers 1952, 251] | **PSam** \*ʕad “white” [Heine 1978, 99].

• **WRift \*ʕag** “to eat” [KM 50]: cf. **Qwd.** ag-, **Asa** ʕag-im- “eat” | **Ma'a** -ʕa [loss of \*-C# reg.] “to eat” | **Dhl.** ʕag- “to eat” (SCu.: Ehret 275) ||| **ECu.:** Yaaku -ek- [k < \*g reg.] “to eat” [Heine 1975, 124; Tosco 1989 MS, 10: Yaaku-Dhl.] ||| **Ar.** ʕǧw: ʕaǧā I “1. allaiter (son enfant), av. acc.”, III “2. nourrir l'enfant d'un autre lait que de celui de sa mère, mettre un orphelin en nourrice” [BK II 186]. Perhaps also **ECu.** \*ʕVg- “to drink” [GT]: **HECu.** \*ag- “to drink” [Hds. 1989, 404] | **HECu.** \*ʕug- “to drink” [Black 1974, 108] | **Dullay** \*ʕuk- “trinken” [AMS 1980, 274].

<sup>14</sup> Cf. e.g. (1) IE \*ais- “begehren, wünschen, (auf)suchen” > i.a. OIndic eṣ- “wünschen erstreben, suchen” ~ OSlavonic iskati “suchen”, iska “Wunsch” (IEW 16; EWA I 270–1); (2) AA \*ʔ-b “to desire” [GT] > i.a. Sem. \*ʔby: Hbr. ʔby qal “vouloir, désirer ardemment”, JPAr. ʔābā “vouloir” || Hadramaut ʔabā “vouloir” (Sem.: DRS 3) ||| **Eg.** jbj “dürsten, durstig sein” (OK, Wb I 61, 8) ||| **WCh.:** Boghom yip, Kir yip “to thirst” (WCh.-Eg.: OS 1990, 90, #48); (3) AA \*l-(w)-b “to love” [GT] > i.a. Sem.: Ar. lwb “dürsten” [Erman] ||| **Eg.** 3bj [reg. < \*lby or \*rby] “wünschen” (MK, Wb I 6–7) ||| **WCh.:** Bole-Tangale \*lab- “любить” [Stl. 1987, 247].

<sup>15</sup> Cf. also **WRift** \*ʕač “to be mature, fully grown” [KM] > **Grw.** ʕāč “to be mature”, **Alg.** ʕač “to be fully grown”, ʕača “mature”, **Brg.** hīʕač “to be strong” (**WRift:** KM 62). Primarily \*ʕ “fat, robust”?

<sup>16</sup> **Eg.** ʕḏ [if < \*ʕg] has been alternatively compared to **ECh.:** Mokilko ʔègéy “Öl” [Lks. 1977, 222] (for **Eg.** Mokilko cf. OS 1989, 135; 1992, 190), which is less convincing.

• **WRift \*ʕāg<sup>w</sup>** “to dig for water, dig a well” [KM 51]: cognate with LECu.: Saho *ōʕoge* “to bury” [Vergari 2003, 142] = *-uʕug-* “to bury” [Welmers 1952, 150] = (Assaorta dial.) *-ʕag<sup>w</sup>-* “seppellire” [CR 1913, 39] = *ʕag* “be-/eingraben” [Rn. 1890, 60] = (Irob dial.) *ɣaga* “begraben” [Rn. 1878, 137] ||| Eg. *ʕd* [reg. < \*ʕg<sup>w</sup>] “hakken, aushöhlen, graben (Fundament, Grab)” (OK, ÄWb I 297) ||| WCh.: Angas-Sura \**hok* [\*h- < \*ʕ- (?), \*-k < \*-g# reg.] “to dig (with a hoe?)” [GT 2004, 157]. AP: PKuliak \**ʕug* “to dig” [Ehret 1981, 99].

• **WRift \*ʕalu**, pl. \***ʕalalē** “pot” [KM 54]: cp. perhaps Eg. *ʕn* [reg. < \*ʕl] “Art Gefäß” (OK, Wb I 187, 12).

• **WRift \*ʕāmiya**, pl. \***ʕāmú** “1. fruit, 2. boil (disease)” > i.a. Brg. *ʕāmiya* “1. fruit, 2. testicle”, PIrq. \**ʕamʕāmimo* “testicle” [KM 55–56]: its primary sense may have been “round, circular object”, cf. Ar. *ʕwm*: *ʕām-at-* “1. gerbe, botte (de blé, de foin), 2. cône formé par le turban roulé en spirale, 3. tête de cavalier coiffée d’un turban, 4. outre gonflé d’air” [BK II 413]. Noteworthy might be Eg. *ʕmm* “ein tierischer Körperteil (von Gans, Fisch, Wurm u.a.): Ei, Roggen (?)” (Med., Wb I 186, 14–16), although it has been recently rendered as “brain”.

• **WRift \*ʕānĉ-ō** “(molar) teeth” [KM 57]: this is hardly an inner innovation derived from WRift \**ʕaĉ-* “to be mature” as KM suggest, but rather cognate with LECu.: Saho *aĉ-a* “back teeth” [Lmb. 1987, 533] ||| Om. \**aĉ-* “tooth” [Bnd. 1988, 145]: Ometo \**aĉ-* & Dizoid \**aĉ/ž-u* & Aroid \**aĉ-i* “tooth” [Bnd. 2003, 122, 219] (Cu.-Om.: Dlg. 1973, 307; Flm. 1969, 26–27; 1974, 90; 1976, 320) ||| Ar. *ʕdĉ* “mordre à qqch., saisir avec les dents, et y enfoncer les dents” [BK II 276] ||| WCh. \**H/ʕaĉ-* “to bite, chew” [Stl. 1991 MS, 7; 1995, 61] < AA \**ʕ-(n)-ĉ* “to bite, chew” [GT]. This set of cognates is probably to be separated from SCu.: Dhl. *ħunĉ-* [*ħ-* < \**ʕ-* would be irreg.] “to chew” [EEN 1989, 27] || HECu. \**inĉ-* “to chew” [Hds. 1989, 413], for which cf. perhaps Eg. *nĉ.t* [*-d-* reg. < \**-ĉ-*] “tooth” (OK) via possible met.

• **WRift \*ʕānta**, pl. \***ʕāntō** “termite hill” [KM 57]: its connection to Sandawe *ʕántá* “calabash with wide mouth” preferred by KM is not too likely. Cp. rather Ar. *ʕuntūt-* “1. petite montagne dans le désert, 2. pénible, difficile à gravir (colline)” [BK II 380].

• **WRift \*ʕay-i**, pl. \***ʕay-ō** “flower, blossom” [KM 65] ||| Eg. *ʕ* “Erzeugnis des Ackers” (GR, Wb I 159, 14), cf. *ʕj* “Kind (?)” (LP, Wb I 169, 5). Any connection to Ar. *ʕayiy-* “trop faible” [BK II 416]?

• **PIrq. \*ʕayl-ā** “wedding dance” [KM 65] ||| Ar. *ʕāyil-at-* “2. fille, femme, 3. famille”, *ʕayyal-* “femme, enfant ou toute autre personne dont l’entretien est à la charge d’un père de famille”, *ʕiyāl-* “famille, femme, enfants et toute la domesticité à la charge d’un père de famille” [BK II 423].

• **SCu. \*ʕĉ-** [\*-tl-] [\*-ĉ<sub>2</sub>- < \*-ĉ-] “unripe, raw” [GT after Ehret 277]: WRift \**ʕĉ* “to be unripe” [KM 102] > Irq. & Alg. *ʕĉ* | Ma’a *išé* (SCu.: Ehret 1980, 277): related to Bed. *yaĉaʔ* “feucht, nass, unreif sein” [Rn. 1895, 241] || LECu.: PSam \**ʕayĉi* “unripe” [Heine]: Somali *ʕayĉin* ~ *ʕĉin*, Rendille *ħĉdi* [*ħ-* < \**ʕ-*] (Sam: Heine 1978, 77) ||| Sem.: Ar. *yaĉĉa* I “être plein de vigueur, de jeunesse (se dit d’un homme, d’une plante)”, II “être dans la bien-être, prospérer”, *yaĉĉ-* “frais, tendre, nouveau, plein de vigueur et de santé (homme), tout jeune, né récemment (veau, etc.)”, *yaĉw*: *yaĉā* “être entier, complet et en bon état” [BK II 473–4, 477] ||| Eg. *ʕd* “wohlbehalten, unversehrt, intakt sein” (MK, Wb I 237–238, so also GHWb 165)<sup>17</sup>, orig. “fresh” (?) [GT] ||| WCh.: Diri *yaĉa* “unripe, wet” [Skn. 1977, 47] < AA \**ʕ-(y)-ĉ* “fresh” [GT].<sup>18</sup>

<sup>17</sup> The comparison of Eg. *ʕd* to Sem.: Ar. *ʕalaġa* (Alb. 1918, 238; Ember 1930, #5.a.14; Clc. 1936, #137; Vrg. 1945, 135, #9.b.2, 146, #24.a.1) or to Sem. \**ʕzz* “to be strong, powerful” (Ward 1962, 412, fn. 3; 1968, 69) cannot be accepted, since the Eg. root was neither \**ʕnd* nor was Eg. *d* a reflex of Sem. \**z*.

<sup>18</sup> Lit.: Hodge 1968, 27 (Eg.-LECu.-Ar.); Dlg. 1987, 209, #104 (LECu.-Ar.-SCu.).

● **WRift \*ʕēt** < \*\*ʕaw-it (med.) “to descend” [KM 101] = \*\*ʕa[(w)h]-it (???) [GT] ||| Ar. ʕwh II: ʕawwaha “2. faire halte, s’arrêter à la fin de la nuit, après avoir voyagé pendant toute sa durée, 3. être arrêté, retenu dans un endroit” [BK II 415]. For the semantic shift cf. Eg. ḥnj “1. niederschweben, sich niederlassen auf, 1. Halt machen bei etwas” (OK, Wb III 287).

● **WRift \*ʕis** “to do” [KM 166] || LECu.: Saho is- ~ iš- (so, no \*ʕ-) “to do, make” [Sasse] = iše “to make, do” [Vergari 2003, 107], Boni as- “to prepare, make” [Sasse] | HECu. \*ass- (var. \*iss-) “to do” [GT, cf. Sasse 1982, 107; Hds. 1989, 51] ||| Sem. \*ʕšy “to make, do” [GT] > Ug. ʕšy G “to make, process, work” [DUL 190], Hbr. ʕšy qal “to make, manufacture” [KB 890] || OSA: Sab. ʕs<sub>1</sub>y “1. to make, do, 2. acquire, buy” [SD 20] = Madhabi ʕs<sub>1</sub>y “1. faire, 2. prendre possession de qqch.” [Arbach 1993, 17]. The Sem.-Alg. etymology was first suggested by A.Ju. Militarev (1986, 74).

● **WRift \*ʕōc** “to shrivel, shrink, dry up” [KM 222] ||| Ar. ʕsw ~ ʕsy “1. se durcir et sécher (se dit des plantes), 3. se durcir, contracter une peau dure par suite du travail (se dit des mains)”, also ʕys IV “sécher, perdre sa verdure (se dit des céréales)” [BK II 258, 420] (WRift \*c = Ar. s < AA \*c reg.).

● **SCu. \*ʕōc** “cheekbone” [Ehret 278] || ECu. \*ʕaḍ- “cheek” [GT]: e.g. Gollango ʕaḍ- “Backe” [AMS 1980, 195] | Konso aḍ-á, Dirayta aḍ-á (pl.), Mosiya aḍ-eta “cheek” (Konsoid: Lmb. 1987, 533, #8.b) ||| Sem. \*ʕaṣ(aṣ)- “bone” [Lsl. 1945, 233].

● **WRift \*ʕōš** “to curse, insult” [KM 222] may eventually derive from AA \*ʕ-š “to call, ask” [GT] > Eg. ʕš “rufen” (MK, Wb I 227, 4) = “to summon” (FD 48) ||| ECu. (?) Yaaku -es- “to beg sy.” (tr.) [Heine 1975, 121] = -is- “to pray” [Ehret 1987, 117, #491] ||| WCh.: Bole ʕeeš- “rufen” [Lks. 1971, 133] | Boghom yi:s “rufen” [Smz.] (WCh.: JI 1993, 140; 1994 II, 58). The Eg.-WCh. parallel was suggested already in OS 1992, 193; Orel 1995, 100, #8. For the semantic shift in WRift cf. the same change of meaning in SCu. \*wāʕ- “to curse, revile” [Ehret] < AA \*w-ʕ “to shout” [GT].<sup>19</sup>

● **SCu. \*ʕora** “ashes” [Ehr.] > WRift \*ʕura “ashes” [KM 310] | Asa wura-to [wu- < \*ʕu- < \*ʕu-] “ashes” [Ehr.] | Ma’a iʕorá “ashes” [Ehr.] (SCu.: Ehr. 1980, 279, #34) || (?) ECu. \*ʕār- “steam, smoke” [Ss. 1982, 161] (S-ECu.: Ehr. 1987, 117, #493) ||| Eg. ʕj “Feuersglut” (LP, Wb I 166, 9).

● **WRift \*ʕuĉ** “1. to wrap, 2. twist, wring” [KM 311]: cf. HECu. \*unĉiĉ- (?) [epenthetic \*-n-?] “to twist” [Hds.]: Hadiya inĉiĉ-, Kambata inĉiĉĉ-, Sidamo hūnĉ- (HECu.: Hds. 1989, 159), which indicates that WRift \*ʕuĉ < \*ʕuĉ here. A reflexation with non-glottalized AA \*-ĉ- has been retained by Ar. ʕanaša I “plier, courber”, III “mettre la bras autour du cou de qqn. et le saisir” [BK II 384].

● **SCu. \*ʕum-** “to stop (intr.)” [GT]: Irq. ʕom- “to stop (intr.)” | Ma’a -ʕúma “to stand, come to stop”, -ʕumáti “to stop (tr.)” (SCu.: Ehret 1980, 278) ||| Eg. wʕm.t [Belova’s law: < \*ʕum] “Verhinderung, Einschränkung” (XVIII., GHWb 185)<sup>20</sup>.

● **SCu. \*ʕur-** “strength” [GT]: Irq. & Alg. ʕuru | Ma’a ʕú [reg. < \*ʕur] (SCu.: Ehret 1980, 279) ||| Eg. ʕ3 [< \*ʕr] “great” (OK, Wb I 161–162).

<sup>19</sup> Cf. Sem.: Ar. waʕwafa “to shout” [Ember 1930, 32] ||| Eg. wʕ3 [< \*wʕʔ] “Böses reden, schwächen, schwächlich reden” (MK, Wb I 279, 14–17) = “to curse” (FD 57) ||| PCu. \*waʕ-/wāʕ- “to yell” [Ehret] > Bed. wuʕ “to cry out” [Ehret] || ECu. \*waʕ- “to shout” [Sasse 1979, 42] || SCu. \*wāʕ- “to curse, revile” [Ehret]: Asa waʕ-am [Ehret] | Dhl. wāʕ- “to insult, curse” [EEN 1989, 43] (SCu.: Ehret 1980, 313; Cu.: Ehret 1987, #585) vs. WRift \*ʕōʕ- (different ablaut pattern of the same root?) “to call, warn” [KM 219] ||| Ch. \*wa “to call” [Nwm. 1977, 23]. The AA comparison was first suggested in Mlt. 1984, 157.

<sup>20</sup> The meaning of the NEg. hapax is corroborated by the regular SCu. reflex. We may postulate an underlying Eg. verbal root \*wʕm “to hinder, keep up, stop”, whose Coptic reflex has already been identified by W. Westendorf (KHW 271) in (S) OYΩOME > OYΩME “einschränken, verringern”, refl. “sich zurückhalten, nachsichtig sein, nachgeben”.

● **WRift \*ʕurufi** “lizard (sp.)” [KM 310]: presumably related to Ar. ʕaḍrafūṭ- (< √ʕḍrf with inetymological \*-ʕ- + additional -ṭ, cf. pl. ʕaḍārifu, dimin. ʕuḍayrifu) “sorte de lézard sur lequel les démons voyagent” [BK II 280].

● **PRift \*baʕ-** “to excel, exceed” [Ehr.] > WRift \*baʕ “to surpass, exceed, win” [KM]: Irq. baʕ- “to excel, exceed” [Ehr.] = bāʕ “to surpass, defeat” [KM], Grw. bāʕ “to surpass, exceed, win” [KM], Alg. baʕ- “to overcome” [Ehr.] = baʕ “to surpass, exceed, win” [KM], Brg. baʕ “to surpass, exceed, win” [KM] | Qwd. baʕ-at- “to increase in size” [Ehr.] (WRift: KM 66; SCu.-Orm.: Ehret 1980, 338) || LECu.: (?) Oromo bay-e “plenty, abundant” [Ehr.] = bāy-ē “many, much” [Hds. 1989, 96] || Sem. \*bɣy “das Maß überschreiten” [GB]: Hbr. bʕy qal “(das Wasser) anschwellen, d.i. überkochen machen (Feuer)”, nifal “anschwellen, sich vorschieben (v. einem Mauerstück)” [GB 106] || Ar. bɣy: bayā “das Maß überschreiten (auch von einer Wunde)” [GB] = “1. sortir des bornes, 2. dépasser ses droits, abuser, opprimer” [DRS 76] < AA \*b-ɣ-(y) “to exceed” [GT].

● **WRift \*boʕ-** “antisocial behaviour” [KM] > Irq. biʕiri “greed”, Alg. biʕa “adultery”, Brg. biʕa “madness, lunacy” (WRift: KM 73) || Sem. \*bʕw > Ar. baʕw- “crime”, baʕā “commettre un crime, un péché”, Thamudic bʕw “commettre un crime”, OSA bʕw “attaquer par surprise, se révolter” (Sem.: DRS 74).

● **WRift \*boʕ-** “black” [GT]: Irq. boʕ-â, Brg. boʕ-i (WRift: Flm. 1969, 22, #6; KM ) || Ar. byṭ ~ bys (root ext. -ṭ vs. -s?): bayṭa “être tacheté de noir, grisâtre”, cf. bays- “couleur noire” [DRS 76: isolated in Sem.] || WCh.: Fyer ʔwû [*\*bʔiH < \*√b-w-ɣ?*] “black” [Jng. 1968, 9, #91] | cf. Bole buʔum “black” [IL in JI 1994 II, 28]: < AA \*b-ɣ “black” [GT].

● **WRift \*daʕarā** “ashes” [KM 78] || ECU. \*darʕ- [met. < \*\*daʕr-?] “ashes” [Sasse 1979, 16] || Sem.: Ar. daʕira I “1. fumer beaucoup (se dit du bois vert, quand il ne brûle pas bien), 2. fumer et ne pas donner de flamme (se dit d’un briquet consistant en une espèce de bois qui s’allume par le frottement)” [BK I 700] = “fumer sans donner de flamme” [DRS 293: isolated in Sem.]. In the light of ECU. \*rʕ- and the Ar. cognate, the derivation of the WRift term from \*daʕ- “to burn, roast” does not seem fully evident.

● **WRift \*dabaʕ** “to circumcise” [KM] = “to deprive” [GT] > Irq. dabaʕ “to lack”, Brg. dabaʕ “to circumcise” (WRift: KM 80) || Ar. daʕaba I “repousser, éloigner” [BK I 698].

● **SCu. \*deʕ-** “to be fat” [Ehret]: WRift \*duʕiya “fat, oil”, pl. \*duʕáy “pieces of fat” [KM 99] > Irq. deʕ-eta “fat, lard” [Ehret] = dééʕēta “fat”, cf. díʕi “oil” [Wtl. 1953] = diʕi “oil, fat” [Mgw. 1989, 112], Burunge dǝʕǝ & dǝgiʕa “Fett” [Mnh. 1906, 331] | Asa daʕ-ara “heavy” [Ehret] | Dhl. deʕʕ-em- “to be fat” [Ehret] = deʕ-em- “to be fat” [EEN 1989, 28] (SCu.: Ehret 1980, 165, #22) || Sem. \*dʕdʕ “to move with heavy steps, with difficulty” [GT]: Ar. daʕdaʕa “courir d’un pas lourd” || Amh. da (da) ʕalā “marcher lentement, avec difficulté, parler difficilement” (Sem.: DRS 289–290) || (???) OEg. \*dʕ [reg. < \*dʕ] attested only late as LEg. tʕ “fett, gemästet (von Ochsen)” (GR, Wb V 325, 15).

● **WRift \*gaʕaw** “to watch, observe” [KM 111] || Eg. wdʕ [reg. < \*gʕwʕ] “offenbar sein” (LP, Wb I 406, 15).

● **WRift \*kiʕ** “to return”, \*kiʕ-is (caus.) “to bring back” [KM 175] || Sem.: Ar. kʕ IV “retenir qqn. et l’empêcher d’aborder qqch.”, kʕkʕ: kaʕkaʕa “2. retenir, contenir qqn. et l’empêcher d’aborder qqch.” [BK II 906, 908].

● **WRift \*kʕafi ~ \*kʕafay** “hare” [KM 187]: perhaps connected to Sem.: Ar. kwʕ: kāʕa “reculer de peur, et s’éloigner de qqch.” [BK II 944].

● **WRift \*liʕ** “to hate” [KM] > Irq. liʕáy (m) “paying back, hate among neighbours, revenge”, Alg. liʕ “to hate”, Brg. liʕ “to be unhappy” (WRift: KM 194) || LECu.: (?) PSomali \*neʕ- [unless < \*neʕeb-] “hassen” [Lmb. 1986, 445] || Eg. nʕw [*\*lʕʕ*] “médisance (?)” (CT IV 20c, AL 78.1995) = “slander (?)” (Osing) > Dem. lʕ “Strafe, Verleumdung” (DG 260:10) = “pun-



ishment, slander” (CED) > (SALB)  $\lambda\lambda$  “Neid, Bosheit, Verleumdung” (KHW) = “envy, slander” (CD 134b) = (SB)  $\lambda\lambda$  vs. (F)  $\lambda\epsilon$  “1. envie, méchanceté, 2. calomnie” (DELC 93; Eg.-Cpt.: KHW 74; Osing 1978, 187) ||| WCh.: Angas ka-la (prefix ka-?) “accusation” [ALC 1978, 23] < AA \*l- $\lambda$  “to slander from hatred” [GT].

• **SCu. \*lo $\lambda$ -** “to bring to” [GT]: Brg. lu $\lambda$ - “to apply”, Alg. lu $\lambda$ - “to apply” | Qwd. lo $\lambda$ - “to give” | DhI. lo $\lambda$ - “to bring to” (SCu.: Ehret 1980, 206) || Agaw: Qwara lē “geben”, caus. lä-š “bringen” [Rn. 1885, 92] ||| Eg. jw $\lambda$  [< \*lw $\lambda$ ] “beschenken mit (m)” (XVIII., Wb I 51, 8) = “to reward” (FD 13) → jw $\lambda$  “Belohnung” (GR, Wb I 51, 10) < AA \*l-w- $\lambda$  “to hand over to” [GT].

• **Pirq. \*na $\lambda$ ay** “child” [KM 214]: cp. perhaps Ar. na $\lambda$ - “faible, débile, infirme (se dit d’un homme)” [BK II 1291].

• **WRift \*ni $\lambda$**  “to move in a circle” [KM] attested in Brg. nīni $\lambda$ -id “to encircle” + WRift \*ni $\lambda$ -im “to dance” [KM 217] > Brg. ni $\lambda$ -im “to have sexual intercourse” has a striking parallel with similar semantic development in Sem. \*n $\lambda$ y > Akk. \*nī/ū D “aufhetzen, aufstacheln, aufputschen, anregen”, ēpiš nu??ūti “jemand, der die Leute mit Witzen o.dgl. anregt, Witzmacher” [Soden 1955, 388–9] || Ar. n $\lambda$ y: na $\lambda$ ā I “3. exciter”, X “1. courir çà et là et porter son cavalier on ne sait où” [BK II 1300] ||| Eg. n $\lambda$ w “sich Gauklerweise herumziehen” (PT 702a, ÜKAPT VI 138) = “to caper with” (AEPT 132) = “to cavort (?)” (Allen 1984, 570) = “bondir (??)” (Jacq 1986, 31) = “herumtollen” (GHWb 395; ÄWb I 600), cf. n $\lambda$ w.tj “fahrender Gaukler” (AÄG 107, §247; ÄWb I 599) > (?) n $\lambda$ w “s’accoupler” (late NK, AL 77.2005) = “sich paaren” (GHWb 395). Cf. also Sem.: Hbr. nw $\lambda$  “to quiver, wave, move unsteadily” || ES: Tna. na $\lambda$ nā $\lambda$ e “to shake, make fall, reverse” (Sem.: Lsl. 1958, 33). The PT hapax and the LEg. verb have been connected by Satzinger (2003, 222). Eventually, the underlying AA root may have primarily signified an excited rhythmical circular motion.

• **Rift \*po $\lambda$ -** “to clean” [Ehr.] > Brg. bo $\lambda$ -em-is- | Qwd. po $\lambda$ - (Rift: Ehr. 1980, 340, #2): presumably a met. of AA \* $\lambda$ -b “to be clean” [GT] > Sem.: Akk. G ebēbu “to be clean, pure”, D ubububu “to clean, purify” [AHW 180–1] ||| Eg.  $\lambda$ b.w “purification, purity” (OK, FD 40), w $\lambda$ b “(to be) pure” (OK, FD 57) = “reinigen, rein sein” (Wb I 280–2) ||| CCh.: PHigi \*y[a] $\lambda$ - [met. < \*Hab-] “to wash” [GT]: Higi-Nkafa ya $\lambda$ u-, Kapsiki ya $\lambda$ u-, Fali-Gili ‘ye $\lambda$ u-, Higi-Ghye ya $\lambda$ u- | Bachama ya $\lambda$ wə “to wash” | Kuseri ubau “to wash” (CCh.: Krf. 1981). A widely accepted AA etymology.<sup>21</sup>

• **SCu. \*wā $\lambda$ -** “to spill out, flow” [Ehret] > WRift \*wa $\lambda$  “to vomit, flood, gush out” [KM 312] | Qwd. wa $\lambda$ a-muko “river” [Ehret] | Ma’a -wa “to urinate” [Ehret] (SCu.: Ehret 1980, 313): cognate with ECu. \*wa/i $\lambda$ - “water” [Sasse 1982, 186] > LECu.: cf. e.g. Afar wa $\lambda$  “river” [Mkr. 1987, 298] | HECu. \*wa $\lambda$ -a “water” [Hds. 1989, 423]. There was also an AA var. root with \*y-, cf. LECu. \*yā $\lambda$ - “to flow away” [Sasse 1982, 192] ||| Eg. j $\lambda$ j “waschen” (OK, Wb I 39)<sup>22</sup>. A.B. Dolgopolsky (1988, 631, #17) affiliated ECu. \*yā $\lambda$ - with the Sem. \*hy $\lambda$  assuming a secondary prefix \*h- (for which, however, cf. rather SCu. \*ho[ $\lambda$ ]- “to pour” [GT] below). It is entirely plausible that PAA \*w- $\lambda$  ~ \*y- $\lambda$  ~ \*h- $\lambda$  were ultimately related.

### SCu. \*h

SCu.	WRift	Qwadza	Asa	Ma’a	Dahalo	Sem.	Eg.	AA
*h	*h	-h-	?-/h- (?)	h-, -h-, - $\emptyset$	h	*h	h > h	*h

<sup>21</sup> For Eg.-Akk.: Alb. 1918, 222, fn. 1; 1923, 67; Holma 1919, 37; Ember 1930, #5.a.17; Blv. 1991, 88, #16; 1993, 53, #16. For Eg.-CCh.: OS 1992, 200.

<sup>22</sup> For the shift of meaning cf., e.g., PIE \*plew- “rinnen, fließen” > i. a. Gk. πλύνω “ich wasche”, OHGerm. flouwen, flewen “waschen, spülen”, Lith. pláuju (caus.) “wasche, spüle” (IEW 835–6).

• **WRift \*bōh** “heap, bundle” [KM 75] ||| WCh.: Waja à búw-à [irreg. b-?] “to pile up, heap up (past)” [Kwh. 1990, 240] < AA \*b-w-ḥ “to heap up” [GT]: remotely cognate to the reflexes of AA \*p-(w)-ḥ “to heap up” [GT] > Eg. pḥ.t “die Garbe des Kornes” (OK, Wb I 533, 11) = “gerbe” (AL 78.1489) = “der Garbenhaufen (des Kornes), Kornmiete” (GHWb 287) ||| NBrb. \*a-ffa < \*\*a-fwaH “stack of sheaves” [GT]: Mgild a-ffa, pl. i-ff-an “stack of grain-sheaves” [Harries 1974, 222] | Iznasen & Rif & Senhazha ta-ffa, pl. ta-ffiw-in “meule de gerbes à dépiquer” [Rns. 1932, 297] | Qabyle ta-ffa, pl. ta-ffw-in “tas de bois” [Dlt. 1982, 189] || CCh. \*pa (?) “to pile up” [GT]: Margi pá “to pile up, fold” [Hfm. in RK 1973, 134, #123, #131, #135] | Bana pá “ramasser, rassembler” [Hfm. in Brt.-Jng. 1990, 86].

• **WRift \*cēḥ-a** “dry dung” [KM]: Irq. ceḥa “dry cowdung” [Ehret 1980, 355], Grw. cēḥa “dry dung” [KM], Alg. cēḥa “cow dung” [KM], Brg. cēḥō “dung of sheep and goats” [KM] (WRift: KM 296) ||| Eg. šḥ.w (pl.) “Dreck” (NE, Wb IV 211, 12).<sup>23</sup>

• **WRift \*cuḥiya** “finch”, \***cuḥ-ā** “finches” [KM 302] ||| Eg. šḥ.t “Art Vogel (der geopfert wird)” (XXII., Wb IV 209, 8).

• **WRift \*caḥ-as** (caus.) “to forbid, prevent” [KM]: Irq. caḥ-ās “to keep cattle herd together, prevent cattle from going astray”, Alg. caḥ-as & Brg. cācaḥ-as “to forbid, prevent” (WRift: KM 2004, 275) ||| Sem.: (?) Ar. ḏḥw II “2. faire paître les bestiaux” [BK II 12] (unless < “pleine jour”) ||| NBrb.: Tamazight **vd**: ta-ḏa (no pl.) “protection, alliance instituée entre deux tribus, surtout contre les vols” [Tf. 1991, 86] (not found in Qabyle, Tuareg).

• **WRift \*cēḥ** “to build, do, prepare”, \*cēḥitu “building, constructing”, cēḥimā “character, nature” [KM 282] ||| Sem.: Ar. šyḥ: I “1. faire qqch. avec le plus grand zèle, être très-sérieusement occupé de qqch.”, III “1. s’appliquer avec assiduité etc.”, IV “2. faire qqch. avec le plus grand soin” [BK II 1295].

• **WRift \*fāfaḥ** < \*\***faḥafaḥ** “to intrigue, persuade” > esp. Brg. hifāfaḥ “to tell or give away secrets, tattle, blab” [KM 103] ||| Sem.: Ar. fḥw II: faḥḥā “2. indiquer qqch. par des allusions, faire une allusion à qqch.”, faḥwā- and fuḥawā?- “sens (des mots, des paroles, d’un discours), intention, ce que l’on se propose dans une phrase, dans un discours”, cf. perhaps faḥfaḥa “3. être sincère dans son amitié” > faḥfāḥ-at- “bavard, loquace” [BK II 551, 549].

• **WRift \*fēḥ-is** (caus.) “to tear, slit” [KM 106] ||| Sem.: Ar. fyḥ: fāḥa “1. être vaste, spacieux (p.ex. la caverne), 2. répandre” > esp. fayāḥ- “incursion, invasion de l’ennemi qui se répand dans toutes les parties du territoire” [BK II 644]. The primary sense of the underlying AA root (\*f-y-ḥ) might have been “to open wide”.

• **WRift \*gaḥ** “to shake off, throw off” > \*gāgaḥ “to let fall, drop off”, \*gaḥ-at “to scold, rebuke” [KM 112] ||| Sem.: Ar. ḡwḥ: ḡāḥa “2. dévier, s’écarter de la ligne droite”, II “déchausser (le pied), c.-à-d. en ôter la chaussure” [BK I 350] ||| Eg. wdḥ [reg. < \*wgḥ < \*gwḥ via Belova’s law] “(den Säugling) entwöhnen” (PT, Wb I 409, 14).<sup>24</sup>

• **WRift \*guḥulay** “knobbed club” [KM 121]: its semantically dubious equation with Bed. gulhe “upper arm” (suggested by KM) cannot be accepted. Instead, I would rather think of a basic sense “huge round knob”. Cp. Ar. ḡaḥl- “(outr) ronde, pleine”, ḡayḥal- “gros, massif, imposant” [DRS] = ḡaḥl- “grande outre à eau”, ḡayḥal- “1. rocher, pierre énorme, 2. grand, énorme (se dit de toute chose)” [BK I 257] || MSA: Jibbali ḡihál- “tonneaux” [DRS], Sgt. ḡhel “devenir fort” [Lsl. 1938, 108] (Sem.: DRS 112).

<sup>23</sup> Is NEg. šḥ.w just the met. of OEg. ḥs [\*ḥās] “Exkremente, Kot” > (S) **zoc** (PT, Wb III 164) ||| (???) Bed. hūš [irreg. -š] “dünnere Kot” [Rn.] = “liquid faecal matter” [Rpr.] (Bed.-Eg.: Zhl. 1932–3, 170)?

<sup>24</sup> Semantically, less likely seems an equation of the Eg. root with Sem.: Hbr. ḡyḥ: goḥī (qal part. with suffix) “der mich hervorzieht” [GB 137] || Ar. ḡaḥḥa “to draw” [Zbr.], cf. also ḡḥw: ḡaḥā I & VIII “arracher, déraciner, extirper” [BK I 257–8] | MSA: Jbl. wgḥ: egēḥ “to snatch, take by force”, ṣagēḥ “to snatch from” [Jns. 1981, 288, 69] (Sem.: Zbr. 1971, #53).

• **SCu. \*h[ay]d-** “to be fat” [GT] = \*hed- [Ehr.] > Irq. ḥar-em-is- [-r- reg. < \*-d-] (caus.) “to fatten” | Dhl. ḥed- “to be fat”, ḥéd̄dune “fat (person)” (SCu.: Ehret 1980, 301, #16) ||| ECu. \*ḥayḍ- “fat” [Sasse 1979, 46]. Irreg. SCu. \*-d- vs. ECu. \*-ḍ-.

• **WRift \*ḥaf-a** “sticks of the roof construction of a house” [KM 146]: cf. Dobase ḥap-e “kleiner Zweig” [AMS 1980, 163].

• **WRift \*ḥaf** “to spread” [KM 146] > cf. esp. Grw. ḥāf “to cover, lay on top”: equated by KM with Oromo ʔāf “to spread”. GT: cf. also Ar. ḥaffa “couvrir de tous côtés d’un voile, d’un rideau (une litière, etc.)” [BK I 455].

• **WRift \*ḥagili** > only Irq. ḥagli “string of beads” [KM 331] ||| Ar. ḥaḡl- ~ ḥiḡl- ~ ḥiḡil- “anneau, boucle d’entrave, entrave” [BK I 384]. KM’s comparison with LECu.: Somali ḥarg-o “rope” may only be valid provided -rg- < \*-lg-.

• **WRift \*ḥam-** “to be lost”, \*ḥam-is (caus. suffix \*-is-) “to extinguish” [KM 147–8] = “to be spoiled” [GT]<sup>25</sup>: Brg. ḥamu “to be spoiled, lost”, ḥam-is- “to spoil, lose the use of”, Alg. ḥam-is- “to get lost, forget” (WRift-LECu.: KM l.c.; Ehret 1980, 334, 379) || ECu. \*ḥam-/\*ḥum- “bad” [Sasse 1979, 38]: cf. PSam \*ḥum “bad” [Heine 1978, 64] ||| Eg. ḥm.t “Übel, Unglück” (GR, Wb III 80, 12) ||| Sem.: (?) Geez ḥamäyā “to speak ill of” [Lsl. 1944, 55–56]. For Eg.-ECu. see Takács 1995, 161, #c; 1996, 90, #c.

• **SCu. \*ḥame** “father’s brother” [Ehret]: Ma’a hame “my uncle” | Dhl. ḥáme “father’s brother” (SCu.: Ehret 1980, 299) || Bed. hamo “father/mother-in-law” [Rpr.] ||| Eg. \*ḥm [palatalized < \*ḥm] attested as šm “father-in-law” (fem. šm.t “mother-in-law”) → Cpt. (S) **ግዕዝ** “father-in-law” (OK, Roquet 1977; Vcl. 1990, 53; Ward 1996, 43, fn. 8) ||| Sem. \*ḥam- “father-in-law” [GT]. The Sem.-Eg. match has already been known.<sup>26</sup> Perhaps Cu. < Ar.?

• **WRift \*ḥāmpu** “wing” [KM 148] was equated by KM with ECu. \*ḥubn- “muscle, limb”, which is both semantically and phonologically (WRift \*-p- ≠ ECu. \*-b-) unlikely. Cp. rather Eg. ḥp “vielleicht ein Wort für Hand (?)” (NK, Wb III 69, 17).

• **WRift \*ḥapē**, pl. \*ḥapapu “soil, ground, earth, land” [KM 149] ||| Sem.: presumably Ar. ḥifāf- “côté (d’une chose)”, ḥāff-at- “côté, bord, marge”, ḥāf-at- “bord, marge, extrémité” [BK I 455–6, 516] ||| Eg. ḥp.tj “1. als Bez. der Welt, 2. (vereinzelt) als äußerste Grenze eines Gebietes”, ḥpḥp “das Äußerste der Welt”, cf. t3 m ḥp.t=f “die Erde in ihrer ganzen Ausdehnung” (GR, Wb III 69).

• **WRift \*ḥaraci** “weather” [KM 150]: cp. perhaps Ar. ḥars- “âge, moment, temps” [BK I 407]. Cf. Hungarian idő “time” ~ idő(járás) “weather”.

• **SCu. \*ḥas-** “tendon (bow string)” [Ehr.]: Irq. ḥasini “spinal cord” | Asa ʔasu-k “tendon (bow string)” | Ma’a lu-hási “tendon (bow string)” | Dhl. ḥáso “tendon (bow string)” (SCu.: Ehret 1980, 300) ||| Eg. ḥs [GT: GW ḥs3] “Faden, Schnur (z.B. zum Aufreihen von Perlen)” (NE, Wb III 166, 4) = “thread, cord, string” (DLE II 139).

• **PIrq. \*ḥasām-a** “problems, trouble” [KM 150]: cf. perhaps Ar. ḥusūm- “malheureuses (nuits)”, ma-ḥsūm- “mal nourri et qui ne grandit pas (enfant)” [BK II 427].

• **PIrq. \*ḥaš** “to worry” [KM 150]: cf. Ar. ḥwš VII “3. faire attention, appréhender, s’inquiéter de qqch.”, ḥyš I “avoir peur, être effrayé”, perhaps also ḥšy III “2. faire attention”, V “3. avoir honte devant qqch.”, VI “éviter qqch., se garder de qqch.” [BK I 513, 526, 436].

• **WRift \*ḥay-** ~ \*ḥaw- “row, family line” [GT] = \*ḥayi “row, line” [KM]: Iraqw ḥai “1. a patrilinear clan, 2. loosely, one’s kinsfolk” [Wtl. 1953, 87] = ḥay “kinsfolk, clan” [Ehret] = ḥayi

<sup>25</sup> WRift \*ḥam- is to be separated from WRift \*hamu-: Alg. hamu “hardship, distress”, Irq. hamusoʔo “unlucky person” (contra Ehret).

<sup>26</sup> Lit.: Erman 1892, 116 (Cpt.-Sem.); Chn. 1947, #127 (Eg.-Sem.-Bed.); Ward 1968, 69 (Eg.-Sem.); Rsl. 1971, 303 (Eg.-Sem.); Blv. 1989, 16, #11 (Eg.-Sem.).



“line, clan” [KM] ||| Eg. wḥj.t ~ wḥw.t [reg. < \*ḥuy/w-] “Familie, Sippe” (MK, Wb I 346, 9) ||| Sem. \*ḥayy- “clan” [GT]: Hbr. ḥay “family, kinsfolk” [KB 309] || Ar. ḥayy- “tribal community, descendants of one father” [Smith/KB] = “tribu (grande subdivision)” [BK I 523]. From the same root derives the isogloss of the following entry. Lit. for the AA etymology: Ember 1911, 93; Vrg. 1945, 138, #12.a.1; Mlt. 1984, 16 (Eg.-Sem.); HSED #1256 (Eg.-Iraqw). The reconstruction of SCu. \*ḥa- “relative, kin” [Ehret 1980, 299] is baseless.

• **WRift \*ḥeʔ-ēs-** (caus. suffix -ēs-) “to finish, bring to an end” [KM 152] ||| Sem.: Ar. ḥwy: ḥawā “réunir, rassembler” [BK I 522] ||| Eg. šwǝj (caus. prefix s-) “to collect, assemble” (MK, FD 258). For the semantic shift in WRift cf. Eg. kmm “to complete” > “to put an end to” (MK, FD 286).

• **ERift \*ḥel-** “to clean” [Ehret]: Asa hil-us- [h reg. < \*ḥ] “to strain, filter” || LECu.: Somali ḥal- “to wash” (Cu.: Ehret 1980, 335) ||| Eg. ḥ3.tj “Bleicher, Wäscher” (OK, GHWb 501; cf. Wb III 7, 7): nomen actoris of an unattested \*ḥ3 [< \*ḥl] “to clean white” [GT]<sup>27</sup> ||| Sem. \*ḥll: Akk. ellu [\*ḥall-u] “clean, pure”, Syr. ḥll pael “to wash, clean”, Aram. of Targum ḥll “abspülen”, pael “waschen, abspülen”. For Sem.-LECu.-ERift see Dlg. 1987, 198, #24; 1988, 631, #18.

• **WRift \*ḥitim** (dur.) “to destroy” [KM 154] ||| Eg. ḥtm “vernichten, beseitigen (Unreines, Durst), vertilgen” (OK, ÄWb I 911). Not yet clear whether the equation of Eg. ḥtm with Ar. ḥtm (supported by a number of Egyptologists<sup>28</sup>) is correct.

• **WRift \*ḥiyaʔa** “male relative, male sibling” vs. \*\*ḥiyaʔaʔō > \*ḥiʔō “sister” [KM 154] > i.a. Irq. ḥiya “brother” [Ehr. 1980, 299] = ḥiyaʔ “brother, cousin” [KM] ||| Sem.: Sqt. ḥiyo “fraternité” [Lsl. 1938, 172]. Ultimately related to AA \*ḥ-y “clan” [GT] (dealt with s.v. WRift \*ḥay- ~ \*ḥaw- “family line” above)?

• **WRift \*ḥōš** “to scratch, rub” [KM 155] || LECu.: Oromo hōsa “to caress, stroke (animals to calm down)” [Gragg 1982, 214] ||| Sem.: Ar. ḥwš III “1. exciter, pousser qqn. contre qqn. ou à qqch.” [BK I 513].

• **WRift \*ḥūmb-āt** “to crawl” [KM 155] ||| Eg. ḥnm.w [nm < \*nb reg.] “crawler”, ḥnmnm “to creep” (CT, DCT 338).

• **Alg. ḥungululay** “kite” [Ehr. 1980, 302, #25 with false Dhl. cognate] ||| Eg. tḥr [< \*kḥl] “Art Falke” (OK, Wb V 384, 12) ||| WCh.: Angas-Sura \*kələŋ < \*kalan ~ \*kelen ~ \*kilin (?) [reg. < \*k-l-ḥ-n] “falcon, hawk” [GT 2004, 171]<sup>29</sup> | Dera kalinže [-že < \*-ge] “kite” [Krf.] | Burma čəŋlèt [< \*kVIVŋ- via met.?] “kite” [Krf.] || CCh.: Pidlimdi kilengərti “hawk” [Krf.] etc. (Ch.: Krf. 1981, #202, #203).

