

ВЕСТНИК РГГУ

RSUH BULLETIN

Ежемесячный научный журнал

Scientific Monthly

Серия «Языкознание»

Linguistics Series

№ 5 (2009)

Москва

Moscow

2009



Институт языкознания Российской Академии наук
Российский государственный гуманитарный университет

Вопросы языкового родства

Международный научный журнал

№ **1** (2009)

ВЯР

Москва 2009

Institute of Linguistics of the Russian Academy of Sciences

Russian State University for the Humanities

Journal of Language Relationship

International Scientific Periodical

N° **1** (2009)



Moscow 2009

Редакционный Совет:

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. КОРПЛАНДТ (Лейден)

А. ЛУБОЦКИЙ (Лейден)

Редакционная коллегия:

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

О. А. МУДРАК

В. Е. ЧЕРНОВ

Advisory Board:

Chairman:

Vyach. Vs. IVANOV (Moscow–Los Angeles, Calif.)

M. E. ALEXEEV (Moscow)

W. BAXTER (Ann Arbor, Mich.)

V. BLAŽEK (Brno)

A. B. DOLGOPOLSKY (Haifa)

M. GELL-MANN (Santa Fe, New Mexico)

F. KORTLANDT (Leiden)

A. LUBOTSKY (Leiden)

V. F. VYDRIN (Saint-Petersburg)

Editorial Staff:

V. A. DYBO (Editor-in-Chief)

G. S. STAROSTIN (Managing Editor)

T. A. MIKHAILOVA (Editorial Secretary)

K. V. BABAEV

S. G. BOLOTOV

V. Ye. CHERNOV

A. V. DYBO

O. A. MUDRAK

Вопросы языкового родства: Международный научный журнал / Рос. Акад. наук. Ин-т языкознания; Рос. гос. гуманитар. ун-т; под ред. В. А. Дыбо. — М.: Изд-во РГГУ, 2009. — № 1. — XII+164 с. — (Вестник РГГУ: Ежемесячный научный журнал; Серия «Языкознание»; № 5).

Journal of Language Relationship: International Scientific Periodical / Russian Academy of Sciences. Institute of Linguistics; Russian State University for the Humanities; Ed. by V. A. Dybo. — Moscow: RSUH Publishers, 2009. — N° 1. — XII+164 p.. — (RSUH Bulletin: Scientific Monthly; *Linguistics* Series; N° 5).

ISSN 1998-6769

<http://journal.nostratic.ru>
journal@nostratic.ru

Гарнитура *Таймс Нью Роман* / *Times New Roman*TM typeface

© 2006 The Monotype Corporation

Дополнительные знаки: С. Г. Болотов / Add-on symbols by S. G. Volotov

Компьютерная верстка: С. Г. Болотов / Typeset by S. G. Volotov

© 2008

ISSN 1998676-9



9 771998 676003

Note from the Editors

Dear friends and colleagues!

It is a great pleasure for us to finally be able to present the first issue of our brand new “*Journal of Language Relationship*”. The *Journal*, jointly issued by the Russian State University for the Humanities and the Institute of Linguistics of the Russian Academy of Science, is a peer-reviewed edition that will be published on a semi-yearly basis and, as its title implies, will be fully dedicated to issues of establishing, verifying, and clarifying various aspects of genetic relationship between the world’s languages and language groups.

The subject is, of course, anything but new — work on genetic relationship lies at the heart of comparative linguistics, and thus, from a historical point of view, at the heart of linguistics as a whole — but even today, remains an inexhaustible source of research, discoveries, arguments, debates, and controversies. What really constitutes genetic relationship? How should one be able to demonstrate such a thing? Which of the existing theories and hypotheses on genetic relationship should be accepted and which ones should be rejected, and based on what criteria? None of these questions have so far been answered to the satisfaction of even most, let alone all, scholars in the field, and we sincerely hope that our *Journal* will, at least to a certain extent, eventually help alleviate some of these issues.

One of the most hotly debated topics in comparative and historical linguistics today is that of language relationship on deep chronological levels — and the question of whether the classic comparative method, still the basic foundation of our science, really has an impenetrable time limit to it, and if yes, what kind of limit it really is. Our *Journal* heartily welcomes publications on this issue, theoretical and methodological as well as those that try to offer mostly practical, data-based answers to such questions. We believe that only a proper combination of both approaches can eventually help overcome what is sometimes seen as a serious lack of mutual understanding between the so-called “lumpers” and “splitters” in historical linguistics.

Since the *Journal* is essentially based in Moscow, it is probably inevitable that, for the initial period at least, much of its space will be dedicated to papers written from the point of view of the so-called Moscow school of comparative linguistics, whose working style for the past fifty years or so could arguably be best described as “applying traditional linguistic methodology to non-traditional language families”, the latter including both relatively short-range, but poorly studied language stocks as well as hypothetical macrofamilies. However, the *Journal* will by no means restrict itself to any kind of one-sided, narrow approach to language comparison, and, as time goes by, we hope to cover all kinds of ground, for which reason we welcome all our colleagues to participate.

In addition, not even the “Moscow school” should be thought of as something rigid and close to internal discussion. One thing that can be said about comparative linguistics, for better or for worse, is that no two people in this field of study hold the same opinions on every detail of it, and, apart from those precious few “axioms” of comparative linguistics that define it as science rather than science fiction, there is no, nor will there ever be, set ideological stance about a great many aspects of its methodology and practical application.

Some of these problems and controversies are discussed in greater detail in the opening article by Vyach. Vs. IVANOV, an attempt to summarize the main goals and perspectives of comparative-historical linguistics as a whole, as well as to illustrate them on the basis of the author’s personal experience with ancient languages of the Near East. Another theoretical overview, outlining the achievements of the joint American-Russian *Evolution of Human Languages* project (conducted by the Santa Fe Institute in New Mexico), comes from M. GELL-MANN, I. PEIROS, and G. STAROSTIN. These achievements, mainly based on the methodology of the Moscow school, range from what is now considered solid scientific hypotheses to preliminary speculations, and will unquestionably stimulate further discussion.

Parts of the traditional comparative methodology that is used to reach these conclusions, however, are taken to task in V. BELIKOV's article, dedicated to the issue of language contacts and their impact on existing schemes of genetic classification; the author's caveats should certainly not be taken lightly by those in favor of advancing radical hypotheses on the world's linguistic prehistory. One of the most common arguments in such hypotheses — that of the personal pronouns as potentially indicative of deep-level linguistic relationship — is discussed in K. BABAEV's paper from a "paradigmatic" point of view.

One of the main standing points of our *Journal* will be to dedicate as much space as possible to practically oriented, "data-heavy" publications, which we see as truly essential to the development of our science. Historical scenarios and conclusions do not carry much weight when not properly illustrated by data, and, to that effect, V. VYDRIN's article on the genetic classification of the South Mande languages and the potential homeland of Proto-South Mande is accompanied by a near-exhaustive list of 100-wordlists for most of the concerned languages, so that his statistical conclusions can be easily checked by even those readers who do not have a serious background in African linguistics. Other data-relying publications include A. MILITAREV's paper on pastoral lexicon in Proto-Afrasian and V. BLAŽEK's paper on some possible traces of an Australian substratum in Dravidian.

Finally, the volume is made even more representative through the inclusion of two papers on, respectively, North Caucasian (by M. ALEXEYEV) and Iranian linguistics (by J. EDELMAN). These will also be of interest to both specialists in the field and general comparative linguists, as the issues raised therein are still actual for other language families as well.

In addition to theoretical and practically oriented papers, as well as a *review* section, we plan to expand our future issues by adding a *discussion* forum, with at least one paper per publication open to comments from scholars in the field, as well as *reports* on various linguistic activities (conferences, seminars, &c.). For now, the most important thing is that the *Journal* be able to reach its target audience — all those interested in comparative-historical issues — and become a modest vehicle for stimulating such interest among the linguistic and general scientific community.

От Редакции

Дорогие коллеги!

Первый номер нашего нового журнала *«Вопросы языкового родства»* наконец вышел из печати. Журнал является рецензируемым и издается совместно Российским государственным гуманитарным университетом и Институтом языкознания РАН. Планируется выпускать по два номера в год. Тематика журнала — сравнительно-историческое языкознание в принятом смысле этого термина, то есть работа по установлению языкового родства, его доказательству и уточнению различных аспектов генетических связей между различными языками мира.

Установление языкового родства — это основная задача сравнительно-исторического языкознания (а сравнительно-историческое языкознание, как известно, положило начало всему дальнейшему развитию науки о языке). Но и по сей день между специалистами-языковедами не достигнуто согласие относительно того, в чем, собственно, состоит эта задача, в каких случаях мы считаем, что родство группы языков достаточно аргументировано, чтобы считать его доказанным. Мы надеемся, что наш журнал внесет некоторый вклад в решение этой проблемы. По крайней мере, он, несомненно, будет способствовать выработке совместимых позиций по этому поводу хотя бы у части исследователей.

Одна из наиболее жарко дискутируемых тем сравнительно-исторического языкознания — возможность работы в области отдаленного родства языков в рамках классической компаративистской процедуры, которая остается общепринятым основанием нашей науки. Есть ли у сравнительно-исторического метода временные ограничения, и каковы эти ограничения? Редколлегия журнала будет приветствовать любые исследования по этой тематике, как теоретические, так и (тем более) практические, основанные на реальном языковом материале. По-видимому, только сочетание обоих этих подходов может помочь установлению взаимопонимания между сторонниками и противниками дальнего сравнения.

Поскольку журнал издается в Москве, можно ожидать, что, по крайней мере, на начальной стадии его существования большую часть его печатной площади будут занимать работы, написанные в рамках Московской школы сравнительно-исторического языкознания. Основная особенность этой школы — так сказать, применение традиционных методов к нетрадиционным языковым семьям, то есть либо сравнительно-исторические исследования по установлению отдаленного родства, либо компаративистическая обработка до сих пор мало изученных в этом отношении, хотя и не очень глубоких по времени языковых групп. Следует, однако, подчеркнуть, что редколлегия ни в коем случае не планирует ограничиваться односторонним подбором статей и рассчитывает со временем при содействии коллег покрыть все области компаративистики.

Кроме того, не следует думать, что даже Московская школа сравнительно-исторического языкознания представляет собой единый монолит в части методологических воззрений. Она вполне открыта для внутренних и внешних дискуссий — исключая лишь незыблемые постулаты, на которых держится сам сравнительно-исторический метод, и благодаря которым лингвистическая реконструкция остается наукой, а не превращается в научную фантастику.

Более детально часть упомянутых вопросов обсуждается во вводной статье Вяч. Вс. Иванова, в которой основные проблемы и перспективы сравнительно-исторического языкознания обобщены и проиллюстрированы примерами из личного опыта автора, работающего с языками древней Передней Азии. Еще одна теоретическая статья — М. Гелл-Манн, И. И. Пейрос, Г. С. Старостин, *«Текущее состояние и перспективы изучения дальнего родства языков»*. В ней подводятся итоги российско-американского проекта *«Эволюция языков человечества»* (Институт Санта Фе). Статья демонстрирует широкий спектр результатов этой работы, основанной на методологии Московской школы сравнительно-исторического языкознания — от солидно обоснованных научных гипотез до импрессионистических рассуждений, — и, несомненно, возбудит дальнейшие дискуссии.

Традиционные постулаты сравнительно-исторического метода подвергаются критике в статье В. И. Беликова, специалиста, известного как постоянного оппонента компаративистов, рассматривающих, по его мнению, только дивергентные процессы в истории языков и не уделяющих достаточного внимания конвергентным процессам. Конечно, его подход оставляет мало шансов сторонникам глубокого сравнения, стремящимся к реконструкции «праязыка человечества». В статье К. В. Бабаева, напротив, обсуждается один из наиболее частых аргументов в гипотезах отдаленного родства — совпадение/различие личных местоимений. Предлагается рассматривать набор личных местоимений в языке как целостную парадигму, что должно повысить доказательность аргумента.

В первом номере журнала естественно некоторое преобладание статей, обсуждающих методологические предпосылки исследования, над самими исследованиями. Мы, однако, хотели бы, чтобы в дальнейшем большее место занимали работы «практического» характера, содержащие в себе, в частности, полный языковой материал, обеспечивающий аргументацию автора. Такие статьи в первом выпуске нашего журнала представлены, в первую очередь, работой В. Ф. Выдрина о классификации языков группы Южные Манде, в которой приводится почти полный набор сводешевских стословников для рассматриваемых языков. Это позволяет оценить авторские выводы даже читателю, слабо знакомому с африканистикой. Две другие такие статьи — работа А. Ю. Милитарева о скотоводческой лексике в праафразийском и работа В. Блажека о возможных следах австралийского субстрата в прадравидийском.

Традиционно изучаемые семьи представлены статьями М. Е. Алексеева о северокавказском аблауте и Дж. И. Эдельман об иранских языках. Эти работы интересны не только для «практической» компаративистики, но и для общего языкознания; в них поднимаются вопросы, с которыми часто приходится сталкиваться специалистам по различным языкам мира.

В дальнейшем мы рассчитываем расширить раздел рецензий и добавить разделы «Хроника» и «Материалы к дискуссии» — последний должен состоять хотя бы из одной статьи с комментариями специалистов в каждом номере.

Будем надеяться, что нашему журналу удастся внести свою лепту в дело возрождения и поддержания интереса к сравнительно-историческим исследованиям не только в профессиональных кругах, но и в широкой лингвистической аудитории.

Table of Contents / Содержание

<i>Note from the Editors</i>	VII
<i>От редакции</i>	IX
Table of Contents / Содержание.	XI
<i>Note for Contributors / Будущим авторам</i>	XII
Articles / Статьи	1–142
Вячеслав Вc. ИВАНОВ. <i>К исследованию отношений между языками</i>	
Vyach. Vs. IVANOV. <i>Certain Issues of Current Research on Relations between Languages</i>	1
Murray GELL-MANN, ПУА PEIROS, George STAROSTIN. <i>Distant Language Relationship: The Current Perspective</i>	
Мюррей Гелл-Манн, Илья Пейрос, Георгий Старостин. <i>Текущее состояние и перспективы изучения дальнего родства языков</i>	13
М. Е. АЛЕКСЕЕВ. <i>Аблаут в глагольных основах нахско-дагестанских языков</i>	
M. ALEKSEYEV. <i>Vowel Gradation in Nakh-Daghestanian Verbal Stems</i>	31
Kirill БАБАЕВ. <i>Once Again on the Comparison of Personal Pronouns in Proto-Languages</i>	
К. В. БАБАЕВ. <i>К вопросу о сопоставлении личных местоимений в реконструированных праязыках</i>	37
В. И. БЕЛИКОВ. <i>Языковые контакты и генеалогическая классификация.</i>	
V. BELIKOV. <i>Language Contacts and Genealogical Classification of Languages</i>	49
Václav ВЛАЖЕК. <i>Dravidian Numerals</i>	
В. БЛАЖЕК. <i>Числительные в дравидийских языках</i>	69
Д. И. ЭДЕЛЬМАН. <i>Некоторые проблемы сравнительно-исторического иранского языкознания</i>	
J. EDELMAN. <i>On certain Problems of Comparative Iranian Linguistics</i>	81
Alexander МИЛИТАРЕВ. <i>Proto-Afrasian Lexicon Confirming West Asian Homeland: Pastoralism</i>	
А. Ю. МИЛИТАРЕВ. <i>Праафразийская скотоводческая лексика как подтверждение переднеазиатской прародины афразийцев</i>	95
Valentin ВYДРИН. <i>On the Problem of the Proto-Mande Homeland</i>	
В. Ф. ВYДРИН. <i>К проблеме определения прародины языковой семьи манде.</i>	107
Reviews / Рецензии	143–162
А. В. ДЫБО / А. V. ДУВО: К. В. БАБАЕВ. <i>Происхождение индоевропейских показателей лица: Исторический анализ и данные внешнего сравнения. 2008</i>	143
George STAROSTIN / Г. С. СТАРОСТИН: Axel SCHUESSLER. <i>ABC Etymological Dictionary of Old Chinese. 2007</i>	155

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

Articles are published preferably in English or Russian, although publication of texts in other major European languages (French, German, &c.) 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.

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

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).

К исследованию отношений между языками

Статья посвящена краткому обзору актуальных вопросов сравнительно-исторического языкознания, удовлетворительное решение которых автору представляется необходимым для дальнейших прорывов в этой области. К таким вопросам, в частности, относятся лексико-семантическая реконструкция; успешное применение морфологического и лексикостатистического критериев для установления границ языковых семей и макросемей; проблемы фонетической интерпретации реконструированных фонологических систем; усовершенствование критериев разграничения между заимствованной и генетически родственной лексикой. Возможные подходы к решению этих и других вопросов проиллюстрированы автором на материале древних и реконструированных языков Передней Азии.

0. За более чем два века существования сравнительного языкознания накоплен не только огромный фактический материал, относящийся к надежно установленным *родственным связям* между языками, но и достаточно разнообразный набор теоретических предположений и допущений, позволяющих ориентироваться в *основаниях* этой области науки. В связи с началом издания предлагаемого вниманию читателей нового журнала, посвященного таким исследованиям, представляется целесообразным обозреть важнейшие из тех проблем, с которыми сталкиваются все, этой областью знания занимающиеся. Ниже кратко говорится об ее теперешнем состоянии, причем внимание обращено на спорные и дискуссионные моменты, особенно нуждающиеся в совместном обсуждении специалистов по разным семьям языков.

1. Основой для *установления родства* двух и более языков (как древнеписьменных, так и современных или известных благодаря применению к двум первым категориям процедур реконструкции) является выявление систематических отношений между множествами *лексических* и *грамматических морфем* и их сочетаний в *словах*, в этих языках засвидетельствованных. Рассмотрим отдельно вопросы, касающиеся лексики и грамматики.

2.1. Сопоставление определенных *частей словарей* двух и более языков остается главным предварительным инструментом установления родства. Исследователи различаются по характеру критериев отбора лексического материала. За последние десятилетия достигнуто существенное продвижение в сравнении базисных (*стословных*, но также и *200-словных*) списков слов в соответствии с идеями *лексикостатистики* и *глоттохронологии*. Ряд уточнений к исследованиям М. Сводеша и его продолжателей (ср. обзор взглядов и прежней литературы [Арапов, Херц 1974]) был предложен в работах, собранных в книге [Старостин 2007], в частности, по новому методу могут сравниваться не слова, а *корневые (лексические) морфемы в текстах определенной длины*¹. Тем не менее вероятно продолжение обсуждений, касающихся, в частности, различий в коэффициенте, определяющем *темп обновления базисного словаря и скорость лексического (а также, возможно, структурного — фонологического и грамматического) расхождения родственных языков* в зависимости от внешних и внутренних факторов эволюции языка. Существенной остается степень географической и исторической изолированности говорящих на данном языке, ср. пример исландского языка и его особой лексической консервативности. Значительный интерес представляет проблема внешних факторов, влияющих на число заимствований и соответственно на темп изменения словаря. Хотя *пиджины и креолизированные языки* представляют в целом особую проблему, тем не менее история языка должна считаться с наличием случаев, где, как в языке недавно найденных клинописных текстов из Катна (Qatna) в Сирии (2 тыс. до н. э.), словарь глагольных основ в целом отличался от словаря имен существительных — один отражал аккадское письменное наследие, другой — происходящую из традиции другой макросемьи лексику хурритского разговорного диалекта (об этом диалекте ср. анализ глосс у [WEGNER 2007: урок 15]; типологически сходное различие по происхождению глагольной и именной частей словаря отмечается в медновском смешанном алеутско-русском креольском языке).

2.2. Другим практически опробованным способом установления родственных связей между языками, место которых в генеалогической классификации оставалось неясным, явилось *массовое лексическое сравнение*, предполагающее значительно больший объем сохранения

¹ См., например, сравнение хеттского словаря с лувийским: [ИВАНОВ 2001].

древней части словаря. Безусловные положительные результаты, достигнутые в свое время на этом пути Дж. Гринбергом (в работах по классификации языков Африки), не позволяют сомневаться в эффективности этого приема. Но теоретические его основания, в том числе, определение границ рассматриваемого множества слов, остаются недостаточно разработанными.

2.3. Значительные успехи современной лингвистической науки в области изучения смысловых сфер языка и референции делают возможным и желательным детальное диахроническое изучение больших *семантических полей*. Это занятие дает проверяемые результаты и часто допускает *фальсификацию* (в смысле философии науки К. Поппера) при сопоставлении денотатов (референтов) слов с данными смежных наук. Но вместе с тем оно является трудоемким и может поэтому замедлить ход соответствующих сравнений словарей в целом. Поэтому такое исследование скорее должно поощряться на более поздних этапах, после предварительного решения общей задачи установления родства.

2.4. Для развития исторического языкознания в целом назрела задача систематического исследования *типологических закономерностей фонетического, грамматического и семантического изменения* (в том числе и изучаемого статистически на максимально широком материале разных семей языков и сопоставления различных документированных периодов в истории и/или различных ареалов в пространственной лингвистике). Отсутствие достаточно проверенных выводов часто заставляет ограничиваться более или менее случайным набором аналогий. Между тем использование компьютерных баз данных открывает перспективы для создания вполне надежных собраний обобщений в этом разделе лингвистики.

3.1. Как показал опыт детального изучения индоевропейской семьи языков и ее отдельных ветвей, по крайней мере для языков такого типа (*синтетических* с хорошо развитой *морфологией*, значительной ролью *фузии* в структуре слова, обычно *многоморфемного* и *многосложного*, и наличием системы *морфонологических* правил, в том числе акцентуационных) наиболее отчетливые результаты в диахронических исследованиях получены при сравнении *грамматических морфем*. Однако по отношению ко многим семьям языков с другими типологическими характеристиками (в частности, *изолирующих* и *аналитических*, с преобладанием слов, совпадающих с *морфемой* и *слогом*) не имеет смысла (во всяком случае, по отношению к непосредственно доступному для наблюдения периоду истории) говорить о морфологии.

3.2. В таких языках исследователь сосредотачивается на морфемах, выделяемых по их *синтаксической* функции и имеющих *грамматическое значение* (обязательное для целых классов сочетаний), но не выделяемых в пределах одного слова. При всем существенном отличии грамматического типа таких языков (в частности, Юго-Восточной Азии, Африки и Центральной Америки) нет оснований предполагать, что для их сравнительно-исторического изучения потребуются какие-либо особые методы². Современное состояние исследования *универсалий языка в общей грамматике* позволяет думать, что при всех локальных различиях и доказанном существовании редких явлений основные принципы лингвистической структуры одинаковы во всех языках мира. Это можно считать проверенным во всяком случае на материале более тысячи языков, см. [Насрепматн & al. 2005], т. е. примерно шестой части всех существовавших к началу нашего века языков; среди них пока не обнаружено ни одного языка с отступлением от предполагаемых универсальных черт³. В диахронии не существует строгой границы между грамматическими типами. Древнекитайский язык, в классическую эпоху представляющий образец языка без морфологии, в предшествующий период характеризовался наличием падежных форм и противопоставленных друг другу по функции типов глагольных переходных и непереходных основ.

4.1. После того, как с помощью сравнения лексических и грамматических морфем и выявления закономерного характера соответствий между ними установлено родство языков, оказывается возможным построение *системы соответствий*. Следует (в соответствии с тем метаязыковым описанием сравнительного метода, которым наука обязана А. Мейе и Ф. де Соссюру) подчеркнуть *строго формальный характер тех операций*, которые производятся на этом этапе (по этой причине они вполне доступны для алгоритмической формулировки и хотя бы частичной компьютериза-

² Этот теоретический вопрос изучался в ряде последних работ выдающегося исследователя многих языков Юго-Восточной Азии Ю. К. Лекомцева, занятия которого были рано оборваны трагической смертью.

³ Обсуждаемые во многих статьях последних лет особенности числительных в пирахан и некоторых других языках Амазонии в целом не дают оснований для вывода о нарушении общетипологических закономерностей ([Иванов 2005; 2007]).

ции, хотя даже и по отношению к наиболее хорошо изученным индоевропейским языкам мы находимся в самом начале такой работы). Устанавливаемые соответствия между языками представляют *абстрактную систему соотношений*, определяющих реконструируемые единицы только как элементы этой формальной системы. Юношеское исследование де Соссюра (в частности, в том его формализованном изложении, которое принадлежит Ельмслеву [1952]) до сих пор остается образцом формальных операций над системой соответствий, позволяющих выявить наиболее раннюю структуру, получаемую из наблюдаемой посредством вычислений типа математических.

4.2. После построения абстрактной системы соответствий и произведенной с ее помощью строго формальной реконструкции оказывается возможным предложить на основе общетипологических соображений в сочетании с другими косвенными данными фонетическую *интерпретацию* полученной абстрактной системы. Четкое разделение реконструируемой абстрактной системы (или нескольких альтернативных систем, которые теоретически могут быть результатом компьютерных операций над соответствиями) и ее интерпретации представляется необходимым для строгого метаязыкового понимания принципов сравнительного языкознания. Реконструированных (человеком-исследователем или компьютером) систем может быть несколько, как и способов их интерпретации (ср., например, современное состояние вопроса о характере серий индоевропейских смычных и фонем, в которые они развились в хеттском и других анатолийских языках). Важнейшим вопросом являются *критерии выбора наиболее адекватной системы* из многих теоретически возможных. Приходится настаивать на том, что в число критериев не должно входить число повторений традиционной точки зрения в учебниках и других пособиях.

5. Рассмотрим вопрос о критериях выбора абстрактной фонологической системы и ее фонетической интерпретации на примере недавно выявленных соответствий между вновь изученным южно- (или западно-) анатолийским *лувийским* языком и ранее сравнительно полно исследованным северно- (или восточно-) анатолийским *хеттским*. Оказалось, что в ряде случаев начальному хеттскому *š-* отвечает клинописное лувийское *t-/d-* (глухой и звонкий согласный в начальном положении в этом виде клинописи не различаются) при отсутствии соответствующего согласного в родственных формах остальных индоевропейских языков:

(I) хеттск. *šakuwa* ‘глаза’: лувийск. *tawa* (с закономерным исчезновением интервокального лабиовелярного): латинск. *oculus*, русск. *око*, древнеанглийск. *ēage* > *eye*, немецк. *Auge*;

(II) хеттск. *šankuwi-š* ‘ноготь’: лувийск. *tamuga*⁴: латинск. *ungui-s*, русск. *ноготь*, древнеанглийск. *nag(e)l* > *nail*, немецк. *Nagel*;

(III) хеттск. *šak(a)n*, *šak(a)nuwant-* ‘масло, жир’ — лувийск. *tain* ‘масло’ ([HOFFNER 1994; 1995]; [КЛОЕКНОРСТ 2008: 702–703]; [CHD, Vol. Š: 35–36]): возможно, хотя и проблематично, сравнение с латинск. *unguen*⁵;

(IV) хеттск. *šehuṛ/n-* ‘моча’: лувийск. *dūr/n-* ‘моча’ (существительные архаического гетероклитического типа: [КЛОЕКНОРСТ 2008: 741–743; 899]): латинск. *ūrīna* ‘моча’ (сравнение с одним из распространенных в разных макросемьях названием ‘воды’ было давно предложено, но не общепринято).

Сравним разные объяснения этих и других подобных им соответствий, предложенные в научной литературе.

5.1. В первом изложении сравнительной фонетики анатолийских языков К. Мелчерта, который исходил из идеи наличия *общеанатолийского* праязыка, развившегося из общеиндоевропейского и затем распавшегося на хеттский, лувийский и другие анатолийские диалекты, было предположено вслед за замечательным словенским лингвистом Б. Чопом ([ČOP 1965: 100–103]) «*нерегулярное*» развитие проанатолийского **š-* в лувийский зубной смычный ([MELCHERT 1994: 274–275]). Слабость этой формулировки определяется тем, что в других случаях общеанатолийскому и хеттскому **š-* отвечает неизменившееся лувийское *š-*, а остальные зубные смычные в лувийских словах индоевропейского происхождения восходят к соответствующим древним зубным/ переднеязычным **t-*, **d-*, **dh-*.