• **SCu. \*kʷah-** “to be tired, have a rest” [GT]: WRift \*kwaḥ-as “to have a rest” [KM 186] > Alg. kwaḥ-as- “to rest” | Ma’a -kwaha “to be tired”, kwahamu “tiredness” (SCu.: Ehret 1980,

<sup>27</sup> Semantically a little bit weaker is the possibility of comparing Eg. \*ḥ3- to Sem. \*ḥwr: Syr. ḥewār “white”, Ar. ḥiḥwarra “to be intensely white” (Sem.: Rabin 1975, 89, #98).

<sup>28</sup> Ar. ḥaṭama “casser, écraser, broyer” [BK I 452] = “to break, destroy” [Ember] > ḥaṭm-at- “(year of) dearth, drought, sterility”, ḥuṭam-at- “a vehement fire that breaks in pieces everything that is cast into it (as name of the fourth stage of Hell)”, ḥuṭām- “what is broken in pieces, perishing” [Lane 594–5]. For Eg.-Ar. see Albright 1918, 238, #68; Ember 1930, #10.a.29, #14.a.26, #25.b.7; Vergote 1945, 143, #21.b.5. Note that the inner Eg. derivation from tm “vollständig sein” by a prefix ḥ- (Sethe 1910, 80, fn. 2; Vycichl 1936, 110) can be regarded as semantically satisfactory. Th. Schneider’s (1997, 204, #75) hazardous equation with Hbr. ṯtm (hapax, meaning obscure) is even worse, not to mention that Eg. ḥ ≠ Sem. \*ṯ.

<sup>29</sup> Attested in Angas “hawk (Hs. širna)” [Flk. 1915, 207] = kələŋ (KS) “falcon sp. (not eaten by the ḥkàn-čan people, who live among all Angas)”, kələŋ ~ ḥkələŋ (K) “falcon”, nkələŋ “Milan” [Jng. 1962 MS, 16, 28], Sura ḥkələŋ “Falke” [Jng. 1963, 78], Mupun nkələŋ “kite, hawk” [Frj. 1991, 44], Kofyar kòeloeng [kələŋ] “hawk” [Ntg. 1967, 18], Mushere kələŋ “hawk” [Dkl. 1997 MS] = kelen “hawk” [Jng. 1999 MS, 7], Chip nkələŋ “kite” [Krf.], Goe-may keleng (so, -e-) “a hawk” [Srl. 1937, 97] = kilin “kite” [Krf.] = kələŋ “kite” [Hlw. 2000 MS, 16].

266) ||| Sem.: Ar. kiḥkiḥ- ~ kuḥkuḥ- “1. vieille femme décrépète, 2. chamelle très-vieille”, kuḥkuḥ- (pl.) “vieilles femmes décrépètes” [BK II 869–870] ||| Eg. kḥkḥ “alt, hinfallig, gebrechlich werden” (XIX., Wb V 138, 10–12).<sup>30</sup>

• **WRift \*kʷah** “1. to throw, 2. forge” [KM 186] ||| Sem.: Ar. kwḥ: kāḥa I “2. plonger qqn. dans l’eau ou le jeter dans la poussière”, III “2. jeter des gros mots, des injures à la tête de qqn.” [BK II 941].

• **WRift \*leḥ-** “to search” [GT]: Irq. lēlēḥ- (sic, -h-!) “to look for” [Wtl. 1953, 91] = leleḥ- “to seek” [Ehr. 1980, 205] = lēlēḥ- “to search, look for” [Mgw. 1989, 117], Alg. leḥ- “to choose” [Ehret] (WRift: Ehr. 1980, 205, #21) ||| Eg. ḥḥj [\*lḥy?] “wünschen, erbitten” (MK, Wb II 288–289) ||| WCh.: perhaps NBauchi \*le?- < \*\*leḥ- (?) [GT]: Warji leʔy- ~ lāʔləʔa ~ leʔəleʔə “to hunt” [Skn. 1977, 27]<sup>31</sup>. From AA \*l-ḥ “to search for” [GT].

• **WRift \*sēḥ-ā** “tsetse-flies”, sg. \*sēḥ-imo [KM 247] ||| (???) Eg. šḥḥ.t “ein Insekt” (Med., Wb IV 210, 13).

• **WRift \*ṣuḥi**, pl. \*ṣuḥáy “deserted homestead” [KM 263]: perhaps related to Ar. šḥy: šaḥā “1. ouvrir, 2. être ouvert”, šaḥan “large, spacieux, ample” [BK I 1200].

• **WRift \*wārah-** “to pass” [KM 315] ||| Eg. w3ḥ [reg. < \*wrḥ] in: w3ḥ-š “vom Passieren des Wassers durch die Rinder” (OK, Wb I 257, 6) = “den See (š) Passieren (durch die Rinder)” (ÄWb I 304).

• **WRift \*yāḥáy** “shower of rain” [KM 328] ||| Sem.: Ar. ḥayā- ~ ḥayāʔ- “pluie” [Lane 681; BK I 523] ||| Eg. ḥw.t “rain” (PT, Wb III 49). For Eg.-Ar. see Ember 1916, 74; 1930, #14.a.15; Mlt. 1987, 105. Eg. ḥw.t “rain” is to be combined also with Eg. ḥwj “fließen” (since MK, Wb III 48).

### Irregular WRift \*ḥ < AA \*h

• **WRift \*ḥāla** “well, waterhole” [KM 147]: Ehr. has h- in the *Anlaut*, whereas KM consistently ḥ-, cf. Irq. ḥalan, pl. haleli “well” [Ehr.] = ḥalāngw ~ ḥala “well” [KM] etc. Its comparison with LECu.: Rendille ḥár “hollow where water collects in the rainy season” [PG 1999, 136] (suggested by KM)<sup>32</sup> seems less probable than that with ERift: Qwd. ḥali-to (sic, ḥ-/x-) “waterhole” [Kohl-Larsen] = \*hali-to [Ehr.] | Dh. helel-ād- “to flow” [Ehr.] (SCu.: Ehr. 1980, 306) < SCu. \*ḥ/h[a]l- “(to flow from a) well” [GT] ||| Sem.: Ar. hll V: tahallala “soudre, jaillir (se dit de l’eau d’une source, des larmes qui coulent des yeux)”, VII “1. être versé par torrents, se répandre (se dit de la pluie), 2. être baigné de larmes (se dit des yeux), 3. répandre, verser, faire tomber de la pluie”, hall-at- and halāl- “première pluie de la saison” [BK II 1434; DRS 415]<sup>33</sup> ||| (?) Eg. hn.w [reg. < \*hl-w] “Welle, Flut” (NE, Wb II 481, 10–11) ||| SBrb.: Ahaggar tā-hāla, Tawllemmet tāla [contracted < \*ta-hala] “small brook, spring” (SBrb.: Mlt. 1991, 165) ||| WCh.: perhaps Goemay hal “watery” [Srl. 1937, 74] < AA \*hal- “to flow (primarily of water from its source?)” [GT].

• **WRift \*ḥaw** “to be weak” [KM 150] ||| Sem. \*hwy: Tigre ḥawā “se faner” [DRS 386].

• **SCu. \*ḥaw-/\*ḥāy-** “husband” [GT] = \*ḥā- (sic) [Ehr.] > WRift \*ḥawata, pl. \*-ē “husband” [KM 150] | Dh. ḥāʔo [-ʔ- reg. < \*-y-], pl. ḥāʔi “husband” [Ehr.] = ḥāʔo, pl. ḥāʔi ~ ḥāʔōma “man,

<sup>30</sup> Cf. perhaps Eg. kḥw “keuchen (vom ermatteten Herzen)” (Med., Wb V 138, 5; GHWb 887) unless it is of onomatopoeic origin.

<sup>31</sup> For the semantic shift “to search for” > “to hunt”, cf., e.g., the entry for Eg. bḥs in EDE II 284.

<sup>32</sup> Cf. also Eg. ḥ3j [reg. < \*ḥly] “fließen, fluten vom Nil” (GR, Wb III 13), ḥ3j.t “bewässerte Stelle an einem Brunnen in der Oase” (NK, Wb III 13, 11).

<sup>33</sup> D. Cohen (DRS 415) listed this Ar. hll under Sem. \*hll III as a distinct root, carefully separating it from Sem. \*hll II with the basic sense “to appear on the horizon”.

husband, judge” [EEN 1989, 26] (SCu.: Ehr. 1980, 299, #2) || Bed. hiyi (comm.), hiy-ob/ot (acc.) “husband, wife” [Rpr.] ||| Eg. hj (written OK h, MK hj, NK h3j) “Gemahl, Gatte” > (SB) 𐩧𐩢𐩨 reflecting old [\*hǿ] (OK, Wb II 475, 10) < AA \*hiy- “husband” [GT], whereby SCu. \*-a/ā- is also irregular. For Bed.-Eg. see Ember 1917, 21; Cohen 1947, #92; IS 1971, 241, #100; Mlt. 1986, 72; OS 1992, 169; HSED #1174.

• **WRift \*ḥawi** “grass (sp.)” [KM 151] || LECu.: Afar hawāy (h-!) “herbs, spices” [PH 1985, 120].

• **WRift \*ḥay-ūt** “to travel” [KM 151]: cp. perhaps LECu.: Saho hay “get moving!” [Vergari 2003, 96].

• **WRift \*ḥōʔ**, pl. \*ḥoʔ “nice, good” [KM 154] > Irg. ḥoḥo “good, beautiful, fine”, Brg. ḥoʔinay “goodness”, Alg. ḥo “goodness” (WRift: Ehret 1980, 302, #28 with false reconstruction) ||| Sem. \*hyʔ [DRS]: Ar. hāʔa “être beau, bien fait, d’une forme élégante, de belle apparence”, haʔuʔa “être beau, bien fait, d’une forme élégante”, hayyiʔ- ~ hayiʔ- “beau, bien fait (homme)”, hayʔ-at- ~ hiʔ-at- “4./2.belle forme” [BK II 1463; Dozy 781; Fagnan 1923, 182] || Sqt. hiyoh “splendide (?)” [Lsl. 1938, 142] ||| NBrb.: Tamazight heyya “être bien, bon, satisfaisant, meilleur, de bonne qualité” [Tf. 1991, 227: not borrowed from Ar.]<sup>34</sup> (Sem.-Brb.: DRS 398).

### SCu. \*H (?)

There is a group of three Eg.-SCu. cognates, where on the SCu. side Ehret did not reconstruct either \*ʕ- or \*h-. SCu. \*H- is attested in the *Anlaut* only and can be isolated on the basis of a regular correspondence of WRift \*ʕ- ~ Dhl. h-. The Eg. cognates have ʕ-, but the SCu. reflexes are not evident: they do not fit within the regular reflexations of either SCu. \*ʕ- or \*h-. Henceforth, these SCu. reflexes point neither to \*ʕ or \*h. Therefore, I have marked this initial “laryngeal” with the symbol \*H-. It is unclear at the moment why we have a second set of correspondences for what we would normally expect to represent SCu. \*ʕ-. Sem. lexical evidence would be needed to decide whether this phenomenon has anything to do with the distinction of \*ʕ vs. \*ɣ in Sem.

SCu.	WRift	Qwadza	Asa	Ma’a	Dahalo	Sem.	Eg.	AA
*H	*ʕ-	ʕ-	?	?	h-	?	ʕ- ~ ḥ- (1×)	*ʕ/*ɣ (?)

• **SCu. \*Ham-** “to nibble” [Ehr.]: Qwd. am- “to chew” | Dhl. ham- “to toss a portion of food in the mouth” (SCu.: Ehret 1980, 297) || Bed. ʔām “gierig fressen, essen” [Rn.] || LECu.: Somali ʕūn [-n# < \*-m#] “to eat” [Abr.] ||| Eg. ʕm “verschlucken” (OK, Wb I 183–184) ||| WCh.: Pa’a ʔḡmma “to eat” [Jng. 1967, 198] || ECh. \*ʔaym- “to eat” [GT]: Somray ʔám | Sokoro áymè | Dangla ēme, Mokilko ʔóòmí, Migama ʔâyme etc. (ECh.: JI 1994 II, 119, 121) < AA \*ʕ-m “to eat” [GT].<sup>35</sup>

• **SCu. \*Har-/\*Hār-** “goat” [GT]: Irg. & Brg. & Alg. & Grw. ara [pl.] “goats” | Qwd. ali-to [\*-r-] “goat” | Dhl. héri ~ éri “goat, sheep” (SCu.: Ehret 1980, 297; 1974, 73; Flm. 1969, 29) ||| Eg. ʕr “goat” (Med., Wb I 208, 10) ||| WCh.: Kulere war “he-goat” [Jng.] || ECh.: Lele ōrē (pl.)

<sup>34</sup> The position of NBrb.: Qabyle e-hwu “plaire, être agréable” [Dlt. 1982, 295] is dubious (Dlt.: < Ar. hwʔ vs. DRS: ~ Sem. \*hyʔ).

<sup>35</sup> Lit. for this AA etymology: Mnh. 1912, 233; Behnk 1928, 139, #16 (Eg.-Somali); Zhl. 1932–1933, 166 (Eg.-Bed.); Chn. 1947, #60 (Eg.-Bed.-Somali); Djk. 1965, 44 (Eg.-Bed.-Somali); Dlg. 1970, 623, #74 (Eg.-Bed.); HSED #1064 (Eg.-Qwd.).





• **SCu.** \***ham**(ā[ḍ])- “to ripen” [GT: ultimately bicons.?] = \***hamāḍ** [Ehr.] > WRift \***hamāḍ** “to ripen” vs. \***hamana** [KM: < \*\***hamāḍ-ina** “ripeness” [KM 130] | Dhl. **hamāḍ-** “to ripen”, **hāmāḍe** “ripe” [EEN 1989, 14; Tosco 1991, 135] (SCu.: Ehr. 1980, 305, #7-#8) ||| Sem.: Tigre √**hmhm** “mûrir, porter de la sémence” [DRS 422]. Cognate to the following root?

• **WRift** \***ham-** “to be warm” [GT] = \***hām-** “to be(come) hot” > \***ham-é**, \***hāmi**, and \***hamēriya** “heat”, \***ham-is** (caus.) “to cook, burn, set fire to” [KM 130–1] (WRift: also Ehret 1980, 299) ||| Eg. **hm** “heiß sein, brennen” (XVIII. Mag., Wb II 489, 15) < AA \***h-m** “to be warm, hot” [GT]. For Eg.-Irq. see HSED #1285 (with false reconstructions). Cf. WRift \***hum-im** “to cook” [KM 144]?

• **WRift** \***hamaĉ** “to stir” [KM 131]: cf. Ar. **hamaša** I “1. réunir, ramasser, rassembler”, VIII “2. se remuer, s’agiter avec bruit, se mêlant les uns aux autres (se dit du peuple rassemblé en foule)” [BK II 1447].

• **WRift** \***haraʔu** “dew” [KM 133] ||| Sem. \***hrʔ** > Tigre **harʔa** (intr.) “tremper dans l’eau, être très humide, s’amollir” [DRS 447]. Cf. also Sem. \***hrr** > Ar. **hurr-** “grande quantité d’eau ou de lait” etc. [DRS 459].

• **WRift** \***har-ā** “at once, at first” [KM 132] || ECU. \***hor-** to be in front” [Sasse 1979, 62].

• **WRift** \***hawʔ-ut** (med.) “to go away” [KM 134] || LECu.: Saho **haw-eṛḥe** “to go away”, **haw-iše** “to take away” [Vergari 2003, 95] ||| Eg. **hwhw** “davonlaufen (von den Füßen)” (MK, Wb II 485, 1).

• **WRift** \***hay** (exclam.) “excuse me, sorry for you” [KM 134] ||| Sem. \***hwy**: Hbr. **hōy**, Syr. **hāwāy** “ah! malheur!” | Ar. **haway-ka** “gare à toi!” etc. (Sem.: DRS 386).

• **WRift** \***hē** “this” [KM 135] = SCu. \***ha** “this”, cf. \***ho** “this referred to”, \***hū** “that” [Ehret 1980, 304, #1 & 309, #39 & 310, #46]: this is a common AA deictic morpheme, cf. ECU.: e.g. Gollango **há-ma** “1. dort(hin), 2. hierher” [AMS 1980, 201] ||| Sem. \***hā** “demonstrative element” [GT] (Sem.: DRS 346) ||| WCh.: Angas-Sura \***ha** “this” [GT 2004, 151], etc.

• **Dhl.** **hēl-** “to seize, catch hold of” [Ehret 1980, 307] = **hēl-** “to hold, seize” [EEN 1989, 15] || LECu.: Sam \***hel-** “to get” [Heine]: Somali **hel-**, Rendille **hel-**, Boni **hel-** (Heine 1978, 63; Sasse 1979, 40) ||| Eg. **h3j** [< \***hly**] “(den Gegner) annehmen” (MK, Wb II 475, 1) ||| WCh.: Angas eel ~ el (hill) “to seize, snatch” [Flk. 1915, 175] = **ʔel** (P) “ergreifen, schnappen” [Jng. 1962 MS] || CCh.: Gisiga **hal** “nehmen (Hirse, Erdnüsse)” [Lks. 1970, 123] < AA \***h-l** “to seize” [GT].

• **WRift** \***hēleʔ-es** (caus.) “to winnow, sift grains by tossing” [KM 135]: cp. Sem.: Tigre √**hyll**: **helälä** “jeter”, **helälib** ~ **helälit wāda** “jeter rapidement l’un après l’autre” [DRS 402].

• **WRift** \***hīmp** “to breathe” [KM 137] ||| Eth.-Sem.: Tigre **hanfəʔa** “haleter, reprendre haleine, se reposer”, cf. **hanfofəyät** “forte tempête” [DRS] ||| SBrb.: Ahaggar **hənəff-ət** “être essoufflé; pousser des petits gémissements plaintifs” [Prs. 1969, 60, #307]. Tigre-Ahaggar: DRS 435.

• **SCu.** \***hir-** “to sew” [GT] = \***hēr-** [Ehr.] > WRift \***hir-it** “to sew” [KM 142] | Ma’a -**hi** “to sew” [Ehr. 1980, 307, #28] (but Dhl. **has huḍ-** “to sew” [Ehr. 1980, 310, #48 with false etymology!]: cp. perhaps Sem. \***hyr**: EDathina **hayyar** “préparer, mettre en ordre”, Palest Ar. **hayyar**: “effiloche la laine pour faire le feutre” (Sem.: DRS 405).

• **SCu.** \***hoʔ-** “to divine” [Ehr.] > Brg. **hoʔ-** “to curse, heal” | Asa **hoʔ-** “to divine” | Ma’a -**hoʔi** “to divine” (SCu.: Ehr. 1980, 310, #45) ||| (?) Ar. **hwʔ**: **hāʔa** “être inquiet, troublé comme qqn. qui a peur, 2. être agile, vif, 3. attaquer qqn.”, **hūʔ-** “inimitié, hostilité”, **mi-hwāʔ-** ~ **ma-hwāʔ-** “tumulte, vacarme des combats” [BK II 1458; DRS 390: isolated in Sem.].

• **SCu.** \***hoʔi-** “to spill” [Ehr.] > Asa **hoʔ-** “to pour” | Ma’a -**hóʔo** ~ -**hoʔomi** “to vomit”, -**hoʔoti** “to purge” (SCu.: Ehr. 1980, 310, #44) ||| Sem. \***h(y)ʔ** “to flow” [Dlğ.] > esp. OSA (Sabaeen) **hyʔ** “1. couler, s’écouler (eau), 2. aller, s’étendre (limites), 3. offrir un sacrifice”,

mhyʕ “libation (?)” [SD 57], Ar. hʕʕ I: haʕʕa “vomir”, hwʕ I: hāʕa “4. avoir un vomissement”, II “causer un vomissement à qqn., faire vomir”, V “s’aider pour vomir, se faire vomir”, hawʕ- ~ huwāʕ- ~ hayʕūʕ-at- “vomissement”, hyʕ I: hāʕa “3. couler et se répandre, 10. vomir” [BK II 1428, 1458, 1467–8] (Sem.: DRS 390, 404, 439).

• **SCu. \*hopi** “stalk” [Ehret] > WRift \*hopi-na “handle” [KM 142]: Brg. hopi-na “handle, haft” | (?) Ma’a ihópi “leaf” (SCu.: Ehret 1980, 309) || LECu.: Oromo hof-a “stalk, stem” [Ehret] = hof-ā “1. kind of spear for throwing, made of sharpened wood only; 2. stalk (millet or corn; looks like sugar cane)” [Gragg 1982, 211] = hoff-ā “stalk” [Bitima 2000, 146] (SCu.-Oromo: Ehret 1980, 309) ||| Eg. hpw.tj “Art Gestell (auf dem die göttlichen Vögel zu hocken pflegen)” (PT, Wb II 489, 8) = “hpwty-pole” (Faulkner, AEPT 259) = “Kletterbaum (auf dem die göttlichen Vögel zu sitzen pflegen)” (GHWb 492; ÄWb I 749) < AA \*hVp- “stalk” [GT].

• **Dhl. holló** “1. with, 2. and” [EEN 1989, 16; Tosco 1991, 135] ||| Sem. \*hl > Hrs. hel, helt “with, in the company of, at”, em-hel “from amongst”, Mhr. hāl “to, with, in the company of, in(to) the presence of, at” (MSA: Jns. 1977, 51; 1987, 155).

• **PRift \*hō-t** “to dwell, live at, in” [KM 143] < \*\*haw-at (refl.) [GT] < SCu. \*haw- “to stay” [Ehret 1980, 381] || Bed. hāy “sein, existieren”, hāy “lebendig” [Rn. 1895, 132] = hai(y) “to be (on the first of the new moon)”, hāy-am (refl.) “to appear (heavenly body only, esp. new moon)” [Rpr. 1928, 200] || LECu.: Somali hay “sein, esse” [Rn. 1902, 211], Rendille -hai “to be” [Zbr. 1975, 72] ||| Sem. \*hwy “to be(come)” [GT] > Akk. ewû, hence emû “devenir” || Hbr. hāyā, BARAM. & Syr. hawā “être, devenir, survenir” etc. (Sem.: DRS 386) ||| SBrb. \*ʕh₂-h₁-h₁ “être dans” [Prs.] = \*i-h[ā] (pf. int.) [GT]: Ahaggar i-hā, Taneslemt & Tawllemmet & Ayr & Ghat i-ha (Prs. 1969, 43, #119). Cf. also the isogloss (AA \*y-w “to be” [GT], var. to AA \*h-w?) of Eg. jw [irreg. < \*hjw?] “sein” (OK, Wb I 42–43) ||| WCh.: Daffo-Butura yò “sein, to be” [Jng. 1970, 223]. For the AA etymology see Ember 1911, 89; 1930, §4.g.2 (Eg.-Sem.); Behnk 1928, 138, #5 (Eg.-Sem.-Bed.); Cohen 1947, #94 (Sem.-Eg.-Brb.-Cu.); Mlt. 1986, 72, #3.a.2.a (Sem.-Cu.); Blz. 1989, 217 (id. + Rendille); Ctv. 1988, 76 (Eg.-Sem.).

• **PRift \*hu?** “to fall” [GT] = PRift \*hu- “to fall” [Ehret 1980, 381] > WRift \*hū? “to fall” [KM 143]: Irg. hū- [Dlg.] = hu- [Ehret], Grw. & Alg. hū- [Dlg.] = hu- [Ehret], Brg. hu- [Ehret] | Asa hu?- [Ehret] (SCu.: Ehret 1980, 381; Dlg. 1973, 243) ||| Sem. \*hww/y “to fall” [DRS 385] ||| SBrb.: Ahaggar i-hi “tomber dans/sur” [Prs. 1969, 43, #120] < AA \*h-w “to fall” [GT].<sup>38</sup>

• **WRift \*hub** “to carry, bring” [KM 143] > i.a. Iraqw & Gorowa hūw “to carry, bring”, hūwa (f) “load”, hūwā “burden, luggage” || (?) Bed. hawīye [-w- reg. < \*-b-] “pack-saddle” [Rpr. 1928, 200] ||| Sem. \*whb “to give” [GT] > i.a. Hbr. yəhab “fardeau”, cf. Ar. ʔuhb-at- “équipement” (Sem.: DRS 508) ||| WCh.: Pa’a hwà-hwa (redupl., m), pl. hwà-hwí “load” [MSkn. 1979, 180] (Pa’a hwV- weakened < NBch. \*hbV-?) < AA \*h-b (orig. \*hub?) “to bring, carry load” [GT].

• **PIrg. \*hučay** “big earthen pot” [KM 145]: cp. Sem.: Palest. Ar. hišš “poterie” [DRS 463] ||| Bed. hayiš “to make pottery-ware” [Rpr. 1928, 201].

• **WRift \*hūnguruʔu** “central pillars of the house” [KM 145]: cp. Sem. \*hngr: Tigre han-gora “lance de bois (jouet d’enfant)” [DRS 431].

• **WRift \*hur-im** (dur.) “to cook” [KM 145] cannot be compared with LECu.: Somali kar “to become boiled” as KM did (WRift \*h- ≠ Somali k-), for phonological reasons (WRift \*h- ≠ ECU. \*k-). Instead, cf. ECU. \*hUr- “to kindle” [GT] > PSam \*huri “schüren (Feuer)” [Heine 1977, 292] = “to kindle” [Heine 1978, 63] | Dullay \*hōr- [GT] > Dobase hor-as- (caus.) “kochen”,

<sup>38</sup> Etymological lit. for this AA widespread AA root: Erman 1892, 114; Ember 1913, 116, #53; 1930, #3.e.2; Alb. 1927, 224; Clc. 1936, #21; Chn. 1947, #93; Vcl. 1958, 374; 1990, 49; Ward 1969, 265, fn. 5; Mlt.-Stl. 1990, 66; OS 1992, 200; HSED #1163.

Gollango hōr- “wärmen” (Dullay: AMS 1980, 162, 202) ||| Sem.: Ar. haraʔa I “4. cuire trop les viandes, au point qu’elles soient en charpie”, hariʔa “être en charpie pour avoir été trop cuit (se dit des viandes)” [BK II 1409] || ES: Tna. harhar bälä “être en flammes” [DRS 451] ||| Eg. h3 [reg. < \*hr] “rösten (in den Bez. für Brotsorten)” (NE, Wb II 475, 9) = “Glut”,  $\zeta q n h3$  “Röstbrot” (GHWb 486) < AA \*h-r “to cook (?)” [GT]. Any connection to ES: Tigre hwr “ripe” [DRS]?

• **SCu. \*loh-** “to carry load” [GT]: WRift \*lōh- “to migrate, carry goods”, \*lōh-is “to carry, move house” [KM 195] > Irq. loh- “to move hose”, loh-is- “to carry load”, Alg. loh-is- “to carry load” | Qwd. loh-is- “to move house” (SCu.: Ehret 1980, 206) ||| Eg. jwh [ < \*lwh] “beladen mit etwas” (Westcar, Wb I 56, 17–19) ||| WCh.: Suroid \*lē “load” [GT]: Sura léε “Last” [Jng. 1963, 72], Mupun léε “load” [Frj. 1991, 33] < AA \*l-[w]-h “to carry load” [GT].

• **SCu. \*wah-** “to drink” [Ehret] > WRift \*wah- “to drink” [KM 313] | Qwd.-Asa wa-t (continuative suffix -t-) “to drink” [Ehret] | Ma’a -wáha “to drink” [Ehret] (SCu.: Ehret 1980, 313, #19) ||| ES \*wh- > Geez wahəha “devenir aqueux”, Amh. wuha “eau”, wāhayyā “devenir aqueux” (ES: DRS 507, 509). For the semantic shift cf. e.g. IE \*ak<sup>w</sup>- > Latin aqua “water” ~ Tokharic AB yok-tsi “to drink”, Hittite ekuzi “(s)he drinks” (IEW 23).

### Irregular WRift \*h < AA \*ḥ

• **WRift \*dabāh** “to search, look for” [KM 80] ||| Eg. dbḥ “bedürfen, (er)bitten” (PT, Wb V 439–440). This new etymology discards all previous suggestions for the affiliation of our Eg. root<sup>39</sup>.

• **WRift \*habaĉ** “to spread, open mouth” [KM 129] ||| Eg. ḥbḍ [reg. < AA \*ḥbĉ] “vom Öffnen des Mundes” (CT, BD, Wb III 67, 8) = “to open the mouth” (CT, DCT 327) = “öffnen (Mund, um Nahrung aufzunehmen), aufsperrn” (CT, ÄWb II 1651)<sup>40</sup>. Note that Eg. ḥpd “öffnen (nur vom Mund gebraucht)” (BD, Wb III 72, 12) represents a secondary variety (variation of -b- ~ -p-, shift of -ḍ > -d<sup>41</sup>). Any connection to Ar. ḥabaĉa I “2. tirer à soi avec force la corde de l’arc, et la lâcher, en sorte qu’elle produise un son” [BK I 368] < common AA \*ḥ-b-ĉ “to spread out (?)” [GT]?

• **PRift \*hac-** “to be full” [GT] = \*haĉ- [Ehret 1980, 81, #2] > WRift \*hac “to be full” [KM 134] | ERift \*haĉ- “to be full” [GT] > Asa haš- “to be full” [Ehret], Qwd. hacumo “much” [Ehret]: a doubly irreg. cognate appears in Ar. ḥašaʔa “remplir, bourrer, farcir de qqch.”, ḥašiʔa “être rempli” (de qqch.)” [BK I 435]. Note that Rift \*h- ≠ Sem. \*ḥ-, neither Rift \*c corresponds to Sem. \*ṣ.

• **WRift \*haĉ** “to stick, fasten, pin” > \*haĉ-im “to set a trap”, \*haĉ-ir-ō “trap” [KM 133–4]: its comparison with ECu. \*ḥiḍ- “to tie” [Sasse 1979, 36, 59] (not “to fasten”, suggested by KM) is unconvincing. Instead, one might be inclined to compare it with ECu. \*haš- “to hold, keep” [Sasse 1979, 39, 62], but the reconstruction of this latter form is not well-founded (the reflexes listed by Sasse belong in fact to distinct AA roots). Cf. perhaps Ar. ḥašša I “8. traquer, cerner, entourer de tous côtés et ne pas laisser d’issue (augibier, à la bête)” [BK I 430]?

<sup>39</sup> Cf. (1) Rn. 1890, 21, n. 10; Müller 1907, 301, n. 1; Clc. 1936, #458; Vcl. 1958, 373; IS 1964, 6, #19; Hodge 1981, 406; SISAJa III, #59, Mlt.-Stl. 1990, 58: ~ Sem. \*-ḍbaḥ- “to sacrifice” [Frz.]; (2) Rsl. 1971, 290: ~ Akk.  $\sqrt{\text{ṣb}}_{3/4}$  > ṣa/ebū “wünschen” [AHW 1073]; (3) Dlg. 1966, 87, #8.16; 1973, 230; 1983, 131 (followed by Blz. 1989, 210): ~ ECu. \*ḍēb- “(to be) thirst(y)” [Sasse 1982, 62].

<sup>40</sup> An alternative comparison with Sem. \*hbṣ > Ar. habaša “frapper” [BK II 1381; DRS 363] would be semantically much weaker.

<sup>41</sup> For these changes cf. e.g. Ward 1975, 64f.; Peust 1999, 135, §3.12.4 and Vycichl 1990, 46.



● **SCu. \*hap-** “to put up, put above, put on top of” [Ehr.] > Irq. hape “roof” | Ma’a -hapári “to cover” | Dhl. hap- “to suspend” (SCu.: Ehr. 1980, 304, #4) ||| Sem. \*h̥p “to hide, shelter” [Zbr.]: i.a. Hbr. h̥pp qal (with ʕal) “to shield” [KB 339], MHbr. h̥pp polel “to bend over, cover” [Jastrow 1950, 492] | Ar. ḥaffa “7. envelopper de qqch., 9. couvrir de tous côtés d’un voile, d’un rideau etc.” [BK I 455] (Sem.: Zbr. 1971, #91) ||| SBrb.: Ahaggar t-êf-it [\*ta-yaf- < \*ta-Haf-] “linceul” [Fcd. 1951–2, 299].

### SCu. \*ʔ

SCu.	WRift	Qwadza	Asa	Ma’a	Dahalo	Sem.	Eg.	AA
*ʔ	*ʔ	∅-, -ʔ-	∅/ʔ-, -ʔ-	ʔ-, -ʔ-, -∅	∅/ʔ-, -w-(?)	*ʔ	j-, -3-	*ʔ

● **WRift \*ʔah-as** (caus.) “to hear” [KM 64]: cognate to PAgaw \*ʔənq̄w- “ear” [Apl. 1989 MS, 4] || LECu.: Saho okk-ā [Rn.] = ʔok̄kw-ā (sic, ʔ-) [CR] = okk-ā “ear” [Zhl.] = okk-a [Bnd.] (Cu.: Crl. 1938 II, 213; Dlg. 1967, 5; 1973, 183; Zbr. 1989, 580, #21) ||| Eg. ʔnh.wj (dual, epenthetic -n-as in Agaw?) [ʔnh < \*jn̄h reg.] “die zwei Ohren” (MK, Wb I 204–5). For Cu.-Eg. cf. Zhl. 1932–33, 166.

● **WRift \*ʔako** “old man”: Irq. & Grw. áako “old man” | Asa akuya “grandparent (in address)” (Ehret 1980, 377) || ECU. \*ʔakāk-/\*ʔākk- “old man, grandfather” [Sasse 1982, 21, 24] (Cu.: Flm. 1969, 11; Apl. 1994, 3) ||| Eg. jk “alt, bejahrt” (BD, Wb I 34, 4).<sup>42</sup>

● **NWRift \*ʔāk-ō** “old man” [KM 51] = PRift \*ʔak-o “old man” [Ehret 1980, 377]: related to ECU. \*ʔakāk-/\*ʔākk- “old man, grandfather” [Sasse 1982, 21, 24] (Cu.: Flm. 1969, 11; Apl. 1994, 3) ||| Eg. jk “alt, bejahrt” (BD, Wb I 34, 4). For AA \*ʔ- ~ Eg. j-, cf. EDE I 81–83.

● **SCu. \*ʔal-** “girl” [Ehr.]: Qwd. elanko, Asa eleto | Ma’a mʔaléta (SCu.: Ehret 1980, 284) ||| (?) Eg. jr.wt [reg. < \*ʔl-wt] (pl.) “Bez. für Frauen” (PT, Wb I 114, 15).

● **SCu. \*ʔani** “I” [Ehret 1980, 283] ||| Bed. ane “I” [Rpr.] || Agaw: Awngi an “I” [Hetzron] || ECU. \*ʔani/\*ʔanu “I” [Sasse 1982, 26] ||| SOm. \*in-ta “I” [Flm. 1976, 315] ||| Eg. jnk “I” (OK, Wb I 101) ||| Sem. \*ʔan- “I” [GT], etc. A well known AA root with abundant literature (cf. EDE I 82).

● **SCu. \*ʔar-** “old (of things)”: Irq. ar-ʔad [Ehret: \*ʔar-ʔar-at- contracted] | Ma’a -ʔeraru | Dhl. ʔār- [irreg. \*y- < \*ʔ-] “to grow old” (SCu.: Ehret 1980, 287) ||| Eg. j3wj [\*ʔrwy] “to be old” (OK, Wb I 28, 13) ||| CCh.: Mwulyen iʔiri, Bachama ʔiyrey “old” (CCh.: OS) < AA \*ʔ-r “to be old” [GT]. For CCh.-Eg. see HSED #84.

● **WRift \*ʔaša** “fire” [KM 61] ||| Sem. \*ʔišš-(at)- “feu” [DRS 35]: irregular correspondence. Normally, WRift \*ʔs̄ = Sem. \*ʔs̄.

● **SCu. \*ʔaši** [\*-hl-] “fat, oil” [GT]: Qwd. aši-to | Dhl. áši (SCu.: Ehret 1980, 285) ||| (?) Eg. js “sort of oil” (OK, Wb I 130, 15; cf. GHWb 101). Eg. -s < AA \*-š̄ in final position? Otherwise, Eg. s ~ SCu. \*-š̄ are not regular.

● **SCu. \*ʔila** “eye” [Ehret]: Irq.-Brg.-Alg. ila | Qwd. ili-to, Asa ʔila-t | Ma’a iʔilá | Dhl. ʔila (SCu.: Ehret 1980, 291) || ECU. \*ʔil- “eye” [Sasse 1979, 5, 22; 1982, 104] || Agaw: Hamir əl, Qemant [y]əl, Bilin ʔəl [irreg. ʔ- < \*ʔ-] “eye” | Awngi əll “eye” (Agaw: Apl. 1991, 20) ||| Eg. jr.t [< \*ʔl-t?] “eye” (OK, Wb I 106–7) ||| CCh. \*ʔil- “eye” [GT]: Lamang (Hitkala) iri & (Waga dial.) ili [Meek] = ílí [Lks.], Vizik ili, iri [Wolff] | Alataghwa il-yia [Wolff] | Buduma yíl [Lks.] = yèl

<sup>42</sup> The Eg. root is to be read jk, BD -3- being purely orthographical (influenced by an association to Eg. j3w “old?”), cf. the “old man” det. of MK jk “quarryman” and jk.w “Steinbruch” (Wb I 39).

[Cyffer], Gulfei el [Wolff], Logone (pl.) ?àl [Bouny] = al (pl.) [Nct./Lks.] (CCh.: Wolff 1971, 65; JI 1994 II, 126–127) < AA \*ʔil “eye” [GT].<sup>43</sup>

• **PRift \*kʷalaʔ-** (sg. stem), **\*kʷaʔal-** (pl. stem) “widow” [GT]: WRift \*kʷaʔalaʔō, pl. \*kʷaʔēli “1. widow, 2. poor or unmarried woman” [KM 186] > Iraqw kʷálôʔo, pl. kʷáʔēli, Gorowa kʷálôʔo, pl. kʷáʔēli, Alagwa kʷálôʔo, pl. kʷáʔēli ~ kʷaʔal-itoʔo, Burunge kʷálôʔo, pl. kʷáʔēli ~ kʷaʔal-itoʔo | Asa kalaʔ-ayi “widow” (SCu.: Wtl. 1958, 26, #116; Zbr. 1978, 375) ||| Eg. ḥ3r.t [reg. < \*ḥʔl-t] “widow” (MK, Wb III 363, 4) < AA \*qʷ-ʔ-l “widow” [GT].

• **SCu. \*roʔ-** “to go away” [GT]: Ma’a -ro “to leave” [Ehret] | Dhl. roʔ- “to go/pass by” [Ehret] = “to go” [EEN 1980, 40] (SCu.: Ehret 1980, 220) || ECu.: Yaaku -rɛʔɛ “to run away” [Heine 1975, 132] = reʔ- “to run away” [Tosco] (Yaaku-Dhl.: Blz.-Tosco 1994 MS, 3) ||| Eg. rwj [-j reg. < \*-ʔ] “fortgehen, verlassen” (OK, Wb II 406–7) = “to go/pass away, depart, leave” (FD 147)<sup>44</sup> < AA \*r-w-ʔ “to leave” [GT].

## Summary

In the light of the lexical parallels examined above, we can safely maintain that West Rift (and, henceforth, South Cushitic) \*ḥ, \*h, \*ʕ, \*h, and \*ʔ represent a very archaic segment of the phoneme inventory, having retained presumably the same articulation they had in Afro-Asiatic.

## Abbreviations of languages

AA: Afro-Asiatic, Akk.: Akkadian, Alg.: Alagwa, Ar.: Arabic, Aram.: Aramaic, (B): Bohairic, BD: Book of the Dead, Brb.: Berber, Brg.: Burunge, C: Central, Ch.: Chadic, Cpt.: Coptic, CT: Coffin Texts, Cu.: Cushitic, Dhl.: Dahalo, E: East(ern), Eg.: Egyptian, ES: Ethio-Semitic, GR: Greek and Roman period, Grw.: Gorowa, Hbr.: Hebrew, Hrs.: Har-susi, Irq.: Iraqw, Jbl.: Jibbali, Mag.: Magical Texts, Med.: Medical Texts, Mhr.: Mehri, MK: Middle Kingdom, MSA: Modern South Arabian, N: North(ern), NE(g.): New Egyptian, NK: New Kingdom, OK: Old Kingdom, Om.: Omotic, OSA: Old South Arabian, PT: Pyramid Texts, Qwd.: Qwadza, (S): Sahidic, S: South(ern), Sab.: Sabaeen, Sem.: Semitic, Sqt.: Soqotri, Ug.: Ugaritic, W: West(ern), XVIII.: 18<sup>th</sup> Dynasty.

## Abbreviations of author names

Abr.: Abraham, Alb.: Albright, AMS: Amborn & Minker & Sasse, Apl.: Appleyard, BK: Biberstein Kazimirsky, Blv.: Belova, Blz.: Blažek, Bnd.: Bender, Brt.: Barreteau, Clc.: Calice, CR: Conti Rossini, Crl.: Cerulli, Dkl.: Diyakal, Dlg.: Dolgopolsky, Dlt.: Dallet, EEN: Ehret & Elderkin & Nurse, Ehr.: Ehret, Eld.: Elderkin, Flm.: Fleming, Frj.: Frajzyngier, Frz.: Fronzaroli, Ftp.: Fitzpatrick, GT: Takács, Hds.: Hudson, Hfm.: Hoffmann, IL: Institute of Linguistics, IS: Illič-Svityč, JI: Jungraithmayr & Ibrizimow, Jng.: Jungraithmayr, Jns.: Johnstone, JS: Jungraithmayr & Shimizu, KB: Koehler & Baumgartner, KM: Kießling & Mous, Krf.: Kraft, Kwh.: Kleinewillinghöfer, Lks.: Lukas, Lmb.: Lamberti, Lsl.: Leslau, Mgw.: Maghway, Mkr.: Mukarovsky, Mlt.: Militarev, Mnh.: Meinhof, MQK: Mous & Qorro & Kießling, Ntg.: Netting, Nwm.: Newman, OS: Orel & Stolbova, PAM: Prasse & Alojaly & Mohamed, PG: Pillinger & Galboran, PH: Parker & Hayward, Prh.: Porhomovsky, Prs.: Prasse, RB: Rapp & Benzing, RK: Reutt & Kogan, Rn.: Reinisch, Rns.: Renisio, Rpr.: Roper, Rsl.: Rössler, Skn.: Skinner, Srl.: Sirlinger, Ss.: Sasse, Stl.: Stolbova, Stz.: Satzinger, Tf.: Taïfi, Vcl.: Vycichl, Vrg.: Vergote, WP: Weibegue & Palayer, Wtl.: Whiteley, Zbr.: Zaborski, Zhl.: Zyhlarz.

<sup>43</sup> The literature on this widespread AA etymology has been most recently surveyed in EDE I 83.

<sup>44</sup> The comparison of Eg. rwj with Brb. \*r-w-l “to flee” (Möller 1921, 196; 1924, 42; Hintze 1951, 84, #429) is not acceptable for semantical reasons. Besides, O. Rössler (1952, 131, #8) correctly identified Brb. \*r-w-l with Ar. hrwl: harwala “schnell gehen”.

## Literature

- ABRAHAM, R. C.: *Dictionary of the Hausa Language*.<sup>2</sup> London, 1962., University of London Press.
- ABRAHAM, R. C.: *Somali-English Dictionary*.<sup>2</sup> London, 1964., University of London Press Ltd.
- AÄG — EDEL, E.: *Altägyptische Grammatik*. Roma, 1955., Pontificium Institutum Biblicum.
- AECT — FAULKNER, R. O.: *The Ancient Egyptian Coffin Texts*. Vol. I–III. Warminster, 1973–8., Aris & Phillips Ltd.
- AEPT — FAULKNER, R. O.: *The Ancient Egyptian Pyramid Texts*. I. Oxford, 1969., Clarendon Press.
- AHW — SODEN, W. von: *Akkadisches Handwörterbuch*. I–III. Wiesbaden, 1965–1981., Otto Harrassowitz.
- ALBRIGHT, W. F.: Notes on Egypto-Semitic Etymology. II // *American Journal of Semitic Languages and Literatures* 34/4 (1918), 215–255.
- ALBRIGHT, W. F.: The Principles of Egyptian Phonological Development // *Recueil de Travaux Relatifs à la Philologie et à l'Archéologie Égyptiennes et Assyriennes* 40 (1923), 64–70.
- AL I–III — MEEKS, D.: *Année lexicographique. Égypte ancienne*. Tome 1–3 (1977–1979). 2<sup>ème</sup> édition. Paris, 1998., Cybele.
- ALC 1978 — Angas Language Committee (in Cooperation with Nigeria Bible Translation Trust): *Shàk nkarij kè shàtok mwa nḍan Ngas. Ngas–Hausa–English Dictionary with Appendix Showing Some Features of Ngas Grammar*. Jos, Nigeria, 1978., Nigeria Bible Translation Trust.
- ALLEN, T.G.: *The Inflection of the Verb in the Pyramid Texts*. Bibliotheca Aegyptia, vol. 2, fasc. 1–2. Malibu, 1984., Undena Publications.
- AMS — AMBORN, H. & MINKER, G. & SASSE, H.-J.: *Das Dullay. Materialien zu einer ostkuschitischen Sprachgruppe*. Berlin, 1980., Reimer Verlag.
- APPLEYARD, D.: The Internal Classification of the Agaw Languages. A Comparative and Historical Phonology // BYNON, J. (ed.): *Current Progress in Afro-Asiatic Linguistics*. Amsterdam, Philadelphia, 1984., John Benjamins. Pp. 33–67.
- APPLEYARD, D.: *Agaw and Omotic Links. The Evidence of the Lexicon*. MS. Paper presented at the 2<sup>nd</sup> International Symposium on Cushitic and Omotic Languages, Torino, November 1989. Proceedings forthcoming.
- APPLEYARD, D.: *Agaw Vocabulary Comparative Notes*. MS. London, 1989. 24 p.
- APPLEYARD, D.: *Preparing a Comparative Dictionary of Agaw*. MS. Presented at the III. Kuschitisten- und Omotistenkongreß (Berlin, March 1994). 4 p.
- APPLEYARD, D.: *A Comparative Dictionary of the Agaw Languages*. Köln, 2006., Rüdiger Köppe Verlag.
- ARBACH, M.: *Le maḍābīen: lexique — onomastique et grammaire d'une langue de l'Arabie méridionale préislamique*. Tome I. Lexique maḍābīen comparée aux lexiques sabéen, qatabānite et ḥaḍramawtique. Thèse de doctorat. Université de Provence Aix Marseille I, Centre d'Aix, septembre 1993. 128 p.
- ÄWb I — HANNIG, R.: *Ägyptisches Wörterbuch I. Altes Reich und Erste Zwischenzeit*. Hannig-Lexica 4. Mainz am Rhein, 2003., Verlag Philipp von Zabern.
- ÄWb II — HANNIG, R.: *Ägyptisches Wörterbuch II. Mittleres Reich und Zweite Zwischenzeit*. I–II. Hannig-Lexica 5. Mainz am Rhein, 2006., Verlag Philipp von Zabern.
- BARRETEAU, D.: *Description du mofu-gudur. Langue de la famille tchadique parlée au Cameroun*. Livre II. Lexique. Paris, 1988., Éditions de l'ORSTOM.
- BARRETEAU, D. & JUNGRAITHMAYR, H.: Les verbes monoradicaux dans les langues tchadiques // JUNGRAITHMAYR, H. & TOURNEUX, H. (eds.): *Études tchadiques. Verbes monoradicaux suivis d'une note sur la négation en haoussa*. Actes de la XII<sup>ème</sup> réunion de Groupe d'Études Tchadiques LACITO-CNRS-PARIS. Paris, 1990., Librairie Orientaliste Paul Geuthner. Pp. 37–214.
- BELOVA, A. G.: Refleksy semitskih sibiljantov v drevneegipetskom // *Meroé* 4 (1989), 9–21.
- BELOVA, A. G.: Struktura semitskogo kornja i semitskaja morfoložičeskaja sistema // *Voprosy Jazykoznanija* 1 (1991), 79–90.
- BELOVA, A. G.: Sur la reconstruction du vocalisme afroasiatique: quelques correspondances égypto-sémitiques // MUKAROVSKY, H. G. (ed.): *Proceedings of the Fifth International Hamito-Semitic Congress*. Band II. Wien, 1991., Afro-Pub. Pp. 85–93.
- BELOVA, A. G.: K voprosu o rekonstrukcii semitskogo kornevogo vokalizma // *Voprosy Jazykoznanija* 6 (1993), 28–56.
- BENDER, M. L.: Proto-Omotic Phonology and Lexicon // BECHHAUS-GERST, M.; SERZISKO, F. (eds.): *Cushitic-Omotic. Papers from the First International Symposium on Cushitic and Omotic Languages*, Cologne, January 6–9, 1986. Hamburg, 1988., Helmut Buske Verlag. Pp. 121–159.

- BENDER, M. L.: *The Mystery Languages of Ethiopia* // MARCUS, H. (ed.): *New Trends in Ethiopian Studies*. Vol. 1. Lawrenceville, 1994., Red Sea Press. Pp. 1153–1174.
- BENDER, M. L.: *Omotic Lexicon and Phonology*. Carbondale, 2003., SIU Printing/Duplicating.
- BIBERSTEIN KAZIMIRSKI, A. de: *Dictionnaire arabe-français*. Paris, 1860., Maisonneuve & Co. Editeurs. (Reprinted in Beyrouth, no date, Librairie du Liban.)
- BITIMA, T.: *A Dictionary of Oromo Technical Terms. Oromo-English*. Köln, 2000., Rüdiger Köppe Verlag.
- BLACK, P. D.: *Lowland East Cushitic: Subgrouping and Reconstruction*. Ph.D. dissertation. 1974., Yale University.
- BLAŽEK, V.: *A New Contribution to Comparative-Historical Afrasian Linguistics* // *Asian and African Studies* 24 (1989), 203–222.
- BLAŽEK, V. & TOSCO, M.: *Between South and East Cushitic: Reconsidering the Position of Dahalo*. MS. Handout for the XI. Afrikanistentag, Köln, 19–21 September 1994. 6 p.
- CALICE, F. von: *Grundlagen der ägyptisch-semitischen Wortvergleichung*. Wien, 1936., Selbstverlag des Orientalischen Institutes der Universität Wien.
- CD – CRUM, W. E.: *A Coptic Dictionary*. Oxford, 1939., Oxford, 1939., Oxford University Press.
- CERULLI, E.: *Studi etiopici. II. La lingua e la storia dei Sidamo*. Roma, 1938., Istituto per l'Oriente.
- CONTI ROSSINI, C.: *Schizzo del dialetto saho dell'alta Assaorta in Eritrea*. Roma, 1913., Tipografia della R. Accademia dei Lincei.
- CT – BUCK, A. de: *The Egyptian Coffin Texts*. Vol. I–VII. Chicago, 1935–61., The University of Chicago Press.
- DALLET, J.-M.: *Dictionnaire qabyle-français. Parler des At Mangellat (Algerie)*. Paris, 1982., SELAF (Société d'études linguistiques et anthropologiques de France).
- DCT – MOLEN, R. van der: *A Hieroglyphic Dictionary of Egyptian Coffin Texts*. Leiden, 2000., E.J. Brill.
- DELIC – VYČIHL, W.: *Dictionnaire étymologique de la langue copte*. Leuven, 1983., Peeters.
- DG – ERICHSEN, W.: *Demotisches Glossar*. Kopenhagen, 1954., Ejnar Munksgaard.
- DIYAKAL, Ph.: *Mushere-English Dictionary*. Collection of words carried out by Mr. Ph. I. D. started on September 10<sup>th</sup>, 1997 under the supervision of Herrmann Jungraithmayr (Univ. of Frankfurt). MS. 390 p.
- DLE – LESKO, L. H.: *A Dictionary of Late Egyptian*. Volume I, II, III, IV. Berkeley, 1982., 1984., 1987., 1989. B.C. Scribe Publications.
- DOLGOPOL'SKIJ, A. B.: *Materialy po sravnitel'no-istoričeskoj fonetike kušitskih jazykov. Gubnye i dental'nye smyčnye v načal'nom položenii* // USPENSKIJ, B. A. (ed.): *Jazyki Afriki. Voprosy struktury, istorii i tipologii*. Moskva, 1966., Nauka. Pp. 35–88.
- DOLGOPOLSKI, A. B.: *La permutation des \*m et \*b initiaux dans les racines couchitiques* // *Il Congrès International des Africanistes. Communications de la délégation de l'URSS*. Moscou, 1967., Nauka. Pp. 3–17.
- DOLGOPOL'SKIJ, A. B.: *Sravnitel'no-istoričeskaja fonetika kušitskih jazykov*. Moskva, 1973., Nauka.
- DOLGOPOLSKY, A. B.: *Semitic and East Cushitic. Sound Correspondences and Cognate Sets* // SEGERT, S. & BODROGLIGETI, A. J. E. (eds.): *Ethiopian Studies Dedicated to Wolf Leslau*. Wiesbaden, 1983., Otto Harrassowitz. Pp. 123–142.
- DOLGOPOLSKY, A. B.: *South Cushitic Lateral Consonants as Compared to Semitic and East Cushitic* // JUNGRAITHMAYR, H. & MÜLLER, W. W. (eds.): *Proceedings of the Fourth International Hamito-Semitic Congress*. Amsterdam, 1987., John Benjamins. Pp. 195–214.
- DOLGOPOLSKY, A. B.: *Semitic and East Cushitic: Initial Laryngeals*. MS. Haifa, around 1987. 3 p.
- DOLGOPOLSKY, A.: *Semitic and East Cushitic: Word-Initial Laryngeals* // TADDESE, B. (ed.): *Proceedings of the Eighth International Conference of Ethiopian Studies, University of Addis Ababa, 1984*. Volume 1. Addis Ababa, 1988., Institute of Ethiopian Studies, Addis Ababa. Pp. 629–637.
- DOLGOPOLSKY, A. B.: *On Lateral Obstruents in Hamito-Semitic* // SHEVOROSHKIN, V. (ed.): *Reconstructing Languages and Cultures*. Bochum, 1989., Brockmeyer. Pp. 99–103.
- DOZY, R.: *Suppléments aux dictionnaires arabes*. Tome I–II. Leiden, Paris, 1881., E. J. Brill, Maisonneuve.
- DRB – NAÏT-ZERRAD, K.: *Dictionnaire des racines berbères*. Leuven & Paris, since 1998, Peeters.
- DRS – COHEN, D.: *Dictionnaire des racines sémitiques ou attestées dans les langues sémitiques*. Fascicules 1–2. Paris & La Haye, 1970–1976., Mouton. Fascicule 3-. Leuven, 1993-, Peeters. With continuous pagination.
- DUL – OLMO LETE, G. & SANMARTÍN, J.: *A Dictionary of the Ugaritic Language in the Alphabetic Tradition*. Part One [ʔ(a/i/u)-k]. Part Two [l-z]. Leiden, 2003., E.J. Brill.
- EDE I – TAKÁCS, G.: *Etymological Dictionary of Egyptian*. Volume One: A Phonological Introduction. Leiden, 1999., E. J. Brill.



- EDE II — TAKÁCS, G.: *Etymological Dictionary of Egyptian*. Volume Two: *b-, p-, f-*. Leiden, 2001., E. J. Brill.
- EG 1927 — GARDINER, A. H.: *Egyptian Grammar*.<sup>1</sup> Oxford, 1927., Clarendon Press.
- EHRET, Ch.: *Ethiopians and East Africans. The Problem of Contacts*. Nairobi, 1974., East African Publishing House.
- EHRET, Ch.: *Ma'a — English Vocabulary*. MS. Los Angeles, California, around 1974 (in Ehret 1980, 390 dated for 1966–74). 82 p.
- EHRET, Ch.: *The Historical Reconstruction of Southern Cushitic Phonology and Vocabulary*. Berlin, 1980., Dietrich Reimer Verlag.
- EHRET, Ch.: *Revising Proto-Kuliak // Afrika und Übersee* 64 (1981), 81–100.
- EHRET, Ch.: *Proto-Cushitic Reconstruction // Sprache und Geschichte in Afrika* 8 (1987).
- EHRET, Ch. & ELDERKIN, E. D. & NURSE, D.: *Dahalo Lexis and Its Sources // Afrikanistische Arbeitspapiere* 18 (1989), 5–49.
- EHRET, Ch.: *Reconstructing Proto-Afroasiatic (Proto-Afrasian). Vowels, Tone, Consonants, and Vocabulary*. Berkeley, Los Angeles, California, 1995., University of California.
- EILERS, W.: *Semitische Wurzeltheorie // Atti del Secondo Congresso Interna-zionale di Linguistica Camito-Semita, Firenze, 16–19 aprile 1974*. Firenze, 1978., Istituto di Linguistica e di Lingue Orientali, Università di Firenze. Pp. 125–131.
- ELDERKIN, E. D. & MAGHWAY, J. B.: *Some West Rift Roots // African Languages and Cultures* 5/1 (1992), 43–64.
- EMBER, A.: *Kindred Semito-Egyptian Words // ZÄS* 49 (1911), 93–94.
- EMBER, A.: *The Etymological Equivalent in Egyptian of the Common Semitic Word for “Life” // Orientalistische Literaturzeitung* 19 (1916), 72–74.
- EMBER, A.: (a) *New Semito-Egyptian Words*. (b) *Some African Words in Old Egyptian // JAOS* 37 (1917), 21.
- EMBER, A.: *Egypto-Semitic Studies*. Leipzig, 1930., The Alexander Cohut Memorial Foundation.
- ERMAN, A.: *Das Verhältnis des Ägyptischen zu den semitischen Sprachen // Zeitschrift der Deutschen Morgenländischen Gesellschaft* 46 (1892), 93–129.
- EWA — MAYRHOFER, M.: *Etymologisches Wörterbuch des Altindoarischen*. Heidelberg, since 1992, Carl Winter.
- FAGNAN, E.: *Additions aux dictionnaires arabes*. Alger, 1923., Jules Carbonel.
- FD — FAULKNER, R. O.: *A Concise Dictionary of Middle Egyptian*. Oxford, 1962., Clarendon Press.
- FITZPATRICK, J. F. J.: *Some Notes on the Kwolla District and Its Tribes // Journal of the Royal African Society* 10 (1910–11), 16–52, 213–22.
- FLEMING, H. C.: *Asa and Aramanik: Cushitic Hunters in Masai-Land // Ethnology* 8/1 (1969), 1–36.
- FLEMING, H. C.: *Omotic as an Afroasiatic Family // Studies in African Linguistics*. Supplement 5 (1974), 81–94.
- FLEMING, H. C.: *Omotic Overview // BENDER, M. L. (ed.): The Non-Semitic Languages of Ethiopia*. East Lansing, 1976., Michigan State University. Pp. 299–323.
- FOUCAULD, Ch. de: *Dictionnaire touareg-français, dialecte de l’Ahaggar*. I–IV. Paris, 1951–2., Imprimerie Nationale de France.
- FOULKES, H. D.: *Angass Manual. Grammar, Vocabulary*. London, 1915., Kegan Paul, Trench, Trübner and Co.
- FRAJZYNGIER, Z.: *A Pero-English and English-Pero Vocabulary*. Berlin, 1985., Dietrich Reimer Verlag.
- FRAJZYNGIER, Z.: *A Dictionary of Mupun*. Berlin, 1991., Dietrich Reimer Verlag.
- GB — GESENIUS, W. (bearbeitet von BUHL, F.): *Hebräisches und aramäisches Handwörterbuch über das Alte Testament*. Unveränderter Neudruck der 1915 erschienenen 17. Auflage. Berlin, Göttingen, Heidelberg, 1962., Springer-Verlag.
- GHWb — HANNIG, R.: *Grosses Handwörterbuch Ägyptisch-Deutsch (2800–950 v. Chr.)*. Mainz, 1995., Verlag Philipp von Zabern.
- GRAGG, G.: *Oromo Dictionary*. East Lansing, Michigan, 1982., Michigan State University.
- HARRIES, J.: *Tamazight Basic Course. Ait Mgild Dialect. Final Report*. Madison, Wisconsin, 1974.
- HEINE, B.: *Notes on the Yaaku Language (Kenya) // Afrika und Übersee* 58/2 (1975), 119–138.
- HEINE, B.: *Bemerkungen zur Boni-Sprache (Kenia) // Afrika und Übersee* 60 (1977), 242–295.
- HEINE, B.: *The Sam Languages. A History of Rendille, Boni and Somali // Afroasiatic Linguistics* 6/2 (1978), 23–115.
- HELLWIG, B.: *Goemai — English — Hausa Dictionary*. MS. Draft. Printed out on 20 August 2000. 42 p.
- HETZRON, R. & TÁLOS, E.P.: *Christopher Ehret’s “The Historical Reconstruction of Southern Cushitic Phonology and Vocabulary”*. A Review Article // *Sprache und Geschichte in Afrika* 4 (1982), 239–250.
- HINTZE, F.: *Zur hamitosemitischen Wortvergleihung // Zeitschrift für Phonetik und Allgemeine Sprachwissenschaft* 5 (1951), 65–87.

- HODGE, C. T.: Some Afroasiatic Etymologies // *Anthropological Linguistics* 10/3 (1968), 19–29.
- HODGE, C. T.: Comparative Evidence for Egyptian Historical Phonology // YOUNG, D. W. (ed.): *Studies Presented to Hans Jakob Polotsky*. East Gloucester, 1981., Pirtle and Polson. Pp. 401–413.
- HOFFMANN, C.: Zur Sprache der Cibak // LUKAS, J. (ed.): *Afrikanistische Studien*. Berlin, 1955., Akademie-Verlag. Pp. 118–143.
- HOLMA, H.: Zur semitisch-hamitischen Sprachwissenschaft // *Zeitschrift für Assyriologie* 32 (1919), 34–47.
- HSED — OREL, V. É. & STOLBOVA, O. V.: *Hamito-Semitic Etymological Dictionary*. Leiden, 1995., E. J. Brill.
- HUDSON, G.: *Highland East Cushitic Dictionary*. Hamburg, 1989., Buske.
- IEW — POKORNY, J.: *Indogermanisches etymologisches Wörterbuch*. Band I. Bern-München, 1959., Francke Verlag.
- ILLIČ-SVITYČ, V. M.: Drevnejšie indoeuropejsko-semitskie jazykovye kontakty // TOPOROV, V. N. (ed.): *Problemy indoeuropejskogo jazykoznanija*. Moskva, 1964., Nauka. Pp. 3–12.
- ILLIČ-SVITYČ, V. M.: *Opyt sravnenija nostratičeskich jazykov (semitohamitskij, kartvel'skij, indoeuropejskij, ural'skij, dravidijskij, altajskij)*. Vvedenie. *Sravnitel'nyj slovar' (b–k)*. Moskva, 1971., Nauka.
- JACQ, Ch.: *Lexique des verbes de mouvement dans les textes des pyramides et les textes des sarcophages*. Paris, 1986., I.M.O., Recherches Égyptologiques.
- JASTROW, M.: *A Dictionary of the Targumim, the Talmud Babli and Yerushalmi, and the Midrashic Literature*. Volume I: ʔ-k, Volume II: l-t. New York, 1950., Pardes Publishing House Inc.
- JOHNSTONE, T. M.: *Ḥarsūsi Lexicon*. London, 1977., Oxford University Press.
- JOHNSTONE, T. M.: *Jibbāli Lexicon*. London, 1981., Oxford University Press.
- JOHNSTONE, T. M.: *Mehri Lexicon*. London, 1987., University of London.
- JUNGRAITHMAYR, H.: *Wörterbuch der Angas-Sprache*. MS. 1962.
- JUNGRAITHMAYR, H.: *Wörterbuch der Goemay-Sprache*. MS. 1962.
- JUNGRAITHMAYR, H.: Die Sprache des Sura (Maghavul) in Nordnigerien // *Afrika und Übersee* 47 (1963), 8–89, 204–220.
- JUNGRAITHMAYR, H.: Materialien zur Kenntnis des Chip, Montol, Gerka und Burrum (Südplateau, Nordnigerien) // *Afrika und Übersee* 48 (1965), 161–183.
- JUNGRAITHMAYR, H.: Specimens of the Pa'a ("Afa") and Warja Languages with Notes on the Tribes of Ningi Chiefdom (Bauchi Province, Northern Nigeria) // *Afrika und Übersee* 50/3 (1967), 194–205.
- JUNGRAITHMAYR, H.: A Comparative Word List of the Ron Languages (Southern Plateau, N. Nigeria) // *Africana Marburgensia* 1/2 (1968), 3–12.
- JUNGRAITHMAYR, H.: *Die Ron-Sprachen. Tschadohamitische Studien in Nordnigerien*. Glückstadt, 1970., Verlag J. J. Augustin.
- JUNGRAITHMAYR, H. & SHIMIZU, K.: *Chadic Lexical Roots. Vol. II. Tentative Reconstruction, Grading and Distribution*. Berlin, 1981., Verlag von Dietrich Reimer.
- JUNGRAITHMAYR, H.: *Lexique mokilko*. Berlin, 1990., Dietrich Reimer Verlag.
- JUNGRAITHMAYR, H. & BARRETEAU, D.: Les verbes monoradicaux dans les langues tchadiques // JUNGRAITHMAYR, H. & TOURNEUX, H. (eds.): *Études tchadiques. Verbes monoradicaux suivis d'une note sur la négation en haoussa*. Actes de la XII<sup>ème</sup> réunion de Groupe d'Études Tchadiques LACITO-CNRS-PARIS. Paris, 1990., Librairie Orientaliste Paul Geuthner. Pp. 37–214.
- JUNGRAITHMAYR, H. & IBRISZIMOW, D.: "Chadic Lexical Roots" — Wege zum Urtschadischen // *IX. Afrikanistentag — Beiträge zur afrikanischen Sprach- und Literaturwissenschaft*. Köln, 1993., publisher not indicated. Pp. 129–148.
- JUNGRAITHMAYR, H. & IBRISZIMOW, D.: *Chadic Lexical Roots. Volume I. Tentative Reconstruction, Grading, Distribution and Comments*. Berlin, 1994., Dietrich Reimer Verlag.
- JUNGRAITHMAYR, H. & IBRISZIMOW, D.: *Chadic Lexical Roots. Volume II. Documentation*. Berlin, 1994., Dietrich Reimer Verlag.
- JUNGRAITHMAYR, H.: *Mushere Vocabulary*. Preliminary version composed on the basis of one part (the first 70 pages) of the material "Mushere Sprichwörter". MS. Frankfurt a/M, 1999. 20 p.
- KB — KOEHLER, L. & BAUMGARTNER, W.: *The Hebrew and Aramaic Lexicon of the Old Testament*. I–V. Leiden, 1994., E. J. Brill.
- KHW — WESTENDORF, W.: *Koptisches Handwörterbuch*. Heidelberg, 1977., Carl Winter Universitätsverlag.
- KLEINWILLINGHÖFER, U.: Aspects of Vowel Harmony in Waja and Tangale-Waja Common Vocabulary // *Frankfurter Afrikanistische Blätter* 2 (1990), 93–106.
- KRAFT, Ch. H.: *Chadic Wordlists*. I–III. Berlin, 1981., Dietrich Reimer Verlag.



- LAMBERTI, M.: *Die Somali-Dialekte. Eine Vergleichende Untersuchung*. Hamburg, 1986., Helmut Buske Verlag.
- LANE, E. W.: *An Arabic-English Lexicon*. I–VIII. London & Edinburgh, 1863–93., Williams and Norgate.
- LAMBERTI, M.: Some Konsoid Etymologies // *Anthropos* 82/4–6 (1987), 529–541.
- LESLAU, W.: *Lexique soqotri (sudarabique moderne), avec comparaisons et explications étymologiques*. Paris, 1938., Librairie C. Klincksieck.
- LESLAU, W.: The Parts of the Body in Modern South Arabic Languages // *Language* 21 (1945), 230–249.
- LESLAU, W.: *Ethiopic and South Arabic Contributions to the Hebrew Lexicon*. Berkeley, Los Angeles, 1958., University of California.
- LESLAU, W.: *Comparative Dictionary of Geʿez (Classical Ethiopic)*. Wiesbaden, 1987., Otto Harrassowitz.
- LUKAS, J.: Die Logone-Sprache im Zentralen Sudan // *Abhandlungen für die Kunde des Morgenlandes* 21/6 (1936).
- LUKAS, J.: Die Sprache des Buduma in Zentralen Sudan // *Abhandlungen für die Kunde des Morgenlandes* 24/2 (1939).
- LUKAS, J.: Die Personalia und das primäre Verb im Bolanci (Nordnigerien). Mit Beiträge über das Karekare // *Afrika und Übersee* 55 (1971), 114–139.
- LUKAS, J.: Tschadische Studien I. Beiträge zur Kenntnis des Mokulu // *Afrika und Übersee* 60 (1977), 1–58, 182–229.
- MAGHWAY, J. B.: Iraqw Vocabulary // *Afrikanistische Arbeitspapiere* 18 (1989), 91–118.
- MEINHOF, C.: Linguistische Studien in Ostafrika. Fortsetzung // *Mitteilungen des Seminars für Orientalische Sprachen* 9 (1906), 278–333.
- MILITAREV, A. Ju.: Sovremennoe sravnitel'no-istoričeskoe afrazijskoe jazykoznanie: čto ono mozet dat' istoričeskoj nauke? // *Lingvističeskaja rekonstrukcija i drevnejšaja istorija Vostoka*. Čast' 3. Moskva, 1984., Nauka. Pp. 3–26.
- MILITAREV, A. Ju.: Jazyk meroitskoj épigrafiki kak istoričeskij istočnik v svete ego genezisa // *Vestnik Drevnej Istorii* 2 (1984), 153–170.
- MILITAREV, A. Ju.: Proišoždenie kornej so značeniem “tvorit', sozdat'” v afrazijskih jazykah // *Pis'mennye pamjatniki i problemy istorii kul'tury narodov Vostoka. XIX godičnaja naučnaja sessija Leningradskogo Otdelenija Instituta Vostokovedenija Akademii Nauk SSSR*. Moskva, 1986., Nauka. Pp. 63–79.
- MILITAREV, A. Ju.: Afrazijsko-indoevropskie leksičeskie svjazi // *Irano-afrazijskie jazykovye kontakty*. Moskva, 1987., Nauka. Pp. 98–108.
- MILITAREV, A. Ju.; STOLBOVA, O. V.: First Approach to Comparative-Historical Phonology of Afrasian (Consonantism) // MUKAROVSKY, H. G. (ed.): *Proceedings of the Fifth International Hamito-Semitic Congress*. Band I. Wien, 1990., Afro-Pub. Pp. 45–72.
- MILITAREV, A. Ju.: Guančskie jazyki. Fonetika // SOLNCEV, V. M. (ed.): *Jazyki Azii i Afriki*. IV, kniga 2. Moskva, 1991., Glavnaja Redakcija Vostočnoj Literatury. Pp. 163–173.
- MOUS, M. & QORRO, M. & KIEßLING, R.: *Iraqw-English Dictionary with an English and a Thesaurus Index*. Köln, 2002., Rüdiger Köppe Verlag.
- MÖLLER, G.: Aegyptisch-libysches // *Orientalistische Literaturzeitung* 24/9–10 (1921), 193–197.
- MÖLLER, G.: Die Ägypter und ihre libyschen Nachbarn // *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 78 (1924), 36–60.
- MUKAROVSKY, H. G.: *Mande-Chadic Common Stock. A Study of Phonological and Lexical Evidence*. Wien, 1987., Afro-Pub.
- MÜLLER, W. M.: Ägyptische und semitische Umschreibungsfragen // *Orientalistische Literaturzeitung* 10 (1907), 299–305, 358–360.
- MÜLLER, W. W.: Beiträge zur hamito-semitischen Wortvergleichung // BYNON, J. & BYNON, Th. (eds.): *Hamito-Semitic*. The Hague, 1975., Mouton de Gruyter. Pp. 63–74.
- NETTING, R. M.: *Kofyar Vocabulary*. MS. 1967.
- NEWMAN, P.: Chadic Classification and Reconstructions // *Afroasiatic Linguistics* 5/1 (1977), 1–42.
- OREL, V. É. & STOLBOVA, O. V.: Čadsko-egipetskie izoglossy v oblasti kul'turnoj leksiki // *Lingvističeskaja rekonstrukcija i drevnejšaja istorija Vostoka*. Čast' 1. Moskva, 1989., Nauka. Pp. 131–136.
- OREL, V. É. & STOLBOVA, O. V.: K rekonstrukcii praafrazijskogo vokalizma. 3–4 // *Voprosy Jazykoznanija* 2 (1990), 75–90.
- OREL, V. É. & STOLBOVA, O. V.: Cushitic, Chadic, and Egyptian: Lexical Relations // SHEVOROSHKIN, V. (ed.): *Nostratic, Dene-Caucasian, Austric and Amerind*. Bochum, 1992., Brockmeyer. Pp. 167–180.
- OREL, V. É. & STOLBOVA, O. V.: On Chadic-Egyptian Lexical Relations // SHEVOROSHKIN, V. (ed.): *Nostratic, Dene-Caucasian, Austric and Amerind*. Bochum, 1992., Brockmeyer. Pp. 181–203.
- OREL, V. É.: Semitohamitskij, sinokavkazskij, nostratičeskij // *Moskovskij Lingvističeskij Žurnal* 1 (1995), 99–116.

- ORMSBY, G.: Notes on the Angass Language // *Journal of the Royal African Society* 12 (1912–1913), 421–424 & 13 (1913–1914), 54–61, 204–210, 313–315.
- OSING, J.: Review of ČERNÝ, J.: *Coptic Etymological Dictionary* // *JEA* 64 (1978), 186–189.
- PARKER, E. M. & HAYWARD, R. J.: *An Afar-English-French Dictionary (with Grammatical Notes in English)*. London, 1985., School of Oriental and African Studies, University of London.
- PEUST, C.: *Egyptian Phonology. An Introduction to the Phonology of a Dead Language*. Göttingen, 1999., Peust & Gutschmidt Verlag GbR.
- PILLINGER, S. & GALBORAN, L.: *A Rendille Dictionary*. Köln, 1999., Rüdiger Köppe Verlag.
- PORHOMOVSKIJ, V. Ja.: *Istoričeskij konsonantizm jazykov kotoko*. Moskva, 1972., Institut Jazykoznanija Akademii Nauk SSSR.
- PRASSE, K.-G.: A propos de l'origine de *h* touareg (tahaggart) // *Det Kongelige Danske Videnskabernes Selskab. Historisk-filosofiske Meddelelser* 43/3 (1969).
- PRASSE, K.-G. & ALOJALY, Gh. & MOHAMED, Gh.: *Dictionnaire touareg-français (Niger)*. Copenhagen, 2003., Museum Tusulanum Press, University of Copenhagen.
- PT — SETHE, K.: *Die altägyptischen Pyramidentexte*. I–II. Leipzig, 1980., 1910., J.C. Hinrichs'sche Buchhandlung.
- RABIN, Ch.: Lexicostatistics and the Internal Divisions of Semitic // BYNON, J.; BYNON, Th. (eds.): *Hamito-Semitic*. The Hague, 1975., Mouton. Pp. 85–102.
- RAPP, E. L. & BENZING, B.: *Dictionary of the Glavdá Language*. Frankfurt am Main, 1968., Bible Society Frankfurt am Main.
- REINISCH, L.: Die Sprache Saho-Irob in Abessinien // *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften, phil.-hist. Classe* 90/9 (1878), 89–142.
- REINISCH, L.: Die Quarasprache in Abessinien. II. Quarisch-deutsches Wörterbuch // *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften. Phil.-hist. Classe* 109/1 (1885), 3–152.
- REINISCH, L.: Das Zalwort vier und neun in den chamitisch-semitischen Sprachen // *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften. Phil.-hist. Classe* 121/12 (1890).
- REINISCH, L.: *Wörterbuch der Saho-Sprache*. Wien, 1890., Alfred Hölder.
- REINISCH, L.: *Wörterbuch der Bedawye-Sprache*. Wien, 1895., Alfred Hölder Verlag.
- REINISCH, L.: *Die Somali-Sprache*. II. Wörterbuch. Wien, 1902., Alfred Hölder Verlag.
- RENISIO, A.: *Étude sur les dialectes berbères des Beni Iznassen, du Rif et des Senhaja de Sraïr*. Grammaire, textes et lexique. Paris, 1932., Éditions Ernest Leroux.
- REUTT, T. E. & KOGAN, E. Z.: Materialy po leksike jazykov margi i bura // *Bespis'mennye i mladopis'mennye jazyki Afriki*. Moskva, 1973., Nauka. Pp. 83–147.
- ROPER, E.-M.: *Tu Beḏawīe. An Elementary Handbook for the Use of Sudan Government Officials*. Hertford, 1928., Stephen Austin & Sons.
- ROQUET, G.: šmt(=j) “ma belle-mère” (6<sup>e</sup> Dynastie) // *BIFAO* 77 (1977), 119–127.
- RÖSSLER, O.: Der semitische Charakter der libyschen Sprache // *Zeitschrift für Assyriologie* 50 (1952), 121–150.
- RÖSSLER, O.: Das Ägyptische als semitische Sprache // ALTHEIM, F. & STIEHL, R. (eds.): *Christentum am Roten Meer*. Band I. Berlin, New York, 1971., Walter de Gruyter. Pp. 263–325.
- SD — BEESTON, A. F. L. & GHUL, M. A. & MÜLLER, W. W. & RYCKMANS, J.: *Sabaic Dictionary (English-French-Arabic)*. *Dictionnaire sabéen (anglais-français-arabe)*. Louvain-la-Neuve, Beyrouth, 1982., Peeters, Librairie du Liban.
- SASSE, H.-J.: The Consonant Phonemes of Proto-East-Cushitic (PEC) // *Afroasiatic Linguistics* 7/1 (1979), 1–67.
- SASSE, H.-J.: *An Etymological Dictionary of Burji*. Hamburg, 1982., Helmut Buske Verlag.
- SATZINGER, H.: Historische ägyptische Phonologie und die afroasiatische Komparatistik (Rezensionsartikel) // *Wiener Zeitschrift für die Kunde des Morgenlandes* 93 (2003), 211–225.
- SCHNEIDER, Th.: Beiträge zur sogenannten “neueren Komparatistik”. Zum Gedenken an Otto Rössler (1907–1991) // *Lingua Aegyptia* 5 (1997), 189–209.
- SCHUH, R. G.: *Bole-Tangale Languages of the Bauchi Area (Northern Nigeria)*. Berlin, 1978., Dietrich Reimer.
- SED I — KOGAN, A. & MILITAREV, A.: *Semitic Etymological Dictionary*. Vol. I. Anatomy of Man and Animals. Münster, 2000., Ugarit-Verlag.
- SETHE, K.: Neue Spuren der Hyksos in Inschriften der 18. Dynastie // *Zeitschrift für Ägyptische Sprache* 47 (1910), 73–86.
- SIRLINGER, E.: *Dictionary of the Goemay Language*. Jos, Nigeria, 1937., Prefecture Apostolic of Jos.

- SISAJa III — D'JAKONOV, I. M.; BELOVA, A. G.; MILITAREV, A. Ju.; PORHOMOVSKIJ, V. Ja.; STOLBOVA, O. V.: Sravnitel'no-istoričeskij slovar' afrazijskih jazykov. Vypusk 3: s-c-ç-ʃ, č-č-ž, š-ê-ê s labialami // *Pis'mennye pamjatniki i problemy istorii kul'tury narodov Vostoka. XVI godičnaja naučnaja sessija Leningradskogo Otdelenija Instituta Vostokovedenija Akademii Nauk SSSR*. Moskva, 1986., Nauka. Pp. 3–46.
- SKINNER, N.: North Bauchi Chadic Languages: Common Roots // *Afroasiatic Linguistics* 4/1 (1977), 1–49.
- SODEN, W. von: Zum akkadischen Wörterbuch. 61–66 // *Orientalia* NS 24 (1955), 136–145.
- SODEN, W. von: Zum akkadischen Wörterbuch. 67–80 // *Orientalia* NS 24 (1955), 377–394.
- STOLBOVA, O. V.: Soglasnye verhnjezapadnočadskih jazykov. Genetičeskie sootvetstvija // MAKAEV, É. A. (ed.): *Istoričko-tipologičeskie i sinhronno-tipologičeskie issledovanija (na materiale jazykov raznyh sistem)*. Moskva, 1972., Institut Jazykoznanija Akademii Nauk SSSR.
- STOLBOVA, O. V.: Opyt rekonstrukcii verhnjezapadnočadskih kornej // *Jazyki zarubežnogo Vostoka. Sbornik statej*. Moskva, 1977., Nauka. Pp. 152–160.
- STOLBOVA, O. V.: Rekonstrukcija konsonantnoj sistemy zapadnočadskih jazykov // *Pis'mennye pamjatniki i problemy istorii kul'tury narodov Vostoka. XIX godičnaja naučnaja sessija Leningradskogo Otdelenija Instituta Vostokovedenija Akademii Nauk SSSR*. Moskva, 1986., Nauka. Pp. 80–115.
- STOLBOVA, O. V.: Sravnitel'no-istoričeskaja fonetika i slovar' zapadnočadskih jazykov // PORHOMOVSKIJ, V. Ja. (ed.): *Afrikanskoe istoričeskoe jazykoznanie. Problemy rekonstrukcii*. Moskva, 1987., Nauka. Pp. 30–268.
- STOLBOVA, O. V.: *Lateral Sibilants in Chadic (Reconstruction) and Their Correspondences in Semitic and Egyptian*. MS. Paper presented at the Symposium on Chadic and Hamito-Semitic, Frankfurt am Main, 6–8 May 1991. 9 p.
- STOLBOVA, O. V.: *Lateral Sibilants in Chadic (Reconstruction) and Their Correspondences in Semitic and Egyptian* // IBRISZIMOW, D. & LEGER, R. (eds.): *Studia Chadica et Hamitosemitica*. Köln, 1995., Rüdiger Köppe Verlag. Pp. 58–64.
- TAÏFI, M.: *Dictionnaire tamazight-français (parlers du Maroc central)*. Paris, 1991., L'Harmattan-Awal.
- TAKÁCS, G.: Egyptian *m3t* “to think out” // *Archív Orientální* 63/2 (1995), 159–161.
- TAKÁCS, G.: The Afrasian Origin of Egyptian *rm* “Fish” (shortened version) // KOVAL', A. I. & VINOGRADOV, V. A. (eds.): *Problemy izučenija jazykov Afriki. Materialy konferencii posvjaščennoj 30-letiju Otdela Afrikanskih Jazykov Instituta Jazykoznanija Rossijskoj Akademii Nauk (Moskva, 4–6 dekabnja 1995 g.)*. Moskva, 1995., Institut Jazykoznanija RAN. Pp. 159–164.
- TAKÁCS, G.: The Afrasian Origin of Egyptian *rm* “Fish” // *Folia Orientalia* 32 (1996), 89–93.
- TAKÁCS, G.: Sibilant and Velar Consonants of South Cushitic and Their Regular Correspondences in Egyptian and Other Afro-Asiatic Branches // LAMBERTI, M. & TONELLI, L. (eds.): *Afroasiatica Tergestina. Papers from the 9<sup>th</sup> Italian Meeting of Afro-Asiatic (Hamito-Semitic) Linguistics, Trieste, April 23–24, 1998*. Contributi presentati al 9<sup>o</sup> Incontro di Linguistica Afroasiatica (Camito-Semitica), Trieste, 23–24 Aprile 1998. Padova, 1999., Unipress. Pp. 393–426.
- TAKÁCS, G.: South Cushitic Consonant System in Afro-Asiatic Context // *Afrikanistische Arbeitspapiere* 61 (2000), 69–117.
- TAKÁCS, G.: Towards Proto-Afro-Asiatic Phonology: Ancient Remnants in South Cushitic, Angas-Sura, and North Bauchi // *Rocznik Orientalistyczny* 54/2 (2001), 55–125.
- TAKÁCS, G.: South Cushitic Sibilant Affricates in a Comparative-Historical Perspective // BENDER, M. L. (chief ed.) & APPELYARD, D. & TAKÁCS, G. (eds.): *Selected Comparative-Historical Afrasian Linguistic Studies in Memory of Igor M. Diakonoff*. Lincom Studies in Afroasiatic Linguistics 14. München & Newcastle, 2003., Lincom Europa. Pp. 143–162.
- TAKÁCS, G.: *Comparative Dictionary of the Angas-Sura Languages*. Berlin, 2004., Dietrich Reimer Verlag.
- TOSCO, M.: *The Classification of Dahalo: Another Perspective*. Paper presented at the 2<sup>nd</sup> International Congress on Cushitic and Omotic Languages, Torino, 16–18 November 1989. 28 p.
- TOSCO, M.: *A Grammatical Sketch of Dahalo*. Hamburg, 1991., Helmut Buske Verlag.
- ÜKAPT I–VI — SETHE, K.: *Übersetzung und Kommentar zu den altägyptischen Pyramidentexten*. I–VI. Glückstadt, Hamburg, 1935–62., J. J. Augustin.
- VERGARI, M. & VERGARI, R.: *A Basic Saho-English-Italian Dictionary*. Asmara, Eritrea, 2003., (publisher not indicated).
- VERGOTE, J.: *Phonétique historique de l'égyptien*. Paris, 1945., Le Muséon.
- VOIGT, R.: Review of EHRET, Ch.: *The Historical Reconstruction of Southern Cushitic Phonology and Vocabulary* // *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 133/1 (1983), 197–198.
- VYCICHL, W.: Über ein *h*-Präfix im Arabischen // *Wiener Zeitschrift für die Kunde des Morgenlandes* 43 (1936), 109–110.

- VYCICHL, W.: Grundlagen der ägyptisch-semitischen Wortvergleichung // *MDAIK* 16 (1958), 367–405.
- VYCICHL, W.: *La vocalisation de la langue égyptienne*. Tome I<sup>er</sup>. La phonétique. Le Caire, 1990., Institut Français d'Archéologie Orientale.
- WARD, W. A.: Some Egypto-Semitic Etymologies // *Orientalia NS* 31 (1962), 397–412.
- WARD, W. A.: Notes on Some Egypto-Semitic Roots // *Zeitschrift für Ägyptische Sprache* 95 (1968), 65–72.
- WARD, W. A.: The Biconsonantal Root \*b3 and Remarks on Bilabial Interchange in Egyptian // *Zeitschrift für Ägyptische Sprache* 102 (1975), 60–67.
- WARD, W. A.: A New Look at Semitic Personal Names and Loanwords in Egyptian // *Chronique d'Égypte* 71 (1996), 17–47.
- Wb — ERMAN, A. & GRAPOW, H.: *Wörterbuch der ägyptischen Sprache*. I–V.<sup>2</sup> Berlin, 1957–1971., Akademie-Verlag.
- WEIBEGUÉ, Ch. & PALAYER, P.: *Lexique lele-français. Sarh, Tchad*, 1982., Centre d'Études Linguistiques.
- WELMERS, W. E.: Notes on the Structure of Saho // *Word* 8/2 (1952), 145–162 & 8/3 (1952), 236–251.
- WHITELEY, W. H.: *Studies in Iraqw*. Kampala, 1953., East African Institute of Social Research.
- WHITELEY, W. H.: *A Short Description of Item Categories in Iraqw (with Material on Gorowa, Alagwa and Burunge)*. Kampala, 1958., East African Institute of Social Research.
- ZABORSKI, A.: Biconsonantal Verbal Roots in Semitic // *Zeszyty Naukowe Uniwersytetu Jagiellońskiego, prace językoznawcze* 5 (1971), 51–98.
- ZABORSKI, A.: Some Internal Plurals in Cushitic // FRONZAROLI, P. (ed.): *Atti del Secondo Congresso Internazionale di Linguistica Camito-Semitica, Firenze, 16–19 aprile 1974*. Firenze, 1978., Istituto di Linguistica e di Lingue Orientali, Università di Firenze. Pp. 369–378.
- ZABORSKI, A.: Review of EHRET, Ch.: *The Historical Reconstruction of Southern Cushitic. Phonology and Vocabulary // Afrika und Übersee* 67 (1984), 142–144.
- ZABORSKI, A.: Der Wortschatz der Bedscha-Sprache. Eine vergleichende Analyse // SCHULER, E. von (ed.): *XXIII. Deutscher Orientalistentag, vom 16. bis 20. September 1985 in Würzburg. Ausgewählte Vorträge*. Stuttgart, 1989., Franz Steiner Verlag. Pp. 573–591.
- ZYHLARZ, E.: Ursprung und Sprachcharakter des Altägyptischen // *Zeitschrift für Eingeborenen-Sprachen* 23 (1932–1933), 25–45, 81–110, 161–194, 241–254.