⁴ Появление *-m-* вместо *-n-* отмечается кроме лувийского также в тохарском соответствии. Лувийск. *-ug-* отражает другую серию смычных по сравнению с (I).

⁵ Эта этимология предполагает инфиксальный характер **-n-* в латин. *unguo*, др.-инд. *a-na-k-* / *añj-*; в хеттск. *šak(a)nu-* аналогичный носовой аффикс выступает после корня, а не перед его последней согласной фонемой: но характер этой последней вызывает сомнения, т. к. в хеттском отражается велярный, возможно из-за диссимилиации индоевропейского лабиовелярного с последующим **-u-* глагольной основы, ср. о формах **sag-* (**saguwant-*: латинск. *unguent-*?) и *sagn-* в хеттском ([HOFFNER 1994; 1995]).

5.2. Мысль одного из первых создателей сравнительной грамматики хеттского языка Э. СТЕРТЕВАНТА ([STURTEVANT 1927: 163]) пересказана автором только что вышедшего этимологического словаря хеттского языка А. КЛУКХОРСТОМ ([КЛОЕКНОРСТ 2008: 705]) следующим образом. СТЕРТЕВАНТ предлагал объяснить подобные случаи комбинацией двух ранее гипотетически предложенных единиц праязыка — начального **s-mobile* (такого начального **s-*, которое в одних индоевропейских языках наличествует, а в других отсутствует без определенных причин) и следующего за этой гипотетической фонемой ларингального **H₃*. В теперешней записи ([КЛОЕКНОРСТ 2008: там же]) в реконструированной для (I) праформе **s-H₃ek^w* сохраняется действие постулата, по которому индоевропейские корни, начинающиеся в исторических языках с гласного, утратили предшествующий ларингальный элемент. Обсуждая эту гипотезу как одну (по его мнению не самую убедительную) возможность, КЛУКХОРСТ замечает, что для (II) в таком случае следовало бы предположить исходную форму **s-H₃ng^h-u-* ([КЛОЕКНОРСТ 2008: 725], где форма объясняется исходя из **s-mobile*). Но в обеих основах судьба первой восстанавливаемой фонемы остается без достаточных объяснений.

5.3. В тех же работах СТЕРТЕВАНТА ([STURTEVANT 1927: 163]) и КЛУКХОРСТА ([КЛОЕКНОРСТ 2008: 705]) обсуждается и возможная этимология первого слова (I), основанная на его сравнении с (Ia) герм. **sef^wan* ‘видеть’ (готск. *sailwan*, др.-исл. *sjā*, др.-фриз. *siā*, др.-сакс. и др.-в.-немецк. *sehan* > *sehen*, др.-английск. *sēon* > *see*) < **sek^wan* ‘следить глазами, следовать’ (латин. *sequor*).

Предполагается сохранение в анатолийском и германском исходного значения глагола среднего залога ([BALDI 1974; 2002: 73, 278, 365, 396]), тогда как остальные индоевропейские диалекты осуществили семантическую инновацию (обратное предположение об архаизме значения

(Ib) в латинском, а также в др.-ирландск. *sechithir*, греч. *ἔπομαι*, др.-инд. *sáce*, не пользуется популярностью, хотя и имеет сторонников в новых работах по дальнему родству).

Сложность проблемы выясняется из наличия в хеттском не только приведенного существительного, но и однокоренного глагола *sakuwai-* ‘видеть’, возможно, родственного как германским формам с тем же значением, так и

(Iв) албанск. *shoh* < **sāk-ska-* < **sēk^w-[e]sk^ʷ-* ([THUMB 1905–1906: 266, fn. 5]; [HULD 1984: 114–115]; [OREL 2000: 100, 148, 184]; особенно следует отметить тождество праалбанской реконструированной формы древне- или среднехеттскому итеративу *sakuwiske-* ‘видеть’ ([ИВАНОВ 1981], примеры употребления: [CHD, Vol. Š: 551]).

Начиная со статьи СТЕРТЕВАНТА и позднее обсуждался вероятный с точки зрения ученых того времени выбор двух возможных объяснений хеттского названия «глаза» — либо из основы этого глагола (которая или сохранила древнее значение, наличествующее в латинском и в других диалектах кроме германских и албанского, или развила новое значение типа представленного в латинском и других диалектах, кроме анатолийских, германских и албанского), либо из основы индоевропейского названия «глаза», которое в этот период реконструировалось с начальным ларингальным. Трудность его сравнения с хеттск. *šakuwa* ‘глаза’ (Pluralia tantum, скорее всего из древней формы двойственного числа, в хеттском исчезнувшего: [ИВАНОВ 2001: 129]; [HOFFNER & MELCHERT 2008, pt. 1: 68. fn. 15]) относится к написанию и фонетике: несдвоенное написание интервокального *-ku-*, казалось бы, противоречит закону СТЕРТЕВАНТА, хотя и может объясняться, например, ленированием подобных фонем уже в ранний период (ср. также отражение индоевропейского звонкого в родственной древнепрусской форме и наличие долгой ступени гласного корня в древнехеттском и праалбанском; ленирование после долгого гласного в анатолийском могло быть фонетически закономерным). Вместе с тем со словообразовательной и семантической стороны соответствие хеттск. *šakuwa* ‘глаза’ индоевропейским названиям «глаза» поддерживается формой

(Iг) хеттского *šakuni-* (в части форм, связанных с этим словом, закон СТЕРТЕВАНТА соблюдается: [КЛОЕКНОРСТ 2008: 702–703]; [CHD, Vol. Š: 58, 77]), которое можно отождествить с названиями «родника, болота», образованными от той же основы в балто-славянском (русск. диалектн. *око* ‘окно, отверстие, наполненное водой, в болоте’) и армянском.

5.4. Поскольку ни одно из приведенных объяснений СТЕРТЕВАНТА не может быть принято безоговорочно, выдвигается другое предположение, основанное на том, что и в (I) и (II), и в ряде других позднее выявленных аналогичных соответствий (III)–(IV) в основе можно предположить исчезновение ленированного велярного согласного ([КЛОЕКНОРСТ 2008: 705–706]). Как и почему это обстоятельство, подробности которого совсем непонятны в (IV), могло влиять на характер первого согласного, остается неясным. КЛУКХОРСТ предлагает не реконструкцию праформ, а описание некоторых их особенностей, быть может, полезное для исторической фонетики, но не объясняющее происхождения праанатолийских форм.

5.5. Несколько предложенных объяснений этих соответствий между хеттским и лувийским исходит из допущения палатализации. Согласно Я. Пухвелю ([PUHVEL 1974; 1975; 1979]), в (i) хеттск. *šakuwi-* < и.-е. **dh₁^o/(a?)gh^w-* > др.-греч. *σοφός* ‘мудрый’ по аналогии с палатализацией **dy/i-* > *ši-*, происходящей в хеттском (*šiw-att* ‘день’: латинск. *diēs*; в самой ранней надписи царя Аниттаса др.-хетт. *šiu-šummi* ‘наш бог’ или имя ‘Нашего-Бога’: [HOFFNER & MELCHERT 2008, Pt. 1: 138, fn. 9]; латинск. *deus*). Однако ни в одном из достоверных примеров палатализирующий элемент не сохранился, и этимология древнегреческого слова остается спорной. Поэтому гипотеза не может быть проверена.

5.6. Вариант идеи палатализации был предложен Ф. Йосефсоном. В общем исследовании разных случаев предполагаемого действия этого процесса в хеттском он предположил в подобных словах наличие палатального ларингального **h^y* ([JOSEPHSON 1979: 100–101]). Но именно палатальный его характер не подтверждается соответствиями.

5.7. Эти идеи палатализации получили развитие в посмертной публикации Джиллиан Харт ([HART 2004]). Она думала (как бы предугадывая гипотезу Клухворста, ее работу не упоминающего) о неслучайности того, что во всех хеттских словах рассматриваемого типа можно восстановить ларингальный *h₃* (предположительно лабиоларингальный), после которого дальше обнаруживается лабиальный *w* или *u*. Соответственно Харт допускала следующую последовательность изменений:

А. В индоевропейском реконструируются ларингальные **h₃e* и **h₃a* > **h₃o*.

В. В пранатолийских формах осуществляется ассимиляция **g^w... y^w* > *y^w... y^w*.

С. Перед концом пранатолийского периода происходит диссимилиация второго **y^w* > **y^u* в окружении лабиальных *w* или *u*.

Д. Лабиальный элемент **h₃* теряется перед **o*.

Большинство этих процессов определяются скорее интеллектуальной модой (реконструкция фонетических особенностей ларингальных) и не подтверждаются бесспорными свидетельствами исторических языков⁶.

5.8. Если исходить из предположения о происхождении хеттского и лувийского из двух самостоятельных индоевропейских диалектов, то нет необходимости в восстановлении общей праанатолийской формы для разбираемых слов, сведение которых на этом уровне к одному источнику оказывается настолько трудным. На формальном языке соответствий можно думать о новой индоевропейской фонеме, реконструируемой на основании отождествления начального хеттск. *š-* = лувийск. *t-/d-* и отсутствия согласного в других диалектах ([IVANOV 2001]). Однако в том, что касается последнего соответствия, сложности могли бы возникнуть с (I, Ia, Ib, Ir). Представляется возможным следующее решение. Примем за возможную фонетическую интерпретацию реконструируемой начальной индоевропейской фонемы ее функционирование как фрикативного спираанта типа **θ* (английск. *th-* в *thin*). В глагольной основе, которая может быть восстановлена в индоевропейской форме **θok^w-e/o-* ‘видеть’ (откуда глагольные словоформы Ia и Ib, а также такие названия «глаза, взгляда» вероятного глагольного происхождения, как др.-ирл. *ro-sc*), этот начальный согласный представлен как **s-*: **θ-* > **s-*; сохранение в германском глухости интервокального согласного позволяет по закону Вернера реконструировать ударение на слоге глагольного аффикса; **θok^w-é/ó-*. В родственных именных формах всех других индоевропейских диалектов кроме хеттского и лувийского эта фонема исчезает: **θ-* > **θ*; озвончение в германском глухого индоевропейского лабиовелярного интервокального согласного позволяет по закону Вернера реконструировать ударение на корневом слоге **θók^w-e/o*. Ударение на начальном корневом слоге в этом индоевропейском названии глаза подтверждается древнехеттским написанием именной формы с possessивным местоименным суффиксом *ša-a-ku-wa-aš-me-et* ([HOFFNER & MELCHERT 2008, pt. 1: 140]). Соответственно в английск. *eye* и *see*, немецк. *Auge* и *sehen* представлены образованные от одного и того же корня именные и глагольные основы, различавшиеся в индоевропейском месте ударения и соответственно в германском характером согласных. Поскольку различие в судьбе начального согласного **θ* обнаруживается в таких словоформах, которые при общности корня различаются категорией (частью речи) и акцентуацией, можно думать, что и эта особенность звуковой формы глагольной или именной формы определялась теми же факторами или другими, с ними же связанными.

⁶ К числу слов, разбираемых Харт, относится также *suwai-*, для которого предполагалось значение ‘птица’, возводившееся к праиндоевропейскому ([SCHINDLER 1969]). Но в последнее время для этого слова предположено совсем другое значение — ‘отказ’ ([KLOECKHORST 2008: 795]). Возможно, что для окончательного решения проблемы, кроме еще нескольких слов с вероятным начальным **s-* *mobile*, как хеттск. *samluwa-* ‘яблоко’: русск. *яблоко*, англ. *apple* ([ГАМКРЕЛИДЗЕ, ИВАНОВ 1984, т. 2]), потребуется еще и привлечение подобных слов с такими же соответствиями в середине слова (хеттск. *hwīš-* ‘жить’: лувийск. *huit-*, [HART 2004]; [KLOECKHORST 2008: 355]).

Восстановленное таким образом соотношение глагола ‘видеть’ и существительного ‘глаз’ типологически сходно с родством основ ‘слышать’ (хеттск. *ištam-aš-* от основы, находящей соответствия в афроазиатском — древнеегипетском и семитском — и картвельском: [ГАМКРЕЛИДЗЕ, ИВАНОВ 1984, т. 2]) и ‘ухо’ (хеттск. *ištam-ana-*), ср. тохарские производные с теми же значениями и суффиксами, образованные от и.-е. **kleu-/w-* ‘слышать’, и ряд других аналогичных образований.

6.1. После того, как на основании соответствий между родственными языками восстановлена абстрактная система праязыка и дана ее интерпретация, учитывающая типологические и другие косвенные данные фонетического, грамматического и лексического характера, наступает следующий этап сравнения. На этом новом этапе, который представляет особый интерес для нашего журнала, *основные методы установления и истолкования соответствий остаются такими же*, как и на предшествующем этапе. Но меняется объект исследования. Лингвист *сравнивает не языки, а восстановленные праязыковые системы*. Разумеется, элементы такого сравнения есть и на предшествующем этапе, когда нередко оказывается нужным проинтерпретировать в более ранние периоды истории языка.

Но существенным шагом в развитии всех сравнительно-исторических лингвистических исследований явилось наблюдение Х. Педерсена, развитое затем в серии работ В. М. Иллич-Свитыча. Они предположили, что предметом более глубокого (по времени) сравнения могут быть не сами языки, засвидетельствованные историческими памятниками или современными описаниями, а результаты праязыковой реконструкции, осуществленной благодаря их сравнению.

6.2. В этом последнем случае очень важным вопросом является *число и характер сравниваемых праязыков*. В наших беседах с ним на эту тему В. М. Иллич-Свитыч не раз подтверждал, что, готовясь к своему основному словарю, он не был заранее уверен в том, что число праязыков, входивших в ностратическую макросемью, действительно ограничено только теми, которые обозначены в заглавии его книги. Он выбрал для сравнения праязыки этих языковых семей (индоевропейской, картвельской, афроазиатской, дравидийской, уральской и алтайской), считая, что для них уже разработаны сравнительно-исторические грамматики и словари. Начало попарным сравнениям этих языков было положено раньше — начиная с книги Ф. Боппа об индоевропейском и картвельском. Ностратическая гипотеза включила результаты таких более ограниченных сопоставлений, продолжавшихся и после обнародования всей широкой картины этой макросемьи в целом. Так, изолированные уральско-индоевропейские сопоставления, успешно начатые Б. Коллиндером, нашли достойное продолжение в большой серии исследований Б. Чопа.

На современном этапе исследования макросемей степень изученности членов каждой из наиболее изученных среди них может диктовать большую или меньшую определенность в предварительных выводах относительно состава макросемьи и конкретного соотношения ее системы с каждым из отдельных праязыков. Это касается, например, намеченного еще в первых трудах А. Б. Долгопольского и потом Дж. Гринберга вопроса о родственных связях нескольких евразийских языков Северо-Востока Азии с ностратическими. Хотя многие сходства эскимосско-алеутского с другими ностратическими бросились в глаза ряду исследователей еще в первой половине XX в., окончательные выводы могли быть сделаны только после детальных сопоставлений, проведенных О. А. Мудраком. То, что бурушаски и баскский скорее всего родственны той большой совокупности языков, из которых в историческом плане лучше изучены северо-кавказские, было давно предположено многими учеными (иногда, как Н. Марр, страдавшими отсутствием сколько-нибудь надежных методов сравнения). Но из области сомнительных догадок в сферу научно проверяемых гипотез подобные предположения перешли совсем недавно по мере разработки детальной исторической фонетики этих языков.

6.3. У исследований отдаленного родства могут быть разные цели. С одной стороны, они представляют огромный интерес для выяснения ранних этапов распространения и развития макросемей. С другой стороны, следует подчеркнуть их *роль в выполнении традиционных задач исторического языкознания*. Установление отдаленного родства для семьи языков означает возможность продлить изучение истории ее праязыка, восстановив самый древний период, ранее бывший доступным лишь для *внутренней реконструкции*, базировавшейся на соотношениях внутри абстрактной системы. По существу наиболее ранним примером этого подхода в истории индоевропейского языкознания явились исследования Г. Мёллера и А. Кюни, открывших афроазиатские соответствия тем «сонантическим коэффициентам», которые предположил де Соссюр в упомянутой выше пионерской работе (ср. очень объективно изложенную историю вопроса в книге [STURTEVANT 1942]).

6.4. Поскольку по данным лексикостатистики, согласующимся с анализом морфологии глагола, несомненно раннее отделение хеттского и лувийского от других диалектов индоевропейского праязыка, возникает вопрос о тех архаизмах, которые еще могли сохраниться в этих анатолийских языках, а в других индоевропейских уже не прослеживаются. С этой точки зрения большой интерес представляют самые ранние свидетельства об анатолийских языках, сохраненные в староассирийских письмах из торговых колоний в Малой Азии начала 2-го тыс. до н. э. Ранний тип сохраненных в этих текстах древнеанатолийских женских собственных имен представляет собой словосложение с последней (второй) основой *-ne/i^k/ga-*; в соответствии с древними принципами индоевропейского словосложения перед этой основой выступало соединительное **-o- > -a-*. За первой именной основой следовала группа морфов *-a-ni^g/ka*. Основа существительного, в древнеанатолийском выступающая как *nika*, в древнехеттском — *nega-* «сестра, кровная родственница *ego* в его поколении», принадлежит к архаичным общим элементам анатолийского (древнехеттского и лувийского) словаря. Передававшийся в хеттской клинописи несдвоенным написанием интервокального смычного *-g/k-* индоевропейский звонкий придыхательный заднеязычный в лувийском исчезал: хеттск. *anna-nega-* «двоюродная сестра» соответствует лувийск. *anna-niya-[mi]* (первая половина сложения представляет собой анатолийское имя «матери» — в несвязанной форме лувийск. *ann-i*, отвечающее ономатопоэтическим формам «детских» имен родства в других языках), хеттск. *neg-na* «брат» — лувийск. *nana-šr-i* «сестра»: к соответствующему лувийскому названию сиблинга — брата и сестры вместе взятых — присоединен последний элемент, обозначающий в индоевропейском женщину. Следовательно, исходная основа, близкая к хеттской, восстанавливается для праанатолийского (и, если считать два анатолийских языка, где она представлена, разными индоевропейскими диалектами, то и для праязыка, из которого они происходят). Точное соответствие обнаруживается в алтайском **nekV* > тунгусо-маньчжурск. **neku* «младший родственник (сестра, брат)» > негидальск. *нэхун* «двоюродный младший брат, двоюродная младшая сестра, племянник, племянница, брат или сестра мужа или жены, младше говорящего; младший, младшая», дравид. *nāg-* «младшая женщина» < ностратич. **nVkV*⁷. Слова с ностратическими этимологиями, отсутствующие в других индоевропейских языках, могут считаться особыми архаизмами анатолийского ([Иванов 2008]).

6.5. При проведении исследований, касающихся отдаленного родства между семьями, едва ли не определяющая роль во многих случаях принадлежит хронологии языковых явлений. Возвращаясь к подробно изученному за последние полвека вопросу об отношениях между картвельскими языками и индоевропейскими, можно заметить, что одна из главных трудностей состоит в отличении бесспорного общеностратического наследия (включающего такие элементы базисного словаря, как, например, название «сердца», и.е. *ker-dh-*) от позднейших культурных заимствований, которые были достаточно древними для того, чтобы успеть войти в общекартвельский словарь. Существенной оказывается возможность привлечения собственно лингвистических-фонетических критериев: в картвельских словах, унаследованных от праностратического, еще сохраняются исходные ларингальные, в поздних культурных заимствованиях из индоевропейского их уже нет.

В уралистике до сих пор не решен спор между учеными, которые, как А. Йоки в его основополагающей монографии и Й. Койвулехто ([KoiVULENTO 1999]), рассматривают основное множество слов, общих для индоевропейского и финно-угорского, как результат заимствования, и представителями точки зрения, считающей их финно-угорским наследием. Едва ли не основную роль в выборе между заимствованием и дальним родством могут играть морфологические сопоставления, в частности, позволяющие восстановить единую парадигму праформ уральского и индоевропейского глагола ([Хелимский 1982]; [KORTLANDT 2001]).

Хотя некоторые (но далеко не все) уже достоверно выявленные макросемьи отличаются сплошным характером распространения на больших пространствах, это нельзя пока признать общей закономерностью (правда, сходный принцип родства, сопряженного с географической близостью, выявляется и в исследованиях по молекулярно-биологическим связям этнических групп населения). Последующие миграции могли создать систему довольно значительной «чересполосицы» в запутанной мозаике праязыков, географически соседствующих, но принадлежавших к разным макросемьям.

Поскольку реальный состав устанавливаемых макросемей (и тем более их пока проблематичных объединений на еще более удаленных хронологических уровнях) пока еще не мо-

⁷ [STAROSTIN-N]; [STAROSTIN, ДУВО, МУДРАК 2003, Vol. 2: 968]; [Цинциус 1982: 248]. Предполагаемая Старостинным праформа фонетически несводима с уральским сравнением, предложенным ранее Чопом, венгерск. *нб* и другими финно-угорскими и самодийскими названиями «женщины».

жет быть известен полностью и часто еще будет уточняться, принимаемые названия (как и введенные много раньше названия семей языков по отношению к более близким хронологическим уровням) нередко остаются пока условными (это относится к терминам типа «евразийский», «бореальный» и т. д.). Но непрерывное изменение названий может внести путаницу. Возможно, что следует остановиться на стандартизированной схеме пока принимаемых названий, а каждое предлагаемое изменение подвергать достаточно широкому обсуждению. Возможно и сохранение вариантов типа Indogermanisch = Indo-European.

7. Исследование заимствованной лексики, в том числе и миграционных терминов культурного характера, при всей исторической значимости получаемых выводов представляет непосредственный интерес для проблематики ближнего и дальнего родства главным образом в связи с необходимостью отделения этой категории общих слов от унаследованных (особняком стоят отмеченные выше проблемы, связанные с креольскими языками и пиджинами, где разграничение двух этих слоев лексики иногда крайне затруднено). Но в целом ряде случаев оказывается возможным использовать *исследование древних культурных терминов для выяснения отношений между членами одной макросемьи или нескольких макросемей, их заимствовавших.*

Рассмотрим в качестве примера один из многочисленных миграционных терминов, относившихся к струнным музыкальным инструментам. В большинстве языков древнего Восточного Средиземноморья и Ближнего Востока именно благодаря значимости этой сферы культуры большинство соответствующих слов кочевало по всему тогдашнему цивилизованному миру, напоминая ситуацию современной международной научной и технической терминологии.

В хаттской основе *-zinar* ‘арфа, лира, струнный музыкальный инструмент’ наблюдается результат палатализации в положении перед *i* начального заднеязычного согласного в исходном слове семитского происхождения, которое (с другим гласным во втором слоге) отражено в основе хеттского имени деятеля ^{LÚ}*ki-nir-tal-la-aš* ‘музыкант-исполнитель, играющий на струнных инструментах’ (см.: [PUNVEL 1997: 182]; [IVANOV 1999]; [SCHUOL 2004: 163–172]), образованного посредством хеттского суффикса деятеля (по происхождению индоевропейского, этимологически соответствующего слав. **-telb-* > русск. *-мель*) от существительного **kinir-*. В хеттских текстах наряду с этим именем деятеля засвидетельствовано и синонимичное с ним слово, образованное посредством того же хеттского суффикса от хаттской основы ^{LÚ}*ḥa-zinir-talla-* ‘играющий на лире-арфе’ ([ПОРКО 1994: 81, 166]; [SOYSAL 2004: 155, 942, 1070]). Эта хаттская основа включает имеющий северо-кавказское происхождение локативный префикс направления *ḥa-*, за которым следует хаттское (из семитского) название ‘лиры’ *zinir* в форме, по огласовке второго слога (*-i-*, а не *-a-*, как в *zinar*) совпадающей с хеттской. Иначе говоря, в этих производных словах хеттская основа отличается от хаттской только начальным согласным: хеттское *k-* вместо хаттского *z-*. Для двуязычных хеттов, знавших (хотя бы в пределах сакрального обрядового словаря) уже (почти что) мертвый священный хаттский (по происхождению северо-кавказский) язык, из равенства хеттских имен деятеля

$${}^{\text{LÚ}}\text{-kinir-talla-} = {}^{\text{LÚ}}\text{ḥa-zinir-talla}$$

должно было следовать синхронное равенство названий инструментов

$$\text{kinir-} = \text{-zinir-},$$

которое находит диахроническое объяснение в происхождении хаттского *zinar-* / *zinir-* из семитского *kinar-*, отраженного в хеттском слове. Заимствование, объясняемое прежде всего музыкальными культурными контактами, можно соотнести с другими семитскими словами, предполагаемыми в хаттском, таких, как название ‘быка’ *milup* и ‘сына’ *binu* (как и рассматриваемые названия музыкальных инструментов, последнее слово имеет соответствие и в адыгском). В данном случае археологические данные о распространении лиры согласуются с гипотезой о западно-семитском происхождении ее названия.

Наиболее раннее письменное свидетельство этого (западно-)семитского слова представлено в Эбле в лексическом списке, датируемом примерно 2340—2300 г. до н. э.:

GI(= ki)-na-ru₂-um/rûm/lum = (шумерское) **BALAG** ([ПЕТТИНАТО 1982: 264, № 572], о позднейшей литературе об этом слове см. [FRANKLIN 2006]).

Археологические музыковедческие открытия последних лет показывают, что лиры разных размеров распространялись из западно-семитской области по всему Древнему Востоку (см.: [FRANKLIN 2007]). Эти музыковедческие выводы совпадают с лингвистическими. Слово, соответствующее др.-евр. *kinnor*, в форме *kinnarum*, встречается около 1700 г. до н. э. в Мари в тексте, где речь идет о мастерах, изготавливающих лиры для царя. Слово заимствуется в хурритский. Оно встречается в хурритских ритуалах из Богазкёя и дает основу для образования

хурритского производного, по смыслу сходного с приведенным хеттским именем деятеля (но использующего собственно хурритские словообразовательные средства): *LÚki-in-na-ru-ḫu-li*. После Амарнского времени, когда при Аменхотепе IV и после него громадные лиры древне-ближневосточного типа входят в употребление, в древнеегипетском засвидетельствовано заимствованное *knrr* ‘лира’. В эфиопском представлена форма *kerar* < **kenar*.

Наряду с подобными формами, сохраняющими исходную согласную фонему (западно-) семитской основы, в языках Древней Передней Азии и сопредельных стран представлены и формы, в которых, как в хаттском, в позиции перед *i* эта фонема палатализуется и превращается в аффрикату или фрикативную. Гласная, вызвавшая палатализацию, могла потом измениться, как в аккадском *zannaru* ‘лира богини Иштар или богини Инанна’, или сохраниться, как в хаттском *zinar*. Подобный *конвергентный* процесс палатализации мог происходить одновременно в разных языках, как можно предположить по отношению к коптскому *sinra* = [*sinera*]. Применительно к др.-арм. *ձւար* [*ʒnar*] ‘пальцевый струнный музыкальный инструмент, лира’ (см.: [АЧАРЯН 1979: 129]); высказывалось предположение, что термин непосредственно восходит к хаттскому ([МКРТСЧИАН 1974: 316]). Но кроме исчезновения гласного первого слога, вызвавшего палатализацию, есть и другие различия между древнеармянской и хаттской формами. Судя по правилам хеттского клинописного письма, начальный согласный хаттского слова был глухой аффрикатором *z* [*ts*]. Древнеармянский начальный согласный *ձ* [*dʒ*] представляет собой звонкую аффрикату. В словах индоевропейского происхождения он обычно продолжает звонкое **g^{wh}*. Для объяснения различия можно предположить, что после утраты палатализирующего гласного первого слога палатализованный согласный (ставший аффрикатором) оказался перед *-n-*, что могло привести к его регрессивной ассимиляции — озвончению. Для типологического сравнения с хеттским представляет интерес наличие в древнеармянском (как и в египетском и коптском) двух разных форм названия струнного инструмента — *ձւար* [*ʒnar*], фонетически близкого к хаттскому *zinar*, и *քւար* [*k^hnar*] (шипковый струнный инструмент), отражающего первоначальную семитскую форму.