В статье утверждается, что богатый инвентарь ларингальных согласных, восстанавливаемый для западнорифтской подгруппы южнокушитских языков, скорее всего носит архаичный характер; поскольку для всех этих согласных находятся однозначные соответствия в семитских и египетском языках, можно предположить, что слова с ларингалами в западнорифтских языках сохраняют старую афразийскую артикуляцию, что, в свою очередь, подчеркивает чрезвычайную важность этих языков для решения задачи реконструкции праафразийского языка. Гипотеза подтверждается на материале многочисленных этимологий.



## Studies in Uralic vocalism I: A more economical solution for the reconstruction of the Proto-Permic vowel system<sup>1</sup>

The article discusses certain issues with the system of Proto-Permic vowels, reconstructed by V. I. Lytkin in his seminal monograph “Historical Vocalism of the Permic Languages”. Several rules of complementary distribution between vocalic correspondences are established, allowing us to postulate a more compact vowel system for Proto-Permic.

*Keywords:* vocalic systems, linguistic reconstruction, comparative phonology, Proto-Permic vocalism, Permic languages, Komi language, Udmurt language.

In this article we intend to reconsider the data that was used by V. I. Lytkin in his classic work on the historical vocalism of the Permic languages [Lytkin 1964] as a basis for his reconstruction of the Proto-Permic vocalism. Our main point of interest are the vowels *\*e*, *\*e̘*, *\*ε*, *\*ò*, *\*o̘* and *\*ö̘*, assumed by Lytkin for the Proto-Permic stage. The Proto-Komi and Proto-Udmurt vocalic reconstructions given below are based on the following set of correspondences between the dialects:<sup>2</sup>

Vocalic correspondences between Komi dialects										
Proto-Komi	*i	*ú	*u	*e̘	*ò̘	*o̘	*e	*ò	*o	*a
Literary Komi	i	ǰ	u	e	ε̘	o	e	ε̘	o	a
Upper Sysola	i	ǰ	u	e	ε̘	o̘	e	ε̘	o	a
Old Permic	i	ǰ	u	e̘	ε̘	o̘	e	ε̘	o	a
Eastern Vyčegda	i	ú	u	e	ò	o	e	ò	o	a
Jažva	i	ø	u	í	úú	ú	e	ò	o	a

Vocalic correspondences between Udmurt dialects									
Proto-Udmurt	*i	*ú	*j	*u	*e	*ò	*o	*a	
Literary Udmurt	i	u	ǰ	u	e	ε̘	o	a	
South-Western dialects	i	ú	ø	u	e	ò	o	a	
Besermyan	i	ø	ø	u	e	e	o	a	

Notes:

- 1) Proto-Udmurt *\*u* in South-Western dialects yields *ú* in roots of the CV type.
- 2) Proto-Udmurt *\*ú* in Besermyan yields *u* before palatalised consonants.

<sup>1</sup> This study was carried out with the support of RGNF foundation, project № 09-04-95202a/F. I am grateful to N. Yu. Zhivlova who helped me with the English translation of this article.

<sup>2</sup> The tables are based on [Lytkin 1964: 231–233].

Let us start with the Proto-Permic phonemes \**e* и \**ɛ̥*. According to Lytkin's reconstruction, this phonemic opposition is preserved in Proto-Komi, where \**e* yields \**e*, and Proto-Permic \**ɛ̥* is reflected as Proto-Komi \**ɛ̥*. In Proto-Udmurt those two vowels merged, yielding \**e* when adjacent to a palatalised consonant and \**o* when adjacent to a non-palatalised consonant.

The analysis of the Permic material shows that in the words that can be etymologised as going back to Proto-Permic (i.e. words which have Udmurt or other Finno-Ugric etymological correspondences) Proto-Komi phonemes \**e* и \**ɛ̥* appear in complementary distribution: \**e* is the regular outcome before resonants (note three examples of \**ɛ̥* before a lost \**j* or \**l̥*) and voiceless obstruents. Before voiced obstruents the regular outcome is \**ɛ̥*, except the position adjacent to \**b* or \**m*, where both reflexes are possible (there are six examples with \**e* and two with \**ɛ̥*). Thus, the difference between the Proto-Komi vowels \**e* and \**ɛ̥* cannot date back to Proto-Permic.

**Before resonants** (PK \**e* — 26 examples, PK \**ɛ̥* (before a lost \**j* or \**l̥*) — 3 examples):

PK \**berd* — PUD \**bord(i)* 'wall; measure of length' [KESK: 39]; PK \**berg-* — PUD \**berg-* 'to turn around' [KESK: 38–39]; PK \**ćerań* — PUD \**ćonari* 'spider' [KESK: 303]; PK \**ćerig* — PUD \**ćorig* 'fish' [KESK: 303–304]; PK \**er-* — PUD \**or-* 'to roar, curse' [KESK: 332]; PK \**jen* (\**jenm-*) — PUD \**in* (\**inm-*) 'god; sky' [KESK: 99]; PK \**kel* (\**kelj-*) 'husband's brother's wife' — PUD \**kalí* 'a form of address of a younger daughter-in-law to an older one' [KESK: 120]; PK \**keł* 'pale' — PUD \**keł-* 'ginger (a horse's colour only)' [KESK: 121]; PK \**ker* (\**kerj-*) — PUD \**kor* 'log, beam' [KESK: 121]; PK \**ker-* — PUD \**kor-* 'to chop' [KESK: 121–122]; PK \**lem* — PUD \**lem* 'glue' [KESK: 158]; PK \**ńeròd* — PUD \**ńored* 'gland' [KESK: 187]; PK \**pelus* — PUD \**polis* 'paddle' [KESK: 218]; PK \**peł* — PUD \**peł* 'ear' [KESK: 218–219]; PK \**perüd* 'quick' (PFP \**pir3*) [KESK: 220; UEW: 732]; PK \**šer-* — PUD \**šer-* 'to laugh' [KESK: 251]; PK \**šeròl* — PUD \**šerel* 'nit' [KESK: 251]; PK \**šermòd* — PUD \**šermet* 'bridle' [KESK: 251]; PK \**šerd-* — PUD \**šordj-* 'to sift' [KESK: 319]; PK \**ver* — PUD \**vor-* 'man' (in Udmurt *vorgoron* 'man') (cf. [KESK: 52]); PK \**ver-* — PUD \**vor-* 'to flicker, ripple' [KESK: 52]; PK \**verd-* — PUD \**vord-* 'to feed, rear' [KESK: 52]; PK \**verm-* — PUD \**vorm-* 'to overcome, win' [KESK: 53]; PK \**veržòm* — PUD \**voržem* '(to ride) on horseback' [KESK: 52–53]; PK \**zel* — PUD \**zol* 'tight, strong' [KESK: 108]; PK \**zer* — PUD \**zor* 'rain' [KESK: 108].

PK \**jem* 'needle' (PU \**äjmä*) [KESK: 99; UEW: 22]; PK \**pemüd* — PUD \**pelmit* 'dark' [KESK: 219]; PK \**oem* — PUD \**vijim* 'brain' [KESK: 51].

### Exceptions:

PK \**ženg-* 'to strike, shoot' — PUD \**žong-* 'to knock, strike' [KESK: 90]; PK \**žėńüd* 'short' (PFP \**čėn3*) [UEW: 620]<sup>3</sup>.

**Before voiceless obstruents** (PK \**e* — 15 examples):

PK \**ćéc-* — PUD \**tek-* 'to jump' [KESK: 304]; PK \**ćepòl-* — PUD \**ćepijl-* 'to pinch' [KESK: 303]; PK \**esk-* — PUD \**osk-* 'to believe' [KESK: 332]; PK \**lećüd* — PUD \**lećit* 'pointed' [KESK: 165]; PK \**pečòr* — PUD \**pušńer*, \**pošńer* 'nettle' [KESK: 220–221]; PK \**pes-* — PUD \**pos-* 'to brake, crumple' [KESK: 220]; PK \**pet-* — PUD \**pot-* 'to come out, to ascend' [KESK: 220]; PK \**rekt-* — PUD \**žokt-* 'to unload' [KESK: 240]; PK \**setòr* — PUD \**suter* 'blackcurrant' [KESK: 272]; PK \**šet-* — PUD \**šot-* 'to give' [KESK: 251]; PK \**šep* — PUD \**šep*

<sup>3</sup> According to UEW, this etymology is questionable; this makes us doubt whether the Komi word can indeed be traced back to the Proto-Permic stage.



‘ear (of grain)’ [KESK: 319]; PK *\*vek-* — PUd *\*vek-* ‘narrow’ [KESK: 51]; PK *\*vešk-* — PUd *\*vešk-* ‘straight, slim’ [KESK: 54]; PK *\*vešt-* ‘to buy out’ — PUd *\*vošt-* ‘to cancel’ [KESK: 54]; PK *\*žesküd* — PUd *\*žoskjt* ‘tight’ [KESK: 90].

### Exceptions:

PK *\*pēs* — PUd *\*pis* ‘log, firewood’ [KESK: 220]; PK *\*vεtl-* — PUd *\*vetl-* ‘to walk’ [KESK: 54].

### Before voiced obstruents (not adjacent to *\*b* or *\*m*) (PK *\*ε* — 16 examples):

PK *\*cég-* — PUd *\*cig-* ‘to break (itself)’ [KESK: 303]; PK *\*εža* — PUd *\*ožo* ‘sod; grass’ [KESK: 331]; PK *\*gεz* (*\*gεzj-*) — PUd *\*gozi* ‘rope’ [KESK: 76]; PK *\*kεž-* ‘to wait (for), hope’ — PUd *\*kož-* ‘to suppose, believe’ [KESK: 120]; PK *\*kεž-* postposition — PUd *\*kož-* ‘alone, by oneself’ [KESK: 121]; PK *\*kεž-* — PUd *\*kož-* ‘turn away (from a road)’ [KESK: 121]; PK *\*lež-* — PUd *\*lež-* ‘to release, let off’ [KESK: 165]; PK *\*pεž-* — PUd *\*pež-* ‘to spring (with regard to a trap, snare or spring)’ [KESK: 218]; PK *\*rεgüd* — PUd *\*žog* ‘soon’ [KESK: 240]; PK *\*rεž-* — PUd *\*žéz-* ‘to rip open’ [KESK: 240]; PK *\*šēd-* — PUd *\*šed-* ‘to hit upon, chance upon’ [KESK: 318]; PK *\*vεž* — PUd *\*vož* ‘green’ [KESK: 49]; PK *\*vεž* — PUd *\*vož* ‘envy, hatred’ [KESK: 49]; PK *\*vεž* — PUd *\*vož* ‘crossroads’ [KESK: 49–50]; PK *\*vεž-* — PUd *\*voš-* ‘(ex)change, replace’ [KESK: 50]; PK *\*vεža* — PUd *\*vožo* ‘holy, saint’ [KESK: 50].

**One exception** is PK *\*šeg* ‘ankle-bone, ankle’ — PUd *\*šeg* ‘sprout, outgrowth (on a tree)’ [KESK: 318].

### Before voiced obstruents (adjacent to *\*b* or *\*m*) (PK *\*e* — 6 examples, PK *\*ε* — 2 examples):

PK *\*bed* (*\*bedj-*) — PUd *\*bodi* ‘stick, cane’ [KESK: 38]; PK *\*leb-* — PUd *\*lob-* ‘to fly’ [KESK: 165]; PK *\*med* — PUd *\*med* ‘let, may, in order to’ [KESK: 171]; PK *\*meg* ‘river bend’ — PUd *\*mog* ‘roundabout’ [KESK: 171]; PK *\*mez-* — PUd *\*moz-* ‘to set oneself free’ [KESK: 171]; PK *\*šebr-* — PUd *\*šob(i)r-* ‘to cover’ [KESK: 318].

PK *\*mēd* — PUd *\*med* ‘pay, payment’ [KESK: 171]; PK *\*nébūd* — PUd *\*nébit* ‘soft’ [KESK: 187].

\* \* \*

Now let us consider Udmurt correspondences to the Proto-Komi phonemes *\*ε* and *\*e*. In Udmurt three types of relevant positions are attested:

- 1) positions where the regular Proto-Udmurt correspondence to Proto-Komi *\*ε* and *\*e* is *\*e*,
- 2) positions where the regular Proto-Udmurt correspondence to Proto-Komi *\*ε* and *\*e* is *\*o*,
- 3) positions where both Proto-Udmurt *\*e* and *\*o* are possible in the place of Proto-Komi *\*ε* and *\*e* with no clear rule of complementary distribution.

1. Positions where the regular Proto-Udmurt correspondence to Proto-Komi *\*ε* and *\*e* is *\*e*:

### Before palatalised consonants (PUd *\*e* — 11 examples):

PK *\*keí* ‘pale’ — PUd *\*keí-* ‘ginger (colour, only of horses)’ [KESK: 121]; PK *\*lećüd* — PUd *\*lećit* ‘pointed’ [KESK: 165]; PK *\*lež-* — PUd *\*lež-* ‘to let, release’ [KESK: 165]; PK *\*peí* — PUd *\*peí* ‘ear’ [KESK: 218–219]; PK *\*pēmūd* — PUd *\*pelmīt* ‘dark’ [KESK: 219]; PK *\*pεž-* —

PUd \*pež- ‘to spring (a trap, snare or spring)’ [KESK: 218]; PK \*rež ~ -ɣ- — PUd \*žež ~ -ž ‘bridle-rein’ [KESK: 240]; PK \*rež- — PUd \*žež- ‘to rip open’ [KESK: 240]; PK \*sež ~ -ɣ- — PUd \*sež ‘clear’ [KESK: 271]; PK \*šed- — PUd \*šed- ‘to hit upon, chance upon’ [KESK: 318]; PK \*vešk- — PUd \*vešk- ‘straight, slim’ [KESK: 54].

**Before m** (PUd \*e — 2 examples):

PK \*čem ~ -ɣ- — PUd \*čem ‘thick, close’ [KESK: 290]; PK \*íem — PUd \*íem ‘glue’ [KESK: 158].

2. Positions where the regular Proto-Udmurt correspondence to Proto-Komi \*ɣ and \*e is \*o:

**Before n and ŋ** (PUd \*o — 5 examples):

PK \*čerañ — PUd \*čonari ‘spider’ [KESK: 303]; PK \*me (\*men-) — PUd \*mon ‘I’ [KESK: 170]; PK \*šen- ~ -ɣ- — PUd \*šon- ‘to wave’ [KESK: 319]; PK \*te (\*ten-) — PUd \*ton ‘you (2<sup>nd</sup> Sg.)’ [KESK: 293–294]; PK \*ženg- ‘to strike, shoot’ — PUd \*žong- ‘to knock, strike’ [KESK: 90].

**Before l** (PUd \*o — 2 examples):

PK \*pelüs — PUd \*polis ‘paddle’ [KESK: 218]; PK \*zel — PUd \*zol ‘tight, strong’ [KESK: 108].

**Before non-palatalised sibilants and affricates** (PUd \*o — 21 examples):

PK \*esk- — PUd \*osk- ‘to believe’ [KESK: 332]; PK \*ež- ~ \*ɣ- ‘to attack’ — PUd \*ož ‘war’ [KESK: 331]; PK \*ɣža — PUd \*ožo ‘sod; grass’ [KESK: 331]; PK \*gɣz (\*gɣzj-) — PUd \*gozi ‘rope’ [KESK: 76]; PK \*kes- ~ -ɣ- — PUd \*kos- ‘to order’ [KESK: 122–123]; PK \*kež- ‘to wait (for), hope’ — PUd \*kož- ‘to suppose, believe’ [KESK: 120]; PK \*kež- postposition — PUd \*kož- ‘alone, by oneself’ [KESK: 121]; PK \*kež- — PUd \*kož- ‘turn away (from a road)’ [KESK: 121]; PK \*mez- — PUd \*moz- ‘to set oneself free’ [KESK: 171]; PK \*pes- — PUd \*pos- ‘to break, crumple’ [KESK: 220]; PK \*pes- ~ -ɣ- — PUd \*pos- ‘to wash’ [KESK: 220]; PK \*pež ~ -ɣ- — PUd \*pož ‘dirty’ [KESK: 218]; PK \*vešt- ‘to buy out’ — PUd \*vošt- ‘to cancel’ [KESK: 54]; PK \*vezj ~ -ɣ- — PUd \*vozj- ‘to stretch, to hang something’ [KESK: 51]; PK \*vež — PUd \*vož ‘green’ [KESK: 49]; PK \*vež — PUd \*vož ‘envy, hatred’ [KESK: 49]; PK \*vež — PUd \*vož ‘crossroads’ [KESK: 49–50]; PK \*vež- — PUd \*voš- ‘(ex)change, replace’ [KESK: 50]; PK \*veža — PUd \*vožo ‘holy, saint’ [KESK: 50]; PK \*žesküd — PUd \*žosküt ‘tight’ [KESK: 90]; PK \*žež ~ -ɣ- — PUd \*žoz(j) ‘grasshopper’ [KESK: 288].

3. Positions where both Proto-Udmurt \*e and \*o are possible in the place of Proto-Komi \*ɣ and \*e:

**Before non-palatalised stops** (PUd \*o — 12 examples, PUd \*e — 12 examples):

PK \*bed (\*bedj-) — PUd \*bodj ‘stick, cane’ [KESK: 38]; PK \*čeg- ~ -ɣ- ‘a plank on the edge of a bench in a house’ — PUd \*čog ‘stake, peg, nail, rack’ [KESK: 289–290]; PK \*leb- — PUd \*lob- ‘to fly’ [KESK: 165]; PK \*meg ‘river bend’ — PUd \*mog ‘roundabout’ [KESK: 171]; PK \*pet- — PUd \*pot- ‘to come out, to ascend’ [KESK: 220]; PK \*reḡüd — PUd \*žog ‘soon’ [KESK: 240]; PK \*rekt- — PUd \*žokt- ‘to unload’ [KESK: 240]; PK \*rektan ~ -ɣ- — PUd \*žokton ‘reel’ [KESK: 240]; PK \*šet- — PUd \*šot- ‘to give’ [KESK: 251]; PK \*šebr- — PUd \*šob(j)r- ‘to cover’ [KESK: 318]; PK \*vetòs ~ -ɣ- — PUd \*votes ‘spider web, snare’ [KESK: 54]; PK \*žeb ~ -ɣ- ‘weak, puny’ — PUd \*žob ‘dirty, muddy, evil’ [KESK: 102].

PK \**ćécć-* — PUd \**tekć-* ‘to jump’ [KESK: 304]; PK \**ćepól-* — PUd \**ćepjíl-* ‘to pinch’ [KESK: 303]; PK \**geg-* ~ *-e-* — PUd \**keg-* ‘to freeze, feel chilly’ [KESK: 75]; PK \**ĺek* (\**ĺekj-*) ~ *-e-* — PUd \**ĺekj* ‘ankle-bone, knucklebone’ [KESK: 158]; PK \**med* — PUd \**med* ‘let, may, in order to’ [KESK: 171]; PK \**męd* — PUd \**med* ‘pay, payment’ [KESK: 171]; PK \**ńębűd* — PUd \**ńębit* ‘soft’ [KESK: 187]; PK \**śępus* ~ *-e-* — PUd \**śępjs* ‘sack’ [KESK: 250]; PK \**śęg* ‘ankle-bone, ankle’ — PUd \**śęg* ‘sprout, outgrowth (on a tree)’ [KESK: 318]; PK \**śęp* — PUd \**śęp* ‘ear (of grain)’ [KESK: 319]; PK \**vek-* — PUd \**vek-* ‘narrow’ [KESK: 51]; PK \**vetl-* — PUd \**vetl-* ‘to walk’ [KESK: 54].

**Before *r*** (PUd \**o* — 15 examples, PUd \**e* — 10 examples):

PK \**berd* — PUd \**bord(i)* ‘wall; measure of length’ [KESK: 39]; PK \**ćerig* — PUd \**ćorig* ‘fish’ [KESK: 303–304]; PK \**er-* — PUd \**or-* ‘to roar, curse’ [KESK: 332]; PK \**ker* (\**kerj-*) — PUd \**kor* ‘log, beam’ [KESK: 121]; PK \**ker-* — PUd \**kor-* ‘to chop’ [KESK: 121–122]; PK \**ńęród* — PUd \**ńęrod* ‘gland’ [KESK: 187]; PK \**śęr* ~ *-e-* — PUd \**śęor* ‘marten’ [KESK: 250]; PK \**śęrd-* — PUd \**śęordj-* ‘to sift’ [KESK: 319]; PK \**ver* — PUd \**vor-* ‘man’ (in Udmurt *vorgoron* ‘man’) (cf. [KESK: 52]); PK \**ver-* — PUd \**vor-* ‘to flicker, ripple’ [KESK: 52]; PK \**verd-* — PUd \**vord-* ‘to feed, rear’ [KESK: 52]; PK \**verd-* ~ *-e-* — PUd \**vord-* ‘to give birth to’ [KESK: 52]; PK \**verm-* — PUd \**vorm-* ‘to overcome, win’ [KESK: 53]; PK \**verzóm* — PUd \**vorżem* ‘(to ride) on horseback’ [KESK: 52–53]; PK \**zer* — PUd \**zor* ‘rain’ [KESK: 108].

PK \**berg-* — PUd \**berg-* ‘to revolve’ [KESK: 38–39]; PK \**beris* ~ *-e-* — PUd \**beriz* ~ *-z* ‘lime-tree’ [KESK: 39]; PK \**perób* ~ *-e-* ‘a moss-covered riverbank in a wood’ — PUd \**pera* ‘a damp black earth’ [KESK: 220]; PK \**ser-* ~ *-e-* ‘to hoe’ — PUd \**ser-* ‘to untangle’ [KESK: 272]; PK \**sera* ~ *-e-* — PUd \**sereg* ‘firebrand’ [KESK: 272]; PK \**seróg* ~ *-e-* — PUd \**sereg* ‘corner’ [KESK: 272]; PK \**śęr-* — PUd \**śęr-* ‘to laugh’ [KESK: 251]; PK \**śęról* — PUd \**śęrel* ‘nit’ [KESK: 251]; PK \**śęrmód* — PUd \**śęrmet* ‘bridle-rein’ [KESK: 251]; PK \**śęrg-* ~ *-e-* — PUd \**śęr-* ‘to move apart’ [KESK: 319].

This “partial complementary distribution” can be explained in two ways: 1) PUd \**e* and \**o*, corresponding to Proto-Komi \**e* и \**ę*, go back to a single Proto-Permic phoneme, and the double reflex of this phoneme before non-palatalised stops and *r* is conditioned by some unknown factor; 2) PUd \**e* и \**o* in the words analyzed above reflect two different Proto-Permic vowels: this opposition is preserved only in Udmurt and only before non-palatalised stops and *r*. The first explanation seems more plausible, because we have no external data which could corroborate the antiquity of the Proto-Udmurt opposition \**e* vs. \**o*.

Now we may turn to the Proto-Permic vowels \**ε* и \**ö*. According to Lytkin’s reconstruction, those Proto-Permic phonemes merged in Komi, yielding Proto-Komi \**ö*. In Udmurt their reflexes are, however, different: Proto-Permic \**ε* yields PUd \**e*, and Proto-Permic \**ö* yields PUd \**o* (but \**a* before \**l* not after a dental) (see [Lytkin 1964: 128–129]).

An examination of correspondences to Proto-Komi \**ö* in Udmurt reveals that here as well the Proto-Udmurt \**e* and \**o* are in a “partial complementary distribution” like the one shown above (but not identical). In this case we can also distinguish between three types of positions:

- 1) positions in which the regular Proto-Udmurt correspondence to Proto-Komi \**ö* is \**e*,
- 2) positions in which the regular Proto-Udmurt correspondence to Proto-Komi \**ö* is \**o*,
- 3) positions in which both Proto-Udmurt \**e* and \**o* are possible in the place of Proto-Komi \**ö* with no clear rule of complementary distribution.

## 1. Positions in which the regular Proto-Udmurt correspondence to Proto-Komi \*ó is \*e:

### Before palatalised consonants (PUd \*e — 11 examples):

PK \*cój ~ -ǫ̇- — PUd \*céj ‘bark beetle’ [KESK: 234]; PK \*kóc ~ -ǫ̇- — PUd \*kec ‘hare’ [KESK: 143]; PK \*kóc- ~ -ǫ̇- ‘rinse out’ — PUd \*kec- ‘to tilt to one side, overturn; to pour out’ [KESK: 143]; PK \*kój ~ -ǫ̇- ‘a worm (in meat)’ — PUd \*kej ‘moth’ [KESK: 139]; PK \*kóz- ~ -ǫ̇- — PUd \*kež- ‘cold’ [KESK: 139]; PK \*lós- — PUd \*leš- ‘to approach’ [KESK: 163]; PK \*póc ~ -ǫ̇- ‘grandmother’ — PUd \*pec ‘old’ [KESK: 230]; PK \*pòim — PUd \*peń (\*peńm-) ‘ashes’ [KESK: 228]; PK \*pól- — PUd \*pel- ‘to blow’ [KESK: 228]; PK \*róc ~ -ǫ̇- ‘lump, knob’ — PUd \*žec ‘humerus’ [KESK: 244]; PK \*vól- ~ -ǫ̇- — PUd \*vel- ‘just now’ [KESK: 66–67].

**One exception** is PK \*lól ~ -ǫ̇- — PUd \*lóló ‘milk’ (a nursery word) [KESK: 164].

### Before s (PUd \*e — 5 examples):

PK \*còskud — PUd \*ceskít ‘tasty’ [KESK: 312]; PK \*mòs (\*mòsk-) — PUd \*mes ‘cow; female’ [KESK: 176]; PK \*òz- — PUd \*est- ‘to burn’ [KESK: 210]; PK \*sòs — PUd \*ses ‘unclean’ [KESK: 264]; PK \*vòs- — PUd \*ves- ‘thin’ [KESK: 68].

## 2. Positions in which the regular Proto-Udmurt correspondence to Proto-Komi \*ó is \*o:

### Before non-palatalised sibilants and affricates, except s (PUd \*o — 12 examples):

PK \*cóc — PUd \*čoš ‘together’ [KESK: 290]; PK \*cóz ~ -ǫ̇- — PUd \*čož postposition ‘during, throughout’ [KESK: 310]; PK \*józ ~ -ǫ̇- — PUd \*joz ‘joint, articulation’ [KESK: 112]; PK \*józ ~ -ǫ̇- — PUd \*jočko ‘raw, uncooked’ [KESK: 112]; PK \*nòšta ~ -ǫ̇- — PUd \*noš ‘again’ [KESK: 195]; PK \*ròž ~ -ǫ̇- ‘slowly’ — PUd \*róžj- ‘to drawl’ [KESK: 198]; PK \*òš (\*òšk-) — PUd \*oš ‘bull’ [KESK: 213]; PK \*òš- — PUd \*oš- ‘to hang’ [KESK: 213]; PK \*òšmòs — PUd \*ošmes ‘well, spring’ [KESK: 213]; PK \*òžùn ~ \*ǫ̇- — PUd \*ozon ‘berth’ [KESK: 210]; PK \*òžùn ~ \*ǫ̇- — PUd \*ožon ‘chisel’ [KESK: 210]; PK \*ròz ~ -ǫ̇- — PUd \*žoz ‘madder’ [KESK: 243].

### Before l (PUd \*a — 10 examples, PUd \*o (after a dental) — 5 examples):

PK \*cól — PUd \*cal ‘quiet’ [KESK: 310]; PK \*cól- ~ -ǫ̇- ‘to cut a slice’ — PUd \*cal- ‘to hew a stake’ [KESK: 311]; PK \*kól (\*kólj-) — PUd \*kal ‘cord’ [KESK: 137]; PK \*kòlt- ~ -ǫ̇- — PUd \*kalt- ‘to catch, to fish out (using a net), to take out with a ladle (e.g. meat from a kettle)’ [KESK: 138]; PK \*mòlp- ~ -ǫ̇- — PUd \*malp- ‘to think’ [KESK: 175]; PK \*ól- — PUd \*al- ‘to prohibit, repress, soothe’ [KESK: 211]; PK \*pól — PUd \*pal ‘half’ [KESK: 227]; PK \*sól- — PUd \*sal- ‘to spit’ [KESK: 270]; PK \*sòla ~ -ǫ̇- — PUd \*sala ‘hazel-grouse’ [KESK: 270]; PK \*vól — PUd \*val ‘horse’ [KESK: 65].

PK \*dól — PUd \*dol ‘all, everything’ [UEW: 518]; PK \*dòlal- ~ -ǫ̇- — PUd \*dolal- ‘to flutter’ [KESK: 96]; PK \*tól — PUd \*tol ‘winter’ [KESK: 283]; PK \*tòlòs ~ -ǫ̇- — PUd \*toles ‘a two-year old ram’ [KESK: 283]; PK \*tòliús — PUd \*tolež ‘moon’ [KESK: 283].

**One exception** is PK \*kòld- ‘kolobok (a kind of pasty or pie)’ — PUd \*kold-, \*kòld- ‘spring wheat’ [KESK: 137–138].

3. Positions in which both Proto-Udmurt *\*e* and *\*o* are possible in the place of Proto-Komi *\*ö* with no clear rule of complementary distribution:

**Before non-palatalised nasals** (PUd *\*o* – 6 examples, PUd *\*e* – 5 examples):

PK *\*gòn* – PUd *\*gon* ‘feather, wool’ [KESK: 80]; PK *\*lòm* ~ -*ò*- – PUd *\*lom* ‘scab, mange, tetter’ [KESK: 162]; PK *\*pòn*- ~ -*ò*- ‘to copulate’ – PUd *\*pon*-, *\*pòn*- ‘to put’ [KESK: 228]; PK *\*ròm* – PUd *\*žom*- ‘colour’ [KESK: 243–244]; PK *\*ròmiž* ~ -*ò*- – PUd *\*žomes*- ‘cud’ [KESK: 244]; PK *\*ròmud* ~ -*ò*- – PUd *\*žomit* ‘twilight’ [KESK: 244].

PK *\*dònz*- ~ -*ò*- ‘constitute one’s share’ – PUd *\*demz*- ‘to hit against, to catch on, to get stuck’ [KESK: 96]; PK *\*kòn* – PUd *\*kenem* ‘hemp’ [KESK: 141]; PK *\*kòn* ~ -*ò*- – PUd *\*ken* ‘female (of a wood- or black grouse)’ [KESK: 141]; PK *\*òmiž* – PUd *\*emež* ‘raspberry’ [KESK: 211]; PK *\*zòm* ‘steep; straight’ – PUd *\*zem* ‘truth, rightly; truthful’ [KESK: 106].

**Before non-palatalised stops** (PUd *\*o* – 13 examples, PUd *\*e* – 11 examples):

PK *\*čòk*- ~ -*ò*- – PUd *\*čok*- ‘to close’ [KESK: 290]; PK *\*gòg* (*\*gògj*-) – PUd *\*gogj* ‘navel’ [KESK: 79]; PK *\*gòp* (*\*gòpt*-) – PUd *\*gop* ‘pit, hole’ [KESK: 80]; PK *\*jòg* ~ -*ò*- – PUd *\*log* ‘outgrowth (on a tree)’ [KESK: 112]; PK *\*kòt*- – PUd *\*kot* ‘wet’ [KESK: 143]; PK *\*nòd* ~ -*ò*- ‘riddle’ – PUd *\*nod* ‘sharpness, quick wit’ [KESK: 194]; PK *\*òk*- – PUd *\*ok*- ‘to gather’ [KESK: 211]; PK *\*òtik* – PUd *\*odig* ‘one’ [KESK: 212]; PK *\*pòd*- – PUd *\*pod*- ‘to close’ [KESK: 227]; PK *\*pòt* ~ -*ò*- ‘winterweed (plant)’ – PUd *\*pot* ‘goose-foot (plant)’ [KESK: 230]; PK *\*ròp*- ~ -*ò*- – PUd *\*žopi* ‘smoke-hole in a bathhouse’ [KESK: 244]; PK *\*sòp*- ~ -*ò*- ‘to eat, gorge oneself’ – PUd *\*sop*- ‘to champ’ [KESK: 264]; PK *\*tòd*- – PUd *\*tod*- ‘to know’ [KESK: 283].

PK *\*čòkt*- – PUd *\*čekt*- ‘to order, command’ [KESK: 290]; PK *\*jòkt*- – PUd *\*ekt*- ‘to dance’ [KESK: 113]; PK *\*kòkan* ~ -*ò*- – PUd *\*kekon* ‘hellebore’ [KESK: 140]; PK *\*mòd*- – PUd *\*med*- ‘to gather’ [KESK: 176]; PK *\*ògür* – PUd *\*egür* ‘coal’ [KESK: 209]; PK *\*òksù* – PUd *\*eksej* ‘prince, tsar’ [KESK: 211]; PK *\*pòk* (*\*pòkj*-) ‘caviar; calves (of the leg)’ – PUd *\*pek(i)la* ‘kidneys; testicles’ [KESK: 228]; PK *\*sòp* (*\*sòpt*-) – PUd *\*sep* ‘bile’ [KESK: 263]; PK *\*sòkud* ~ -*ò*- – PUd *\*šekjt* ‘hard, difficult’ [KESK: 270]; PK *\*sòp* ~ -*ò*- – PUd *\*šep* ‘widow(er)’ [KESK: 270]; PK *\*tòb*- – PUd *\*teb*- ‘to wrap’ [KESK: 283].

**Before *r*** (PUd *\*o* – 8 examples, PUd *\*e* – 11 examples):

PK *\*gòrd* – PUd *\*gord* ‘red’ [KESK: 80]; PK *\*jòr* ‘pen, enclosure’ – PUd *\*joros* ‘place, surroundings’ [KESK: 113]; PK *\*kòr* – PUd *\*kor*- ‘taste’ [KESK: 141]; PK *\*kòrt* – PUd *\*kort* ‘iron’ [KESK: 142]; PK *\*sòr*- ~ -*ò*- ‘ease oneself, go to the john’ – PUd *\*šor*- ‘loo’ [KESK: 271]; PK *\*šòr* – PUd *\*šor* ‘middle’ [KESK: 322]; PK *\*šòr*- – PUd *\*šor*- ‘to cut (a piece, a slice of bread)’ [KESK: 322–323]; PK *\*šòrt* – PUd *\*šort* ‘yarn’ [KESK: 323].

PK *\*bòr* – PUd *\*ber* ‘rear part, back’ [KESK: 41]; PK *\*dòra* – PUd *\*dera* ‘canvas’ [KESK: 96]; PK *\*gòr* (*\*gòrj*-) – PUd *\*gerj* ‘wooden plow’ [KESK: 80]; PK *\*gòròd* – PUd *\*gerd* ‘knot’ [KESK: 81]; PK *\*kòrt*- ~ -*ò*- – PUd *\*kert*- ‘to tie, bind’ [KESK: 142]; PK *\*kòrtum* ~ -*ò*- ‘rent’ – PUd *\*kertim* ‘tribute’ [KESK: 142]; PK *\*pòruš* – PUd *\*peres* ‘old’ [KESK: 229]; PK *\*pòrtc*- – PUd *\*pertc*- ‘to untie’ [KESK: 229]; PK *\*sòr* ~ -*ò*- – PUd *\*ser* ‘vein, blood-vessel’ [KESK: 264]; PK *\*sòr*- ~ -*ò*- ‘to be delirious (in one’s sleep); to chatter, babble, talk nonsense’ – PUd *\*serek* ‘hot, hot-tempered, anxious’ [KESK: 264]; PK *\*tòr*- – PUd *\*ter*- ‘to fit, to fit into’ [KESK: 284].



It is evident that, despite some particularities, the status of the Udmurt correspondences to the Proto-Komi vowel *\*o* is similar to that of Udmurt correspondences to Proto-Komi *\*e* and *\*e̞*. This means that it is not possible to accept V.I. Lytkin's reconstruction, where in one case the difference between Udmurt vowels *\*o* and *\*e* is regarded as secondary, and in the other case it is traced back to the Proto-Permic period. It is necessary either to consider this difference as Proto-Permic in both cases or to regard it as always secondary, arising under the influence of the following consonants and some other yet unknown factors. We believe that the second solution is more appropriate, because, as well as in the previous case, we do not have at our disposal any external data that could confirm the antiquity of the Proto-Udmurt contrast between *\*e* and *\*o*. It must be noted that, although P. Sammallahti, following V. I. Lytkin, preferred to reconstruct here two different Proto-Permic phonemes, at the same time he pointed out that they resulted from a "non-conditional" split of a single Pre-Permic phoneme *\*o* [Sammallahti 1988: 527]. As we have shown above, this split was at least partially conditioned by the phonetic environment.

Let us turn to the Proto-Permic phonemes *\*o̞* and *\*ö̞*. According to Lytkin, they fell together in Komi as Proto-Komi *\*o̞*. The difference between the two phonemes remained in Proto-Udmurt, where Proto-Permic *\*o̞ > \*u*, and Proto-Permic *\*ö̞ > \*ü*. Lytkin also takes into consideration the secondary shift *\*u > \*ü* before palatalised consonants in Proto-Udmurt. Thus, the assumed opposition between Proto-Permic *\*o̞* and *\*ö̞* remained intact only before non-palatalised consonants. However, examination of the relevant evidence<sup>4</sup> allows to conclude that in this position a complementary distribution between *\*o̞* and *\*ö̞* can be shown as well. The CVC-type stems regularly contain *\*ü*, while CVCVC-type stems always show *\*u*. Only a handful of examples of other stem types (CV, CVCi/i, CVCo) is available; they do not warrant the reconstruction of an opposition between *\*o̞* and *\*ö̞* in Proto-Permic.

#### CVC-type stems (PUd *\*ü* – 14 examples):

PK *\*døn* – PUd *\*dün* 'clean' [KESK: 95]; PK *\*døz* 'crocery' – PUd *\*düz* 'vat' [KESK: 94]; PK *\*gørz-* 'to cry' – PUd *\*gür bõrd-* 'to sob' [KESK: 78]; PK *\*gørzõd-* – PUd *\*gürz-* 'to belch' [KESK: 78]; PK *\*gøz* – PUd *\*küz* 'couple' [KESK: 77]; PK *\*kør* 'husk' – PUd *\*kür* 'lime-bast' [KESK: 133]; PK *\*køs* – PUd *\*küs* 'waist' [KESK: 134]; PK *\*løl* – PUd *\*lül* 'spirit' [KESK: 160]; PK *\*ør* – PUd *\*ür* 'pus' [KESK: 206]; PK *\*ør-* 'tear' – PUd *\*ür-* 'tear apart' [KESK: 207]; PK *\*oraš-* – PUd *\*ür-* 'to weed' [KESK: 206]; PK *\*pøž* – PUd *\*püž* 'sieve' [KESK: 223]; PK *\*šøm* 'coal' – PUd *\*šüm* 'soot' [KESK: 321]; PK *\*tøš* – PUd *\*tüš* 'beard' [KESK: 282].

**One exception** is PK *\*ørt* 'ghost' – PUd *\*urt* 'spirit' [KESK: 208].

#### CVCVC-type stems (PUd *\*u* – 6 examples):

PK *\*cøriud* – PUd *\*cürjt* 'hard' [KESK: 310]; PK *\*kømiž* 'spring onion' – PUd *\*kumiž* 'wild garlic' [Lytkin 1964: 47]; PK *\*kømiul* 'rind' – PUd *\*kumel* 'the bark on a piece of bast' [KESK: 132]; PK *\*køpürt-* – PUd *\*kupjrt-* 'bend' [KESK: 133]; PK *\*nøžtòm* 'ugly' – PUd *\*núžtem* 'naughty' [KESK: 188; Lytkin 1964: 55]; PK *\*øzür* – PUd *\*uzjř* 'rich' [KESK: 203].

<sup>4</sup> We do not adduce here those words in which Komi data show fluctuation between reflexes of Proto-Komi *\*o̞* and *\*u*, since we cannot exclude that such words contained Proto-Permic *\*u* or *\*ü* rather than *\*o̞* or *\*ö̞*.



**CV-type stems** (PUd \**u* – 2 examples):

PK \**śq* – PUd \**śu* ‘hundred’ [KESK: 252]; PK \**vq*- – PUd \**vu*- ‘arrive’ [KESK: 63].

**CVCi/i-type stems** (PUd \**ü* – 3 examples, PUd \**u* – 1 example):

PK \**gq̄b* – PUd \**gübi* ‘mushroom’ [KESK: 76]; PK \**pq̄n* – PUd \**püni* ‘dog’ [KESK: 224]; PK \**śqr* ‘pole’ – PUd \**śüri* ‘perch, beam’ [KESK: 253].

PK \**qz* – PUd \**uzi* ‘wild strawberry’ [KESK: 203].

**CVCo-type stems** (PUd \**u* – 2 examples, PUd \**ü* – 1 example):

PK \**kq̄k*- ‘to hoe’ – PUd \**kukćo* ‘hoe’ [KESK: 129]; PK \**pq̄da* – PUd \**pudo* ‘cattle’ [KESK: 223].

PK \**kq̄ča* – PUd \**küčo* ‘many-coloured’ [Lytkin 1964: 53].

The revised set of correspondences between Komi and Udmurt can be interpreted in the following way:

Proto-Permic	*i	*ü	*j̄	*u	*e	*ó	*ę	*o	*ä	*ǔ	*a	*ɔ
Proto-Komi	*i	*u	*ü	*u	*ē, *e	*ó	*ę̄	*o	*ö	*ǔ̄	*a	*a
Proto-Udmurt	*i	*ü	*j̄	*u	*e, *o	*ü, *u	*j̄	*u	*e, *o, *a	*ö	*a	*u

## Notes:

1) Udmurt \**u* of any origin changes to \**ü* before palatalised consonants.

2) Udmurt \**u* and \**ü* from Proto-Permic \**ü*, \**u*, \**ó* и \**o* may yield \**j̄* under uncertain conditions. Following V. I. Lytkin and P. Sammallahti, we do not think that in those cases it is necessary to reconstruct separate Proto-Permic phonemes, first of all, because of the great number of parallel forms with \**u* and \**j̄* in Udmurt (see [Lytkin 1964: 21]).

The system of the Proto-Permic vocalism in our reconstruction proves to be less complicated than the ones assumed by Lytkin and Sammallahti, since there are only three vowel heights and no reduced vowels:

	front	central	back	
	non-labialised	labialised	non-labialised	labialised
high	i	ü	j̄	u
mid	e	ó	ę	o
low	ä	ǔ	a	ɔ

### Abbreviations

PFP — Proto-Finno-Permic; PK — Proto-Komi; PU — Proto-Uralic; PUd — Proto-Udmurt

### Literature

KESK — В. И. ЛЫТКИН, Е. И. ГУЛЯЕВ. *Краткий этимологический словарь коми языка*. М., 1970.

Lytkin 1964 — В. И. ЛЫТКИН. *Исторический вокализм пермских языков*. М., 1964.

Sammallahti 1988 — P. SAMMALLAHTI. Historical phonology of the Uralic languages // SINOR D. (ed.). *The Uralic Languages: Description, History and Foreign Influences*. Leiden e. a., 1988. P. 478–554.

UEW — K. RÉDEI. *Uralisches Etymologisches Wörterbuch*. Bd. I–III. Budapest, 1986–1991.

В статье рассматриваются некоторые прапермские гласные, реконструированные В. И. Лыткиным в его фундаментальной монографии «Исторический вокализм пермских языков». Установленные нами правила дополнительного распределения позволяют предложить более экономичную реконструкцию прапермского вокализма.

### Типологическая схожесть языков как метод изучения языковой эволюции<sup>1</sup>

Существующие типологические классификации языков опираются на один или небольшое число признаков. С появлением больших типологических баз данных, таких как The World Atlas of Language Structures или «Языки мира», становится возможным строить классификации, учитывающие сотни и тысячи признаков одновременно. Для построения классификаций могут быть применены различные математические алгоритмы, что создает предпосылки создания объективной классификации. В статье приводятся первые результаты в этом направлении. К числу наиболее интересных неожиданных результатов можно отнести типологическое различие языков Африки и Евразии, а также типологическую близость чукотско-камчатских языков с языками синокавказской макросемьи. Предложено несколько (близких) вариантов классификации языков Северной Африки, Европы и Северной и Центральной Азии. Обсуждается возможность использования этого подхода для установления сверхдального родства языков, основанная на двух соображениях. Во-первых, близость одновременно по сотням признаков статистически не может быть результатом случайного совпадения. Во-вторых, по крайней мере, некоторые грамматические свойства являются очень стабильными и несут информацию о языковых состояниях древности.