Иначе обстоит дело с адыгскими музыкальными терминами, которые очень близки к хаттским. Соответствующее хаттск. *zinar* ст.-адыгейск. (отраженное в источниках первой половины XIX в.) *пшиинэ* < **р^чḫə(i)na* ‘струнный музыкальный инструмент’ засвидетельствовано в качестве первой части в двух сложных словах, которые с точностью до порядка слов совпадают с хаттскими.

1) Ст.-адыгейск. *пшиинэ ишхуо* ‘арфа’ < **р^чḫə(i)na-šx^{wa}* ‘струнный музыкальный инструмент + большой’ по С. Л. Николаеву и С. А. Старостину, в западно-кавказском адыгейск. и кабард. *-šx^{wa}*, убыхск. *šx^{wa}* < зап.-кавказск. **č^həḫ^{wa}* < пра-обще-сев.-кавказск. **č^həq^wV* ‘большой’ ([НИКОЛАЕВ, STAROSTIN 1994: 386]; дальнейшие старостинские сопоставления предполагают родство с сино-тибетским **čök* (~ *č-*), енисейским **su(?)K-* (?), бурушаски **śóq-*, баскским **aško* ([STAROSTIN-SC]); сохранение фрикативного из архаичной начальной аффрикаты в данном случае отличает адыгский от хаттского, в котором от начального комплекса осталась только вторая фонема). Поскольку для этого прилагательного в адыгском стала обязательной конечная позиция, изменился порядок элементов по сравнению с хаттским *ḫun-zinar* ‘большой струнный музыкальный инструмент; арфа’, где соответствующее хаттское прилагательное, восходящее к тому же северокавказскому, занимает первое место. Праадыгск. **čx^{wa}-* + **р^чḫə(i)na* ‘большой струнный инструмент, арфа’ содержит в качестве второго элемента слово, соответствующее хат. *zinar*. Можно предположить, что словосложение **čx^{wa}-* + **р^чḫə(i)na* ‘большая лира, арфа’ существовало в диалектах северо-кавказского еще до отделения хаттского от праадыгского — не позднее начала III тыс. до н. э., когда на Ближнем Востоке есть и первые археологические образцы больших лир.

2) Ст.-адыгейск. *энэ пшиинэ* ‘лира, арфа’, кабардинск. *энэ пшиинэ* ‘пальцевый инструмент, аккордеон, балалайка’ < **ʔapa* + **р^чḫə(i)na*; по С. А. Старостину, в западно-кавказском адыгейском и кабардинском первом слове могла произойти диссимилятивная дезабруптивизация; исходное **НарV* восстанавливается по соответствию цезского (**ʔapa* ‘лапа’) с абхазо-абазинским **р^а*, **на-р^а* ‘рука’ ([НИКОЛАЕВ, STAROSTIN 1994: 545]), к которому присоединяется хаттское *ipri-zinar* ‘малый (ручной, пальцевый) струнный музыкальный инструмент; лира’. Как и в первом словосложении, во втором последний элемент исторически тождествен хаттскому.

Следующие фонетические изменения могут быть предположены в адыгском в отличие от более архаической хаттской формы, ближе стоящей к семитскому источнику. В начале слова в праадыгском губной смычный **p-* (в соответствии с общими правилами исторической фонетики этих языков) возник перед аффрикатором (ставшей позднее фрикативным согласно общefonетической закономерности), ср. аналогичное фонетическое соответствие между хеттским *tešha-* /

zašha ‘сон’ (из хаттского: [ИВАНОВ 2001: 232 и 283]) и праабхазо-кабардинским **pšā-há, *pšā-hāpa*, адыгск. *pšəhəp*, кабардинск. *pšəhəpa, pšəh* ‘сон’ ([NIKOLAEV, STAROSTIN 1994: 512–513]). Аффриката в хаттском (в отличие от древнеармянского) была глухой, но в соответствии с общекавказской реконструкцией С. Л. Николаева и С. А. Старостина она была придыхательной. Фрикативный *š*, в который в адыгском превращается аффриката, фонетически близок ко второй половине аффрикаты *z-* [ts] в хаттском. Гласный первого слога сохранился, как в хаттском (в отличие от древнеармянского). Конечный *-r* был потерян в праадыгском, а в хаттском сохранился.

Иначе говоря, в хаттском и праадыгском представлены одни и те же сложения заимствованного семитского слова со словами северо-кавказского (туземного) происхождения. Во всяком случае этот относительно поздний период истории слов после заимствования из семитского в праадыгском переживался одновременно с хаттским. Вероятно поэтому, что и промежуточный источник заимствования, где осуществилась палатализация, для них был общим. Палатализация, общая для адыгского и хаттского, произошла либо (скорее всего) в их общих источниках, представлявших собой пока точно не восстановленные близкие друг к другу северо-кавказские диалекты III-го тыс. до н.э., заимствовавшие слово из семитского, либо в другом языке (например, лувийского типа), откуда это слово могло быть заимствовано и в праадыгский, и в хаттский.

Хаттский язык сохранялся в качестве священного языка религиозных и придворных рядов в столице Хаттуссе и некоторых недавно раскопанных провинциальных центрах Хеттской Империи, но к середине II-го тыс. до н. э., скорее всего, уже был мертвым. Его общность с другими северо-западно-кавказскими языками относится ко времени не позднее III тыс. до н. э. К этому времени можно отнести и распространение этих музыкальных терминов.

8. Перечисленными вопросами (которые иллюстрированы достаточно случайными примерами — их легко было бы заменить или умножить) совсем не ограничено число тех актуальных проблем исторического языкознания и лингвистической таксономии, которые должны быть в центре внимания нашего журнала. Мы надеемся, что лингвисты разных стран и направлений примут участие в их деловом обсуждении на основе данных все увеличивающегося числа изучаемых языков.

Могут представить интерес начавшиеся во многих лабораториях молекулярной биологии исследования генетических признаков связей между группами населения в сопоставлении с выводами лингвистов. Сходство общих принципов передачи информации и искажения кодов и сообщений при их передаче в этих двух бурно развивающихся областях современного знания несомненно. Насколько удастся наладить деловое сотрудничество исследователей в этих и смежных с ними областях (избегая попыток искусственного сближения результатов) — покажет будущее.

Хотелось бы особенно пожелать успеха тем, кто занят сравнением исчезающих языков, скорость умирания которых в некоторых областях земного шара все увеличивается. Для исторического языкознания важно в случае этих языков как можно скорее выявить хотя бы список самых неотложных вопросов, на которые нужно еще успеть получить ответ до тех пор, пока жив хотя бы один активно на нем говорящий (вместе с тем нельзя не предупредить специалистов относительно растущей опасности появления лже-информантов, нередко появляющихся именно тогда, когда уходит на тот свет последний из говоривших на данном языке: это одно из проявлений того рыночного использования исчезновения языков, которое распространяется в самые последние годы).

Необходимость сохранения памяти о недавно еще существовавшем языковом разнообразии человечества, которое стремительно уменьшается с каждым поколением (если не с каждым двумя неделями — такова сейчас средняя скорость ухода языков!), представляется одной из главных задач лингвистики. Наш журнал был бы рад внести посильный вклад в осуществление этой задачи.

Л и т е р а т у р а

- АРАПОВ, ХЕРЦ 1974 — М. В. АРАПОВ, М. М. ХЕРЦ. *Математические методы в исторической лингвистике*. М.
- АЧАРЯН 1979 — Հ. Հ. ԱՇԽՆԱՆ [Р. Я. Ачарян]. *Հայերեն արմատական րիտիրան* [Этимологический корневой словарь армянского языка]. Հ. Դ. [Т. IV]: Պ–Ֆ. Երևան [Ереван].
- ГАМКРЕЛИДЗЕ, ИВАНОВ 1984 — Тамаз В. ГАМКРЕЛИДЗЕ, Вячеслав Вс. ИВАНОВ. [Մ. ԶԱՄԵՐԱԼԻԹԻՔ, Յ. ՈՅՆԵՐՅՈ ~ Thomas V. GAMKRELIDZE & VJACHESLAV VS. IVANOV]. *Индоевропейский язык и индоевропейцы. Реконструкция и историко-типологический анализ праязыка и протокультуры*. Тбилиси: Издательство Тбилисского Университета [იბდოევროპული ენა და იბდოევროპელები. წინარეებისა და წინარეკულტურის რეკონსტრუქცია და ისტორი-

- ულტიპოლოგიური ანალიზი. თბილისი: თბილისის უნივერსიტეტის გამომცემლობა ~ *Indo-European and the Indo-Europeans. A Reconstruction and Historical Typological Analysis of a Protolanguage and a Proto-Culture*. Tbilisi: Publishing House of the Tbilisi State University].
- ЕЛЬМСЛЕВ 1952 — Луи ЕЛЬМСЛЕВ. Метод структурного анализа в лингвистике // *Acta Linguistica*, Vol. VI, fasc. 2–3 [статья первоначально автором была написана и напечатана по-русски, переиздание в СССР: *История языкознания XIX–XX веков* / Сост. В. А. ЗВЕГИНЦЕВ. М., 1965; стр. 102–109 [было несколько русских переизданий]].
- ИВАНОВ 2001 — Вяч. Вс. ИВАНОВ. *Хеттский язык* (2-е изд.). М.: «УРСС».
- ИВАНОВ 2005 — Вяч. Вс. ИВАНОВ. Типология языков бассейна Амазонки. II. *Числительные и счет* // *Вопросы языкознания*, № 5.
- ИВАНОВ 2007 — Вяч. Вс. ИВАНОВ. К антропологии числа // *АБ • 60. Сборник статей к 60-летию Альберта Кауффуловича БАЙБУРИНА* (= *Studia Ethnologica*, Вып. 4), СПб: Европейский Университет в Санкт-Петербурге; стр. 32–51.
- ИВАНОВ 2008 — Вяч. Вс. ИВАНОВ. Анатолийские личные имена и слова в староассирийских текстах XX–XVIII вв. до н. э. — древнейшие саидетельства об индоевропейских языках // *Вопросы языкознания*, № 2.
- СТАРОСТИН 2007 — Сергей А. СТАРОСТИН. *Труды по языкознанию*. М.: «Языки славянских культур».
- ХЕЛИМСКИЙ 1982 — Евгений А. ХЕЛИМСКИЙ. *Древнейшие венгерско-самодийские языковые параллели: Лингвистическая и этногенетическая интерпретация*. М.: «Наука».
- ЦИНЦИУС 1982 — В. И. ЦИНЦИУС. *Негидальский язык*. Л.
- BALDI 1974 — Philip BALDI. Indo-European *sek^w. // *Journal of Indo-European Studies*, 2; pp. 77–86.
- BALDI 2002 — Philip BALDI. *The Foundations of Latin*. Berlin–New York: MOUTON–DE GRUYTER.
- CHD — Hans G. GÜTERBOCK, Harry A. HOFFNER, Theo P. J. VAN DEN HOUT (eds). *The Hittite Dictionary of the Oriental Institute of The University of Chicago*. Chicago, Vol. **L–N**: fasc. 1 (*la-* to *-ma*: 1980), fasc. 2 (*-ma* to *mi(ya)ḫu(wa)nt-*: 1983), fasc. 3 (*mi(ya)ḫu(wa)nt-* to *nai-*: 1986), fasc. 4 (*nai-* to *nutarnu-*: 1989) [http://oi.uchicago.edu/pdf/chd_l-n.pdf]; Vol. **P**: fasc. 1 (*pa-* to *parā*: 1994), fasc. 2 (*parā* to ^(UZU)*pattar*: 1995), fasc. 3 (^(UZU)*pattar* to *putkiya-*: 1997) [<http://oi.uchicago.edu/pdf/CHDP.pdf>]; Vol. **Š**: fasc. 1 (*šā-* to *šaptamenzu*: 2002 [pp. 1–208: <http://oi.uchicago.edu/pdf/CHDS.pdf>]), fasc. 2 (*šaptamenzu* to *-ši-*: 2005 [pp. 209–332: http://oi.uchicago.edu/pdf/CHD_S2.pdf]).
- ЏОП 1965 — Bojan ЏОП. Sur une règle phonétique de la langue louvite // *Linguistica*, 7; pp. 99–123.
- FRANKLIN 2006 — John Curtis FRANKLIN. The Wisdom of the Lyre: Soundings in Ancient Greece, Cyprus and the Near East [Die Weisheit der Leier: Studien zum Altertum Griechenlands, Zyperns und des Nahen Ostens] // Ellen HICKMANN, Arnd Adje BOTH & Ricardo EICHMANN (eds.) / *Musikarchäologie im Kontext. Archäologische Befunde, historische Zusammenhänge, soziokulturelle Beziehungen*. Vorträge des 4. Symposiums der Internationalen Studiengruppe Musikarchäologie im Kloster Michaelstein, 19.–26. September 2004 [Music Archaeology in Contexts: Archaeological Semantics, Historical Implications, Socio-Cultural Connotations. Papers from the 4th Symposium of the International Study Group on Music Archaeology at Monastery Michaelstein, 19–26 September, 2004] (= Serie *Studien zur Musikarchäologie* [Studies in Music Archaeology], V = *Orient-Archäologie*, 20 / Herausgegeben von: Deutsches Archäologisches Institut, Orient-Abteilung). Rahden (Westfalen); pp. 379–397 [<http://www.kingmixers.com/Franklin PDF files copy/WisdomLyre.pdf>].
- FRANKLIN 2007 — John Curtis FRANKLIN. The Global Economy of Music in the Ancient Near East // J. GOODNICK WESTENHOLZ, ed. / *Sounds of Ancient Music*. Library of Archeology, Jerusalem: KETER Press; pp. 27–37 [<http://www.kingmixers.com/Franklin PDF files copy/GlobalEconBLM.pdf>].
- HART 2004 — Gillian R. HART. Some Problems in Anatolian Phonology and Etymology // *Indo-European Perspectives: Studies In Honour of Anna MORPURGO DAVIES* / Edited by J. H. W. PENNEY. Oxford Linguistics: Oxford University Press; pp. 341–354
- HASPELMATH & al. 2005 — Martin HASPELMATH, Matthew DRYER, David GIL, & Bernard COMRIE [eds.]. *World atlas of language structures*. Oxford: Oxford University Press,.
- HOFFNER 1994 — Harry A. HOFFNER Jr. The Hittite Word for ‘Oil’ and its Derivatives // *Historische Sprachforschung*, 107; pp. 104–112.
- HOFFNER 1995 — Harry A. HOFFNER Jr. Oil in Hittite Texts // *The Biblical Archaeologist*, Vol. 58, N^o. 2 [Special Issue: *Anatolian Archaeology: A Tribute to Peter NEVE*] (June, 1995); pp. 108–114.
- HOFFNER & MELCHERT 2008 — Harry A. HOFFNER Jr. & H. Craig MELCHERT. *A Grammar of the Hittite Language*. Part 1. *Reference Grammar*. Winona Lake (Indiana): EISENBRAUNS.