*Keywords:* типологические классификации, типологические базы данных, математические алгоритмы кластеризации, гипотезы сверхдального родства

#### 1. Введение

В любой научной дисциплине классифицируются изучаемые объекты. Хорошо известна таксономическая классификация всех живых существ в биологии. Предмет лингвистики — языки также могут быть классифицированы, причем разными способами. Наиболее продвинутой является генеалогическая классификация языков. Другой подход к классификации — по степени похожеści структуры языков вне зависимости от их происхождения — менее развит. Видимо, первой серьезной попыткой классификации языков по структуре является морфологическая классификация, построенная во второй половине 19-го века в работах Шлегеля, Гумбольдта и Шлейхера [Кузнецов 1954]. Эта классификация сохраняет свое значение и поныне, однако, она учитывает лишь один аспект языковой структуры — способ соединения морфем, так что языки, попавшие по этой классификации в один класс, могут радикально отличаться друг от друга в других аспектах.

---

<sup>1</sup> Работа выполнена при поддержке РФФИ, грант № 10-06-00087а.

Современная типология акцентирует внимание на корреляциях логически независимых типологических признаков и выделяет пучки таких признаков, которые также могут послужить основой для классификации языков. Детально изученный пример — порядок глагола (V) и прямого дополнения (O). Этот параметр может принимать разные значения: VO, OV, свободный порядок. Оказалось, что порядок V и O строго коррелирует с наличием в языке предлогов или послелогов и целым рядом других параметров, связанных с порядком слов и морфем [Comrie 1989]. Но и в этом случае, порядок глагола и прямого дополнения не является доминирующим, определяющим все или хотя бы многие свойства языка.

Остается неясным, возможна ли холистическая классификация языков, разбивающая все языки на несколько непересекающихся групп, так, что внутри групп языки типологически однородны, а между языками разных групп имеются существенные типологические различия, причем, по широкому спектру признаков, охватывающему все основные языковые уровни. Такая классификация будет невозможна, если окажется, что языки равномерно заполняют все пространство возможных комбинаций типологических свойств, так что любые границы между группами языков будут чисто условны. Априори наиболее вероятным представляется наличие нескольких центров сгущения языков, между которыми располагаются языки с промежуточными свойствами. Группы языков вокруг центров сгущения, если таковые будут найдены, и образуют типологические типы в холистической классификации.

К попытке начать развивать такой подход можно отнести предложение Трубецкого [Трубецкой 1987] трактовать индоевропейские языки, как ставшие родственными в результате конвергенции, когда они приобрели 6 общих признаков: отсутствие сингармонизма, консонантизм начала слова не беднее консонантизма середины и конца слова, наличие приставок, наличие аблаутных чередований гласных, наличие чередования согласных в грамматических формах (sandhi), аккузативность. При последовательном развитии этого подхода можно было бы попытаться найти аналогичные описания и для других семей. Так тюркские языки можно было бы охарактеризовать как агглютинативные, с финальным положением глагола, сингармонизмом и какими-то еще свойствами.

Этот подход оказался забыт по ряду причин. Прежде всего, отвергается базовая идея Трубецкого о причинах родственности индоевропейских языков. Далее, были обнаружены языки, которые обладают всеми 6 вышеперечисленными свойствами, но не относятся к индоевропейской семье. Наконец, в рамках подхода, связывающего группу языков с набором характеристических свойств, просто не было получено каких-либо значимых результатов, которые бы привлекли внимание широкого круга лингвистов.

Важный момент, на который следует обратить внимание, состоит в следующем. До последнего времени у лингвистов фактически не было возможности сравнивать языки одновременно по десяткам и тем более сотням и тысячам различных грамматических свойств. Человеческий мозг не способен учитывать одновременно такое число абстрактных переменных. Ситуация радикально изменилась с появлением больших типологических баз данных, таких как The World Atlas of Language Structures [WALS] и «Языки мира» [Поляков & Соловьев 2006], а также адекватных средств обработки больших массивов грамматических данных.

На повестку дня вновь может быть поставлен вопрос о холистической классификации языков в вышеописанном смысле, причем классификация языков может быть осуществлена на основе одновременного учета сотен и более грамматических свойств с применением объективных математических и компьютерных методов анализа.

Если таким способом будет получено некоторое разбиение языков на классы типологически близких языков, то на повестку дня встанет вопрос: чем объясняется эта бли-

зость — общностью происхождения или сближением структуры языков за счет заимствований при длительных ареальных контактах. Теоретически схожесть может оказаться и случайной, однако, если речь идет о сотнях и тысячах признаков, то крайне маловероятно, чтобы столько параметров (большинство из которых независимы друг от друга) приняли близкие значения.

Если типологически близкие языки находятся в далеко отстоящих друг от друга регионах и отсутствуют какие-либо данные (археологические, генетические) об их контактах в древности, то заимствования между ними также будут маловероятны, что позволит предположить общность происхождения. Таким образом, изучение типологической близости языков с помощью типологических баз данных и математических алгоритмов может послужить дополнительным методом в изучении языкового родства.

Вероятно, пионерскими работами в этом русле является статьи Б. Комри и М. Сисоу [Comrie & Cysouw 2006; Cysouw & Comrie 2009]. В данной статье будут приведены первые результаты, полученные в этом направлении.

## 2. Данные и методы

В настоящее время существует целый ряд типологических баз данных, современный обзор приведен в [Everaert & Musgrave 2009]. Все они, кроме двух, являются специализированными, т. е. посвящены определенным аспектам структуры языков. Лишь две базы данных — WALС и «Языки мира» — охватывают все основные разделы грамматики.

WALS — законченный в 2005 г. крупный международный проект, выполненный под руководством М. Хаспельмата, Б. Комри и др. Он включает компьютерную базу данных, доступную на CD-дисках и через Интернет, а также бумажное издание. WALS содержит данные по 2560 языкам (из всех семей) и 142 признакам, из которых 128 грамматические. Каждый признак может принимать от 2 до 9 значений, в среднем, примерно, 5. Некоторым недостатком WALS является то, что не все языки описаны по всем признакам, более того, в среднем язык в WALS имеет всего около 45 признаков.

Две отличительные особенности придают проекту большую значимость. Во-первых, для каждого признака построена карта Земного шара, на которой кружочками разного цвета обозначены языки с различными значениями выбранного признака. Хотя идея графического изображения географического распределения признаков предлагалась и ранее, но впервые она была реализована столь масштабно. Во-вторых, база данных снабжена чрезвычайно удобным интерфейсом, позволяющим легко ориентироваться в огромном массиве информации — WALS включает огромное число справочных статей по различным признакам и языкам. Поисковые средства позволяют находить нужную информацию по комбинации признаков, генерировать новые карты с заданными свойствами.

База данных «Языки мира» создана в Институте языкознания РАН по материалам одноименной серии монографий, доступна в Интернете по адресу [www.dblang2008.nagod.ru](http://www.dblang2008.nagod.ru) и подробно описана в [Поляков & Соловьев 2006].

Она содержат описания 315 языков Евразии по 3821 признаку, относящихся к следующим сферам языка: фонетика, морфология, синтаксис. Все признаки бинарные. Бинарность представления означает, что для каждого языка и для каждого признака в базе данных фиксируется только наличие или отсутствие этого признака в языке, но не степень его проявления. Таким образом, с математической точки зрения БД представляет собой прямоугольную бинарную матрицу размером 315×3821, содержащую более миллиона бит информации. Разделяемый WALS и «Языки мира» общий принцип — рав-

ноправие всех признаков, т. е. признаки не делятся на более и менее важные, не вводится никаких весовых коэффициентов.

В базе данных «Языки мира» представлены следующие языковые семьи и языковые сообщества: австроазиатские, австронезийские, алтайские, афразийские, индоевропейские, кавказские, палеоазиатские, синотибетские, уральские, хуррито-урартские языки, а также языки изоляты: айнский, нивхский, бурушаски, шумерский, эламский. Языки Европы, Северной и Центральной Азии описаны очень полно. Австроазиатские, австронезийские, синотибетские группы представлены только отдельными языками.

Программная оболочка позволяет осуществлять не только поиск, но и другие содержательные операции. Для проведения количественных типологических исследований важной является операция сравнения, позволяющая по двум выбранным языкам найти все признаки (и подсчитать их число), по которым эти языки совпадают и по которым они отличаются.

Число признаков, по которым два языка различаются, называется расстоянием между ними (метрика Хемминга или городская метрика). Ясно, что чем расстояние меньше, тем языки типологически ближе друг к другу. Например, расстояние в базе данных «Языки мира» между белорусским и македонским языками (оба славянские) равно 243, а между белорусским и бирманским — 401.

Полученные числовые данные могут быть использованы для выделения кластеров типологически близких языков. Для этих целей применяются методы кластерного анализа, наиболее популярными из которых являются филогенетические алгоритмы. Они активно используются в биологии при построении систематики живых организмов. Пионером в области применения филогенетических алгоритмов в лингвистике является, видимо, Т. Варноу, начавшая их использовать еще в 1997 г. [Warnow 1997].

Алгоритм NeighborNet, применяемый в данной работе, описан в [Bryant et al. 2005]. Он строит звездообразные структуры, в которых классифицируемые объекты располагаются по окружности так, что языки с меньшим расстоянием располагаются ближе друг к другу. Существуют и другие подходы, один из которых будет затронут в разделе 4.

### 3. Типологическая классификация языков с применением филогенетических алгоритмов

#### 3.1. Данные WALS.

В работе [Comrie & Cysouw 2006] изучается типологического разнообразия языков Новой Гвинеи. Выбрано 48 наиболее полно описанных языков этого региона, представляющих практически все традиционно выделяемые языковые группы Новой Гвинеи. Для сравнения их с языками всего мира следующим образом выбрано 47 языков. Сначала выбрано 150 наиболее полно описанных в WALS языков, не более чем по одному из каждой языковой группы, сопоставимой с романскими языками по возрасту. Затем из них случайным образом выделено 47. Далее к выбранным 95 языкам применен алгоритм NeighborNet. Результат приведен на рис. 1. Жирным шрифтом выделены языки Новой Гвинеи.

Хорошо видно, что языки Новой Гвинеи не образуют компактную группу, а рассредоточены практически по всему языковому пространству. Языки Новой Гвинеи Maybrat, Abun, Arapesh близки к африканским языкам Hausa, Sango и Luvale соответственно, язык Wahgi — к языку американских индейцев Yaqui, язык Imonda — к языку австралийских аборигенов Kayardild, языки Hua, Awtuw — к азиатским языкам Burmese и Bu-



rushaski соответственно и так далее. Другими словами, типологическое разнообразие языков Новой Гвинеи вполне сопоставимо с типологическим разнообразием языков всего мира.

Дает ли диаграмма рис. 1 какую-либо информацию о возможной холистической типологической классификации языков? Ясно видно деление всех языков на 2 группы: левая часть диаграммы (от Zulu до Guarani) и правая часть (от Marind до Gooniyand), между которыми располагается большой зазор, указывающий на большую разницу в типологических свойствах языков этих групп. Легко видеть, что языки левой части диаграммы характеризуются порядком VO, а правой — OV. В [Comrie & Cysouw 2006] авторы обращают внимание на это деление, указывая, однако, что это вполне может быть и артефакт описания. И действительно, в WALS непропорционально много признаков посвящено порядку слов: 17 из 142 признаков, т. е. 12%. Это могло послужить причиной такого деления, т. к. порядок глагола и прямого дополнения, как уже отмечалось, коррелирует с другими признаками, связанными с порядком слов.

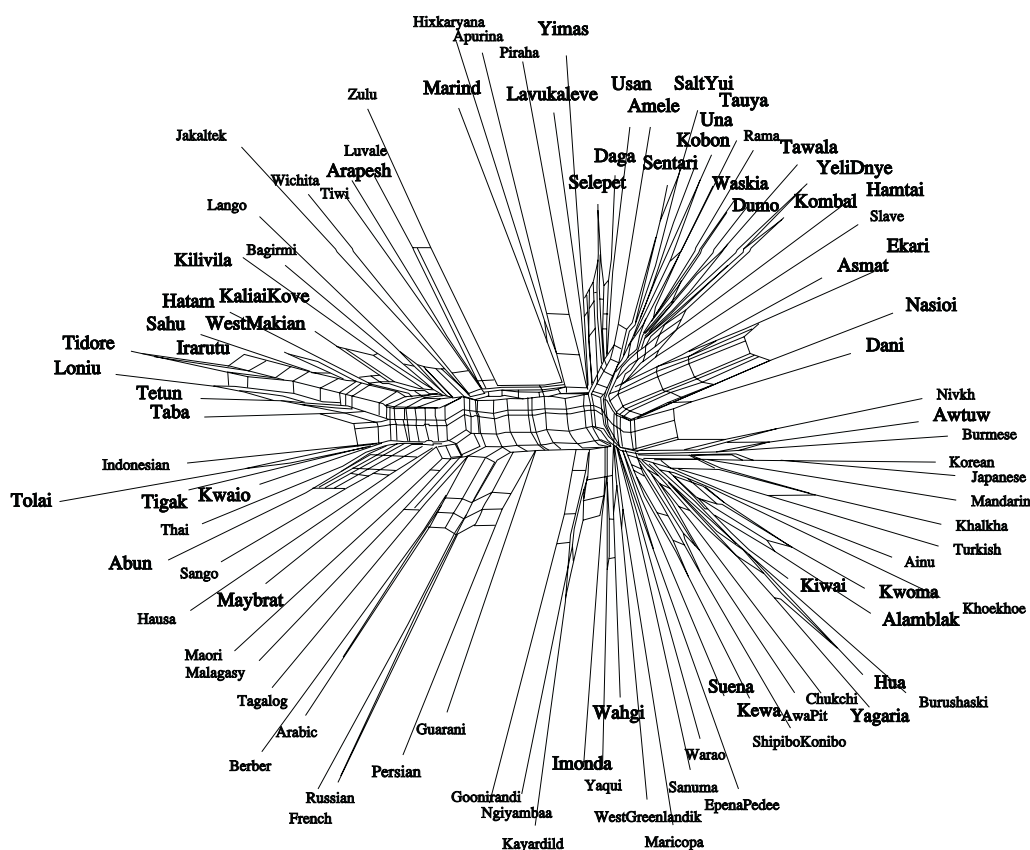


Рис.1. Языки Новой Гвинеи vs. языки остального мира (цит. по: Comrie, Cysouw 2006)

В работе [Cysouw & Comrie 2009] сравнивались языки Африки и Евразии. С помощью алгоритма NeighborNet построена аналогичного вида диаграмма, приведенная на рис. 2.

На ней представлено по одному языку из большинства языковых групп (уровня ветви индоевропейской семьи) этих континентов, включая изоляты, как самостоятельные группы. Здесь ситуация совсем иная — языки Африки (выделены жирным шрифтом), кроме Khoekhoe, группируются отдельно, языки Евразии отдельно. Таким образом, существуют явные типологические различия на континентальном уровне. Причины этого пока не вполне ясны. Но одно из очевидных возможных объяснений состоит в том, что в пределах континентов действуют механизмы конвергенции за счет заимствований.

На этой диаграмме также отражается деление по признаку порядка слов. Африканские языки делятся на 2 группы: справа от Beja до Harar Oromo языки имеют порядок слов VO, вверху от Middle Atlas Berber до Supyire — порядок OV (с несколькими исключениями). Языки Евразии образуют два больших кластера: внизу от Hindi до Persian имеют порядок VO (с несколькими исключениями), слева от Turkish до Mundari — OV. Brahui и Hungarian располагаются посередине между этими группами.

В целом же авторы этих статей отмечают, что представленные данные не выявляют четко выраженных кластеров языков, так что остается большая свобода исследователя в построении типологической классификации.

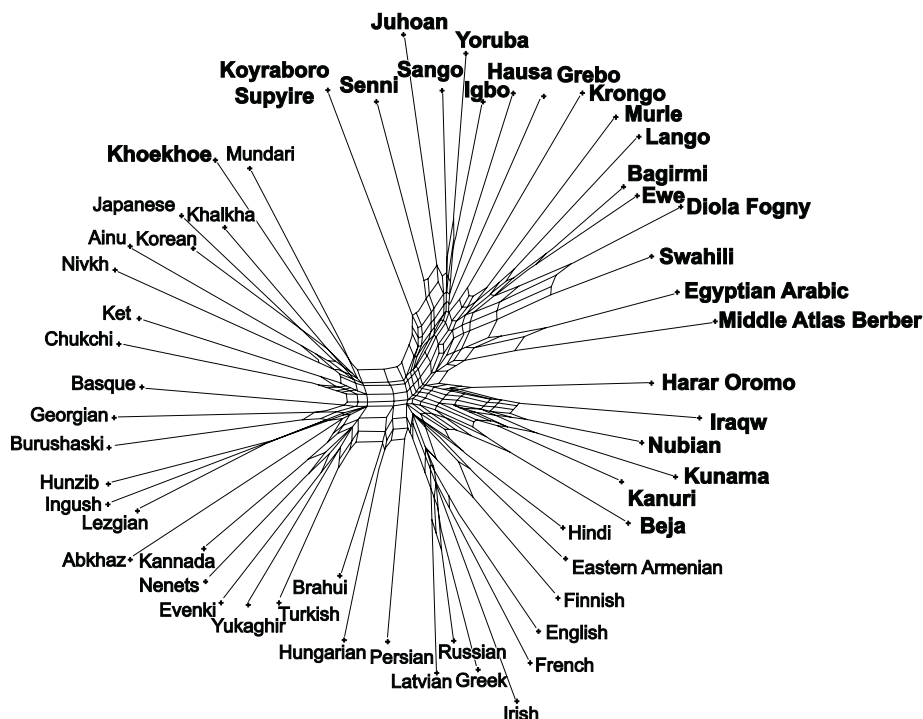


Рис. 2. Языки Африки и Евразии (цит. по: Cysouw, Comrie 2009, p. 194)

### 3.2. Данные «Языки мира».

Для оценки влияния порядка глагола и прямого дополнения мы перепроверим эти результаты с помощью базы данных «Языки мира». Основная идея состоит в применении к тем же языкам того же алгоритма и сравнение результатов. К сожалению, база данных «Языки мира» не содержит описания столь многих, как WALS, языков Африки. В ней на настоящий момент присутствуют лишь афразийские языки. Зато большинство используемых на диаграмме 2 языков Евразии в базе данных «Языки мира» присутствует. В нескольких случаях отсутствующие языки заменены на близкородственные: Eastern Armenian на Armenian, Egyptian Arabic на Modern Arabic, Middle Atlas Berber на Zenaga, Hunsib на Avar, Khalkha на Halh Mongolian, Mundari на Khmer. Результат применения алгоритма кластеризации показан на диаграмме 3.

Опишем основные структурные особенности диаграммы рис. 3.

1. На этой диаграмме расположение языков более равномерное, чем на предыдущих; нет никакого четкого деления на кластеры. Это можно рассматривать как косвенный аргумент в пользу гипотезы моногенеза и/или подтверждение значительного вклада заимствований в уменьшение расстояний между группами языков, расходящимися в ходе представляемой в древовидной форме эволюции.

2. Нет никакого зазора между языками с порядками слов VO и OV. Таким образом, это деление, прослеживающееся на предыдущих диаграммах, является, как отмечалось выше, скорее всего, результатом непропорционально большого числа признаков в WALS, связанных с порядком слов. Более сбалансированный набор признаков базы данных «Языки мира» не обнаруживает этой дихотомии (см. рис. 3).

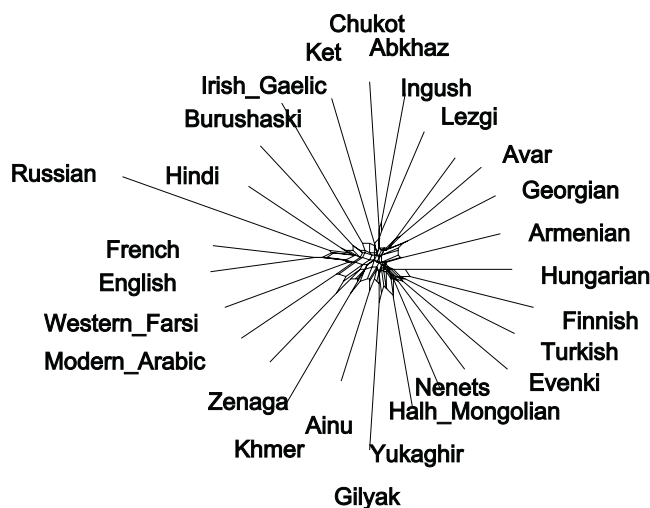


Рис. 3. Языки Африки и Евразии по данным серии «Языки мира»

3. Некоторые языки размещаются на диаграммах этого типа очень странным образом. Например, на рис. 3 ирландский язык (кельтская ветвь индоевропейской семьи) размещен между кетским и бурушаски. Эти языки не являются в действительности типологически близкими, они не являются родственными и располагаются географически очень далеко друг от друга. Вероятной причиной такого размещения является недостаточность данного метода графического представления информации для совершенно точного отображения типологической близости между всеми парами языков. С математической точки зрения база данных «Языки мира» дает представление языков точками в 3821-мерном пространстве признаков. В то же время диаграмма, которую строит NeighborNet, эквивалентна 1-мерному представлению (по окружности). Естественно, что при сворачивании 3821-мерного пространства в 1-мерное могут иметь место искажения.

4. Африканские языки арабский и зенага (афразийская семья) на диаграмме, размещаются рядом друг с другом. Но так как данных слишком мало — всего 2 языка с африканского континента, то трудно говорить о типологическом разделении языков Африки и Евразии на основе этих данных.

5. Проанализируем возможное группирование других языков. Слева четкую группу образуют русский, французский и английский языки, разделяющие как общее происхождение в составе индоевропейской семьи, так и общее географическое положение на европейском континенте. В [Comrie & Cysouw 2006] Б. Комри и М. Сисоу также отмечают типологическую близость русского и французского, выявляемую на диаграмме рис. 1.

Рядом с этими языками располагаются хинди и западный фарси (персидский), что отражает их принадлежность к индоевропейской семье. При этом хинди и персидский находятся на периферии группы индоевропейских языков, что соответствует их более раннему отделению по сравнению с европейской ветвью [Gray & Atkinson 2003] и более далекому географическому положению. Интересно, что хинди и персидский расположены не рядом друг с другом, как можно было бы ожидать на основании их родства в рамках индоиранской ветви, а по разные стороны от европейской части индоевропей-

ской семьи. Точно такое же расположение зафиксировано и на диаграмме рис. 2 по данным WALS. Это указывает на то, что индийские и иранские языки действительно имеют значительные типологические различия, большие, чем славянские, германские и романские. В любом случае, не учитывая ошибочно расположенного ирландского языка, слева на диаграмме языки сгруппировались по общему происхождению.

Справа-внизу располагаются шесть языков, относящихся к алтайской и уральской семьям. Языки этих семей действительно имеют много общих типологических черт. Гипотеза о родственности этих языков поддержана, в частности, в работах [Starostin 2007; Грунтов и др. 2006]. Противники этой гипотезы считают, что похожие черты языков этих семей могут быть следствием ранних (на стадии протоязыков) заимствований. В настоящее время не представляется возможным сказать, какая из этих двух гипотез верна. Можно считать, что эти языки сгруппированы по ареально-генетическому принципу.

Очень интересным является группирование шести языков справа-вверху от абхазского до армянского. Все они располагаются на Кавказе, относясь к 4 различным семьям: индоевропейской, картвельской, нахско-дагестанской и абхазо-адыгской. Таким образом, здесь явно имеет место ареальное группирование. Можно отметить, что грузинский и армянский находятся на одном конце этой группы ближе к урало-алтайским языкам, что коррелирует с их генеалогической общностью в соответствии с ностратической гипотезой [Дыбо 1978]. В то время как абхазский и ингушский языки находятся на другом конце, ближе к кетскому, в соответствии с гипотезой о генеалогической близости северокавказских и енисейских языков [Старостин 2007].

Два языка внизу: гилак (нивхский) и айну являются изолятами. Их генетические связи не установлены, но географически они близки друг к другу — расположены на Дальнем Востоке. Между ними и урало-алтайским филумом (но ближе к первым) расположен юкагирский, который в некоторых работах трактуется как изолят [Николаева & Хелимский 1997], в некоторых относится к уральской семье [Nichols 1992]. Он также располагается на Дальнем Востоке. Т. е. это размещение языков имеет видимые ареальные основы. В этой же части диаграммы находится и кхмерский язык (австроазиатская семья), не имеющий явных ареальных или родственных связей с другими языками на диаграмме. Но так как алгоритм все равно должен был его куда-то поместить, то это место рядом с другими азиатскими изолятами кажется наиболее удачным.

Наконец, в верхней части диаграммы расположено еще несколько изолятов — бурушаски, кетский и чукотский. Бурушаски расположен рядом с хинди, к которому он и географически близок, так что здесь возможны контактные влияния. В [Эдельман 1997] отмечены значительные заимствования в бурушаски из индоиранских языков. Кетский язык расположен близко к группе северокавказских языков, что коррелирует с уже упомянутой гипотезой о дальнем родстве синокавказских и енисейских языков в пределах бореальной макросемьи.

Чукотский язык сам по себе не изолят, а входит в небольшую чукотско-камчатскую семью, но генеалогические связи этой семьи не установлены. Так что чукотский язык не имеет установленных родственных связей ни с каким из языков, представленных на данной диаграмме. Трудно себе представить и ареальные контакты между чукотско-камчатскими и северокавказскими языками, с которыми он оказался рядом на рис. 3, в обозримом прошлом.

Между тем чукотско-камчатские языки действительно имеют ряд общих типологических черт с енисейскими и северокавказскими языками, что отмечалось ранее. Это двусторонняя агглютинация, наличие категории рода, полиперсональный глагол и некоторые другие [Володин 1997]. Обнаруженная типологическая близость чукотского языка с абхазским и кетским может служить основание для выдвижения гипотезы о

сверхдальнем родстве чукотско-камчатской семьи и бореальной макросемьи. Отметим, что на диаграмме рис. 2 по данным WALS чукотский язык расположен рядом с кетским, что дает другой аргумент в пользу этой гипотезы.

В итоге, выделяются, хотя и не столь явно, 5 основных кластеров языков по типологической близости: индоевропейский, урало-алтайский, кавказский (возможно, с чукотским и кетским), дальневосточный (несколько изолятов) и афразийский.

Посмотрим, выделяются ли найденные типологические группы и на диаграмме рис. 2 по данным WALS. На ней также четко видны афразийский (справа, в составе большого африканского кластера), индоевропейский (внизу-справа), кавказский (слева) и дальневосточный (вверху-слева) кластеры. В составе последнего присутствуют также японский и корейский языки (их нет в базе данных «Языки мира»). Неожиданно алтайские и уральские языки не образовали одной группы, оказались разбросаны по разным частям диаграммы. Например, типологически близкие Turkish и Khalkha (алтайская семья) оказались на разных концах диаграммы. Это является проявлением недостатков алгоритма классификации. На диаграмме рис.1 эти языки оказались рядом.

В некоторых случаях, вероятно, на классификации сказываются и определенные проблемы самой WALS. В работе [Polyakov et al. 2009] проводилось сопоставление аналогичных классификаций на основе WALS и «Языки мира» и показано, что классификация на основе «Языки мира» точнее отражает ареально-генеалогические связи. В [Polyakov et al. 2009] высказано мнение, что данные WALS сильно зашумлены, в первую очередь, пробелами в данных. Все же, в первом приближении, обнаруженная типологическая классификация подтверждается данными WALS.

#### 4. Алгоритм ординации

Как отмечалось выше, алгоритм NeighborNet строит, по сути, одномерное представление близости языков. Другие математические средства позволяют создавать иные представления, в том числе 2-мерные, располагая языки на плоскости так, что типологически близкие языки размещаются близко друг к другу. В числе этих средств алгоритм ординации [R-Project]. В [Поляков & Соловьев 2006] построена ординация для множества языков из нижеприведенного списка (по техническим причинам отсутствуют языки с номерами 1, 25 и 33) с использованием базы данных «Языки мира».

2 венгерский	14 бурятский	27 эстонский	40 испанский
3 финский	15 азербайджанский	28 македонский	41 итальянский
4 ассамский	16 вепсский	29 немецкий	42 галисийский
5 дари	17 хантыйский	30 бенгальский	43 абхазский
6 ительменский	18 турецкий	31 румынский	44 белорусский
7 португальский	19 бирманский	32 лезгинский	45 болгарский
8 грузинский	20 армянский	34 корякский	46 датский
9 бурушаски	21 багвалинский	35 персидский	47 нивхский
10 аккадский	22 агульский	36 таджикский	48 шугнанский
11 норвежский	23 могольский	37 чукотский	49 польский
12 английский	24 калмыцкий	38 туркменский	
13 исландский	26 монгольский	39 татарский	



Результаты приведены на рис. 4. Линии разделяют возможные кластеры. Слева располагаются европейские языки. Внизу — иранские (в эту область попали также бурушаски и аккадский). В центре-вверху — плотная группа из алтайских, уральских, кавказских и некоторых индоиранских языков. Это образование можно назвать центрально-азиатским кластером. В нем языки нескольких семей сближаются, очевидно, вследствие ареальных контактов. Маленькую группу образуют три языка с номерами 34, 37 и 43 — корякский, чукотский и абхазский. Наконец, языки №19 и 47 (бирманский и нивхский) — изоляты в данной выборке.

Как видим, метод ординации дает несколько иное, укрупненное деление на кластеры типологически близких языков. Иранские языки, хотя и выделены на рис. 4 в отдельный кластер, но он располагается рядом с кластером европейских языков, так что здесь вполне правомерно объединение их в единый кластер индоевропейских языков. Вновь абхазский язык сближается с языками чукотско-камчатской семьи, что дает очередное подтверждение гипотезы об их сверхдальнем родстве.

Полученная методом ординации классификация коррелирует, но не совпадает с приведенными в предыдущем параграфе.

## 5. Заключение

Целью статьи является не столько представление конкретных результатов, которые еще далеки от окончательных, сколько очертить общий путь, двигаясь по которому можно получить новые типологические классификации языков. На этом пути должны быть решены многие методологические проблемы. Было показано, что математические методы и алгоритмы не являются идеальными, выдаваемые ими результаты являются лишь приближенными. В этом направлении в последние годы ведется интенсивная работа: в [Wichmann & Saunders 2007] проведено сопоставление различных филогенетических алгоритмов, в [Поляков & Соловьев 2006] сравнивались различные меры близости, в [Polyakov et al. 2009] — различные источники данных.

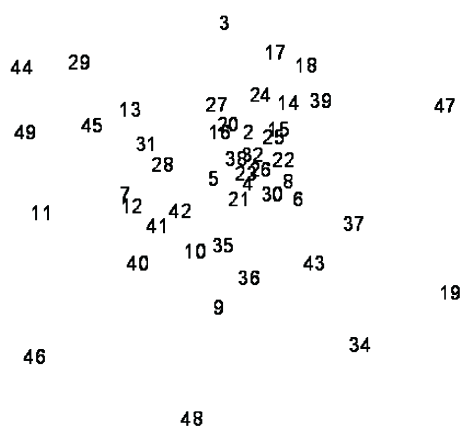


Рис. 4. Ординация тестового множества языков

Несмотря на указанные трудности, даже первоначальные результаты внушают оптимизм. Полученные разными методами типологические классификации коррелируют между собой. Типологическая близость во всех упоминавшихся в статье кластерах может быть объяснена общностью происхождения или ареальными контактами. В первом приближении на пространстве Северной Африки, Европы, Северной и Центральной



Азии можно выделить следующие основные кластеры типологически близких языков: индоевропейский, урало-алтайский, кавказский, дальневосточный и афразийский. Некоторые из полученных результатов вполне ожидаемы, как например, выделение в особую группу индоевропейских языков. Насколько эта классификация удачна — покажут будущие исследования (см. рис. 4).

На этом пути могут быть получены и новые неожиданные результаты, например, типологическая близость чукотско-камчатской семьи с синокавказской макросемьей. Эта близость может быть следствием сверхдальнего родства (на временной глубине порядка 15 тыс. лет до н. э.), не устанавливаемого классическим сравнительно-историческим методом. Типологические черты языков, по крайней мере, некоторые, являются более стабильными, чем лексемы [Nichols 2007], что позволяет выявлять сверхдальнее языковое родство или древние языковые контакты.

В целом все направление с использованием больших типологических баз данных и математических методов анализа данных представляется чрезвычайно перспективным и может привести к формированию новых разделов типологии и получению общезначимых результатов.

А. В. Дыбо  
РГГУ, Москва

К сожалению, следует констатировать, что в статье В. Д. Соловьева с самого начала имеется логическая ошибка. Эта же ошибка, конечно, имеется и у Н. С. Трубецкого, которого призывает вспомнить автор. По определению, родственные языки — это, как известно, языки, восходящие к общему языку-предку. Вот примерные основания этого определения:

Между звуковой оболочкой слова и его значением нет необходимой связи, т. е. в общем случае минимальный знак языка (морфема) является символом (в пирсовском смысле). Тогда однотипное сходство внешних оболочек при относительном тождестве значений большого количества знаков между разными языками — т. е. гипотеза о случайном совпадении здесь неприемлема — означает, что эти знаки в разных языках суть отображения одного и того же набора знаков. Если эти знаки относятся к «малопроницаемым» подсистемам языка, т. е. гипотеза о массовом заимствовании неприемлема, то следует считать, что рассматриваемые языки являются отображениями одного и того же языка. Этот прототипический язык может быть построен как язык-посредник, содержащий глубинные формы, из которых по непротиворечивому набору правил могут быть построены формы представленных языков. Построенный таким образом глубинный язык

(реально обычно строятся какие-то фрагменты языковой системы) называется праязыком данной группы языков, а группа называется родственными языками. В частности, доказательство родства некоторого множества языков фактически заключается в построении представительного фрагмента праязыковой системы (и, соответственно, в предъявлении этого фрагмента и набора правил вывода поверхностных языковых форм). Интерпретация такой ситуации — историческая, т. е. генетически родственными языками называются те два или несколько языков, которые представляют собой результаты различных линий эволюции одного и того же языка, существовавшего раньше и тем самым являющегося для этих языков праязыком.

Таким образом, языки по определению не могут стать родственными в результате конвергентного развития. То сходство, которое может между ними образоваться в последнем случае, нельзя называть языковым родством во избежание терминологической путаницы. Некоторые авторы называют его сродством. Т. е., если бы Н. С. Трубецкой дал себе труд как-то формализовать свою задачу, он должен был бы написать, что собирается показать, что отношения между индоевропейскими языками являются не родством, а сродством.

Теоретически возможна, конечно, и такая модель развития, при которой носители нескольких (абсолютно неизвестных нам) языков в результате тесных контактов между собой развили некоторый общий язык, не являющийся по определению генетическим потомком какого-либо из тех, прежних языков, или же являющийся генетическим потомком нескольких из них, из которого в дальнейшем дивергентным путем развились языки-потомки — индоевропейские языки. Трудностью здесь является то, что на практике специалистам по сравнительно-историческому языкознанию еще не довелось столкнуться с чистыми случаями того, чтобы принципиально нельзя было определить, генетическим потомком какого из контактировавших языков является данный язык, и отнести сходства с остальными языками к заимствованиям. Т. е., в надежных случаях никогда не удастся построить для одной группы родственных языков несколько разных, но равно представительных праязыковых подсистем. Отмечу, что проблема генеалогической принадлежности креольских языков и пиджинов неоднократно дискутировалась между нами с В. И. Беликовым, и последнему пока не удалось привести пример языка «смешанного происхождения», для которого с помощью сравнительно-исторического анализа нельзя было бы определить принадлежность к определенной языковой семье.

Типологические сходства по определению (см. выше) не являются доказательными для языкового родства. Они могут показывать либо сродство языков, образовавшееся в результате контактов, либо сохранившиеся от праязыкового состояния сходства, либо — третья возможность, которая, как мне кажется, недостаточно учитывается автором статьи, — сходства, обусловленные структурной близостью человеческих языков, чисто типологического характера (пример: левое ветвление побуждает к отсутствию префиксов, или что-нибудь подобное), или — четвертая возможность, также недостаточно учитываемая автором, — случайно возникать в разных языках в силу совершенно различных процессов (пример из области материальных сходств — новогреческое и полинезийское название глаза, греч. *мати*, полинез. *мата*). Четвертая возможность для типологических сходств чисто арифметически значительно более вероятна, чем для материальных (которые учитываются при доказательстве родства). В самом деле, типологические признаки обычно принимают небольшое число значений. Возьмем бинарный (левое-правое ветвление) или кватернарный (эргативный — номинативный — ...) признак. У каждого языка есть только одна из двух или одна

из четырех возможностей, вероятность случайного попадания 0,5 или 0,25. Если же взять возможность совпадения названия глаза из четырех букв, то легко подсчитать, что при принятии алфавита, допустим, из 40 букв, вероятность случайного совпадения между двумя языками будет 0,00000390625. Сравнительный же метод требует значительного числа таких совпадений.

База данных «Языков мира», по которой ведется подсчет, состоит из 3821 бинарного признака. Чтобы получить вероятность совпадения, равную случаю с одним четырехбуквенным «глазом», нужно совпадение около 18 бинарных признаков. Если для доказательства родства требовать совпадения 600 четырехбуквенных слов с заданными значениями (что, в общем, соответствует обычной практике сравнительно-исторического языкознания), — и именно для таких случаев считать совпадение неслучайным, то бинарных признаков потребуется около  $18 \times 600 = 10800$  признаков. Впрочем, и такое совпадение не будет демонстрировать родства — по определению, а покажет только неслучайность сходства.

Читателю предлагается оценить степень случайности типологических расстояний между любыми двумя выбранными языками по базе «Языков мира» (македонским и белорусским — 245, т. е. совпадают 3576 признаков или бирманским и белорусским — 401, т. е. совпадают 3420 признаков).

Из свойств диаграммы на рис. 3 выше, первое — «1. На этой диаграмме расположение языков более равномерное, чем на предыдущих; нет никакого четкого разделения на кластеры. Это можно рассматривать как косвенный аргумент в пользу гипотезы моногенеза и/или подтверждение значительного вклада заимствований в уменьшение расстояний между группами языков, расходящимися в ходе представляемой в древовидной форме эволюции».

Конечно, ничего про моногенез или заимствования, как я уже объясняла, это свойство не говорит, а говорит о том, что подобранные для описания языков бинарные типологические признаки обладают высокой степенью взаимной независимости, что хорошо говорит о составителях схемы описания в «Языках мира».

Вторая особенность и ее причины автором, видимо, проинтерпретированы правильно.

Дальнейшее описание кластеризации по алгоритму NeighborNet показывает в основном свойства этого алгоритма. То же, в, общем, можно сказать о плоскостном представлении близости языков.

В заключении работы говорится: «Даже первоначальные результаты внушают оптимизм. Полу-

ченные разными методами типологические классификации коррелируют между собой. Типологическая близость во всех упоминавшихся в статье кластерах может быть объяснена общностью происхождения или ареальными контактами».

У меня как у компаративиста приведенные примеры классификаций ни в малой степени не вызывают оптимизма (заметим на полях, что, например, армянский язык не попал в индоевропейский кластер ни по какой классификации, а ведь это даже не хеттский; разброд наблюдается и среди относительно близких генетически индоиранских языков), а типологическая близость в кластерах может быть объяснена еще и случайностью, как упоминалось выше.

Добавлю, что утверждение «Типологические черты языков, по крайней мере, некоторые, являются более стабильными, чем лексемы» основано на работе Дж. Николс, результаты которой крайне сомнительны из-за особенностей выборки анализируемых языков и чрезвычайно узкой признаковой базы сопоставления.

Итак, по-моему, пока количество типологических признаков для классификации языков мира, по неслучайности равносильной генетической, явно недостаточно, а то, что полученная на достаточном количестве признаков классификация совпадет или как-то будет скоррелирована с генетической классификацией, данное предварительное исследование не показывает.

Г. С. Старостин  
РГГУ, Москва

Статья В. Д. Соловьева посвящена достаточно «горячей» теме в современной лингвистике — значимости (потенциальной) сравнительных данных, накопленных в закромах лингвистической типологии, для решения спорных вопросов в области генеалогической классификации языков. Специально подчеркиваю слово «спорный», хотя в самом тексте статьи оно и не упоминается, потому что там, где с задачей классификации успешно и непротиворечиво справляется классическая компаративистика, никакие альтернативные методы, по сути, не требуются, что имплицитно признает и автор статьи: так, отмечаемый им факт «странного» типологического сближения ирландского языка с кетским и бурушаски на рис. 3 не предлагается (слава богу!) интерпретировать как серьезный аргумент в пользу ирландско-кетского родства, а лишь как свидетельство определенного несовершенства компьютерного алгоритма (положим, что так).

Работа по сближению типологии и компаративистики ведется уже давно; начало ее можно, по видимому, отсчитывать со ставших классическими (вне зависимости от того, заслуженно или нет) работ Дж. Николс 1990-х гг., в последнее время же построение «историко-типологических» компьютерных моделей переживает своеобразный бум — так, совсем недавно на эту же тему была опубликована очередная статья Р. Грея [Greenhill et al. 2010].

Объяснить такую неожиданную популярность можно тремя факторами. Во-первых, это устойчивое представление о том, что сравнительный метод как основной способ установления генетического родства языков «изжил» себя как парадигма, по крайней мере там, где речь касается временной глубины, превышающей 6—8 тысяч лет (концепция, активно пропагандируемая, в том числе, и Дж. Николс), и вызванное этим в чем-то, пожалуй, даже благородное желание помочь «застрявшим» компаративистам, неспособным своими силами проникнуть в столь глубокое прошлое. Во-вторых, это стремление к структурированию и осмыслению самого массива типологических данных, в том числе и с точки зрения тех исторических связей, которые его образуют. Наконец, в-третьих, это своего рода синкретическая тенденция применить к лингвистической таксономии те же методы, которые давно и успешно применяются в других областях, прежде всего — в биологии.

Безусловно, хотелось бы надеяться, что со временем методика анализа данных, описываемая как в дискуссионной работе, так и в других аналогичных исследованиях, будет усовершенствована и станет реальным подспорьем для компаративиста. К сожалению, на данном этапе можно только констатировать факт абсолютной практической неприменимости полученных результатов, несмотря на несколько натянутый оптимизм, демонстрируемый автором.

Зададимся простым вопросом: действительно ли полученные результаты можно считать, хотя бы в каком-то одном отношении, *полезными* для решения спорных вопросов сравнительно-исторического языкознания? Диаграмма, приведенная на рисунке 3, схематически суммирует данные по целому ряду языков Евразии, потенциально связанных между собой дальним родством (в рамках таких гипотез, как ностратическая, сино-кавказская и ряд других), причем здесь возможны два типа ситуаций — (а) члены одной макросемьи, отделившись друг от друга, продолжают существовать на географически смежных территориях, (б) оказываются в ситуации отсутствия каких-либо контактов в результате дальней миграции одной или нескольких составляющих первоначального этноса. Какие *новые* выводы относительно этих двух типов ситуаций позволяет нам сделать диаграмма?

По ситуациям типа (а) — очевидно, никаких. Алгоритм «грамотно» располагает в непосредственной близости друг от друга языки уральской и алтайской семей, но, как отмечает автор, диаграмма позволяет утверждать лишь то, что «эти языки сгруппированы по ареально-генетическому принципу». Но что такое «ареально-генетический принцип», как не признание того факта, что отмеченные типологические сходства могут объясняться либо контактами, либо общим происхождением, либо и тем и другим, т. е., по сути, не добавляют к нашим текущим представлениям об урало-алтайских языках ничего нового?

Тот факт, что абхазский и нахско-дагестанские языки располагаются рядом друг с другом, что могло бы косвенно подтверждать идею общесевернокавказского родства, полностью обесценивается тем, что неподалеку от них рядом друг с другом расположены языки хинди и бурушаски. Последнее объясняется контактными влияниями индоарийских (а также дардских) языков на бурушаски, но что в таком случае мешает нам считать и абхазо-дагестанское «соседство» результатом интенсивных языковых контактов?

По ситуациям типа (б) результаты оказываются лишь маргинально более ценными. *Допустим*, что типологическая близость языков, устанавливаемая по данному алгоритму, не является продуктом случайного развития (момент сам по себе чрезвычайно спорный, но здесь, по крайней мере, можно легко понять, как усовершенствовать алгоритм за счет подключения еще большего количества бинарных или многозначных признаков), а всегда отражает либо конвергенцию, либо родство, либо и то, и другое. В этом случае факт расположения

кетского языка в непосредственном соседстве от северокавказских, казалось бы, должен был порадовать любого сторонника сино-кавказской гипотезы; поскольку енисейские языки расположены вне зоны возможных контактов с кавказскими языками, речь может идти только об опции родства.

Однако что делать с оказавшимися на том же конце диаграммы чукотско-камчатскими языками? Автор говорит о возможности «сверхдального родства чукотско-камчатской семьи и бореальной макросемьи», но это довольно расплывчатое замечание, т. к. в гипер-гипотетическую «бореальную макросемью» в том виде, в котором она сегодня представляется в рамках Московской школы компаративистики, по сути, входят все без исключения языки, представленные на диаграмме, что совершенно не объясняет особую близость чукотского языка именно к кавказско-енисейской ее части. Может быть, соседство чукотского с кетским следует понимать, наоборот, как ареальное влияние енисейских языков, ранее имевших более широкое распространение, на чукотско-камчатскую семью? Но в таком случае почему и енисейские языки, в свою очередь, не могли подвергнуться влиянию какого-нибудь древнего «кавказоидного» субстрата, повлиявшего на их типологическую структуру?

Отметим, наконец, что ситуации типа (б) на представленных диаграммах вообще возникают *крайне редко*. Как правило, рядом друг с другом предпочитают располагаться языки одного и того же ареала, что особенно наглядно видно на рис. 2, где языки Африки, несмотря на признаваемое всеми специалистами значительное генетическое разнообразие этого региона, четко противопоставлены языкам Евразии (диаграмма по алгоритму Б. Комри и др., но, скорее всего, схожие результаты были бы получены и автором статьи при условии подключения африканских данных). Нет ни малейшего сомнения в том, что, если бы к работе алгоритма были подключены данные по Юго-Восточной Азии, сино-тибетские языки встали бы рядом с австроазиатскими и тайскими, т. к. по своим типологическим характеристикам они почти неотличимы (хотя к генетическому происхождению этих языков они и не имеют никакого отношения).

Не подлежит сомнению, что диаграммы, полученные по методике как автора работы, так и Б. Комри и других исследователей, в целом неслучайны и имеют право на существенную интерпретацию: они представляют собой способ построения формальной, объективной типологической классификации языков, являясь значительным шагом



вперед по сравнению с «ручными», серьезно ограниченными классификациями типологов предыдущего поколения. Но для того, чтобы иметь возможность спроецировать их в историческую плоскость, необходимо опираться в первую очередь на такие типологические признаки, которые являются в языке наиболее устойчивыми, т. е. чрезвычайно редко (в идеале — никогда не) меняют свое значение на противоположное под ареальным влиянием окружающих их языковых семей.

Факт самого существования таких признаков, однако, не бесспорен. В статье он обозначен лишь в паре мест, и то скорее вскользь — в резюме говорится, что «некоторые грамматические свойства являются очень стабильными и несут информацию о языковых состояниях древности», а в конце утверждается, что «типологические черты языков, по крайней мере, некоторые, являются более стабильными, чем лексемы». Никаких конкретных примеров таких типологических черт, однако, не приводится; вместо этого читатель отсылается к презентации Дж. Николс на соответствующую тему, но и там было бы тщетно обнаружить такие примеры, тем более что, по утверждению самой Дж. Николс, «typology... for purposes of discriminating genealogical from other relatedness... has weak resolution at all time depths». Вместо «типологических» черт она предпочитает делать упор на «полу-типологические» («semi-typological characters») — формального определения для этого понятия не предлагается, но, судя по приводимым примерам (личные показатели в алгонкинских языках, маркеры рода и числа в афразийских языках, местоименная парадигма  $n : m$  в «американских языках»), речь идет об элементарном распределении фонетически сходных грамматических морфов в языках мира, что, строго говоря, уже не имеет непосредственного отношения к типологии как таковой (отсюда и оговорка «semi-typological», на наш взгляд, не очень успешная), а представляет со-

бой, по сути, разновидность гринберговского метода массового сравнения, ограниченного грамматическим материалом.

Поиск «устойчивых» типологических признаков — задача чрезвычайно интересная (хотя плодотворное ее разрешение сомнительно), но как раз про нее в статье ничего не говорится; меж тем, до тех пор, пока точного определения таких признаков не существует, генетическая интерпретация получаемых типологических схем неправомерна. Даже «рекомендательное» использование их как потенциальных индикаторов родства, судя по представленным результатам, даст примерно сопоставимое количество истинных и ложных «наводок», и в этом смысле практическая ценность предлагаемой методики близка к нулю.

Думается, что, при условии как дальнейшего усовершенствования самого алгоритма анализа, так и расширения типологической базы данных и привлечения материала максимально большого числа языков, можно будет создать прочную основу для построения единой формальной *ареально-типологической* классификации языков мира — задачи в целом не менее важной, чем построения классификации генетической. Но остается совершенно непонятным, для чего вообще настаивать на специальной значимости данной работы для выяснения генетических связей между языками.

Возможно, в каком-то смысле это объясняется неправильно проведенной аналогией между лингвистикой и биологией; настоящим лингвистическим аналогом классификационных признаков в биологии все-таки должны быть не абстрактные структурные особенности того или иного языка, а его морфемный инвентарь (или хотя бы то, что у Дж. Николс называется «полутипологическими» признаками). Только на таком материале представляется возможным получить сколь-либо существенные сведения относительно дальнего родства между теми или иными языковыми семьями.

В. Д. Соловьев

Казанский государственный университет

Прежде всего, следует отметить, что упоминание Н. С. Трубецкого в статье — не более чем историческая реминисценция. Оно, по сути, несущественно и может быть полностью удалено из статьи

без какого-либо ущерба для понимания. Видимо, я неудачно представил материал, необоснованно сместив акцент на работу Трубецкого. Родство в статье понимается, конечно же, в классическом

смысле. Ни определение понятия «родство», ни справедливость древовидной модели эволюции языков не являются предметом данной статьи. Главный тезис статьи: созданные несколько лет назад большие типологические базы данных могут быть полезны и для установления родства.

Разумеется, исследования в этом направлении находятся в самом начале, и именно сейчас дискуссия, организованная редакцией журнала, является очень полезной. Далее я остановлюсь на двух моментах: методологии исследований и причинах, побудивших использовать типологические (как синоним будет использоваться слово 'грамматические') данные в компаративистике.

В дискуссии выявились четыре методологические проблемы предлагаемого подхода. Это: объем имеющихся типологических данных, стабильность грамматических признаков, влияние заимствований, корректность используемых филогенетических алгоритмов.

Общая идея оценки объемов лексических и грамматических данных, предложенная А. В. Дыбо, конечно, правильна. Однако приведенные в ее реплике подсчеты очень приблизительны и могут быть уточнены. Вероятность случайного совпадения двух 4-буквенных слов в двух языках оценивается ей как 0,00000390625, исходя из полного перебора всех комбинаций букв. При этом учитываются такие «слова», как *шцйь* и даже *ъьъь*. Вряд ли в каком-либо языке мира используются такие комбинации букв для обозначения какого-либо понятия. Ясно, что следует подсчитывать не все вообще сочетания букв, а лишь приемлемые в человеческих языках. Сделать это не просто, но отмечу, что, если взять русский язык, то получится следующая картина. Общее число комбинаций из 4 букв в 33-буквенном алфавите — 1185921, в то же время в максимально полном 144-тысячном словаре русского языка, используемом в компьютерной программе Aspell ([http://ru.wikipedia.org/wiki/GNU\\_Aspell](http://ru.wikipedia.org/wiki/GNU_Aspell)), зафиксировано только 1987 четырехбуквенных слова, т. е. примерно в 600 раз меньше числа всех комбинаций.

Далее, А. В. Дыбо в своих подсчетах ориентируется на сопоставление 600 слов в сравниваемых языках. Однако многие компаративистские исследования оперируют с 200-, 100-, 40- (проект [ASJP]) и даже 35-словными (Яхонтов) списками, и полученные на этом пути результаты не отбрасываются как заведомо неверные. Если учесть эти соображения, то окажется, что уже объем грамматических данных, накопленных в базе данных «Языки мира», значительно превосходит объем лексических дан-

ных. Это гарантирует невозможность случайной близости языков.

Следующий вопрос — о стабильности типологических признаков. Это ключевой вопрос развиваемого подхода. Данную точку зрения разделяют и участники дискуссии. Г. С. Старостин пишет: «Поиск “устойчивых” типологических признаков — задача чрезвычайно интересная (хотя плодотворное ее разрешение сомнительно), но как раз про нее в статье ничего не говорится; меж тем, до тех пор, пока точного определения таких признаков не существует, генетическая интерпретация получаемых типологических схем неправомерна».

Данная область развивается чрезвычайно быстро и даже за то краткое время, пока моя статья лежала в редакции журнала, появилось несколько важных работ по стабильности типологических признаков [Greenhill et al. 2010; Wichmann & Holman 2009; Соловьев & Фасхутдинов 2009], которые нельзя оставить без внимания и которые в значительной степени проясняют ситуацию.

Важными являются два вопроса: какова скорость изменения грамматических признаков по сравнению с лексическими и какие из грамматических признаков более стабильны. Отметим, во-первых, упомянутую Г. С. Старостиным работу Р. Грея [Greenhill et al. 2010], в которой показано, что скорость изменения типологических признаков примерно такая же, как и лексических. Независимая оценка темпов эволюции типологических признаков приведена в монографии С. Вихмана и Е. Холмана [Wichmann & Holman 2009], в которой суммированы результаты предшествующих работ этих авторов, а также Николс, Камхолца и др. Если в 100-словном списке Сводеша за 1000 лет меняется в среднем 14 слов, то из признаков в базе данных WALS за то же время меняется в среднем 19, т. е. скорость изменения типологических признаков, хотя и несколько выше, чем лексических, но сопоставима с последней. Любопытно, что если исключить изменения признаков в результате заимствований, то результаты вновь оказываются близкими. Согласно С. А. Старостину [Старостин 1989], число исконных (не обусловленных заимствованиями) замен в 100-словном списке равно 5 за 1000 лет, согласно же [Wichmann & Holman 2009], исконных типологических замен — 7. Уже эта близость значений показывает, что использование грамматических признаков уместно.

Казалось бы, можно сделать вывод, что использование типологических черт не дает преимуществ по сравнению с лексическими. Однако следует иметь в виду, что для лексического уровня языка



стабильность подсчитывалась для ядра лексики — наиболее стабильных слов, обозначающих важнейшие концепты человеческой культуры, в то время как для грамматики были взяты все признаки, включенные в базу данных, а не только самые стабильные. Таким образом, напрашивается следующий шаг — выделение ядра грамматики — наиболее стабильных признаков.

В литературе дано несколько строгих определений стабильности грамматических признаков, позволяющих оценить стабильность количественно. Наиболее проработанными являются независимо предложенные подходы Е. Масловой [Маслова 2004] и С. Вихмана, Е. Холмана [Wichmann & Holman 2009]. В работе [Wichmann & Holman 2009] приведен список высокостабильных признаков из WALS (с количественной оценкой стабильности). Возникает естественный вопрос, является ли метод определения стабильности из [Wichmann & Holman 2009] корректным.

Мною в статье [Соловьев & Фасхутдинов 2009] на материале нашей базы данных проведено сопоставление четырех различных методов: Е. Масловой, С. Вихмана и Е. Холмана, одного метода, восходящего к идеям Николс, и предложенного мной оригинального метода оценки стабильности на основе аппроксимации с помощью филогенетического алгоритма Maximal Parsimony [Semple & Steel 2003] числа изменений значений грамматического признака в ходе эволюции. Оказалось, что все 4 метода подсчета дают коррелирующие результаты. В работах [Соловьев & Фасхутдинов 2009; Polyakov et al. 2009] показано, что стабильность, подсчитанная по данным «Языки мира», коррелирует со стабильностью, подсчитанной по данным WALS, а также хорошо согласуется и с результатами, опубликованными ранее в типологической литературе.

Например, к высокостабильным все методы подсчета отнесли такие признаки как: наличие рода, наличие предлогов, оппозиция инклюзив/экс-клюдив и др. Таким образом, в настоящее время уже имеются четкие процедуры количественной оценки стабильности признаков, выделена группа высокостабильных признаков. Особо хочу подчеркнуть, что все предложенные методы оценки стабильности могут, как и все остальное в этом мире, критиковаться и улучшаться, но, во всяком случае, они есть, опубликованы в открытой печати и допускают проверку.

Третья методологическая проблема — учет заимствований — является очень тяжелой (и для классического сравнительно-исторического метода тоже), исследования в этом направлении плани-

руются. То же самое касается и усовершенствования математических алгоритмов анализа данных.

Как можно оценить текущий статус развиваемого метода? Действительно ли имеет место «факт абсолютной практической неприменимости полученных результатов», как считает Г. С. Старостин?

Проведем такой мысленный эксперимент. Представим себе, что в некоей параллельной Вселенной наука развивалась несколько по иному, чем у нас, и сравнительно-исторический метод был придуман не в Европе, а в экваториальной Африке. После классификации местных языков африканские лингвисты обратили бы свое внимание на слабо описанные языки Европы и Азии. Если бы до начала применения к этим языкам строгих процедур сравнительно-исторического метода (на опыте нашей Вселенной нам известно, что построение генетической классификации индоевропейских языков ведется уже 200 лет усилиями десятков тысяч лингвистов, и она до сих пор не завершена) наши гипотетические лингвисты имели возможность познакомиться с данными по типологической близости языков из приведенной в статье диаграммы 3, то они сэкономили бы массу времени.

Они могли бы высказать гипотезу о родстве хинди, русского, французского, английского, персидского языков (которая в дальнейшем была бы подтверждена строгими методами) или весьма правдоподобную гипотезу об урало-алтайском родстве (которая, как мы знаем, существовала в свое время в нашей исторической лингвистике, хотя сейчас ее мало кто признает), и они бы не тратили время на проверку гипотез, скажем, о русско-абхазской семье или франко-кхмерско-ненецкой. Типологическая диаграмма пропускает небольшое число правильных гипотез (индоевропейская принадлежность армянского и кельтского), но отсекает тысячи неверных. Можно предположить, что лингвисты параллельной Вселенной сочтут этот метод важным первым шагом в установлении родства языков. Возвращаясь в нашу Вселенную, мы обнаруживаем, что в точности такая оценка дается (в том числе и Г. С. Старостиним [Starostin G. 2009]) методу массового сравнения Гринберга. Мне кажется, что статус этих двух методов довольно близок.

Почему в статье метод, основанный на типологической близости, не используется для получения новых результатов, например, не применен к какому-либо сложному случаю, вроде построения генетической классификации папуасских языков? Если бы мы поступили так, то полученный результат было бы трудно оценить, в виду отсутствия

общепризнанной классификации папуасских языков. Применяв новый метод к хорошо изученным языкам, мы имеем возможность протестировать его. Вопрос следует поставить не так: что нового дает диаграмма 3, а насколько она близка к надежно установленной классификации?

Наконец, перейдем к последнему и самому принципиальному вопросу: почему делается попытка использовать типологические данные для установления родства языков, почему не ограничиться построением ареально-типологической классификации (как предлагает Г. С. Старостин). Причем ведь такая попытка делается не только мной, и не только Николс. Можно привести целый ряд работ других авторов (обратим внимание, в частности, на публикации С. Вихмана, например, [Wichmann & Saunders 2007]).

Г. С. Старостин выделяет три фактора несколько неожиданной популярности «историко-типологических моделей», и с ним можно в целом согласиться. Однако есть и еще один, на мой взгляд, ключевой момент, коррелирующий с первым из этих факторов, но отличающийся от него.

Широко распространенной является точка зрения, что сравнительно-исторический метод ограничен временным диапазоном 6–8 тыс. лет. Совершенно вне зависимости от того, верно это утверждение или нет, оно, на мой взгляд, крайне выгодно для сравнительно-исторического метода, поскольку содержит пресуппозицию, что уж на меньших-то временных отрезках у этого метода проблем нет. Посмотрим, действительно ли это так. Какова степень достоверности результатов, полученных сравнительно-историческим методом на малой временной глубине, и, главное, какова разрешающая способность метода, т. е. какую долю реальных языковых генетических общностей он способен установить? Рассмотрим тюркские языки — это хорошо описанные языки, к которым сравнительно-исторический метод применяется уже давно, причем ведущими специалистами в этой области. Возраст семьи — всего порядка 2 тыс. лет [СИГТЯ 2002].

В Лингвистическом энциклопедическом словаре [ЛЭС] сообщается о 6 классификациях тюркских языков Богородицкого, Радлова, Корша, Самойловича, Рамстедта, Баскакова. В [Языки мира. Тюркские] описываются также классификации Бенцинга и Менгеса, Маслова, упомянуты классификации Березина, Ильминского, Аристова, Катанова. Вероятно, большинство из этих классификаций уже устарели, поэтому мы обратимся к современным «каноническим» источникам [Языки мира. Тюрк-

ские; Бурлак & Старостин 2001, Gordon 2005] и самым последним публикациям ведущего российского коллектива тюркологов [Wichmann & Saunders 2007; СИГТЯ 2006] и посмотрим, совпадают ли эти классификации.

Халаджский язык. В [Бурлак & Старостин 2001] отнесен к самостоятельной ветви, в [Gordon 2005; СИГТЯ 2002, стр. 4] — к огузским языкам (южным в другой терминологии), в [СИГТЯ 2002, стр. 726] — к карлукским языкам.

Чулымский язык. В [Бурлак & Старостин 2001] отнесен к сибирским, в [СИГТЯ 2002] — к кыргызским, в [Gordon 2005] — кыпчакским (западным), в [Языки мира. Тюркские] — к тюркским(!).

Очень странной является генеалогическая классификация некоторых диалектов, приведенная в книге «Языки мира. Тюркские языки». В [Изидинова 1997] крымскотатарский литературный язык, а также его степной диалект отнесены к кыпчакским языкам, при этом южный диалект — к огузским. Следуя определению родства, приведенному в реплике А. В. Дыбо: «По определению, родственные языки — это, как известно, языки, восходящие к общему языку-предку», приходим к парадоксальной ситуации. С одной стороны, диалект — это «разновидность данного языка» [ЛЭС]. Все диалекты языка восходят к общему праязыку, от которого они отделились обычно не более чем несколько сот лет назад. Если исходно крымскотатарский язык являлся кыпчакским, то непонятно, как его южный диалект мог сменить свою генеалогическую принадлежность с кыпчакской на огузскую.

С другой стороны, если указанные диалекты восходят к протокыпчакскому и протоогузскому языкам, разделившимся более 1,5 тыс. лет назад [СИГТЯ 2002], то как они могли настолько сблизиться, что стали вариантами одного языка? А. В. Дыбо не исключает, что «Теоретически возможна, конечно, и такая модель развития, при которой носители нескольких ... языков в результате тесных контактов между собой развили некоторый общий язык». Может быть, это как раз такой случай?

Эта же классификация крымскотатарских диалектов приведена и в [Бурлак & Старостин 2001]. Более того, такая ситуация не является чем-то исключительным, из ряда вон выходящим. Урумский язык, согласно [СИГТЯ 2002, стр. 5], обладает чертами как огузских, так и кыпчакских языков. У узбекского языка в [Ходжиев 1997] выделено три наречия: кыпчакское, огузское и карлукское!

Представляется, что сравнительно-историческому методу не хватает фундаментальной монографии (на русском языке), в которой подобные неяс-

ные моменты нашли бы ясное теоретическое объяснение. В англоязычной литературе дело обстоит несколько лучшим образом. Несколько серьезных монографий и обзорных статьи выпустил, пожалуй, ведущий на Западе защитник сравнительно-исторического метода в его наиболее строгой форме Л. Кемпбелл. Его последняя монография [Campbell & Poser 2008], рецензировавшаяся в журнале Вопросы языкового родства [Starostin G. 2009], однако, вопреки желанию автора, наносит серьезный удар по всему сравнительно-историческому методу.

В какой-то мере, подводя итоги развития этого метода, монография заставляет читателя поставить вопрос: каковы конечные цели и задачи метода, что с его помощью можно сделать? Что может дать сравнительно-исторический метод: полное описание дерева эволюции языков семей со всеми разветвлениями (пусть и на ограниченной временной глубине) или только сами семьи?

В монографии Кемпбелла задача сравнительно-исторического метода, фактически, сводится к ответу на вопрос — родственны два заданных языка или нет, т. е. результатом будет описание только семей (неструктурированного множества языков, входящих в семью), но не структуры дерева эволюции семьи. Если это так, то в исторической лингвистике обнаруживается большая лакуна, которую займут другие методы, в том числе и «историко-типологический».

Если же цель сравнительно-исторического метода состоит все же в описании полного дерева эволюции, то давайте посмотрим, в какой мере эта цель достигнута. В [Бурлак & Старостин 2001] тюркские языки делятся на 7 равноправных ветвей, никак не объединяемых в более крупные кластеры. Но могло ли быть так, чтобы народ, являвшийся носителем прототюркского языка, вдруг одновременно распался на 7 народов, разбежавшихся в разные стороны? Здравый смысл подсказывает, что это абсурд. Даже одновременное деление протоязыка на 3 языка-потомка представляется достаточно маловероятным. Неясно, что может заставить народ разделиться одновременно на несколько частей, двинувшихся в разные стороны. Разумеется, в истории имели место серьезные социально-политические изменения, которые приводили к распаду языковых общностей. Например, падение Римской империи привело к региональной дифференциации латыни и, в итоге, к появлению романских языков. Однако такого рода события являются далеко не одномоментными. Римская империя разделилась на Западную Римскую импе-

рию и Восточную Римскую империю в 395 г., а Восточная Римская империя распалась затем только почти через сто лет — в 476 г. [Неронова 1983]. Романские же языки окончательно обособились лишь в IX веке [ЛЭС; Неронова 1983]. Неизвестны строго доказанные случаи одновременного распада народа на несколько частей с формированием разных языков. В случае распада прототюркского логично предположить, что мы просто не можем определить точный порядок отделения ветвей.

Если семья состоит из  $n$  языков, то полное дерево эволюции будет состоять из  $n - 1$  внутренней (т. е. не концевой) вершины, отражающих все генетические подгруппы семьи (промежуточные протоязыки и случаи их деления на 2 потомка). Деревья именно с таким характером ветвления строят большинство современных филогенетических алгоритмов [Semple & Steel 2003].

В [Бурлак & Старостин 2001] перечислено 36 тюркских языков с 8 внутренними вершинами (вершина всего дерева — прототюркский и семь ветвей). Таким образом, в этой классификации идентифицировано лишь 23% вершин из полного дерева. В [Gordon 2005] классифицировано 40 тюркских языков в дереве с 12 внутренними вершинами, т. е. идентифицировано лишь 32% от всех вершин.

В рамках сравнительно-исторического метода построено и полное дерево эволюции тюркских языков. В [СИГТЯ 2002] приведено дерево, включающее 41 тюркский язык с бинарным ветвлением вершин за исключением одного-единственного случая тернарного ветвления (это исключение не в счет). К сожалению, первые же сравнения его с данными других публикаций показывают, что оно вряд ли претендует на статус надежно установленного. Например, тувинский и тофаларский в этом дереве отнесены к одной ветви с якутским, а в [Бурлак & Старостин 2001] они входят в группу сибирских языков, в которой нет якутского. На этом дереве башкирский находится на одной ветви с ногайским, а татарский на другой ветви, а в [Gordon 2005], наоборот, башкирский с татарским находятся на одной ветви, а ногайский на другой.

Если перейти к более высокому уровню, то мы обнаружим, что в дереве из [СИГТЯ 2002] от прототюркского первым отделяется чувашский язык, затем сибирская ветвь и только потом обособляются кыпчакские, карлукские и огузские ветви. В другой же монографии [СИГТЯ 2006], написанной практически тем же коллективом авторов, что и [СИГТЯ 2002], сначала от прототюркского также отделяется чувашский, но далее реконструируется иная картина эволюции. На стр. 771 приводится

дерево, в котором карлукско-кыпчакские языки располагаются на одной ветви с сибирскими, а огузские располагаются отдельно.

Таким образом, хотя с помощью сравнительно-исторического метода сгенерировано много гипотез о структуре дерева эволюции тюркских языков, на роль надежно установленных может претендовать лишь 6–7 группировок языков, т. е. в любом случае не более четверти от общего числа.

Может быть, тюркские языки недостаточно изучены? Давайте посмотрим, как обстоят дела с самым изученным семейством — индоевропейскими языками. В [Бурлак & Старостин 2001] в классификацию включено порядка 190 языков, в дереве 36 внутренних вершин — это около 20%. В генеалогической классификации [Gordon 2005] присутствует 439 индоевропейских языков и 91 внутренняя вершина — опять примерно 20%.

Итак, разрешающая способность сравнительно-исторического метода находится где-то в районе 20%, т. е. он позволяет идентифицировать (даже на малых временных глубинах) только пятую часть языковых генетических общностей. Если исследования будут продолжаться такими темпами, то для построения полного дерева индоевропейских языков потребуется еще лет 800. Впрочем, складывается впечатление, что темпы получения новых результатов здесь замедлились — основные генетические группы индоевропейских языков Мейе знал еще сто лет назад [Meillet 1903].

Вероятно, водораздел возможностей сравнительно-исторического метода проходит не между «до» и «после» рубежа в 6–8 тыс. лет назад, а между четко выделяющимися по каким-то историческим причинам генетическими группами (восточно-славянские, тюркские, ностратические и т. д.) и нечеткими группами (сибирские?, романо-кельтские?, словенско-сербохорватские? и т. д.), плохо поддающимися сравнительно-историческому методу, независимо от их возраста.

Это еще один фактор, стимулирующий развитие новых методов, прежде всего, с привлечением новых данных. Невозможно спорить, что генетическая близость является одной из причин типологической близости и, следовательно, диаграммы типологической близости содержат генетический сигнал (как сейчас принято говорить). Вопрос состоит в том, как эту информацию оттуда извлечь и как использовать.

С пуском Большого адронного коллайдера в ЦЕРНе у физиков появилась надежда на построение, как ее называют, Теории Великого объединения (включающей все известные фундаментальные физические силы). Если такая теория будет создана, то она явится обобщением классических работ Гелл-Манна по квантовой хромодинамике и слабому взаимодействию. Может быть, и в лингвистике пришла пора Великого объединения всех накопленных данных для совершения прорыва в увлекательнейшей области эволюции языка?

## Литература

- Бурлак & Старостин 2001 — БУРЛАК С. А., СТАРОСТИН С. А. *Введение в лингвистическую компаративистику*. М.: УРСС, 2001.
- Володин 1997 — ВОЛОДИН А. П. Палеоазиатские языки. В сб. *Языки мира. Палеоазиатские языки*. М.: Индрик, 1997. С. 8–11.
- Грунтов и др. 2006 — ГРУНТОВ И. А., ДЫБО А. В., КОРМУШИН И. В., РЕШЕТНИКОВ К. Ю. Палеокультура, прародина и внешние связи языков алтайской семьи в древности. В сб. *Этнокультурное взаимодействие в Евразии*. Кн. 2. М.: Наука, 2006. с.30–40.
- Дыбо 1978 — ДЫБО В. А. Ностратическая гипотеза (итоги и проблемы) // *Известия АН СССР, Сер. литературы и языка*, т. 37, №5, 1978.
- Изидинова 1997 — ИЗИДИНОВА С. Р. Крымскотатарский язык // *Языки мира. Тюркские языки*. М.: Издательство «Индрик», 1997. С. 298–308.
- Кузнецов 1954 — КУЗНЕЦОВ П. С. *Морфологическая классификация языков*. М., 1954.
- ЛЭС — *Лингвистический энциклопедический словарь*. М.: «Советская энциклопедия», 1990.
- Маслова 2004 — МАСЛОВА Е. Динамика типологических распределений и стабильность языковых типов // *Вопросы языкознания*. № 5. 2004. С. 3–16.
- Неронова 1983 — НЕРОНОВА В. Д. Вторжение варваров и крушение Римской империи // *История древнего мира. Упадок древних обществ*. М.: Гл. ред. восточной литературы, 1983. С. 239–273.
- Николаева & Хелимский 1997 — НИКОЛАЕВА И. А., ХЕЛИМСКИЙ Е. А. Юкагирский язык. В сб. *Языки мира. Палеоазиатские языки*. М.: Индрик, 1997. 155–168.



- Поляков & Соловьев 2006 — ПОЛЯКОВ В. Н., СОЛОВЬЕВ В. Д. *Компьютерные модели и методы в типологии и компаративистике*. Казань: КГУ, 2006.
- СИГТЯ 2002 — *Сравнительно-историческая грамматика тюркских языков. Региональные реконструкции* / Ред. Поцелуевский Е. А.. М.: Наука, 2002.
- СИГТЯ 2006 — *Сравнительно-историческая грамматика тюркских языков. Пратюркский язык-основа* / Ред. Поцелуевский Е. А.. М.: Наука, 2006.
- Соловьев & Фасхутдинов 2009 — СОЛОВЬЕВ В. Д., ФАХСУТДИНОВ Р. Ф. Методика оценки стабильности грамматических свойств // *Известия РАН. Серия литературы и языка*. Т. 68. № 4. 2009.
- Старостин 1989 — СТАРОСТИН С. А. Сравнительно-историческое языкознание и лексикостатистика // *Лингвистическая реконструкция и древнейшая история Востока*. Ч. 1. М., 1989. С.3—39.
- Старостин 2007 — СТАРОСТИН С. А. Гипотеза о генетических связях сино-тибетских языков с енисейскими и северокавказскими языками // С. А. СТАРОСТИН. *Труды по языкознанию*. М.: Языки славянских культур, 2007. С. 265—282.
- Трубецкой 1987 — ТРУБЕЦКОЙ Н. С. Мысли об индоевропейской проблеме // ТРУБЕЦКОЙ Н. С. *Избранные труды по филологии*. М., 1987.
- Ходжиев 1997 — ХОДЖИЕВ А. П. Узбекский язык // *Языки мира. Тюркские языки*. М.: Издательство «Индрик», 1997. С. 426—436.
- Эдельман 1997 — ЭДЕЛЬМАН Д. И. Бурушаски язык // *Языки мира. Палеоазиатские языки*. М.: Индрик, 1997. С. 204—220.
- Языки мира. Тюркские — *Языки мира: Тюркские языки*. М.: «Индрик», 1997.
- ASJP — BROWN, Cecil H., Eric W. HOLMAN, Søren WICHMANN, and Viveka VELUPILLAI. 2008. Automated classification of the World's languages: A description of the method and preliminary results. *STUF — Language Typology and Universals* 61.4. 285—308.
- Bryant et al. 2005 — BRYANT, D., FILIMON, F. and GRAY, R. Untangling our past: Languages, Trees, Splits and Networks // *The Evolution of Cultural Diversity: Phylogenetic Approaches*. Editors: R. MACE, C. HOLDEN, S. SHEN-NAN. Publisher: UCL Press, 2005, pp. 69—85.
- Campbell & Poser 2008 — Lyle CAMPBELL & William J. POSER. *Language Classification: History and Method*. Cambridge: Cambridge University Press. 2008.
- Comrie & Cysouw 2006 — COMRIE B., CYSOUW M. New Guinea through the eyes of WALS. *Language and Linguistics in Melanesia*. <http://email.eva.mpg.de/~cysouw/publications.html>. 2006.
- Comrie 1989 — COMRIE B. *Language Universals and Linguistic Typology*. Chicago: The University of Chicago Press. 1989.
- Cysouw & Comrie 2009 — CYSOUW M., COMRIE B. How varied typologically are the languages of Africa? // Rudie BOTHA & Chris KNIGHT. (eds.). *The Cradle of Language*. Oxford, 2009.
- Everaert & Musgrave 2009 — EVERAERT M., MUSGRAVE S. (eds.). *The Use of Databases in Cross-Linguistic Studies*. Berlin, 2009.
- Gordon 2005 — *Ethnologue: Languages of the World*. Gordon R. G. Jr. (ed.). SIL International: Dallas, 2005, <http://www.ethnologue.com>.
- Gray & Atkinson 2003 — GRAY R., ATKINSON Q. Language-tree divergence timer support the Anatolian theory of Indo-European origin // *Nature*. V. 426. 2003. p. 435—439.
- Greenhill et al. 2010 — GREENHILL S. J., ATKINSON Q. K., MEADE A., GRAY R. D. The shape and tempo of language evolution // *Proceedings of the Royal Society: Biological Sciences* 2010, Apr. 7, [http://simon.net.nz/files/2010/04/Greenhill\\_et\\_al2010-preprint.pdf](http://simon.net.nz/files/2010/04/Greenhill_et_al2010-preprint.pdf).
- Meillet 1903 — MEILLET F. *Introduction à l'étude comparative des langues indoeuropéennes*. Paris. 1903.
- Nichols 1992 — NICHOLS J. *Linguistic Diversity in Space and Time*. Chicago: University of Chicago Press. 1992.
- Nichols 2007 — NICHOLS J. Typology in the service of classification. [http://aalc07.psu.edu/papers/jn\\_typol\\_class3.pdf](http://aalc07.psu.edu/papers/jn_typol_class3.pdf). Stanford, 2007.
- Polyakov et al. 2009 — POLYAKOV V., SOLOVYEV V., WICHMANN S., BELYAEV O. Using WALS and Jazyki mira // *Language Typology*. V. 13. 2009. P. 135—165.
- R-Project — R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2006. <http://www.R-project.org>.
- Semple & Steel 2003 — SEMPLE C., STEEL M.: *Phylogenetics*. New York. Oxford University Press, 2003.

- Starostin 2007 — S. STAROSTIN. Indo-European among other language families problems of dating, contacts and genetic relationships // С. А. СТАРОСТИН. *Труды по языкознанию*. М.: Языки славянских культур. 2007. С. 806—820.
- Starostin G. 2009 — G. STAROSTIN. Book review: Lyle Campbell & William J. Poser. *Language Classification: History and Method, 2008* // *Вопросы языкового родства*. № 2009, с. 158—174.
- WALS — *The World Atlas of Language Structures* / HASPELMATH M., DRYER M., GIL D., COMRIE B. (Eds.) Oxford, 2005.
- Warnow 1997 — WARNOW T. Mathematical approaches to comparative linguistics // *Proc. Natl. Acad. Sci. USA*. 1997, v. 94.
- Wichmann & Holman 2009 — WICHMANN S., HOLMAN E. W. *Temporal stability of linguistic typological features*. Lincom: Muenchen. 2009.
- Wichmann & Saunders 2007 — WICHMANN S., SAUNDERS A. How to use typological database in historical linguistic research // *Diachronica*, 2007, v. 24, №2.

Known typological classifications of languages, as a rule, are based upon one or just a few select features. With the appearance of large typological databases such as The World Atlas of Language Structures or “Languages of the World”, it has become possible to build classifications that simultaneously account for hundreds, if not thousands, of features. In building these classifications, one can employ various mathematical algorithms, thus raising their level of objectivity. The present paper discusses the first results achieved in this field. Some of the most interesting and least expected discoveries include the typological differentiation between the languages of Africa and Eurasia as well as typological proximity between Chukchee-Kamchatkan and Sino-Caucasian languages. The paper offers several possible variants for the classification of the languages of Northern Africa, Europe, North and Central Asia. It also discusses the possibility of using this methodics for the establishment of distant relationship between languages, based on two considerations: (a) simultaneous matches between hundreds of features cannot be coincidental; (b) at least some grammatical features are quite stable and capable of preserving information on language stages that go back very deep in time.



### *The Indo-European Elements in Hurrian*

by Arnaud Fournet & Allan R. Bomhard.

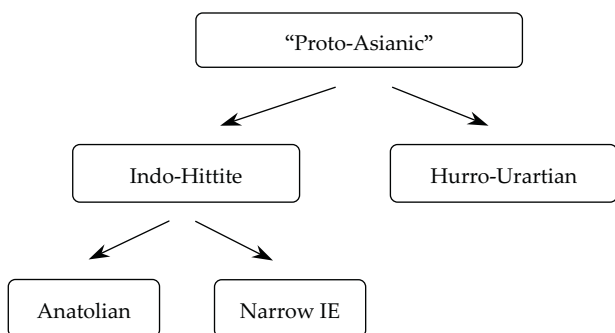
La Garenne Colombes / Charleston, 2010. 166 pp. W/o ISBN

0. The following review is based on two essential premises:

- 1) Publications dedicated to the Hurrian language are relatively scarce, meaning that any professional investigations in this field are welcome.
- 2) As a scientist, I suppose that any ideas have the right to life as long as their author can justify them by means of the scientific method.

1. The book under review is written by Arnaud Fournet, a French linguist specializing in comparative research on languages represented by limited textual corpora, and Allan Bomhard, an American macro-comparativist, whose papers and monographs on the Nostratic problem are well-known. As is noted in the Preface, Fournet generally takes the responsibility for Hurrian data, whereas Bomhard is responsible for IE matters.

The central statement of the authors (pp. 147–159) is that Hurrian (*scil.* Proto-Hurro-Urartian?) is a close relative of the Proto-IE language. F&B do not articulate it explicitly, but they probably imply that Hurrian is a member of the Nostratic macrofamily and, within this macrofamily, Hurrian is closer to IE than, e.g., Proto-Uralic. Their conception can thus be summarized in the form of the following tree:



The book under review consists of several chapters. Ch. 1 is a general introduction. Ch. 2–6 (pp. 15–78) provide a grammatical sketch of Hurrian, accompanied by suggested IE etymological grammatical *comparanda* and excurses into the comparative grammar of IE. Ch. 7 (pp. 79–106) contains the Hurrian vocabulary – a wordlist compiled from various sources (on which see below). Ch. 8 (pp. 107–146) is the etymological dictionary of Hurrian with IE cognates of Hurrian roots listed throughout. The last Ch. 9 (pp. 147–160) describes the grammatical features of the postulated “Proto-Asianic” language.

2. **Synchronic portions.** The book contains both a compilative synchronic grammar of Hurrian and a number of comparative ideas which are interspersed with each other, making the task of the reviewer somewhat difficult. In this section I provide some remarks on the synchronic portions of the book (although, running ahead, F&B’s comparative studies, from my point of view, require more attention).

p. 2. The Hurrian loanword in Sumerian *tabira* ‘metallurgy’ is to be read *tibira*, *tabira* ‘sculptor’, *scil.* ‘metal furniture-maker, coppersmith’ (= Akk. *qurqurru* ‘metal-worker, esp. coppersmith’), the more etymologically correct variant *tabira* is mostly attested in personal names, see Wilhelm 1988: 50 ff.; ePSD; CDA: 291; CAD G: 137 ff.

p. 3. The suggested etymology of PN *Kikkuli* (fn. 4) seems totally improbable.

p. 4. The authors note that, up to now, two main sources of the Hurrian language have been the Mitanni letter and the Boghazkoi Bilingual. I suppose that Ugaritic vocabularies (the bilingual fragment of HAR-*ra hubullu* series, the quadrilingual version of S<sup>a</sup> Vocabulary and the trilingual version of S<sup>a</sup> Vocabulary) must be added to this list, since they provide a lot of important and partly exclusive lexicographic information. The first two of these have been taken into

account by F&B, since these vocabularies were already known to E. Laroche and, therefore, their data were included into Laroche's *Glossaire* (GLH). But it should be more correct, however, to use Huehnergard's new edition of the quadrilingual *S<sup>a</sup> Voc.* (Huehnergard 1987/2008) instead of the older one by Laroche. The trilingual *S<sup>a</sup> Voc.* contains some Hurrian *hapax legomena* belonging to the basic vocabulary (e.g., 'new', 'louse', 'rain', 'smoke', 'tooth'; see the edition in André-Salvini & Salvini 1998; André-Salvini & Salvini 1999), but this text apparently remains unknown to F&B.

Generally speaking, F&B restrict themselves to two main lexicographical sources, as is noted on p. 79: Laroche's *GLH* (1980) and the data of Boghazkoi Bilingual from Catsanicos 1996 and Neu 1988. Other sources became mapped out; in particular, it concerns the multivolume series *Corpus der Hurritischen Sprachdenkmäler*, on which F&B make the following unjustified remark (p. 4): "these monographs seem to have an extremely limited diffusion in libraries".

p. 9 ff. It might be more correct to explicitly discriminate between a formal synchronic *phonological* description (which does indeed permit to postulate two rows of obstruents like /t~/d/ or /t~/t<sup>h</sup>/) and a more sophisticated *phonetico-phonological* description which contains three rows (e.g., t: ~ t ~ d) and reveals some features of Proto-Hurro-Urartian phonology.

p. 10. The inventory and phonetic realisation of Hurrian labial consonants is a rather intricate question, since during the millennium that attested Hurrian was spoken there were a number of Hurrian dialects which not only used different cuneiform orthographies, but could also undergo different phonetic changes. Maybe some of F&B's considerations about /f/ will become useful in future discussions, but I suppose that any extended discussion about Hurrian labial sounds must begin with a table like the one offered by Diakonoff (Diakonoff & Starostin 1986: 14). In any case, Wilhelm's (2004: 99) or Khačikyan's (Хачикян 2010b: 131) analyses seem more correct linguistically.

p. 18. "As discussed above, there is no real support for the existence of [...] /o/ in the available documentation". This is a very strange statement. First, the only remark "above" about /o/ that I have managed to find is a short paragraph at the bottom of p. 14 which does not explain the problem. Next, the second passage about /o/ is situated "below" — on p. 27. Over just a few sentences, F&B attempt to show that in the Mitanni letter the cuneiform sign U meant [ʔu], while the

sign Ú was plain [u]. I am not able to fully understand their conception, because in the same paragraph the authors actually gloss Mitanni Ú as [ʔu]. The traditional view, however, is that Mitanni U covers [o], while Mitanni Ú = [u] & [ʔ] (proposed already by F. Bork and E. Speiser in the 1930s). It is possible that the orthographic system of the Bo.Bil. represents the same opposition U ~ Ú, cf. Wilhelm 1992: 124 f., although this requires an additional investigation. In any case, I see no reasons to reject the traditional interpretation of Mitanni U as [o] and Mitanni Ú as [u] / [ʔ].

Unfortunately, such criticisms can be easily multiplied.

**3. Etymological portions.** It is well known that Hurrian is not an isolated language, but has a close relative — the Urartian language. It is also reliably established that Urartian is not a direct descendant of Hurrian, but that these languages represent two separate branches of a common proto-language (Proto-Hurro-Urartian), see, e.g., Хачикян 2010a. One serious methodological flaw of F&B is that they exclude Urartian data from the comparison, do not attempt to reconstruct Proto-Hurro-Urartian forms or grammatical features, and confine themselves to Hurrian data from *GLH* and *Bo.Bil.* This stands in sharp contrast, e.g., with Diakonoff & Starostin's (1988) approach.

The second dramatic failure of the authors is related to the question: what do we mean when we say that the genetic relationship between language X and Y is "proven"? It is not always stated explicitly, but intuitively understood by professional comparativists that two languages can be considered genetically related, if there exists (1) an appreciable number of etymological matches between their basic vocabularies, and (2) an appreciable number of etymological matches between their main grammatical exponents (number, case, person), see Campbell & Poser 2008: 4; Бурлак & Старостин 2005: 7–24. Following Бурлак & Старостин 2005 (*pace* Campbell & Poser 2008) I believe that the former condition is strong, while the latter can serve as additional proof. If our comparison is based on cultural words (e.g. 'a k. of vessel') or roots with abstract semantics (like 'to break' or 'to swell'), it is easy to "establish" a genetic relationship between any two languages among the world with any system of phonetic correspondences that one likes.<sup>2</sup> Bellow we

<sup>1</sup> Or even one row /t/, if one treats intervocalic geminates as biphonemic combinations /t+/t/.