- HULD 1984 — Martin. E. HULD. *Basic Albanian Etymologies*. Columbus (Ohio): Slavica.
- IVANOV 1999 — Vyacheslav IVANOV. An Ancient Name of The Lyre // *Archiv Orientalní*, vol. 67, № 4; pp. 585–600.
- IVANOV 2001 — Vyacheslav V. IVANOV. Southern Anatolian and Northern Anatolian as Separate Indo-European Dialects and Anatolian as a Late Linguistic Zone // *Greater Anatolian and the Indo-Hittite Family* / Ed. R. DREWS (= *Journal of Indo-European Studies*. Monograph Series. № 38). Washington (D.C.): Institute for the Study of Man; pp. 131–183.
- JOSEPHSON 1979 — F. JOSEPHSON. Assibilation in Anatolian // *Hethitisch und Indogermanisch*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck; SS. 91–103.
- KLOEKHORST 2008 — Alwin KLOEKHORST. Etymological Dictionary of the Hittite Inherited Lexicon // *Leiden Indo-European Etymological Dictionary Series* / ed. A. LUBOTSKY. Leiden: BRILL.
- KOIVULEHTO 1999 — Jorma KOIVULEHTO. Verba mutata. Quae vestigia antiquissimi cum Germanis aliisque Indo-Europaeis contactus in linguis Fennicis reliquerint // *Suomalais-Ugrilaisen Seuran Toimituksia*, 237. Helsinki: Suomalais-Ugrilainen Seura.
- KORTLANDT 2001 — Frederik KORTLANDT. *The Indo-Uralic Verb*. Leiden [https://openaccess.leidenuniv.nl/bitstream/1887/1949/1/344_139.pdf] (~ [KORTLANDT 2002]: *The Indo-Uralic Verb* // R. BLOKLAND & C. HASSELBLATT (eds.) / *Finno-Ugrians and Indo-Europeans: Linguistic and literary contacts* (= *Studia Fenno-Ugrica Groningana* 2). Maastricht: SHAKER; pp. 217–227 [www.kortlandt.nl/publications/art203e.pdf]).
- MELCHERT 1994 — H. Craig MELCHERT. *Anatolian Historical Phonology*. Amsterdam–Atlanta (GA).
- MKRTSCHIAN 1974 — N. MKRTSCHIAN [Ն. ՄԿՐՏՅԱՆ]. Neue hethitisch-armenische lexikalische Parallelen // *Acta Antiqua Academiae Scientiarum Hungaricae*, XXII.
- NIKOLAEV, STAROSTIN 1994 — S. L. NIKOLAEV, S. A. STAROSTIN. *A North Caucasian Etymological Dictionary*. Moscow: “Asterisk” [= NCED].
- OREL 2000 — Vladimir OREL. *A Concise Historical Grammar of the Albanian Language*. Leiden: BRILL.
- POPKO 1994 — M. POPKO. Zippalanda: Ein Kultzentrum im hethitischen Kleinasien // *Texte der Hethiter*, Heft 21. Heidelberg: Carl WINTER.
- PUHVEL 1974 — Jaan PUHVEL. On Labiovelars in Hittite // *Journal of the American Oriental Society*, Volume 94, Number 3: July–September 1974; pp. 291–295.
- PUHVEL 1975 — Jaan PUHVEL. Greek Attestations of Indo-European **dhyagh^w* // *Incontri linguistici*, 2; pp. 129–134.
- PUHVEL 1979 — Jaan PUHVEL. Some Hittite Etymologies // *Florilegium Anatolicum. Melanges offerts à Emmanuel Laroche*. Paris: DE BOCCARD; pp. 297–304.
- PUHVEL 1997 — Jaan PUHVEL. Hittite Etymological Dictionary. Words beginning with *K*. // *Trends in Linguistics*, 14. Berlin–New York: MOUTON–DE GRUYTER.
- SCHUOL 2004 — Monika SCHUOL. Hethitische Kultmusik. Eine Untersuchung der Instrumental- und Vokalmusik anhand hethitischer Ritualtexte und von archäologischen Zeugnissen [Hittite Cultic Music. A Study of Instrumental and Vocal Music on the Basis of Hittite Ritual Texts and Archaeological Evidence] // *Orient-Archäologie*, 14 / Herausgegeben von: Deutsches Archäologisches Institut, Orient-Abteilung. Rahden (Westfalen).
- SCHINDLER 1969 — J. SCHINDLER. Die indogermanischen Wörter für “Vogel” und “Ei” // *Die Sprache*, 15; SS. 144–167.
- SOYSAL 2004 — O. SOYSAL. Hattischer Wortschatz in hethitischer Textüberlieferung // *Handbook of Oriental Studies*, Section One. *The Near and Middle East*. Vol. 74. Leiden–Boston: BRILL.
- STAROSTIN-N — S. A. STAROSTIN. *Nostratic Etymology* at <http://starling.rinet.ru>.
- STAROSTIN-SC — S. A. STAROSTIN. *Sino-Caucasian Etymology* at <http://starling.rinet.ru>.
- STAROSTIN, DYBO, MUDRAK 2003 — Sergey A. STAROSTIN, Anna V. DYBO, Oleg A. MUDRAK. *An Etymological Dictionary of Altaic Languages*. HdO Vol. 8-1, 8-2, 8-3. Leiden: BRILL, 2003 [= EDAL].
- STURTEVANT 1927 — Edgar H. STURTEVANT. The *e*-Perfect in Hittite // *Language*, Vol. 3; pp. 163–168.
- STURTEVANT 1942 — Edgar H. STURTEVANT. *Indo-Hittite Laryngeals*. Baltimore.
- THUMB 1905–1906 — A. THUMB. Zur Vorgeschichte von Got. *paida* // *Zeitschrift für Deutsche Wortforschung*, 7; SS. 261–267.
- WEGNER 2007 — Ilse WEGNER. *Eine Einführung in die hurritische Sprache*. 2nd edition. Wiesbaden: HARRASSOWITZ.

S u m m a r y

The article represents a brief survey of certain actual problems of comparative-historical linguistics, the solution of which, in the author’s opinion, is necessary to achieve further breakthroughs in this field. This includes such issues as lexical and semantic reconstruction; a proper application of morphological and lexicostatistical criteria to determine the limits of language families and macrofamilies; phonetic interpretation of the reconstructed phonological systems; and finding better criteria to distinguish between borrowed and inherited lexicon. Some approaches to solving these and other issues are demonstrated by the author on the basis of ancient and reconstructed languages of the Near Eastern region.

Murray GELL-MANN
(Santa Fe Institute)

Ilia PEIROS
(Santa Fe Institute)

George STAROSTIN
(Russian State University for the Humanities)

Distant Language Relationship: The Current Perspective

The idea that it may be possible to use data from modern languages in order to trace their origins back ten thousand years or even further, let alone even partially reconstruct the shape of their remote ancestors, remains a subject of discussion. Our article, partially based on results of research conducted by a joint team of Russian and American scholars within the *Evolution of Human Languages* project, discusses the methodological basis of this approach and the various ways of overcoming the most common obstacles associated with it. In the second part of the article we summarize the main results of the project and suggest further specific directions for exploring the question of deep-level linguistic relationships and the prehistory of human languages.

The study of genetic relationships among the world's languages has, for the last two hundred years, steadily remained one of the most fascinating and productive areas of research in linguistics. Yet, despite great advances achieved in the field during that time, even more remains to be done, due primarily to the immense quantity of the material, widely varying degrees of our knowledge about different linguistic areas of the world, and a lack of proper coordination of efforts between specialists.

This is why, in an attempt to overcome at least some of these problems, *Evolution of Human Languages (EHL)*, an international program of the Santa Fe Institute, was launched in 2001. The primary goal of the present paper, then, is to outline what we consider the major achievements of the *EHL* program, along with a brief description of its methodological basis and various problems encountered along the way.

1. General background.

One of the most important concerns of *EHL* is that all the work conducted within the project should rest on a properly defined and well-grounded foundation, one that would acknowledge and respect all the achievements of traditional comparative historical linguistics, yet at the same time give clear instructions as to how these achievements should be used in further research on deeper level relationships between languages — research that is, for the most part, still in its early stages and the standards for which have not yet been carved in stone.

This means that it is perhaps reasonable to begin our article with a very brief summary of some of the axioms of comparative linguistics (even though, no doubt, all of them will be familiar to the majority of our readers), after which we shall add to those some of the (usually) less emphasized methodological assumptions as they are shared by *EHL* participants.

It is commonly recognized today that all languages are subject to historical change; regardless of the environmental or social context, no human language can be completely immune to this process. One inevitable type of change occurs when languages are passed down from parents to children: this includes modifications within the phonological system, shifts of meanings for various morphemes, loss of certain morphological and lexical oppositions and emergence of new ones, &c. This process of “vertical” transmission is in stark contrast with the “horizontal” process of borrowing features from other languages. Since both types of changes can occur independently in different speech communities using a particular language, that language may eventually split into several different ones, each being a direct descendant of the original one (the “proto-language”).

Given that the splitting process frequently leads to the appearance of an entire family of related languages in the place of just one descendant, this process is best illustrated by a tree-type representation¹.

¹ Two major alternative models should be mentioned: the “wave theory” and “language convergence”. The “wave theory” talks mainly about possible causes of language changes (wave-like spread of innovations) and thus

Each family tree consists of a number of nodes connected by arrows. Nodes represent (proto-)languages, while arrows are used to represent relations between the nodes. Therefore, while studying any particular language family, we need to address two different questions: “which nodes will we identify? (grouping)” and “how are they linked to each other in the tree? (branching)”.

With the exception of a tiny handful of ancient languages attested in writing, modern languages are our only source of information on previous stages of language development. Reconstruction of these stages based on a scientific study of modern languages is the main task of comparative linguistics.

The major principles of comparative linguistics, after half a century of preliminary research, were explicitly and systematically formulated by the Neogrammarian school in the second half of the XIXth century on the factual basis of the Indo-European language family. Today, despite occasional criticism from the theoretical side, they continue to remain as the backbone of this field of study. For the most part, these principles concern reconstruction of proto-languages through a comparison of their descendants that necessarily involves discovery and description of systematic patterns of language change. Primary among these is the presence of regular (recurrent) phonetic correspondences between compared lexical items. Over the last century or so, the comparative method has been successfully applied to a whole set of families whose languages vary widely in phonological and grammatical structure, area of dispersal, lifestyle of the speaker community, and cultural background. We therefore accept the universal nature of the comparative method².

2. Language families: the three types.

Expanding beyond general textbook information on basic premises of comparative linguistics, we may classify all well-established and hypothetical linguistic families³ into three types, depending on the level of transparency of relationships exhibited by their daughter-languages. This classification has both a practical / intuitive origin and a more formalized one, and we believe it to be of serious use for macrocomparative purposes.

In practical terms, type A would be associated with compact, well-defined families (which we may simply call “groups”) like Germanic, Polynesian, or Turkic. As a rule, it is intuitively evident for specialists and native speakers alike that such families have arisen not very long ago, and it is usually not very difficult to obtain a proto-language reconstruction for each of them in strict accordance with the comparative method. These reconstructions normally constitute the most reliable results in historical linguistics.

Type B families (below called “stocks”) consist of languages whose relationship is not intuitively evident for speakers, so that special research is needed to demonstrate it. Nevertheless, a strong case often can be made for genetic relationship even without resorting to strict phonological reconstruction, since the languages still preserve a large number of the common forms and features of their proto-language. The best-known case of such a stock (usually called “family”) is Indo-European; other examples would include, for instance, Turkic, Dravidian, or Uto-Aztecan. The number of reliable reconstructions for this type of family is significantly smaller than for those of type A.

As a rule, the best results at this level are achieved not through direct comparison of modern languages, but in two steps: (i) reconstructing proto-languages for all the type A groups within the larger family and (ii) using these reconstructed “intermediate” proto-languages as the base material for reconstructing the proto-language of the larger family, thus increasing the level of transparency⁴.

does not contradict the “tree model”. The language convergence model claims that new languages can arise through convergence of formerly unrelated idioms. In most known cases, “convergence” can be simply interpreted as massive-scale borrowing from language *A* to language *B*, without affecting its genetic affiliation. For a good discussion of the subject see, e. g., [DIMMENDAAL 1995].

² For a more detailed explanation of the comparative method and the way it can be applied to different language families any number of general handbooks on historical linguistics may be consulted. Some of the best include, in English, [HOCK & JOSEPH 1996]; [CAMPBELL 2004], in Russian, [BURLAK & STAROSTIN 2005].

³ The term ‘family’ is extended here to signify any type of genetic group of languages, families of any age and internal diversity, superfamilies, and even super-super-families. For a detailed discussion, see [PEIROS 1997].

⁴ One notable exception from this procedure is Indo-European. Although the family (= stock) unquestionably belongs to Type B, its reconstruction was effectuated in the 19th century not on the basis of intermediate reconstructions of Proto-Slavonic, Proto-Germanic, Proto-Indo-Aryan, &c. but on the basis of historically attested ancient languages (Old Indian, Ancient Greek, Classical Latin, &c.); these were used as entities close to proto-language status for their respective branches.