<sup>2</sup> Cf. E. Helimski's review of Bomhard's *Towards Proto-Nostratic: A New Approach to Comparison of Proto-Indo-European and Proto-Afroasiatic* (Helimski 1987/2000, more detailed in Хелимский 1989/2000), where Helimski, using Bomhard's methods, has established such "regular correspondences" as IE  $\theta$ - ~ Semitic \*b- and IE \*t- ~ Semitic \*b.

will see whether F&B get caught in this pitfall or not. As for grammatical exponents, such comparisons are standardly monoconsonantal; therefore, one must demonstrate some kind of isomorphisms on paradigmatic sets of grammatical exponents between two languages, since isolated monophonemic comparisons are possible between various unrelated languages.

p. 32. The adjectival suffix =*and* can hardly be postulated for Hurrian: *pis=and=* is a finite verbal stem (Wegner 2007: 186 f.), while *aw=and=alli* can theoretically be a Hittite-Luwian formation based on the Hurrian root.

p. 33. The verbal suffix =*Všk* is scarcely attested (in two stems only?), and its meaning and function are unknown.

-*K-* is, in reality, a rather common “global” diminutive morpheme attested in various families across Eurasia (e.g., in Proto-Altaic or Proto-North Caucasian).

p. 37. Additional examples for the hypothetical \**Hn* > Hurr. *nm* are needed. Hurr. =*nV=* is a nominal morpheme whose function has not been determined; “article” is its somewhat jargon-like designation in secondary literature (see, e.g., the discussion in Wegner 2007: 61). Its plural absolutive form is =*na*, not \*\*=*nna*. These facts make the comparison with the IE demonstrative pronoun \**no*, \**ne*, \**H<sub>1</sub>o-no* (with a fictitious IE laryngeal) totally improbable. This is a significantly typical example of F&B’s etymological approach; in the future, I will not comment upon every such case.

p. 38. Here for the first time, a Hurrian ghost morpheme “Accusative -*n(i)*, -*ne*, -*an*” appears. Of course, on the following pages this “accusative” in -*n* will play a vital role in the comparison with Proto-IE, but I have no idea of how the author came by this enigmatic Hurrian case exponent.<sup>3</sup>

p. 41 ff. The author undertakes the comparison of two case/number systems — Hurrian and Proto-IE. Out of twelve slots only two Hurrian morphemes have reliable IE cognates: the Hurr. plural exponent =*aš* (~ IE pl. \*-s) and the Hurr. ergative exponent =*š* (~ IE nom. \*-s). All other Hurrian case markers remain without IE etymology.

No. 1: the IE deictic stem \**ol-* does not have any plural semantics.

No. 4: the Hurrian accusative in -*n(i)* does not exist. The IE accusative ending is \**-m*, not \*\*=*-n*. The Greek and Hittite accusatives in -*n* are the result of the development \**m#* > \**n#* which is regular for these languages (but it is not the case of Hurrian, where final *m* is not prohibited). The Gothic accusative form *hanan*

‘rooster’ is formed from a *n*-stem and cannot prove that the Proto-Germanic accusative ending was \**-n*.

No. 7: An enigmatic “Hittite enclitic particle -*t/da*” ‘towards’ is quoted, which does not exist.

No. 11: There is no absolutive marker -*i* with a locative meaning in Hurrian. Such locative forms from Bo.Bil. actually contain the exponent -*ni* which is not the “article” in the absolutive case, but the same case marker -*ni* as in the ablative-instrumental case (see, e.g., Wilhelm 1993: 105 ff.).

This is by no means an exhaustive list of criticisms; summing up, the Hurrian-IE nominal grammatical comparison is a complete failure.

p. 55 ff. The Hurrian-IE personal pronouns comparison is likewise unsuccessful, since the authors do not manage to quote even a single unequivocal parallel between Hurrian and Proto-IE personal pronouns.

p. 66 ff. The same concerns verbal personal exponents. No persuasive matches between Hurrian and Proto-IE.

Thus, the monophonemic grammatical comparison does not prove the Hurrian-IE relationship. Let us now look at the root etymologies (ca. 125 entries) proposed by the authors (p. 107–143). Below I will not analyze all of the etymologies. For one thing, these include connections between phantom Hurrian and phantom Proto-IE forms; e.g., Hurr. *hab=an* (p. 113) ‘to set moving’ (whose correct meaning is ‘to go towards smth.’) is compared with IE \**H<sub>2</sub>ep-* / \**H<sub>2</sub>ebh-* ‘to move quickly, to run, to flow; (flowing or running) water, river, stream, current’, but the quoted verbal meaning is unattested in IE, while the ‘water’ semantics has nothing to do with the Hurrian meaning. Another group consists of connections with highly unlikely semantic shifts. E.g., Hurr. *na(w)* (p. 121) ‘to graze, feed’ is compared with IE *neyo-* ‘new’. In a third type of cases, the authors arbitrarily segment IE roots into a monoconsonantal nucleus and a “root extension”. E.g., Hurr. *ašuhi* ‘pine-tree’ is compared with local IE \**osp-* ‘aspen, poplar’ which is analyzed as \**os(-p-)* by F&B. In the fourth group I would place those etymologies whose proposed IE cognates are isolated within the IE family, i.e. attested in one subbranch only.<sup>4</sup> E.g., Hurr. *keb=an* (p. 111) ‘to bring’ (whose correct translation is rather ‘to send’) is compared with Germanic \**gebanan* ‘to give’ (maybe related to Lat. *habēō* ‘to hold, to have’, OIr. *gaibid* ‘to take’); Hurr. *mane*, =*ma/me* ‘3<sup>rd</sup> p. sg. pronoun’ is compared with some dubious Celtic *m*-forms with the meaning ‘here’. A lot

<sup>3</sup> Note that the “article” =*nV=* is not used in the absolutive singular (see, e.g., Wilhelm 2004: 107).

<sup>4</sup> See Ringe 1999 on the theoretical easiness to find isolated *comparanda* for any CVC-root when many languages are compared simultaneously.

of etymologies consist of forms with a vague general semantics like ‘to destroy’ etc.

Nevertheless, F&B do quote a few comparisons that could make an unexperienced reader believe that Hurrian is indeed somewhat related to Proto-IE: e.g., Hurr. *ašš* ‘to be seated’ ~ IE *\*es-* ‘to sit’.

As I have already mentioned above, relationship between languages must be proven by root etymologies which belong to the basic vocabulary. The core of the basic vocabulary is the so-called Swadesh 100-wordlist. The Swadesh list is not homogenous, but its entries possess different degrees of stability. In particular, it is possible to single out the 50 most stable words, and this shortened 50-wordlist is a very suitable tool for a quick language relationship check (see Staros-

tin G. 2010 for detail). E.g., there are 23 matches between Ancient Greek and Old Indic in the 50-wordlist (ὄστέον ~ *asth-* ‘bone’, κύων ~ *švan-* ‘dog’ and so forth). Correspondingly, the Proto-IE ~ Proto-Uralic comparison yields from 12 to 14 matches on the same list (*\*ued-* ~ *\*wete* ‘water’, *\*l/nomen* ~ *\*l/nime* ‘name’ and so on), depending on acceptance/rejection of a couple questionable etymologies.<sup>5</sup>

Let us now look at the Hurrian 50-item wordlist compiled by the reviewer (for the general principles of the compilation process now see Kassian et al. 2010). Index figures that accompany the English entries designate the relative index of stability as calculated by S. Starostin (Старостин 2007a; Starostin G. 2010; ‘we<sub>1</sub>’ is the most stable word, ‘night<sub>50</sub>’ — the least stable one).

Meaning	Hurrian word	Textual comments
ashes <sub>38</sub>	<i>šalmi</i>	Ugar.C. (Huehnergard 1987/2008: 378). Cf. the verbal form <i>šalm=ul=ušt=ed=a</i> (Msk.) ‘to burn(?)’ (André-Salvini & Salvini 1998: 15).
bird <sub>33</sub>	<i>eradi</i>	Ugar.C., Bogh. (Wegner 2007; Haas & Wegner 2004: 341; André-Salvini & Salvini 1998).
black <sub>48</sub>	<i>timeri</i>	Bo.Bil. (Catsanicos 1996: 227). Attested in the collocation <i>timeri eše</i> = Hitt. <i>tankui tegan</i> ‘the Underworld’, lit. ‘black earth’ (not ‘dark earth’, since Hitt. <i>tankui</i> = Sum. GIGGI/GE <sub>6</sub> = Akkad. <i>šalmu</i> mean ‘black’).
blood <sub>20</sub>	<i>zurgi</i>	Msk., Bogh. (GLH: 309; Wegner 2007).
bone <sub>34</sub>	—	
die <sub>13</sub>	—	It is possible that <i>ull=ul</i> (Bogh.) means ‘to die’, an intransitive(?) stem of <i>ull-</i> ‘zerstören’. Cf. GLH: 279; Salvini & Wegner 2004 (ChS 1/6): 42, 184.
dog <sub>16</sub>	<i>erwi, erbi</i>	Bo.Bil., Bogh. Ugar.C. (Catsanicos 1996; GLH: 83; André-Salvini & Salvini 1998)
drink <sub>15</sub>	—	
dry <sub>24</sub>	—	Cf. the participle <i>šib=a</i> (Bo.Bil.) which theoretically may mean ‘dried-up, dried-out’, Catsanicos 1996: 275; Neu 1996 (StBoT 32): 316–319.
ear <sub>32</sub>	<i>nui, nuwi</i> (probably also <i>nuhi</i> )	Bo.Bil. Bogh. (Catsanicos 1996: 273; Wegner 1995: 121–2). The Hurr. word <i>leli-</i> means ‘a k. of beverage (vel sim.)’ (not ‘ear!’) and does not seem inherited in view of the initial <i>l-</i> .
eat <sub>25</sub>	<i>ul</i>	<i>ul=an=o=m</i> ‘(the dog) began to eat (the bread)’ (Bo.Bil., Catsanicos 1996: 216). The Hurr. verb refers to a dog and could therefore mean ‘to devour’ rather than the neutral ‘to eat’ (of humans). But since it is translated by the basic Hitt. verb <i>ed-</i> ‘to eat (of humans)’ rather than by Hitt. <i>karap-</i> ‘to eat (of animals)’, I tentatively consider <i>ul-</i> to be the default Hurr. verb ‘to eat’.
egg <sub>47</sub>	—	
eye <sub>4</sub>	<i>ši</i>	Bo.Bil., Bogh. Attested with the “article”: <i>šī=na</i> (Catsanicos 1996: 270; Wegner 1995: 124–5). The previous candidate for ‘eye’ <i>paši-</i> actually denotes ‘mouth’. The noun <i>furi</i> (< <i>fur-</i> ‘to see, to look’) rather means something like ‘look’, not ‘eye’ <i>per se</i> (Wegner 2007).
fire <sub>7</sub>	<i>tari</i>	Bo.Bil., Mit., Bogh., Ugar.C. (Catsanicos 1996: 231; GLH: 257; Huehnergard 1987/2008: 378; André-Salvini & Salvini 1998).
foot <sub>43</sub>	<i>uri, ur=ni</i> ‘foot, leg, thigh’	Bo.Bil., Bogh., Ugar.C. (Wegner 2007; Huehnergard 1987/2008: 379; André-Salvini & Salvini 1998). The similar noun <i>ugri-</i> (Bogh.) is attested in the meaning ‘leg of a table’ (GLH: 277).
hair <sub>27</sub>	—	

<sup>5</sup> Calculated by G. Starostin, pers. comm., as part of our ongoing research on the Preliminary Lexicostatistical Tree of the world’s languages (within the “Evolution of Human Language” project, supported by the Santa Fe Institute)

Meaning	Hurrian word	Textual comments
hand <sub>11</sub>	<i>šuni, šummi</i>	<i>šuni</i> — Bo.Bil. (Wegner 2007: 219; Catsanicos 1996: 200), <i>šummi</i> — Bogh., Ugar.C. (GLH: 242). The stem <i>šummi</i> might be derived from <i>šuni</i> with the nominal suffix <i>-me</i> (on which see Wegner 2007: 59), if <i>mm</i> < <i>nm</i> .
head <sub>49</sub>	<i>pahi</i>	Bogh., Mit., Mari, Ugar.C., Ugar.A. (GLH: 192; André-Salvini & Salvini 1998).
hear <sub>45</sub>	<i>haš</i>	Tišatal, Bo.Bil., Bogh., Mit., Ugar. (Wegner 2007; GLH: 95; Catsanicos 1996: 267). Another verb — <i>šalh</i> (Mari, Bo.Bil., Bogh.) — seems to have the meaning ‘to listen (to), hearken’ (Wegner 2007; Catsanicos 1996: 267).
heart <sub>14</sub>	<i>tiša</i>	Bo.Bil., Bogh., Mit., Ugar.C. (GLH: 266; Catsanicos 1996: 199).
horn <sub>44</sub>	—	
I <sub>3</sub>	<i>iš</i> = [abs., erg.], <i>šo</i> = / <i>šu</i> = [obl.]	Wegner 2007: 82.
kill <sub>42</sub>	—	Cf. <i>uw</i> ‘to slaughter (ox etc.)’, <i>šur</i> ‘to slaughter (small cattle)’ (Catsanicos 1996: 234).
leaf <sub>41</sub>	—	
louse <sub>17</sub>	<i>aphe</i>	Ugar.C. (André-Salvini & Salvini 1998; André-Salvini & Salvini 1999).
meat <sub>46</sub>	<i>uzi</i>	Bogh., Ugar.C. (GLH: 291).
moon <sub>18</sub>	<i>kušuh</i>	Name of the Moon god. Mari, Emar, Nuzi, Msk., Ugar., Bogh. (GLH: 156).
mouth <sub>31</sub>	<i>paši, faši</i>	Mari, Bogh., Ugar.C., Ugar.A. (Wegner 2007; GLH: 197; André-Salvini & Salvini 1998).
nail <sub>19</sub>	—	
name <sub>10</sub>	—	
new <sub>23</sub>	<i>šuhe</i>	Ugar.C. (André-Salvini & Salvini 1998; Huehnergard 1987/2008: 378).
night <sub>50</sub>	—	
nose <sub>29</sub>	<i>punhi, puhhi</i>	Bogh. ( <i>punhi</i> ), Ugar.C. ( <i>puhhi</i> ), see Wegner 1995: 123–4; André-Salvini & Salvini 1998.
not <sub>30</sub>	= <i>u</i> (also = <i>wa</i> , = <i>ud</i> , = <i>wur</i> )	The basic Hurr. negation morpheme is <i>u</i> which is attested as = <i>u</i> in ergative indicative forms of the 1 <sup>st</sup> & 2 <sup>nd</sup> p., = <i>ud</i> in ergative indicative forms of the 3 <sup>rd</sup> p. (Bo.Bil., probably Ugar.C.), = <i>wur</i> in the stem <i>mann-</i> ‘to be’ (Bo.Bil. only), = <i>wa</i> in non-indicative (desiderative) forms of the 1 <sup>st</sup> & 3 <sup>rd</sup> p. Although in the 3 <sup>rd</sup> p. <i>u</i> -based morphemes began to be superseded by the old prohibitive particle <i>ma</i> already in Bo.Bil., Bo.Bil. still retains ergative forms in = <i>ud</i> . The antiquity of the morpheme <i>u</i> is proved by Urartian data, where <i>ui</i> and = <i>ure</i> are negative exponents in transitive indicative forms (see Yakubovich forthc.).
not <sub>30</sub>	= <i>kkV</i>	= <i>kkV</i> is the basic Hurr. negative exponent in non-ergative (i.e. intrans. & antipass.) indicative forms. This morpheme seems lost in the Urartian branch.
one <sub>21</sub>	<i>šukko</i>	Bo.Bil., Mit., Bogh. (Wegner 2007: 81).
rain <sub>39</sub>	<i>išena</i>	Ugar.C. (André-Salvini & Salvini 1998).
smoke <sub>36</sub>	<i>hiuri</i>	Ugar.C. (André-Salvini & Salvini 1998). The well attested form <i>hubr=ušhi</i> ‘incense burner’ may be derived from the same root, although the vowel alternation is not clear.
star <sub>40</sub>	—	The Hurr. word <i>zuzuhe</i> (Ugar.C.) means something like ‘container, box, vessel’ rather than ‘star’ (see Huehnergard 1987/2008: 26, 54, 377) and seems to be a variant of Hurr. <i>zizzuhi</i> ‘cruche’.
stone <sub>9</sub>	—	
sun <sub>35</sub>	<i>šimigi</i>	Name of the Sun god. Msk., Mit., Ugar., Bogh. (GLH: 232).
tail <sub>26</sub>	—	
thou <sub>5</sub>	<i>fe</i> =	Wegner 2007: 82.
tongue <sub>8</sub>	<i>irde</i>	Bogh. (Wegner 1995: 117–8).
tooth <sub>22</sub>	<i>šini=ni</i>	Ugar.C. (André-Salvini & Salvini 1998: 16–17).
tree <sub>37</sub>	<i>tali</i> ‘tree; wood’	Bo.Bil., Msk., Ugar.C., Bogh. (GLH: 253; Catsanicos 1996).
two <sub>2</sub>	<i>šin(i)</i>	Mari, Mit., Bogh., Ugar.A. (Wegner 2007: 81; GLH: 234).



Meaning	Hurrian word	Textual comments
water <sub>28</sub>	<i>šīwe</i> [archaic?], <i>šīye</i>	Mari ( <i>šīwe</i> ), Bo.Bil., Ugar.A., Ugar.C. ( <i>šīye</i> ), Bogh. ( <i>šīye</i> , <i>šīwe</i> ) (GLH: 230; Wegner 2007; Cat-sanicos 1996).
we <sub>1</sub>	<i>ša(=tti=lla)</i> [abs.], <i>šie(=š=)</i> [erg.], <i>ša(=š=)</i> [obl.]	Wegner 2007: 83.
what <sub>12</sub>	—	<i>Pace</i> Хачикян 2010b: 138, the Hurrian interrogative stem <i>iya</i> (Bo.Bil., Mit.) is not attested in the meaning ‘what?’.
who <sub>6</sub>	<i>abe</i> , <i>aw=</i>	Bo.Bil., Mit. (Wegner 2007: 85; Хачикян 2010b: 138).

Only six items on this list are etymologized by F&B. Five of these etymologies are rather weak and feeble:

- 1) *iš-* [abs., erg.], *šo-/šu-* [obl.] ‘P (p. 116) ~ IE \*eǵh- / \*eǵh- ‘P. Phonetically and morphologically unconvincing.
- 2) *nui* ‘ear’ (p. 123) ~ IE \*neuks- ‘to smell (trans.)’, attested in Slav. and Germ. only.<sup>6</sup>
- 3) *tali* ‘tree; wood’ (p. 130–1) ~ IE \*dhal- ‘to bloom; to be leafy’ as reconstructed in Pok. 234. It is very likely, however, that this reconstruction covers a whole number of etymologically unrelated roots with different vocalism and meanings in several IE subbranches. Celt. \*dol-V- ‘leaf’ (Matasović 2009: 102 f.), Celt. \*dal-n- ‘come into being, turn into something’ (Matasović 2009: 88 f.). Germ. \*đeljaz ‘dill’, OHG *tolde* ‘foliage’ (Orel 2003: 70). Greek θάλλω, whose primary meaning may be something like ‘to spring forth with or from moisture’ (Lowenstam 1979). Arm. *dalar* ‘green, fresh’ (Martirosyan 2010: 231). Alb. *dal* ‘to go out’ (Orel 1998: 54). It is possible that some of these roots could eventually be etymologically united, but I am not sure that a uniform root \*dhVl- with the general semantics of ‘plant’ is truly reconstructible for the Proto-IE level.
- 4) *timeri* ‘black’ (p. 133–4) ~ IE *dhem-* ‘dark’, attested in Celtic (Middle Irish) and Germanic only, see Matasović 2009: 95 (F&B quote Mlr. *deim* ‘dim, dark’ as OIr. *dem* ‘black, dark’). Can hardly be projected onto the Proto-IE level.
- 5) *zurgi* ‘blood’ (p. 143) ~ IE \*dhor- ‘to gush forth, to burst forth, to spurt’. Semantically unconvincing.

The sixth etymology could be accepted as an additional comparison between languages whose genetic relationship has already been established:

- 6) *šīwe*, *šīye* ‘water’ (p. 128–9). Several roots of the shape sV $\mu$ - with a general semantics of ‘moisture’ can indeed be found within IE languages

(although F&B merge all of the IE forms in one unified \*sew-/sow-/su-): IE \*sī $\mu$ e- ‘to rain’ (Toch., Greek, Alb.)<sup>7</sup> ~ IE \*sī $\mu$ - ‘juice; sea’ (Balt. & Germ. only) ~ IE \*sai $\mu$ - (~ -o-) ‘juice; to squeeze the juice’ (Toch., In.-Ir., Germ., Celt.), see Piet.dbf.

Not a single entry on the Hurrian 50-item wordlist has a direct semantic match with the corresponding entry for Proto-IE. Therefore I must conclude that the Hurrian-IE root comparison also fails.

**4. Conclusions.** Fournet and Bomhard have not managed to demonstrate the relationship between Hurrian and IE. I suppose that it should be clear to everybody who is familiar with Hurro-Urartian and IE languages that these two families are genetically unrelated. The philological, esp. Hurritological level of the book under review is also not particularly high.

The genealogical attribution of Hurro-Urartian remains an open question. The most natural assumption, in view of the geographical distribution and typological similarity, would be to include HU into the East Caucasian (Nakh-Dagestanian) stock of the North Caucasian linguistic family. This was originally proposed as early as the second half of the 19<sup>th</sup> c.; more recently, this idea was further developed by some Soviet authors — J. Braun, G. Klimov, Ju. Dešeriev and I. Diakonoff (see, e.g., Diakonoff 1971: 161 ff.; Дьяконов 1978), after which the monograph *Hurro-Urartian as an Eastern Caucasian Language* by Diakonoff and S. Starostin appeared (Diakonoff & Starostin 1986). In the last few decades, however, there has been some very serious progress in the North Caucasian linguistics as compared to the 1980s. The fundamental reconstruction of the Proto-North Caucasian language by S. Nikolaev and S. Starostin was completed and published as NCED (1994). Comparable progress was made in the reconstructions of the Proto-Yeniseian<sup>8</sup>

<sup>7</sup> Or IE \*suH-/sHu-, if the Hitt. stems *šukha-*, *išhuwa-* ‘to throw, strew, pour out’ are indeed related.

<sup>8</sup> See Старостин 1982/2007 and Yenet.dbf, based on Старостин 1995 and Werner VWJS with additions and corrections.

<sup>6</sup> The reconstruction of the variant \*neus- for some Germanic forms is probably unnecessary.



and Proto-Sino-Tibetan<sup>9</sup> languages. Finally, the theory of a Sino-Caucasian linguistic macrofamily, which unites North Caucasian, Basque, Yeniseian, Burushaski and Sino-Tibetan families, has partially been substantiated.<sup>10</sup>

It is currently obvious that HU can hardly be a member of the East Caucasian stock. On the other hand, there is a number of rather striking isoglosses between HU and Sino-Caucasian.<sup>11</sup> Thus, it is very likely that HU is an extinct member of the Sino-Caucasian macrofamily.<sup>12</sup> It is interesting that, lexicostatistically, HU seems to be closer to the Yeniseian family rather than to the North Caucasian or Sino-Tibetan ones. A similar situation is observed with another Asia Minor extinct language — Hattic, which also possesses some exclusive isoglosses with Proto-Yeniseian (see Kassian forthc.), although there is no specific visible relationship between HU and Hattic.

In any case, further etymological studies of HU will be possible only after a Hurro-Urartian dictionary has been compiled and a Hurro-Urartian reconstruction (based on such a dictionary) is accomplished. As of

today, we do not even possess synchronic Hurrian and Urartian thesauri.

## Literature

<sup>9</sup> See *Sibet.dbf*, based on Peiros & Starostin 1996, but seriously improved.

<sup>10</sup> For the comparative phonetics of the Sino-Caucasian macrofamily see S. Starostin's SCC (this work was not finished and therefore remains unpublished). The highly preliminary Sino-Caucasian etymological dictionary by S. Starostin is available as *Scet.dbf*. Some other papers by the same author, dedicated to the Sino-Caucasian problem, can be found in Старостин 2007 (both in Russian and English).

<sup>11</sup> Cf., e.g., the following entries of the 50-wordlist (SCauc. forms are quoted after the Tower of Babel project databases):

Hurr. *iš-* [abs., erg.], *šo-/šu-* [obl.] 'I' (1<sup>st</sup> p. sg. pronoun) ~ SCauc. \*zV 'I' > NCauc. \*zō, erg. \*ʔez(V), gen. \*ʔiz(V), obl. \*zā- 'I', Yen. \*ʔaʒ 'I', Burush. \*ʒa 'I' (see the discussion in *Caucet.dbf* sub \*zō about the distribution of stem variants within the paradigm).

Hurr. *fe-* (< \*we-) 'thou' (2<sup>nd</sup> p. sg. pronoun) ~ SCauc. \*wV 'thou' > NCauc. \*uō 'thou' (the direct stem), Yen. \*ʔaw (/ʔu) 'thou', Burush. \*u-n 'thou'.

Hurr. *time=ri* 'black' ~ SCauc. \*rVmV́ 'black, dark' > STib. \*rVm 'dark' ~ Yen. \*tum- 'black'.

Hurr. *zur=gi* 'blood' ~ SCauc. \*čāŁwV́ > NCauc. \*čāŁwV 'vital force (vel sim.)', Yen. \*sur 'blood', Basque \*i-seřdi 'sweat; sap'

Hurr. *aphe* 'louse' ~ SCauc. \*ʔamkV́ 'louse, small insect' > NCauc. \*ʔamkV 'a k. of insect, vermin, worm', STib. \*mōŋ / \*mōk (~ -ū-) 'midge, moth', Yen. \*ʔake (~ x-) 'louse'. The comparison seems very likely, although the Hurrian denasalization *mk* > *pk* requires additional examples. Note that the cluster *mK* seems synchronically unattested in proper Hurrian words. An interesting parallel is the name of the Sumerian goddess *ḫdamkina* which can sometimes be spelled as *ḫtap-ki-in-na* in Hurrian contexts (GLH: 70; van Gessel 1998 1: 438). An assumption of further fricativization *pk* > *ph* also looks natural, since the cluster *pk* is probably unattested in proper Hurrian words.

<sup>12</sup> At the beginning of the 2000s, S. Starostin himself tended to lean towards the same conclusion.

- André-Salvini & Salvini 1998 — B. ANDRÉ-SALVINI & M. SALVINI. Un nouveau vocabulaire trilingue sumérien-akkadien-hourrite de Ras Shamra // *Studies in the Culture and Civilization of Nuzi and the Hurrians* 9 (1998). P. 3–31.
- André-Salvini & Salvini 1999 — B. ANDRÉ-SALVINI & M. SALVINI. La colonne I du vocabulaire S<sup>a</sup> trilingue RS 94-2939 // *Studi Micenei ed Egeo-Anatolici* 41/1 (1999). P. 145–148.
- CAD — *The Assyrian Dictionary of the Oriental Institute of the University of Chicago*. Chicago, 1956—.
- Campbell & Poser 2008 — L. CAMPBELL & W. J. POSER. *Language Classification: History and Method*. Cambridge University Press, 2008.
- Catsanicos 1996 — J. CATSANICOS. L'apport de la bilingue de Hattusa à la lexicologie hourrite // *Amurru 1: Mari, Ébla et les hourrites*. Paris, 1996. P. 197–296.
- Caucet.dbf — *North Caucasian etymological database* by S. L. NIKOLAYEV & S. A. STAROSTIN (= NCED) // Available online at Tower of Babel Project.
- CDA — J. BLACK, A. GEORGE, N. POSTGATE. *A Concise Dictionary of Akkadian*. 2<sup>nd</sup> ed. Wiesbaden, 2000.
- Diakonoff 1971 — I. M. DIAKONOFF. *Hurrisch und Urartäisch* (Münchener Studien zur Sprachwissenschaft, Beiheft, N.F., 6). München, 1971.
- Diakonoff & Starostin 1986 — I. M. DIAKONOFF, S. A. STAROSTIN. *Hurro-Urartian as an Eastern Caucasian Language* (Münchener Studien zur Sprachwissenschaft, Beiheft, N.F., 12). München, 1986.
- ePSD — *Electronic Pennsylvania Sumerian Dictionary Project* (<http://psd.museum.upenn.edu/epsd/index.html>).
- GLH — E. LAROCHE. *Glossaire de la langue hourrite*. Paris, 1980.
- Haas & Wegner 2004 — V. HAAS & I. WEGNER. Das Gegenwortpaar 'wahr' und 'falsch' im Hurritischen // D. GRODDEK & S. RÖBLE (eds). *Šarnikzel. Hethitologische Studien zum Gedenken an Emil Orgetorix Forrer*. Dresden, 2004. P. 339–344.
- Helimski 1987/2000 — E. HELIMSKI. A "new approach" to Nostratic comparison // E. A. ХЕЛИМСКИЙ. *Компаративистика, уралистика: Лекции и статьи*. Moscow, 2000. P. 476–480 [First publ.: *Journal of the American Oriental Society* 107/1 (1987). P. 97–100.]
- Huehnergard 1987/2008 — J. HUEHNERGARD. *Ugaritic Vocabulary in Syllabic Transcription* (HSS 32). 1<sup>st</sup> ed.: Atlanta, 1987; 2<sup>nd</sup> repr. ed.: Winona Lake, 2008, with addenda on pp. 375–406.
- Kassian forthc. — A. KASSIAN. Hattic as a Sino-Caucasian language // *Ugarit-Forschungen*. Forthcoming.
- Kassian et al. 2010 — A. KASSIAN, G. STAROSTIN, A. DYBO, V. CHERNOV. The Swadesh wordlist. An attempt at semantic specification // *Journal of Language Relationship*, No. 4 (2010) [the present volume]. P. 46–89.
- Lowenstam 1979 — S. LOWENSTAM. The meaning of IE \*dhal- // *Transactions of the American Philological Association* 109 (1979). P. 125–135.
- Martirosyan 2010 — H. K. MARTIROSYAN. *Etymological dictionary of the Armenian inherited lexicon*. Brill, 2010.
- Matasović 2009 — R. MATASOVIĆ. *Etymological Dictionary of Proto-Celtic*. Brill, 2009.

- NCED — S. L. NIKOLAYEV, S. A. STAROSTIN. *A North Caucasian Etymological Dictionary*. Moscow, 1994 [reprinted: 3 vols. Ann Arbor: Caravan Books, 2007.]. Available online as Caucet.dbf.
- Neu 1988 — E. NEU. *Das Hurritische. Ein altorientalische Sprache in neuem Licht*. (AAWLM 1988 no. 3). Mainz/Stuttgart, 1988.
- Neu 1996 — E. NEU. *Das hurritische Epos der Freilassung I: Untersuchungen zu einem hurritisch-hethitischen Textensemble aus Hat-tuša* (Studien zu den Boğazköy Texten 32). Wiesbaden, 1996.
- Orel 1998 — VI. OREL. *Albanian etymological dictionary*. Brill, 1998.
- Orel 2003 — VI. OREL. *A handbook of Germanic etymology*. Brill, 2003.
- Peiros & Starostin 1996 — I. I. PEIROS, S. A. STAROSTIN. *A Comparative Vocabulary of Five Sino-Tibetan Languages*. 6 vols. Melbourne, 1996.
- Piet.dbf — *Indo-European etymological database* by Sergei NIKOLAYEV // Available online at Tower of Babel Project.
- Pok. — J. POKORNY. *Indogermanisches etymologisches Wörterbuch*. Bern, 1959.
- Ringe 1999 — D. RINGE. How hard is it to match CVC-roots? // *Transactions of the Philological Society* 97/2 (1999). P. 213–244.
- Salvini & Wegner 2004 — M. SALVINI & I. WEGNER. *Die mythologischen Texte* (ChS 1/6). Roma, 2004.
- SCC — S. A. STAROSTIN. *Sino-Caucasian*. Unfinished MS, the middle of the 2000s. Available online at Tower of Babel Project.
- Scet.dbf — *Sino-Caucasian etymological database* by S. A. STAROSTIN // Available online at Tower of Babel Project.
- Starostin G. 2010 — G. S. STAROSTIN. Preliminary lexicostatistics as a basis for language classification: A new approach // *Journal of Language Relationship*, No. 3 (2010). P. 79–116.
- Stibet.dbf — *Sino-Tibetan etymological database* by S. A. STAROSTIN (= Peiros & Starostin 1996, but with serious improvement) // Available online at Tower of Babel Project.
- Tower of Babel project — Etymological database project “The Tower of Babel”, headed by S. A. Starostin. In affiliation with the “Evolution of Human Languages” (EHL) Project at the Santa Fe Institute. Available online: <http://starling.rinet.ru/>.
- van Gessel 1998 — B. VAN GESSEL. *The Onomasticon of the Hittite Pantheon* (HbOr I/33). Leiden, 1998–2001.
- Wegner 1995 — I. WEGNER. Die hurritischen Körperteilbezeichnungen // *Zeitschrift für Assyriologie* 85 (1995). P. 116–126.
- Wegner 2007 — I. WEGNER. *Hurritisch. Eine Einführung*. 2<sup>nd</sup> rev. ed. Wiesbaden, 2007.
- Werner VWJS — H. WERNER. *Vergleichendes Wörterbuch der Jenseit-Sprachen*. 3 vols. Wiesbaden, 2002.
- Wilhelm 1988 — G. WILHELM. Gedanken zur Frühgeschichte der Hurriter und zum hurritisch-urartäischen Sprachvergleich // V. HAAS (ed.). *Hurriter und Hurritisch* (Xenia, 21). Konstanz, 1988. P. 43–67.
- Wilhelm 1992 — G. WILHELM. Hurritische Lexikographie und Grammatik. Die hurritisch-hethitische Biingue aus Boğazköy // *Orientalia NS* 61 (1992). P. 122–141.
- Wilhelm 1993 — G. WILHELM. Zur Grammatik und Lexicon des Hurritischen // *Zeitschrift für Assyriologie* 83 (1993). P. 99–118.
- Wilhelm 2004 — G. WILHELM. Hurrian // R. D. WOODARD (ed.). *The Cambridge Encyclopedia of the World's Ancient Languages*. Cambridge: Cambridge University Press, 2004. P. 95–118.
- Yakovovich forthc. — I. YAKUBOVICH. Morphological Negation in Urartian // *Aramazd. Armenian Journal of Near Eastern Studies*. Forthcoming.
- Yenet.dbf — *Yeniseian etymological database* by S. A. STAROSTIN (= Starostin 1995 & Werner VWJS with additions and corrections) // Available online at Tower of Babel Project.
- Бурлак & Старостин 2005 — С. А. БУРЛАК, С. А. СТАРОСТИН. *Сравнительно-историческое языкознание*. 2-е изд. М., 2005.
- Дьяконов 1978 — И. М. ДЬЯКОНОВ. Хуррито-урартский и восточнокавказские языки // *Древний Восток* 3. Ереван, 1978. С. 25–38.
- Старостин 1982/2007 — С. А. СТАРОСТИН. Праенисейская реконструкция и внешние связи енисейских языков // Старостин 2007. С. 147–246. [First publ.: Керский сборник. Л., 1982. С. 144–237]
- Старостин 1995 — С. А. СТАРОСТИН. Сравнительный словарь енисейских языков // *Керский сборник (Studia Ketica)*. Вып. 4. М., 1995. С. 176–315.
- Старостин 2007 — С. А. СТАРОСТИН. *Труды по языкознанию [Works in Linguistics]*. М., 2007.
- Старостин 2007а — С. А. СТАРОСТИН. Определение устойчивости базисной лексики [Defining the Stability of Basic Lexicon] // Старостин 2007. С. 827–839.
- Хачикян 2010а — М. Л. ХАЧИКЯН. Хуррито-урартские языки // *Языки мира. Древние реликтовые языки Передней Азии*. М.: РАН, Институт языкознания, 2010. С. 118–125.
- Хачикян 2010б — М. Л. ХАЧИКЯН. Хурритский язык // *Языки мира. Древние реликтовые языки Передней Азии*. М.: РАН, Институт языкознания, 2010. С. 126–148.
- Хелимский 1989/2000 — Е. А. ХЕЛИМСКИЙ. К оценке надежности индоевропейско-семитских лексических сопоставлений // Е. А. ХЕЛИМСКИЙ. *Компаративистика, уралистика: Лекции и статьи*. М., 2000. P. 481–486 [First publ.: *Палеобалканистика и античность*. М., 1989. P. 13–20.]

*Аспекты компаративистики 4.*

Под ред. Г. С. Старостина.

*Orientalia et Classica: Труды Института восточных культур и античности,*  
выпуск XXVIII. Москва: РГГУ, 2009. 262 с. ISBN 978-5-7281-1110-8

Начиная с 2005 года, сборники *Аспекты компаративистики* являлись основным печатным органом Московской компаративистической школы, группирующей вокруг Центра компаративистики Института восточных культур и античности Российского государственного гуманитарного университета. С появлением периодического журнала *Вопросы языкового родства* они приняли на себя более специализированную роль сборников материалов конференций, организуемых Центром компаративистики. Рецензируемый сборник содержит расширенные доклады участников Чтений памяти С. А. Старостина, проводившихся в марте 2008 года. Он объективно отражает круг интересов друзей, коллег и учеников Сергея Анатольевича Старостина — от исследований в области дальнего языкового родства, являющихся «фирменным продуктом» Московской компаративистической школы, до сравнительных грамматик традиционных языковых семей, этимологии и этнолингвистической реконструкции. Статьи, вошедшие в сборник, написаны на русском и английском языках.

Учитывая широкий географический охват статей рецензируемого сборника, едва ли найдется исследователь, обладающий филологической подготовкой по всем обсуждаемым в нем языковым семьям. Рецензент не претендует на оценку научного вклада каждой отдельной статьи с точки зрения узкого специалиста в соответствующей области. Высказывая конкретные замечания там, где это уместно, мы ограничиваемся в большинстве случаев изложением идей авторов и их обсуждением с точки зрения общей методологии исторической лингвистики.

Русскоязычная часть сборника открывается работой А. А. Бурыкина «Идеи С. А. Старостина и дальнейшие перспективы алтаистики». Данная статья содержит информацию историко-биографического характера и, по признанию самого автора, частично ориентирована на будущих историков науки. К сожалению, лингвистические разделы статьи, отражающие взгляды автора на общеалтайскую реконструкцию, написаны без опре-

деленного плана и вызывают вопросы методологического характера. В частности, приведен ряд примеров вокалических соответствий эвенк. *и* ~ монг. *a/e*, призванных обосновать связь между эвенкийским *bi* и монгольским *ba* 'мы' (стр. 14–16). С точки зрения общих принципов реконструкции, было бы правильнее оперировать сегодня пратунгусо-манчжурскими и прамонгольскими формами в качестве промежуточного звена в аргументации. В ряде случаев изложение нетривиальных научных гипотез не сопровождается ни их обоснованием, ни ссылками на предшествующую литературу. Так, например, обстоит дело с объяснением тюркских «квазисоответствий *r~z*, *l~š*» как результатов позиционно распределенных упрощений в группах консонантных кластеров (стр. 21). Учитывая тему доклада, вызывает удивление отсутствие ссылок на алтайский этимологический словарь [Starostin et al. 2003].

Работа А. И. Когана «О ряде фонетических изменений в языке Кашмири и их относительной датировке» относится к наиболее глубоко проработанным статьям настоящего сборника. Автор рассматривает три группы звуковых переходов: изменения в вокализме, приведшие к развитию кашмирского умлаута, циклическое передвижение кашмирских аффрикат, приведшее к исчезновению их церебрального ряда, и эволюцию кашмирских сибилантов, также имевшую своим следствием исчезновение церебрального  $\zeta$ . Все эти изменения разительно отличают Кашмири от большинства других дардских языков, а последние два сближают фонетический облик Кашмири с обликом большинства индоарийских языков, не принадлежащих к дардской группе. Используя данные кашмирских заимствований из иранского и персидского языков, а также типологические аргументы, автор доказывает недавний характер этих процессов. Применение внутренней реконструкции позволяет ему восстановить пракашмирскую фонетическую систему, не отличавшуюся сколько-нибудь заметным образом от усредненного дардского фонетического профиля. Поскольку несо-

вершенное овладение вторым языком часто приводит к фонетической интерференции, можно высказать предположение, что перестройка звуковой системы кашмири была связана с выходом носителей этого языка за пределы собственно дардского ареала, сопровождавшимся языковой ассимиляцией индоарийцев — аборигенов Кашмирской долины.

Темой работы П. М. Кожина «Средиземноморское «койнэ» в первом тыс. до н. э.» является социолингвистический анализ функционирования финикийского языка. Автор подробно останавливается на двух надписях, иллюстрирующих статус финикийского как средства межрегионального общения. Это финикийско-лувийская билингва из Каратепе (юго-восточная Анатолия) и финикийско-этрусская билингва из Пирги (тирренское побережье Италии). Лувийский и этрусский языки являлись автохтонными в соответствующих регионах, чего нельзя сказать о финикийском языке, очевидно принятом на вооружение по соображениям престижа. Автор, впрочем, придерживается несколько другой точки зрения в отношении билингвы из Пирги, считая, что ее этрусский текст упоминает некую народность или область *eniāsa pulutḫva*, где «был в обиходе финикийский язык и доминировал культ Астарты» (стр. 66). К сожалению, подобный перевод *eniāsa pulutḫva*, не являющийся общепринятым, не доказан в работе Кожина и, в любом случае, не подтверждается финикийским текстом билингвы. К списку цитируемой вторичной литературы по финикийской социолингвистике следует добавить Lipiński, 2004.

С нашей точки зрения, вопрос о статусе финикийского языка в раннем железном веке не следует отделять от вопроса об этногенезе финикийцев. Территория прибрежного Леванта, на которой в начале первого тысячелетия до н. э. процветали финикийские торговые княжества, подверглась в начале 12 века до н. э. нашествию «народов моря». Хотя проблема языков, на которых говорили эти завоеватели, не может быть решена окончательно за отсутствием письменных памятников, египетские источники изображают «народы моря» как союз племен, происходящих из различных частей средиземноморья, откуда следует, что они были разнородны также и в этническом отношении. При подобных обстоятельствах разумно предположить, что язык местного семитского населения, из которого постепенно развился финикийский, послужил для завоевателей языком межэтнического общения. Постепенно финикийский язык мог распространить этот статус также и на некото-

рые группы «народов моря», которые не осели на территории Финикии, с той разницей, что для них он не стал первым языком.

В этой связи следует отметить связь «народов моря» с теми регионами, откуда происходят надписи из Пирги и Каратепе. В числе племен, вторгавшихся в Египет на исходе бронзового века, упоминаются денъен и эквеш. Эти этнонимы связываются с топонимами Дануна и Хиява, обозначающими равнинную Киликию в финикийских и лувийских текстах. Сама надпись из Каратепе именуется местную правящую династию «домом Мопса», отсылая тем самым к известному персонажу греческой мифологии, якобы жившему во время Троянской войны, и мигрировавшему в Киликию из западной Анатолии [Oettinger 2008]. Еще одним враждебным племенем, известным из египетских источников, является турша/тереш, обычно отождествляемое с тирсенами/этрусками. Об активности предков тирсенов в восточном средиземноморье свидетельствуют, например, надписи с острова Лемнос на языке, близком к этрусскому [Beekes 2003].