To use an example from Dravidian languages, a successful direct comparison between, for instance, Kui *siṅg* and Toda *ūz* ‘five’ would hardly be imaginable; however, if the former is first traced back to Proto-Gondwan **say-* and the latter to Proto-South Dravidian **ay-* (with both *-iṅ* and *-z* elements going back to old numeric suffixes), the comparison becomes easier to interpret; all that is needed is to verify the regularity of the correspondence “Gondwan **s-* : Proto-South Dravidian \emptyset ”, which is, indeed, regular [BURROW & EMENEAU 1984: 245].

In addition, relying on intermediate proto-languages usually leads to a drastic increase in the number of suggested etymologies: there will always be more matches between Proto-Germanic and Proto-Slavic, for instance, than there will be for modern German and Russian, due to the inevitability of lexical loss in individual daughter languages. A relevant form may be lost in Russian, but retained in Polish and, say, Bulgarian and thus reconstructed for the proto-language.

Finally, Type C, for which the name “superfamily” or “macrofamily” is often reserved, represents a situation in which comparison between attested (usually modern) languages may yield so little evidence that any hypothesis of their being related risks falling within the limits of chance coincidence. With thousands of years of phonetic change obscuring potential regularity of correspondences, and, even more importantly, thousands of years of gradual semantic and lexical replacement drastically reducing the numbers of true cognates and exact semantic matches, credible reconstructions of proto-languages for superfamilies cannot usually be significantly advanced by comparison of modern languages. However, one may always try to increase the level of transparency by employing reconstructions of intermediate stages ($A > B > C$).

Thus, in a situation similar to that of the previous example, hardly anyone would propose to compare directly Hungarian *víz* ‘water’ with Greek *hydōr* id., due to lack of obvious phonetic similarity. In addition, even if we made a wild guess, proposing that the correspondence between Hungarian *v-* and Greek *h-* is regular, we would not be able to support our proposition with a significant number of equally strong lexical matches between Hungarian and Greek.

However, if, through systematic reconstruction, the Hungarian form is first traced back to Proto-Uralic (Type B family) **wete*, and the Greek form to Proto-Indo-European (also type B) **wedōr*, the reconstructed forms, being much closer to each other phonetically, provide a strong incentive for further comparison. In addition, operating with Proto-Indo-European and Proto-Uralic rather than Hungarian and Greek allows us to demonstrate the regularity of this correspondence in a much more satisfactory manner; cf. such additional comparisons as PIE **uegh-* ‘to carry’ : PU **wiye* id., PIE **uisu-* ‘all’ : PU **weńce* id., PIE **uonk-* ‘crooked’ : PU **wanjka* id., &c. (in these and other examples the Greek and/or Hungarian equivalents are either missing or show different reflexes of the original reconstructed phoneme **w-*).

The basic principle underlying this and many other such examples is as follows: if the overall number of cognates, their degree of phonetic similarity, and the degree of transparency of regular phonetic correspondences between proto-languages X and Y all show a significant increase compared to the same features of any given daughter language of X and any given daughter language of Y, this is a strong signal in favour of genetic relationship between X and Y.

This semi-intuitive understanding of different types of families is conveniently formalized on the basis of such methods as *lexicostatistics* and *glottochronology*. The former measures degrees of relationship between languages based on the relative number of cognates in their basic lexicon (most conveniently represented by the so-called “SWADESH wordlist” of 100 items); the latter carries this procedure even further by trying to calculate approximately the absolute age of separation between related languages⁵. Thus,

⁵ A detailed discussion of glottochronology (frequently defined as the method of dating the age of language separation based on the assumption of a constant rate of change within the basic lexicon) lies beyond the scope of this paper, and acceptance or rejection of this method is generally irrelevant to the main points that the current article is trying to emphasize. It should, nevertheless, be stated that the actual formula that is being consistently applied within *EHL* is not the original one, first suggested by MORRIS SWADESH in the early 1950s and heavily criticized by many specialists in historical linguistics. Instead, we use a significantly revised version of it put forward by S. STAROSTIN. The revised variant, in contrast with the SWADESH formula, works well on all historically attested cases (this has been demonstrated in [STAROSTIN 2000]) and has so far yielded credible results for all analyzed families (for some of the figures, see, e. g., [STAROSTIN 2003]). Further methodological improvements, such as introducing individual replacement rates for each word on the 100-wordlist ([STAROSTIN 2007]), were gradually introduced by STAROSTIN until his demise in 2005; today, the work is carried on by some of his disciples, e. g. [VASILYEV & MILITAREV 2008].

Several serious objections against the methods of lexicostatistics and glottochronology are worth a brief mention. One is based on misleading percentages of cognates obtained while comparing Icelandic with other Scandinavian lan-

each of the three family types can be roughly associated with (a) certain ranges of percentages of lexical matches between related (modern) languages and (b) a certain range of ages.

Type	Percentage of matches on the SWADESH 100-wordlist	Approximate age ⁶
A	≈ 45%–95%	≈ 2,000–4,000 KYA ⁷
B	≈ 10%–45%	≈ 4,000–7,000 KYA
C	less than 10%	older than 7,000 KYA

In reconstructing a proto-language we first start with attested languages, thus moving in the opposite direction from actual historic development. However, it is only in this way that we can preserve and account for all the relevant information maintained in recorded languages. Relatively simple reconstructions of type A families are used to advance to more ancient stages (type B); with a sufficient number of type B reconstructions at our disposal, we have the right to begin the study of an even deeper, Type C, relationship, where “B-reconstructions” perform the same function as modern languages for the first step of the research.

It is necessary to specify here what we mean by “reconstruction”, since the term can be used differently in various kinds of historical research. An actual reconstruction consists of several steps:

- (i) formulation of the initial relationship hypothesis. This is usually done on an intuitive basis through observing certain kinds of similarities between a number of languages — primarily lexical and morphological (if the languages in question *do* have morphology);
- (ij) accumulating a representative list of similar-looking morphemes in the languages under investigation (“similar-looking” first and foremost presupposes straightforward phonetic similarity between the morphemes, but may also signify a preliminary, sketchy notion of patterns of correspondence). At this stage such a collection may include all sorts of comparisons: borrowings, chance resemblances, forms of common origin, even mistakenly transcribed data (particularly for languages where such data are the only available kind);
- (iij) performing a reconstruction of the phonological system of the protolanguage by tracing recurrent patterns in the accumulated material; these might suggest rejecting some look-alikes or adding previously neglected comparisons;
- (iv) reconstructing the morphemic inventory of the proto-language by sifting through the list of accumulated lookalikes and modifying it according to one’s current understanding of the languages’ historical phonologies;
- (v) reconstructing (wherever possible) the grammar of the proto-language.

It must be stressed that the order of these stages cannot be changed: phonological correspondences can only be established on the basis of morphemic look-alikes, while reconstructing morphological structure is impossible without a list of reconstructed grammatical morphemes. The minimum amount of information about any given proto-language must necessarily include a reconstructed phonological system accompanied by a representative list of reconstructed morphemes (both lexical and grammatical). In other words, a list of *segmental units* presented for the proto-language is always more substantial evidence for its existence than any proposals on its typological structure, means of word formation, &c., which represent the final stage of research.

3. Problems of linguistic grouping.

Of the several hundred language families mentioned in general linguistic literature, some are well-studied, with credible proto-language reconstructions, while others may be rejected as errone-

guages [BERGSLAND & VOGT 1962]; this issue has been eliminated by S. STAROSTIN [2000], who showed that such errors are usually generated by large amounts of borrowings in the wordlist, and that elimination of borrowings from the calculations leads to much better results. The idea of diversity of retention rates put forward by R. BLUST on the basis of Austronesian, (see, for example, [BLUST 2000]), has been demonstrated as unnecessary by one of the authors of the present article [PEIROS 2000]. Also, BLUST’s observation that existing lexicostatistical classifications of Austronesian, such as [DYEN 1965], do not confirm the mainstream views on this family, is not particularly relevant, since so far, neither the lexicostatistical, nor the generally accepted classification of this family have yet received proper justification.

⁶ We crudely estimate the margin of error of our datings at an order of about 10%.

⁷ Thousands of years ago.

ous or unconvincing. A list of more or less accepted language families is given in *Ethnologue* [GORDON 2005], which can be used here as a temporary starting point for further discussion.

The number of families identified within Eurasia is not very large⁸; it includes mainly B-type families like Indo-European, Uralic, Dravidian, Sino-Tibetan, Tai-Kadai, &c. There are also a few smaller A-type families that require further comparative study. Some languages of Eurasia are treated as isolates, without any known genetic connection⁹. No C-type families in Eurasia are accepted by *Ethnologue*.

When it comes to classification of African languages, the same source gives us quite a different picture. Here it lists but four “families”, all of them C-type: Khoisan, Afroasiatic, Niger-Congo, and Nilo-Saharan. The Khoisan family is the smallest and, through recent research, probably, slightly better understood than the others¹⁰ (although its external connections are hard to define because of unique phonetic features such as the so-called ‘click’ sounds, making it hard to compare even the intermediate reconstructions with the other African taxa). There is also a large amount of information about the Afroasiatic superfamily, including lists of etymologies and major phonological correspondences, but much further work is needed to raise it to expected standards, especially within its Chadic, Cushitic, and Omotic branches. Both the Niger-Congo and Nilo-Saharan superfamilies were identified by Joseph GREENBERG ([1966]) based on his “mass comparison” method (see below). Unlike GREENBERG’s hypotheses on “Amerind” (see below), his identification of African superfamilies has in large part been accepted, despite the lack of conclusive comparative results.

The Indo-Pacific region remains linguistically the worst known part of the world. Most of the languages spoken here (both Austronesian and non-Austronesian = Papuan) are not well-documented. For the Papuan languages it has so far only been possible to group them into a dozen families, ranging from small A-type ones to major C-type ones like the Trans-New Guinea superfamily, plus a significant number of isolates. None of these families has been extensively studied based on the comparative method. All the languages of Australia are grouped by *Ethnologue* in a single C-type super-family¹¹, even though this plausible suggestion cannot yet be firmly backed up by the comparative method either.

The most diverse situation, according to *Ethnologue*, is found in the Americas. Here it lists 79 groups¹², most of them formed by a small number of languages. Only 18 families (all of them A or B-types) have more than 15 languages each. For some of these larger families there are reconstructions (including etymological dictionaries).

Table 1. A summary of *Ethnologue* groupings

Regions	Number of languages in families				126
	large (40+)	medium (15–39)	small (2–15)	isolates	total
Eurasia	8 (B-type)	2 (B-type)	7 (A/B-types)	5	22
Africa	3 (C-type)	1 (C-type)	—	—	4
Indo-Pacific	3 (B/C types)	3 (B/C types)	8 (A-type)	6	20
Australia	1 (C-type)	—	—	—	1
Americas	8 (B-type)	10 (A/B type)	36 (A-type)	25	79

These groupings are based on a number of reasons, ranging from solid comparative evidence to affiliations based on little more than ‘general’ agreement rather than facts. This brings up the question: is there actually a way to evaluate various grouping proposals according to the same standard?

One common reproach directed towards traditional comparative linguistics is that it lacks an objective methodology of language classification¹³. With more and more linguistic stocks added to the

⁸ Very similar classifications are found in *Encyclopedia Britannica* and other respected sources.

⁹ The list of isolates will be longer if we add extinct languages such as Sumerian, Etruscan, and others.

¹⁰ See SANDS [1998], where the idea of all Khoisan languages with the exception of Hadza belonging to a single family is deemed realistic, and G. STAROSTIN [2003] and [2008], offering lexical and lexicostatistical evidence for such a relationship.

¹¹ The extinct languages of Tasmania, however, must be classified differently ([CROWLEY & DIXON 1981]).

¹² CAMPBELL [1997] gives even more groups and strongly rejects most attempts to search for deeper connections.

¹³ The most commonly used method is that of ‘shared innovations’, but it has not been formalized enough to resolve numerous complicated cases of relationship, nor is it possible to use ‘shared innovations’ as a universal classificatory standard for all of the world’s languages. A painfully obvious example of the limitations of this

list of families subject to comparative research, the problem of finding strict criteria for classification becomes even more immediate, as answers are required for such questions as ‘which languages can be put in one family?’ (absolute classification) and ‘which groups form the family and how they are related to each other?’ (relative classification).

One potential answer is to rely on a properly formalized and standardized application of lexicostatistics, already mentioned above. This solution is frequently chosen by specialists in African languages and, to a lesser extent, those in the Pacific region. It is also embraced by all members of the *EHL* project. However, application of lexicostatistics to a given language family cannot be fully successful without preliminary identification of cognates — meaning that lexicostatistical classifications can be of two types: (a) preliminary, based on limited sets of historical data and used only in the exploratory stages of research and (b) advanced, based on a long process of etymological research on compared languages.

4. *Etymological dictionaries.*

We have already pointed out that related languages always possess morphemes of common origin, inherited from their common ancestor. If no such morphemes are found, this may indicate that: (i) the languages studied are related on such a remote level that one cannot find those morphemes, or (ii) they are not related at all (even though there is no certain method to prove lack of relationship). The first criterion, then, is to use morphemes of potentially common origin as a tool to evaluate various grouping proposals. The logic is straightforward: if a grouping proposal is supported by a sufficient etymological list (EL) of morphemes sharing a common origin, we accept it; otherwise, we remain unconvinced.

One commonly asked but rarely answered question is how many actual etymologies we need to constitute a satisfactory EL. To show that the morphemes in question may truly share a common origin and not be just look-alikes, we need to establish regular phonetic correspondences between the languages. At least 300–400 proto-morphemes are needed, on average, to support and illustrate any given phonological reconstruction. This approximate number — reflecting the idea that each phoneme of our reconstructed proto-language has to be illustrated by a statistically significant number of examples — can be seen as the lower limit for our EL.

Ideally, however, for each proposed family we need to have a full etymological dictionary, containing all the morphemes that can be reconstructed for its proto-language. Etymological dictionaries of well-studied A-type families usually have about 2–3 thousand proto-morphemes, with reflexes found in at least two daughter languages. The size of etymological dictionaries for B and C-type families depends mainly on the level of transparency of the relations among their daughter families and the number of proto-morphemes known for each of them. Dictionaries of Indo-European, Altaic, Uralic, and several other well-studied B-type families have more than 2,000 entries; by induction we can assume that this number should be expected for any proposed family regardless of its age (although it depends on such a technical factor as the number of languages / branches used for reconstruction).