В статье Т. А. Михайловой обсуждается этимология галльской формы *banon* ‘жен’ (род. мн.), встречающаяся в надписи из Ларзака. Для пракельтского языка реконструируются, в частности, следующие падежные формы от существительного ‘женщина, жена’: им. ед. *benā*, род. ед. *bnās*, им. мн. *bnās*, род. мн. *banōn* (стр. 69). Исходя из этой парадигмы, можно считать *banon* тривиальным аналогическим образованием, возникшим под влиянием падежных форм с нулевой степенью аблаута в корне. Автор, однако, не удовлетворяется таким решением и предлагает видеть в этой форме родительный падеж от незасвидетельствованного *\*banona* «предположительно — ‘жена’ в широком смысле или же, напротив ‘женщина’ в уничижительном смысле». В качестве обоснования этой гипотезы приводятся параллели из других кельтских языков: валл. *banon* ‘королева, молодая женщина’ и корн. *benep* ‘невеста’. Предложенное автором решение представляется противоречащим «бритве Оккама», поскольку надпись из Ларзака содержит другие формы базовой общекельтской лексемы со значением ‘женщина, жена’, но не дает независимых оснований для реконструкции производного *\*banona*. К тому же, предложенная интерпретация не снимает формальных проблем, поскольку *\*banona*, как любой другой вторичный отыменной дериват, не должно реконструироваться с корневым аблаутом (ср. галльское личное имя *Banonias* (род. ед.) [Matasović 2009: 61]).



Статья Ю. А. Норманской продолжает развивать выдвинутую ранее тем же автором гипотезу, согласно которой разноместное ударения в мокшанском диалекте мордовского языка можно проецировать на прафинно-волжский языковой уровень. Согласно этой гипотезе, позиция по отношению к реконструированному ударению являлась одним из решающих факторов, определявших развитие гласных в отдельных финно-волжских языках. В данной работе автор обращается к саамскому материалу, сосредотачиваясь на рефлексах финно-угорских фонем \*e и \*ä. По нашему мнению, статья содержит достаточно примеров, иллюстрирующих корреляцию между нетривиальным распределением рефлексов ф.-у. \*e в мордовском и саамском языках (группа 1: морд. i ~ саам. \*ε, \*ē; группа 2 ~ морд. e/ä ~ саам. \*e). Поэтому если корреляция между мокшанским ударением и отражением \*e доказана для мордовского, то она, видимо, доказана и для саамского. Что касается гипотезы о связи двоякого отражения саамского \*ä с местом мокшанского ударения, она обосновывается противопоставлением одного и трех примеров (стр. 83, 84—85) и поэтому не может считаться доказанной.

Работа С. А. Яцемирского «Labyrinthos: суффикс -nth- в минойском и тирренском языках» привлекает интересный языковой материал, но удивляет своими далеко идущими выводами. Автор обнаруживает большое количество форм, записанных линейным письмом А, в которых последней силлабограммой является -te-, и отождествляет этот элемент с греческим субстратным суффиксом -ivθo-. С другой стороны, в этрусском и ретийском также представлен суффикс -te, основное значение которого — указание на принадлежность к какой-либо общности или происхождение из города или местности. На основании этого факта, а также постулируемого типологического сходства между этрусским и «минойским» языками, Яцемирский приходит к выводу об их генетическом родстве. Подобный вывод представляется нам методологически неправомерным. Для объединения языков в одну группу недостаточно сопоставления изолированных суффиксальных морфем, значение одной из которых, к тому же, является неустановленным за отсутствием дешифровки языка линейного письма А. Хочется надеяться, что автору в ходе его дальнейшей работы удастся найти более сильные аргументы в пользу своей гипотезы.

Отдельно хотелось бы сказать об этимологии греческого λαβύρινθος ‘лабиринт’. Соглашаясь с

неприемлемостью возведения данного термина к лидийскому λάβρυς ‘топор’, известному в передаче Плутарха, нельзя вместе с тем поддержать предложенное автором сопоставление с глоссой Гесихия λάβριος ‘яма’. Нет никаких независимых оснований полагать, что прототипом легендарного критского Лабиринта, места убийства Минотавра Тесеем, являлось именно подземное сооружение. Гипотеза об идентификации Лабиринта греческой мифологии с дворцом царей Кносса, раскопанным Эвансом, по нашему мнению, остается заманчивой, но единственное, что можно сказать о Лабиринте с большой долей уверенности — это то, что мы имеем дело с критским именем собственным. Поэтому его когнаты следует, в первую очередь, поискать среди других топонимов. С формальной точки зрения, наиболее привлекательным остается сравнение гр. λαβύρινθος с карийским топонимом Λάβραυνδα. Оба имени собственных, вероятно, принадлежат к одному и тому же слою доиндоевропейской субстратной лексики, но λαβύρινθος оказалось заимствовано в древнегреческий язык непосредственно, а Λάβραυνδα — через посредство анатолийских языков. Возможная этимология слова для ‘лабиринта’, опирающаяся на данное сравнение, обсуждается в работе Yakubovich 2002.

Англоязычная часть сборника открывается статьей К. Бабаева «The Problem of clusivity in Indo-European vs. Nostratic», в которой справедливо критикуется один из многочисленных примеров «экзотизации» языковой реконструкции. Несмотря на отсутствие инклюзивных местоимений (‘мы с тобой’) в подавляющем большинстве языков, объединяемых в ностратическую макросемью, а также в большинстве подгрупп индоевропейской семьи, предпринимались попытки реконструировать оппозицию по инклюзивности/экслюзивности как для индоевропейского, так и для ностратического праязыков. Подобная практика может отчасти объясняться подспудным желанием сделать реконструкции более интересными для типологов и общих лингвистов, но главным образом вызвана объективным наличием более чем одного кандидата на местоимение первого лица множественного числа в соответствующих праязыках. Для опровержения реконструкции инклюзивных местоимений необходимо выявить ошибочных кандидатов на ‘мы’, а также предложить сценарии их вторичной эволюции.

Оставляя в стороне решение данной проблемы для ностратического праязыка, можно показать вторичный характер форм на *mV-* по отношению

*wei-* ‘мы’ в индоевропейском (ср. стр. 115–16). В двух группах, армянской и балто-славянской, проникновение *\*mes* ‘мы’ в местоименную парадигму можно объяснить синтагматической аналогией с первичным глагольным окончанием 1 мн. *\*-mes/ \*-mos*. Подобное тривиальное развитие могло происходить по отдельности в каждой из семей и, как сообщил нам Александр Китаев, находит прямую параллель в сефардском диалекте испанского языка, где мы находим *mosotros* ‘мы’ вместо кастильского *nosotros*. Тохарские клитики =*m* (А) / =*me* (Б) употребляются для всех трех лиц множественного числа в косвенных падежах и, по видимому, родственны хеттской клитике =*mas* ‘вам, вас; им’, никогда не употреблявшейся в первом лице, причем хеттская семантика является более архаичной [Yakubovich 2010: 191, сн. 48]. Что касается хеттской клитики =*mes*, она значит ‘мои (мн.)’ и не имеет отношения к рассматриваемой проблеме.

Эрудированная работа Вацлава Блажека посвящена обзору существующих гипотез относительно происхождения имени богини Афины и города Афин. Обсуждаются этимологии на базе догреческого субстрата, хурритского, хеттского, лувийского, древнеегипетского и семитских языков. Наиболее многообещающими признаются сравнение с хурр. *adani* ‘табурет (vel sim.)’, предполагающее примат топонима по отношению к теониму, и сопоставление с именем угаритской богини-воительницы Анат. Отмечается, что эти два объяснения не противоречат друг другу, и возможно греческая форма возникла в результате синкретического влияния хурритского и семитских языков (стр. 140). Однако с исторической точки зрения, учитывая, что рассматриваемое имя собственное засвидетельствовано уже в микенскую эпоху, наиболее вероятной из упомянутых в статье гипотез является субстратная. К сожалению, эта гипотеза не сулит скорого установления этимологии Афин(ы), поскольку язык субстрата континентальной Греции не может быть дешифрован за отсутствием текстов.

В любом случае, следует согласиться с мнением Блажека о неудачности лувийских этимологий Афин(ы). Лувийское (*a)tani* ‘враг’ (стр. 128) следует читать /*aguni-*/, что является рефлексом более раннего *\*al(i)wanni-* [Yakubovich 2009, см. ниже]. Лувийское *\*atni-*, упомянутое на стр. 131 как вариант *atri-* ‘персона, душа; статуя’, следует читать *tani-* ‘душа’, причем эти два существительных этимологически не связаны друг с другом [van den Hout 2002]. Возможно, однако, что корень лувийского

(COR) *á-ta-na-sa-ma-* ‘мудрость’ является результатом их контаминации.

Исследованию полинезийской метрики посвящена совместная работа А. И. Давлетшина и А. В. Козьмина. Сравнивая дисперсии длины строки поэтического текста, измеряемой в морях, в трех различных языках, авторы пришли к выводу, что стихосложение на языке Тонга является наиболее регулярным, гавайское стихосложение занимает промежуточное положение, а стихи на языке рапануи являются наиболее свободными в метрическом отношении. Данное распределение коррелирует с данными исторической фонетики: вокалическая система тонга оказывается наиболее консервативной, в гавайском языке обнаруживаются вторичные долготы, а язык рапануи нейтрализовал оппозицию между историческими долгими и краткими гласными. На этом основании можно выдвинуть гипотезу о модели силлабического стихосложения, унаследованной из общеполнезийского состояния, которая подверглась различной степени эрозии в языках данной группы по лингвистическим причинам.

Большая статья А. С. Касьяна «Anatolian lexical isolates and their external Nostratic cognates» представляет собой попытку практического подхода к интеграции индо-хеттской и ностратической гипотез. Если исходить из того, что анатолийский и позднеиндоевропейский праязыки существовали в одно и то же время и являлись двумя ветвями индо-хеттской семьи, то логично предположить, что каждый из них обладал лексическими архаизмами, которые были утеряны в другом праязыке. Для выявления подобных архаизмов в анатолийской лексике и отличения их от позднейших инноваций необходимо выйти за пределы индо-хеттской общности и рассмотреть специфические сходения между анатолийским словарем и лексическим материалом неиндоевропейских ностратических языков. Исследование Касьяна ставит своей целью собрать все мыслимые параллели между анатолийскими лексемами без надежных индоевропейских этимологий и их ностратическими когнатами. При этом, этимологии разбиты на две группы по степени их вероятности. Рискнем предположить, что данный максимальный список будет постепенно сокращаться по мере уточнения ностратической реконструкции.

Касьян отвергает гипотезу Э. Рикен [apud Yakubovich 2009] о заимствовании хетт. *alwanzahh-* ‘заколдовать’ из лув. *ali-wa-ni-* ‘враг’ как по фонетическим, так и по семантическим соображениям, предлагая вместо этого ностратическую этимоло-



гию для хеттского глагола (стр. 155). Однако, согласно личному сообщению самого А. С. Касьяна, анатолийский иероглиф бронзового века \*416 = <alī>, вероятно, имеет также чтение <alu>, поскольку в железном веке его рефлекс \*319 мог употребляться для передачи /lu/, /ru/. Поэтому для ранне-лувийского 'враг' можно предложить теперь чтение *al(u)wanni-*, которое, кстати, лучше согласуется с его отражением как *alunni-*, *arunni-* в позднелувийском. Что касается семантики, черная магия могла восприниматься, как метод отчуждения объекта ритуала от сакрального мира, и превращения его во врага богов. Сходное по функции действие в изменившейся этнографической ситуации описывается русскими глаголами *чмырить*, *чмарить*, представляющими собой деривационную параллель к хетт. *alwanzahh-*.

Работа Никиты Круглого-Энке состоит в основном из подробной росписи рефлексов ностратических вибрантов в эскимосо-алеутских языках. Целью работы является проверка гипотезы о существовании особого ностратического вибранта \**r*<sub>1</sub>, чьи рефлексы неотличимы от \**r* во всех ностратических языках кроме японского, где \**r*<sub>1</sub> якобы отражается как *-(n)t-* в интервокальной позиции. Выясняются, что данные эскимосо-алеутских языков не подтверждают такой гипотезы. Хотя для данной группы также можно было бы реконструировать вибрант \**r*<sub>1</sub> на основании соответствия эск. \**j* ~ ал. \**ǰ*, гипотетическое \**r*<sub>1</sub> эскимосо-алеутской реконструкции не обнаруживает корреляции с гипотетическим \**r*<sub>1</sub> реконструкции ностратической. Следует отдать дань уважения научной честности автора, признающего отрицательные результаты проведенного эксперимента.

Заключительная статья Георгия Старостина посвящена поискам исторического объяснения возникновения контраста между дентальными, альвеолярными и ретрофлексными рядами согласных в дравидийских языках. Как известно, подобный контраст не характерен для какой-либо другой семьи, относимой к ностратической макросемье. Автор справедливо отмечает, что ссылка на иноязычный субстрат в дравидийском не имеет статуса объяснительной гипотезы, поскольку противопоставление между тремя рядами корональных согласных наблюдается и в унаследованной лексике. В статье по очереди обсуждаются возможные ностратические соответствия для отдельных дравидийских согласных корональной зоны. Для сонантов немаркированным признается дентальный ряд, сонанты альвеолярного ряда рассматриваются как возможный результат вто-

ричной палатализации, а ретрофлексные сонанты возводятся к ностратическим палатальным. Напротив, для смычных немаркированным признается ретрофлексный ряд, альвеолярный ряд также объясняется через вторичную палатализацию, а дентальный ряд выводится из рефлексов ностратических \**lT* и \**rT*. Последнее из предлагаемых соответствий, вероятно, вызовет наибольшую дискуссию, поскольку, как отмечает и сам автор, именно \**lT* и \**rT* отражаются как ретрофлексные смычные в санскрите. Для объяснения того, как одни и те же кластеры могут, как вызывать, так и блокировать ретрофлексивизацию, желательно построение артикуляторной или акустической модели.

Подведем итоги. Несомненным достоинством рецензируемого сборника является творческий (не компилятивный) характер всех собранных в нем публикаций. Что касается научного уровня этих публикаций, он варьирует от статей, которые сделали бы честь любому профессиональному журналу, до работ, написанных на любительском уровне. Это, разумеется, естественно для материалов конференции, публикуемых без предварительного реферирования. Будем надеяться, что по мере упрочения традиции Старостинских чтений будет создана система отбора докладов для последующей публикации, что, в свою очередь, повысит престиж конференции.

## Литература

- BEEKES, Robert S. 2003. *The Origin of the Etruscans*. Amsterdam: Akademie van Wetenschappen.
- VAN DEN HOUT, Theo. 2002. Self, Soul, and Portrait in Hieroglyphic Luvian // *Silva Anatolica. Anatolian Studies Presented to Maciej Popko on the Occasion of His 65th Birthday*. P. TARACHA (ed.). Warsaw: Agade. Pp. 171–86.
- LIPÍŃSKI, Edward. 2004. *Itineraria Phoenicia*. Leuven: Peeters.
- MATASOVIĆ, Ranko. 2009. *Etymological Dictionary of Proto-Celtic*. Leiden: Brill.
- OETTINGER, Norbert. 2008. The Seer Mopsos (Muksas) as a Historical Figure // *Hiitites, Greeks, and Their Neighbors in Ancient Anatolia*. B.-J. COLLINS, M. BACHVAROVA and J. RUTHERFORD (eds.). Oxford: Oxbow. Pp. 63–6.
- STAROSTIN, Sergey, Anna DYBO, and Oleg MUDRAK. 2003. *Etymological Dictionary of the Altaic Languages*. Leiden: Brill.
- YAKUBOVICH, Ilya. 2002. Labyrinth for Tyrants // *Studia Linguarum 3 (Memoriae A. A. Korolëv dicata)*. A. S. KASSIAN and A. V. SIDEL'TSEV (eds.). Moscow: Languages of Slavonic Culture. Pp. 93–116.
- YAKUBOVICH, Ilya. 2009. The Luvian Enemy // *Kadmos 47(1/2)*: 1–19.
- YAKUBOVICH, Ilya. 2010. *Sociolinguistics of the Luvian Language*. Leiden: Brill.

*The Journal of Indo-European Studies.*  
Vol. 36, 2008.

### No 1—2

Журнал открывается объемной и довольно спорной работой известного археолога Дэвида Энтони, посвященной проблеме выделения прото-германской прародины и, шире — соотношению лингвистических и археологических данных (D. W. Anthony. A New Approach to Language and Archeology: The Usatovo Culture and the Separation of Pre-Germanic, pp. 1—51). В общем эта проблема старая и до сих пор не разрешенная<sup>1</sup>; тем более интересным кажется заявленный автором «новый подход» к ней. В начале автор формулирует необходимость определить собственную позицию относительно предполагаемой прародины индоевропейцев и заявляет о том, что является сторонником теории Ямной культуры, разделяя позицию Гимбутас, Мэллори, Адамса и др.; прародину он локализует (в пространстве и во времени) в «Понтийско-Каспийских степях между 4500 и 2500 до н. э.» (р. 6). Анатолийская теория, отодвигающая выделение индоевропейцев в середину 6-го тыс. до н. э., по его мнению, противоречит распространению обозначений *колеса* и *колесницы* в и.-е. диалектах. Далее Д. Энтони выдвигает предположение, что на северо-востоке данного региона, предположительно — на юге Урала, происходили длительные и регулярные контакты и.-е. общности с уральскими народами, чем и объясняются заимствования в уральском. Позднее, примерно в середине 3-го тыс. до н. э., выделилась и мигрировала в Европу западная ветвь и.-е., где она способствовала смене языка у местного, субстратного населения (никак не атри-

буированного у Энтони). Основной причиной смены языка автор считает наличие социального престижа у завоевателей в глазах местного населения (в работе приводятся аналогичные феномены, относящиеся к более позднему времени). На данном этапе автор предполагает выделение диалекта, называемого им пре-германским (в отличие от протогерманского), в котором он видит своего рода континуум, переходный этап от и.-е. к языку, распавшемуся еще позднее на отдельные диалекты германского. Знакомый с трудами В. Блажека и С. А. Старостина, автор признает, что первыми от и.-е. общности отделились анатолийские языки, затем — тохарские, затем — кельтские (однако предлагает при этом все же придерживаться итало-кельтской теории, отмечая, однако, возможность постоянных контактов и ранних заимствований между группами). Следующей, по его мнению, отделилась греко-армянская группа и, таким образом, индоиранские, балто-славянские и пре-германские языки образовали, по его подсчетам, основывающимся в основном уже на археологических данных, некий диалектный континуум, находящийся в постоянном контакте как бы внутри себя. Эти группы диалектов, как пишет Энтони, локализовались еще в районе Понт-Каспий, откуда собственно и начались их миграции на запад и на север в 4-м тысячелетии до н. э. Более локальной зоной выделения западной ветви и.-е. языков он считает место культур Триполье-Кукутень, носителей которых он идентифицирует как пре-германцев (с элементами других пре-диалектов). Время распада так называемого «позднего Триполья» Д. Энтони датирует 2500—2300 гг. до н. э. и связывает этот процесс с активными миграциями на запад. При этом носители поздне-Трипольской культуры, сконцентрированные в районе между Днестром и Прутом, во многом подверглись интеграции степной культуры Усатова, и именно там,

---

<sup>1</sup> См., например, краткий обзор в статье: Фалилеев А. И. Лингвистическая атрибуция носителей культуры Триполье/Кукутень. О доверии к лингвистическим построениям археологов // *Acta Linguistica Petropolitana*. Т. IV, часть I, СПб., 2008. Ср. также: Дыбо В. А. Язык — этнос — археологическая культура (несколько мыслей по поводу индоевропейской проблемы) // Сравнительно-историческое исследование языков: современное состояние и перспективы. М., 2003.

по его мнению, зародился пре-германский диалект. Именно поэтому, как полагает автор, прямого соотнесения следов материальной культуры, как и данных антропологии, которые могли бы быть непосредственно соотнесены с предками германцев, нет и не может быть, поскольку данный протоэтнос формировался как слияние нескольких этносов и культур. Большую роль в распространении носителей пре-германского языка на территории Европы Д. Энтони отводит доместикации лошади в данном регионе, которая использовалась там (по данным археологии) не только как тяговое животное, но и для верховой езды. В статье уделено значительное место перечислению преимуществ верховой езды как для скотоводства (пастух верхом на лошади может обслуживать значительное число голов крупного рогатого скота), так и для перемещений значительных групп населения, а также, естественно, во время ведения войны. В то же время, как он полагает, именно там были заложены основы социальных взаимоотношений, базирующихся на «взаимном договоре»: как между членами общины (гость-хозяин), так и между человеком и божеством. Выделение своего рода «элиты всадников», носителей и.-е. диалекта, заставило основное степное население (доиндоевропейское) перейти на этот «более престижный» как экономически, так и социально язык. Собственно археологический материал в работе представлен описанием курганных захоронений, которые, как он полагает, были заимствованы частью пре-германцев от степной усатовской культуры. Вопрос об изначальной этнической принадлежности носителей усатовской культуры в статье не ставится.

Ставшая традиционной для этого издания кельтская тематика представлена на этот раз одной работой — статьей Дж. Карни «Недуг уладов: перспектива обмена» (J. Carney. The Pangs of the Ulstermen: An Exchangist Perspective, pp. 52—66). Автор фиксирует свое внимание на том, что герой ирландского эпоса Кухулин почему-то оказывается неподверженным недугу, наложенному на всех уладов оскорбленной ими богиней. После краткого обзора существующих интерпретаций этого феномена Дж. Карни предлагает собственную, отметим — довольно убедительную: Кухулин, который до этого добровольно согласился исполнять функции сторожа дома кузнеца Кулана (вместо убитого им пса) тем самым автоматически перешел в «третью» дюмезилевскую функцию и оказался вне полагающейся ему по рождению группы «воинов», на которых проклятие богини и было наложено.

Автор статьи об отражении мифа о потопе в древнеиндийской традиции Г. Дюкер (G. Ducœr. Passing through Flood Waters in Vedic Thought, pp. 67—77) старается показать древность и аутентичность этого мотива и находит его следы в «Ригведе». Несмотря на то, что в эксплицитной форме миф о потопе представлен лишь в «Шатапатхабрахмане» и традиционно считается заимствованным из ассиро-вавилонской традиции, Дюкер предлагает видеть его имплицитные упоминания уже в гимнах Ригvedы. Так, в метафорических описаниях спасения человека от невзгод мира посредством его *перенесения* через них (ср. устойчивая метафора в русск. пер.: «Как на лодке через реку») в статье предлагается видеть архаические представления о спасении человека от вод на предоставленном божеством корабле. Аналогичным образом автор рассматривает и описание Сомы, которая уподобляется океану, потокам, бурному морю и проч. Миф о спасении человечества посредством перенесения через воды в Ведической традиции представлен в метафорической форме еще шире и, по мнению автора, кодируется основой *pr-*, к и.-е. *\*per-* ‘над, сверх, через’. Таким образом, автор приходит к выводу, что «изложенный в «Шатапатхабрахмане» миф восходит непосредственно к ведической традиции о спасении человечества» (р. 74).

Небольшая, но содержательная заметка Дж. Грегпина (J. Greppin, More Material on the Urartian Substratum in Armenian, pp. 79—87) основывается на работах И. Дьяконова и С. Старостина, согласно которым хуррито-урартские языки родственны лезгинским (точнее — входят в группу нахско-дагестанских — Т. М.). Автор предлагает список из восьми дагестанских лексем, имеющих фонологические и семантические параллели в древнеармянском. По его мнению, они являются ранними заимствованиями в армянский из утраченного пласта языка Урарту.

Статья А. Клокхорста «Некоторые индо-уральские черты в хеттском» (A. Kloekhorst, Some Indo-Uralic Aspects of Hittite, pp. 88—95) посвящена уточнению стадильной датировки выделения анатолийской группы из и.-е. семьи языков. Автор акцентирует свое внимание на двух морфологических аспектах (постопозитивные местоимения-клитики при имени и реконструируемая форма личного местоимения 2 sg.), которые находят параллели в языках уральской семьи (по определению автора — «сестры индоевропейской»). Так, наличие в хеттском безударных форм личных possessивов, которые для и.-е. являются скорее эксклюзивной чертой (possessивные местоименные

клитики в языках Балканского союза автор считает поздней инновацией) и которые также принимают падежную флексию (по схеме: 1 sg.: NOUN + ENDING + mV + ENDING, e. g. *attan-man* > *attamman* 'моего отца', асс. sg.) находит параллели в языках уральских (ср., например, саамск.: *ač'ča-m* 'мой отец'). Т. е. в его формулировке, приведенной выше и.-е. схеме соответствует уральская: NOUN + CASE SUFFIX + mV (р. 92). Эта параллель, как пишет Клокхорст, «слишком поразительна, чтобы быть случайной». Не менее ярким кажется ему и то, что и для хеттского, и для прото-уральского форма личного местоимения 2 sg реконструируется как *\*ti* (для хеттск. *\*tiH*), тогда как для других и.-е. языков она реконструируется как *\*tu*. На базе изложенного автор делает вывод, что уральские языки, как и анатолийские, сохранили архаические черты, присущие индоевро-уральской общности, тогда как все остальные и.-е. языки возникли в результате совместных параллельных инноваций (как их называет автор — CI: Common innovations, р. 94).

В небольшой заметке, посвященной предположительной этимологии греч. βεῦδος 'женская рубаха, туника', А. Лубоцкий на базе анализа контекстов, где встречается лексема, предлагает видеть в ней старое фригийское заимствование: < *bevδος* 'статуя, образ' (к. и.-е. *\*bʰeudʰ-os-* 'восприятие, зрительный образ', см. pp. 96—98).

К. Монетт посвящает свое детальное исследование образу Иного мира в «Беовульфе», «Шахнаме» и ирландском «Похищении быка из Куальнге» (С. Monette. Heroes and Hells in *Beowulf*, the *Shahnameh* and the *Táin Bó Cúailnge*, pp. 99—147). Автор обращает внимание на сходство в описании водяных монстров — противников героя — в указанных памятниках, а также в традиционных описаниях самого поединка протагониста с зооморфным противником. Однако в статье не делается, казалось бы, традиционный же вывод об общих генетических корнях мотива, но предлагается совершенно иной, смелый подход: мотив змееборчества представляет собой относительно позднее заимствование из античной традиции (в основном — из мифов о Геракле, широко распространенных как в Европе, так и в Азии в период раннего Средневековья).

В статье Б. Огибенина предлагается описание нового словарного индекса для санскрита (New

Type Word Indexes for Use in Sanskrit Philology, pp. 148—162). После перечисления существующих «традиционных» (serial indexes) словарей санскрита и их мотивированной оценки автор предлагает новый тип словаря-указателя: обратный словарь (reverse index). Словари такого типа, как пишет автор, также существуют, однако группировка материала в них, по его мнению, далека до совершенной. Б. Огибенин предлагает не выделять в обратном словаре разделы по частям речи, но помещать слова в их изоморфо-фонологической форме, что позволит сделать такой словарь одновременно лексикологическим и грамматическим. К тому же в словаре такого типа предлагается «раскрытие» сандхи от собственно текстового вхождения лексемы до его исходной формы на уровне морфемных швов. Данный принцип, по его мнению, особенно актуален для изучения морфологии (и этимологии) буддийского санскрита (см. р. 162).

Журнал завершается довольно подробным, одновременно лингвистическим и этнографическим исследованием А. Поручича (A. Poruciuc. Two Romanian terms (*tureci* and *cioareci*) based on Old Germanic designations of leg-coverings, pp. 163—184), посвященным не только собственно германским заимствованиям в румынском (для обозначения видов обуви), но и шире — истории костюма варварских обществ периода Империи, значению кожаных штанов для верховой езды, заимствованиям деталей одежды германцев и кельтов римлянами с сохранением соответствующей лексемы и проч., в основном — на базе анализа распространения и семантической вариации так называемого кельто-германского корня *\*braca*, обозначающего кожаное покрытие для ноги (как штаны, так и обувь). Анализ заявленных в названии работы основ также представлен на широком фоне фонетических и семантических вариаций, в котором собственно этимология и непосредственный источник предполагаемого заимствования совершенно теряется. Однако нельзя не согласиться с общей идеей автора о том, что на территории Дакии могли существовать не дошедшие до нас германские диалекты, а также с тем, что семантика лексем, обозначающих детали одежды и обуви, в значительной степени варьирует и иногда в ее основе лежит не собственно объект, но название материала, из которого он сделан.

## № 3—4

Второй выпуск журнала в основном посвящен смежной теме — различным аспектам реконст-

рукции индоевропейской мифологии и поэтики; собственно языкознанию в нем уделено гораздо



меньше места. Приведем названия статей, говорящие сами за себя:

E. R. Anderson and V. Norouzalibeik. *Father-Son Combat: An Indo-European Typescene and its Variations* (pp. 269—332);

J. Fisher. *Asterisking Ennius: The Annales of Quintus Ennius and the Indo-European Tradition* (pp. 333—355);

E. Lyle. *The Marriage and Recovery of the Young Goddess: Story and Structure* (pp. 356—370);

A. F. Erol. *Analyzing the Relationship between the Crescent Shaped Shield and the Amazons* (pp. 411—427);

C. Anghelina. *The Immovable Olympus: Olympus as Zeus' Seat* (pp. 428—441);

J. Colarusso. *The Hunters (Indo-European proto-myths: The Storm God, The Good King, The Mighty Hunter)* (pp. 442—463).

На этом фоне особый интерес представляет короткая, но яркая полемическая заметка С. Вансеверен, посвященная проблеме обнаружения «нового» индоевропейского языка (S. Vanséveren. *A "New" Ancient Indo-European Language? On Assumed Linguistic Contacts between Sumerian and Indo-European "Euphratic"*, pp. 371—382). Работа направлена против серии исследований Г. Витакера (G. Whittaker. *Traces of an Early Indo-European Language in Southern Mesopotamia // Göttingen Beiträge zur Sprachwissenschaft*, 1, 1998; *Word Formation in Euphratic // J. Clackson, B. Olsen (eds.), Indo-European Word Formation, Copenhagen, 2004; The Sumerian Question: reviewing the issues // W. H. Van Soldt (ed.) Ethnicity in Ancient Mesopotamia*, 48<sup>th</sup> RAI. Leiden, 2005 и др.). В свое время, обнаружив в традиционно считающемся изолятом шумерском языке ряд лексем, имеющих сходство с индоевропейскими, Г. Витакер предположил наличие в Южной Месопотамии в 4 тыс. до н. э. не оставившего свидетельств и.е. языка, который мог послужить источником для соответствующих заимствований (он насчитывает более 200 подобных корней). По его мнению, данный язык принадлежал скорее к западному типу и.е. языков, имел ларингалы и уже сформировал женский род. Как пишет С. Вансеверен, «теоретически факт обнаружения «нового» и.е. языка возможен, однако конкретное исследование каждый ряд сталкивается с неразрешимыми трудностями» (р. 371)<sup>2</sup>. По мнению Г. Ви-

такера, особого внимания в шумерском заслуживают отдельные «аномальные» слова, противоречащие традиционной фонотактике: лишенные вокалического сингармонизма, имеющие кластеры, а также представленные вариативно, что может говорить о возможном заимствовании. В своей критике данных идей, как мы позволим себе заметить, С. Вансеверен, возможно, слишком строга. Так, например, предложенное возведение *nun/nin* 'царица' и *gan* 'женщина' к и.е. *\*g<sup>w</sup>en-/\*g<sup>w</sup>neh<sub>2</sub>*, по ее мнению, не могло дать столь вариативный анлаут, но ср. при этом галльск. *geneta, gnata, neta, nata* 'девушка' (Т. М.). Аналогичным образом автор «не может понять», каким образом, согласно Витакеру, и.е. *\*penk<sup>w</sup>e* могло дать шумерск. *kinga* 'пять', как бы забывая о лат. *quinque*... Возражает автор и против предложенного Витакером отождествления шумерского *peš* 'рыба' с и.е. *\*pisk-i-*, утверждая, что данная основа представлена только в латыни, а также германских и кельтских языках, и что само базовое понятие «рыба» на и.е. уровне не реконструируется (р. 375). Это действительно так, однако привлечение славянских данных дает целый разброс продолжений данной и.е. основы для обозначения видов рыб (родо-видовой сдвиг — явление, достаточно частое в диахронической семантике). Кроме того, не следует забывать, что при реконструкции «евфратского» Витакер говорил лишь о заимствованиях в шумерском. Впрочем, на полемику мы не имеем права. Признавая наличие некоторого сходства шумерских и и.е. основ, С. Вансеверин согласна, что при желании здесь можно действительно увидеть и.е. заимствования, но не из эфемерного евфратского, а «из самого протоиндоевропейского» (р. 379), еще не распавшегося на диалекты. Что же касается «нового» индоевропейского языка, то его реконструкция представляется ей «хрупким картонным домиком» (р. 380).

Статья П. Бросмана посвящена анализу фонотактических условий, при которых на протоиндоевропейском уровне (еще до выделения анатолийских языков) был возможен переход *iw > im* (P. Brosman. *PIE iw > im: a Reaffirmation*, pp. 383—400).

В. Сайерс обращается к предположительной этимологии английских слов *dog* и *cur* 'дворняжка' (W. Sayers. *The Etymology of English dog and cur*, pp. 401—410). В первом он предлагает видеть позднее бриттское заимствование из композита *da* 'хороший' и *ci* 'собака' (в ленированной форме *-gi*). Второй термин представляет больше трудностей для реконструкции ввиду его распространения также в скандинавской зоне и в нидерландском. Автор выдвигает осторожное предположе-

<sup>2</sup> Сошлемся в данной связи на недавнюю попытку Дж. Кука реконструировать другой изолят — тартессийский, как один из кельтских языков (J. T. Koch. *Tartessian. Celtic in the South-west at the Dawn of History*. Aberystwyth, 2009).

ние, что это может быть сложным образом трансформированный латинизм: из *canis curialis*, однако сам же говорит о сомнительности подобной реконструкции. Статья завершается выводом, что реконструкция названий такого важного для культу-

ры понятия, как *собака*, должно проводиться с учетом разнообразия использования этого domesticiрованного животного — на охоте, для охраны дома, для охраны разных видов скота и проч. (р. 408).



Пятые традиционные чтения памяти С. А. Старостина  
Москва, РГГУ, 2—3 апреля 2010 г.

5-я ежегодная конференция «Чтения памяти С. А. Старостина» состоялась 2—3 апреля 2010 года, как обычно, под эгидой Института восточных культур и античности РГГУ. «Старостинские чтения» ежегодно организуются сотрудниками Центра компаративистики Института восточных культур и античности в память об основателе Центра и его первом руководителе Сергее Анатольевиче Старостине (1953—2005), с именем которого неразрывно связан расцвет т. н. Московской школы сравнительно-исторического языкознания.

По традиции Чтения открыли директор Института И. С. Смирнов и руководитель Центра компаративистики, чл.-корр. РАН В. А. Дыбо. В конференции принимали участие сотрудники Института, аспиранты и студенты РГГУ, а также наши коллеги из МГУ и других учебных и научных организаций Москвы и Санкт-Петербурга.

Основной тематикой Чтений традиционно являются доклады по истории языков и языковых семей мира, их близнему и дальнему генетическому родству, спорным вопросам их исторической фонетики, грамматики, лексикологии. Именно эти сообщения образовали основной блок конференции — 20 из 24 докладов.

Отдельные сообщения были сгруппированы вместе по принципу сфокусированности на той или иной языковой семье или языковом регионе (индоевропеистика, уралистика, африканистика и т. д.). Естественно, больше всего докладов относилось к индоевропейской семье (П. В. Башарин, «Субстратные заимствования в гилянском»; В. А. Дыбо, «О прагерманском ударении»; А. И. Коган, «О критериях выделения индоарийских заимствований в языке кашмири»; Т. А. Михайлова, «Еще раз о поло-возрастной стратификации, или куда делась индоевропейская «дочь» в островном кельтском»; И. С. Якубович, «Этимология и внутренняя форма хеттского *anīya* ‘делать’»). За ней следуют афразийская (Л. Е. Коган, «Диахроническая стра-

тификация базовой лексики современных южно-аравийских языков»; С. В. Лёзов, «Грамматическая реконструкция правосточноарамейского и прарамейского языков»; В. Я. Порхомовский, «Традиционные грамматики и глубокая морфологическая реконструкция» и О. В. Столбова, «О некоторых вопросах чадской лексикостатистики»); уральская (М. А. Живлов, «Два правила из истории мордовского вокализма»); Ю. В. Норманская, «К вопросу о финно-угорской прародине»; К. Ю. Решетников, «Несколько новых угорских этимологий»); алтайская (А. В. Дыбо, «К проблемам огузской морфонологии: процессы конца слова и генетическая классификация»; Е. В. Коровина, «К реконструкции тунгусо-маньчжурских названий растений»); северокавказская (А. С. Касьян, «Некоторые мысли по поводу лингвистической атрибуции майкопской археологической культуры»; О. А. Мудрак, «Связь ударения и окончаний имен в аварском»). Другие доклады касаются еще трех семей (койсанской — Г. С. Старостин, «Новые данные по лексикостатистике койсанских языков»; манде — В. Ф. Выдрин, «Лингвистические данные к ранней этноистории Фута-Джаллона»; полинезийской — А. В. Козьмин, «Реконструкция одного протополинезийского текста»).

Но большинство докладов было легко расклассифицировать и по их теоретической значимости. Первое место в этом отношении заняли текущие проблемы лексикостатистики — метода определения степени языкового родства по количеству совпадений в базисной лексике сравниваемых языковых семей. Они обсуждались на материале языков чадской семьи в докладе О. В. Столбовой, койсанской (бушмено-готтентотской) семьи в докладе Г. С. Старостина, и семьи манде в докладе В. Ф. Выдрина. Активную дискуссию вызвал доклад Л. Е. Когана о базисной лексике языков Южной Аравии, демонстрирующей весьма «необычное» поведение с точки зрения стандартных ожиданий лексикостатистики.

Другой метод генетической классификации, по пучкам «связанных» диахронических правил (автор категорически возражает против названия его методом общих инноваций), рассматривался в докладе А. В. Дыбо «К проблемам огузской морфологии: процессы конца слова и генетическая классификация». Рассмотрены были разноуровневые, но связанные между собой процессы фонетического изменения и парадигматических унификаций, которые позволяют для каждого из огузских языков построить однозначную временную последовательность правил. На основании сопоставления этих последовательностей предлагается новая классификация огузских языков, с ранним выделением азербайджанского.

Тема грамматической реконструкции и ее роли в языковой классификации была представлена докладами К. В. Бабаева («К вопросу о диахроническом развитии серий личных показателей в языках мира»), С. В. Лезова («Грамматическая реконструкция правосточноарамейского и праарамейского языков») и В. Я. Порхомовского («Традиционные грамматики и глубокая морфологическая реконструкция»).

Проблемы сравнительно-исторической акцентологии были рассмотрены двумя участниками конференции. В докладе В. А. Дыбо «О прагерманском ударении» было показано, что косвенные данные о месте прагерманского ударения (чередования в связи с законом Вернера и рефлексы по закону Хольцмана) указывают на его близость к прабалтославянскому. В частности, удается восстановить деление германских глаголов на баритонированные и окситонированные, и эти группы обнаруживают хорошую материальную корреляцию с соответственно балтославянскими неподвижноударными и подвижноударными. Проблемам акцентуации был также посвящен доклад О. А. Мудрака «Связь ударения и окончаний имен в аварском». Автор предложил частичное распределение аварских типов акцентных парадигм по сегментному облику конца основы (выделив таким образом три парадигмы) и — в ответ на уточняющие вопросы — указал собравшимся на необходимость дальнейших занятий прааварской акцентуацией.

В основном к акцентологической проблематике относились также положения А. В. Циммерлинга («Порядок клитик и полноударных элементов предложения»).

Проблемы реконструкции сегментной фонетики рассматривались только в одном докладе:

М. А. Живлов, «Два правила из истории мордовского вокализма». Докладчик рассмотрел судь-

бу прафинно-угорских гласных  $*\ddot{a}$  и  $*\bar{e}$  перед  $*e$  второго слога в мордовских языках. Для всех контекстов кроме интервокального  $*-k-$  и  $*-\gamma-$  формулируется, что в именах ПФУ  $*\ddot{a}-e$ ,  $*\bar{e}-e >$  эрзя  $e$ , мокша  $\ddot{a}$ , а в глаголах ПФУ  $*\ddot{a}-e$ ,  $*\bar{e}-e >$  эрзя  $i$ , мокша  $i$ . Для случаев перед интервокальным  $*-k-$  и  $*-\gamma-$  рефлексы другие: эрзя  $i$ , мокша  $i$ ; эрзя  $e$ , мокша  $\ddot{a}$ ; эрзя  $i \sim e$ , мокша  $i$ ; эрзя  $u \sim o$ , мокша  $u$ .

Вопросы методологии и практики реконструкции прародин языковых семей были затронуты Ю. В. Норманской для уральской семьи («К вопросу о финно-угорской прародине») и Е. В. Коровиной для алтайской семьи («Еще раз о тунгусо-маньчжурской прародине: названия деревьев»). В докладе Ю. В. Норманской рассматривалась возможная локализация финно-угорской прародины на материале лексики, касающейся природного окружения (растений, рыб, животных и др.). Было отмечено, что очень важно соотносить языковые данные не с современным распространением этих видов, а с данными на предполагаемое время распада, поскольку ареалы в историческое время могли измениться и изменяться. Показано, что согласно языковым данным прафинно-угры должны были занимать ареал обитания лосося, уклеи и соболя, что попадает на северо-западе на верховья Печоры и Северный Урал, а на юго-востоке — на север Южного Предуралья. Проблеме локализации прародины был также посвящен доклад Е. В. Коровиной «К реконструкции тунгусо-маньчжурских названий растений»; результаты исследования указывают на среднее течение Амура (северо-восток Бурейского района).

Близкой теме — проблеме соотнесения археологических культур и этносов — был посвящен доклад А. С. Касьяна «Некоторые мысли по поводу лингвистической атрибуции майкопской археологической культуры». В нем утверждалось, что носители майкопской культуры говорили на северокавказском праязыке и пришли на территорию Крыма с северо-запада. Это подтверждается тем, что прасеверокавказцам было известно коневодство, скотоводство, агрокультура, текстильное производство и металлургия, что в точности соответствует характеристикам майкопской культуры.

Много докладов было посвящено проблемам выделения и анализу заимствований. Это доклад П. В. Башарина «Субстратные заимствования в гиланском», где предлагался ряд вероятных заимствований северокавказского происхождения. А. И. Коган в своем докладе «О критериях выделения индоарийских заимствований в языке кашмири» говорил о разноуровневых и разноязычных груп-

пах индоарийских заимствований в этом дардском языке, выделяемых на основании типов фонетических соответствий. Сюда же можно отнести доклад С. А. Яцемирского «Основные минойские и тирренские заимствования в греческом и способы их выделения».

Наконец, такая классическая тема, как разработка отдельных этимологий, также занимала участников конференции, причем использовался материал различных семей. Доклад Т. А. Михайловой «Еще раз о поло-возрастной стратификации, или куда делась индоевропейская «дочь» в островном кельтском» был посвящен проблеме семантических переходов. В нем на примере анализа кельтской лексики со значением 'дочь' было показано, что они имеют не однонаправленный, а спиральный характер, поскольку кельтское слово претерпело в своем развитии следующее изменение значений: девочка > дочь > служанка > дочь. И. С. Якубович в докладе «Этимология и внутренняя форма хеттского *anīya* 'делать'» проанализировал этот глагол как приставочное производное с помощью анатолийской приставки *\*an-* «с» от глагола *iya* «делать». В докладе К. Ю. Решетникова «Несколько новых угорских этимологий» был проэти-

мологизирован ряд обско-угорских слов (*\*waysa* «тонкий», *\*wejā* «засыпать», *\*ōsar* «злой дух») и несколько слов, относящихся к венгерской анатомической лексике (*váz* «железа», *mörvény*, *mörigy* «костяк»).

Помимо этих и других докладов, в целом не выходящих за рамки вопросов исторического языкознания, финальная часть конференции была — опять-таки в соответствии с уже сложившейся традицией, стремящейся максимально полным образом отразить разносторонность научных интересов С. А. Старостина — отдана под ряд докладов смешанного характера. Так, А. В. Козьмин представил результаты любопытного исследования из области сравнительной текстологии («Реконструкция одного протополинезийского текста»). По преимуществу текстологическим был также доклад П. М. Кожина «Некоторые темные места в «Слове о Полку Игореве»». Наконец, С. А. Крылов с помощью своего доклада («Инвентаризация лингвистических знаний с помощью металингвистических баз данных в интегрированной информационной среде StarLing») напомнил участникам Чтений о немаловажных заслугах С. А. Старостина в области компьютеризации лингвистической науки.