Apart from quantity, reliable ELs also have to meet certain qualitative requirements. These are based on the well-known (although approximate) division of the lexicon into ‘basic’ and ‘cultural’, as well as the empirically and theoretically backed assumption that ‘basic lexicon’ is generally less subject to processes of horizontal transmission than ‘cultural lexicon’. Therefore, if two or more languages are related, they will always have etymologies that belong to the sphere of basic lexicon.

The process of phonological reconstruction is so well defined in comparative linguistics that if two linguists have the same systems of phonological correspondences and lexical data, they are usually expected to get the same proto-form (or, at least, be able to agree upon one). Semantic reconstruction is different: we still lack a formal procedure to evaluate proposed semantic connections. Historical dictionaries of an individual language, be it English, French or a less familiar language, contain many non-trivial semantic changes (e. g. English *clean* = German *klein* ‘small’) some of which would never be accepted by linguists unless recorded in literature; likewise, etymological dic-

method is the still unresolved issue of the tree of Indo-European languages. While there is little question about the consistency of its ‘small’, basic branches — Celtic, Germanic, Slavic, etc. — ‘shared innovations’ do not seem to work where larger sub-groupings of Indo-European are concerned (such as, e. g., the very hypothetical Italo-Celtic). Likewise, ‘shared innovations’ have not managed to resolve numerous problems with the internal classification of such families as Uralic, Dravidian, Austronesian, and not a few others.

tionaries of language families frequently feature unusual semantic changes that some scholars may consider controversial (cf., for example, Nenets *sārmik* ‘wolf’ = Selkup *suurem* ‘bird’, out of a Proto-Uralic root with a presumed original meaning ‘wild animal’ [RÉDEI 1988: 490]). The balance between ‘straightforward’ and ‘loose’ semantic reconstruction is a good indication of the quality of any given EL: a reliable one always has a large proportion of straightforward semantic comparisons.

Finally, if a certain proto-stem is attested in all the daughter languages, it is usually attributed to the proto-language level; if not, we need to decide which chronological level it represents. If very few of the etymologies are represented widely enough in the daughter languages, this increases the probability that we may actually be dealing with language contact or even chance resemblances.

Etymologies that meet these conditions can be called ‘core’, loosely defined as etymologies that

(a) belong to the basic lexicon;

(b) satisfy the proposed system of phonological correspondences, either fully or partially (in the latter case, accompanied by a reasonable explanation of the irregularly corresponding parts of the forms; this remark is introduced specially to account for the fact that even in well-established families of A or B types occasional irregularities are practically inevitable — e. g., Latin *quattuor* and Proto-Germanic **fiðwor* ‘four’ do not correspond to each other ideally, but can still be counted as a ‘core’ etymology since the latter form quite likely is the result of a phonetic dissimilation from **hwiðwor*);

(c) constitute either a perfect semantic match (i. e. the meanings of the compared words are the same) or a near-perfect one (the meanings are not the same, but the semantic development is trivial and frequently attested in languages, e. g. ‘water’ → ‘rain’, ‘belly’ → ‘stomach’, ‘eye’ → ‘see’, &c.);

(d) undoubtedly belong to the proto-language level, i. e. are represented in most or at least non-contacting daughter languages / branches.

The more core etymologies we find in an EL, the more reliable it becomes; and with this growing reliability we can allow ourselves to include more numerous ‘peripheral’ etymologies, ones that are based on the same system of correspondences as the core ones but either do not belong to the basic lexicon or violate the ‘near-perfect semantics’ rule. The ultimate proof of relationship, however, always lies with the ‘core’ part.

The ‘core etymology’ criterion is not always precisely formulated in various handbooks on historical linguistics, but probably no one would argue that there is even a single well-studied language family for which it does not work. And when properly applied to such linguistic situations as we find between English and French, Tamil and Old Indian, Chinese and Japanese, Turkic and Mongolic, it rarely, if ever, fails to indicate the difference between traces of old contacts and genetic inheritance.

However, this criterion has so far remained unapplied to many lesser known families that do not have a long tradition of scientific research. The *Ethnologue* classification recognizes many of these not because their proto-languages have been reconstructed and core etymologies established, but because scholars have used various similarities (lexical, grammatical, even typological) as indications of possible groupings. Such an approach is typical, for example, of the language families of New Guinea, where preliminary identification was not accompanied by proper comparative work¹⁴. As a result, further study suggests that some of these groupings have been mapped out incorrectly, e. g. the Sepik-Ramu family [FOLEY 2005], whereas others are better supported by data.

The immense Austronesian family with its twelve hundred languages represents an intermediate situation: some of its members are tied together on solid comparative grounds, while others (mainly in West Oceania) are added on the basis of unsystematic similarities. Such a basis may be useful at a very early stage of research, but the problem is that unsystematic similarities can be traced between almost any languages, and only a transition to systematic phonetic correspondences and ‘core’ etymologies guarantees a convincing relationship/classification proposal.

5. Language families and linguistic tradition.

From a purely practical point of view, it helps to distinguish between what we have tentatively named ‘classic’ and ‘non-traditional’ language families. The first type includes families with a long his-

¹⁴ At this stage of research it is practically impossible to find hard evidence in favor of Torricelli, Left May, and some other groupings given in *Ethnologue*.

tory of comparative studies, like Germanic, Semitic, or Turkic. Current specialists in Germanic linguistics, for example, have access to a vast amount of information, not just literature on Proto-Germanic, but also etymological and dialectal dictionaries of modern languages, detailed descriptions of histories of individual languages, and, in many cases, native informants. Only a meticulous study of this literature and the major languages of the family enables scholars to find a new topic for research and investigate it.

A ‘non-traditional’ family never enjoys the same level of attention. Some or even most of its languages are poorly known, and historical research here is usually carried out by one or two scholars (who often have to collect their own data in the field before using them for their reconstructions). Each of these scholars is doing research equivalent to that of dozens of specialists in ‘classic’ families; yet at the same time they are expected to produce results of the same quality, which is often not possible. This is not to suggest that much looser requirements should be applied to the investigation of ‘non-traditional’ families. But one should always remember that their study is conducted in a less favorable environment and must be treated accordingly.

Today, Indo-European comparative linguistics is still — for good reason — seen as the best standard to which one can hold comparative research done on the basis of other language families, ‘classic’ and ‘non-traditional’ alike. Yet there are also certain types of research for which it makes little sense to hold them up to such high standards.

Imagine, for instance, a situation when someone asked to evaluate a 17th century map of North America responds in the following manner: ‘Since this map is plagued with mistakes and there are much more precise recent maps available, this one is only valuable as an antique.’ This answer is at best incomplete, since it forgets to mention that

- at the time of its charting, this map was the best of its kind;
- many of the locations are mapped quite precisely;
- it may have been used as a source for further geographical discoveries, &c.

The same historical approach is normally applied to an ‘obsolete’ piece of linguistic research. We evaluate it according to its impact on the respective field and not just according to the current state of linguistic knowledge. It would certainly look odd if, for example, a modern reference to R. CALDWELL’s trail-blazing *Comparative Grammar of Dravidian Languages* ([1856]) were to simply list all of its (quite numerous) errors, but failed to mention that it was actually the first major publication on Dravidian historical linguistics, which opened this field of research to subsequent generations of linguists who were later able to improve on its achievements.

The same logic must be applied to evaluating historical research done for ‘non-traditional’ language families. It would not be consistent, for example, to compare reconstructions of Proto-Arawakan by E. MATTESON or D. PAYNE ([MATTESON 1972], [PAYNE 1991]) based on modern languages fresh from the field with, e. g., J. POKORNY’s dictionary of Indo-European ([1958]), based on at least a century of linguistic and philological studies. Yet even with that impressive background, POKORNY’s dictionary is still known to contain factual mistakes and dubious etymologies, with scores of linguists working to improve it; but are there currently even two linguists in the world working on Proto-Arawakan?

Finally, the same reservations that should be made for poorly studied and ‘traditionless’ groups apply equally well to research on Type C families. Proto-languages for these families are derived from reconstructions of Type B and Type A families which themselves may be incomplete and partially incorrect (just as work on recently transcribed ‘minor’ languages can be), and research on them is usually carried out by just a tiny handful of specialists (as is the case with Arawakan). From a technical standpoint, this inevitably decreases the general quality of the output; yet there is still no denying its significance for further generations of researchers.

Within the linguistic community, there are two approaches to evaluating research on ‘suspicious’ families of these kinds. We may provisionally call those approaches ‘hypercritical’ and ‘bona fide’. The former is perhaps best illustrated by a series of publications criticizing the Altaic theory (the hypothesis of a genetic relationship among Turkic, Mongolian, Manchu-Tungus, Korean, and Japanese languages) by authors such as G. CLAUSON, G. DÖRFER, and, more recently, J. JANHUNEN, S. GEORG, and others. This kind of criticism usually aims at weaker, less easily defensible parts of the theory and, upon discarding them, uses induction to carry the skepticism over to its stronger sides. It is possible, for instance, to concentrate one’s attention on several etymologies containing mistakes on the part of the authors

(usually mixing together *factual mistakes* and *alternate hypotheses*, so that it is hard for the general reader to tell one from the other) and omit the better etymologies from the discussion altogether¹⁵.

What is forgotten in the process of such criticism is that the same procedure can easily be used to discredit even commonly accepted, ‘traditional’ theories of genetic relationship, essentially bringing comparative research to a standstill altogether. The first effects of this may already be observed in such works as [MARCANTONIO 2002], in which the author presumes to discredit the well-proved theory of Uralic (Fenno-Ugric and Samoyed) relationship. Fortunately, the work met more criticism than appraisal from specialists in the field, but it should be noted that the author is a professional linguist, and that the book was issued by no less than the Philological Society series at Cambridge. It is alarming to imagine that this may, indeed, be but the first sign of things to come; surely there are plenty of things that can be criticized about Indo-European as well.

Evaluations and reviews that we call ‘bona fide’, on the other hand, are based on assuming that:

- the publication may contain important new data;
- its author(s) is/are not intentionally trying to deceive the linguistic community;
- new approaches, especially in the domains not fully explored by traditional comparative linguistics, should not be prohibited unless they lead to conclusions that directly contradict the basic methodology of the science;
- factual mistakes may undermine, but not necessarily invalidate the major conclusions;
- if these conclusions go against the views of the ‘mainstream’, this does not justify their immediate rejection.

This by no means equates a ‘bona fide’ review with one that willingly endorses the proposed genetic relationship; the only serious difference is that ‘bona fide’ reviewers may be willing to weigh the positive evidence and explicitly evaluate it along with the negative. Yet such a position is not frequently found in the literature.

As a typical example we might mention the treatment of Joseph GREENBERG’s Amerind hypothesis [GREENBERG 1987]. Several books and dozens of articles were then written in response, contradicting the proposal that all native languages of the Americas that are not Eskimo-Aleut or Na-Dene belong to just one Amerind super-family. The authors gave theoretical reasons why they think GREENBERG’s method is wrong¹⁶; listed his numerous typos and mistakes; argued — in many cases, quite convincingly — that certain languages should actually be assigned to different low-level families, &c. (a representative example is [CAMPBELL 1997]). Yet at the same time it still remains unexplained why there are so many similar-looking words found, for instance, in GREENBERG’s Equatorial family. Does this mean that all of his Equatorial comparisons have no value whatsoever? Is the same also true for other major families within “Amerind”? How can we provide a reasonable explanation for lexical similarities among the major families listed by GREENBERG? Shouldn’t we carefully investigate each proposed connection in more depth instead of bluntly ascribing all of them to ‘chance’ and ‘contact’ factors? No ‘bona fide’ review of this pioneer work has ever been published, thus leaving the Amerind problem unresolved and the evidence in favour of either “Amerind” as a whole or subdivisions of “Amerind” as separate families still waiting for a proper treatment.

6. *Evolution of Human Languages Project.*

In the light of this state of affairs, the main goal of *Evolution of Human Languages (EHL)*, has been stated as a thorough scientific investigation of degrees of genetic relationship among the world’s languages. Its principal unique feature is the creation of a comprehensive network of etymological data based on the following general principles:

¹⁵ For a detailed discussion of such methods, see [DYBO & STAROSTIN 2007].

¹⁶ The Amerind hypothesis does not follow the methodology advocated within this publication; instead, GREENBERG applies a method that eschews explicit intermediate reconstructions in favour of direct language comparison, even within Type C families. We consider this method useful at times for forming preliminary hypotheses on genetic relationship, which then have to be strengthened through a careful application of the comparative method as described above. This, however, does not mean that objective evaluation of GREENBERG’s method and results should focus strictly on their negative aspects.

- (i) we strongly believe that, in the highly computerized world of today, etymological data must be stored in electronic databases and not only in the old fashioned published format;
- (ij) etymological information should be easily and instantly available to every scholar, meaning that the databases should have public access;
- (iij) the best outcome is usually achieved when at least two scholars are working on the same topic, and active interaction between researchers specializing in different language families is of crucial importance as well;
- (iv) even the ‘raw’ and preliminary results of etymological research should be open for public inspection and suggestions for improvement (according to the ‘bona fide’ strategy), although creation and editing of the databases themselves should, of course, be restricted to specialists.

The general database network can be described as follows¹⁷:

1. It is a hierarchical structure, reflecting the step-by-step method of comparative reconstruction. Starting at the lowest levels (Germanic, Turkic, South Dravidian, &c.), the databases are then systematically linked to higher level ones, gradually progressing from A-type families to B- and C-type taxa.

2. All the databases follow one standard unified format (proto-form with approximate meaning; reflexes and meanings in daughter languages; commentaries and bibliography), although minor variations are possible.

3. The hyperlink system works in both directions, so it is possible to search for related forms in various distantly related modern languages, e. g. find potential cognates between Chinese and Chechen, or Finnish and Japanese (see below).

4. Efforts are undertaken to make each database as comprehensive as possible by incorporating all the lexical (and sometimes grammatical) material from the compared languages that is historically relevant. Naturally, the size and reliability of each database are directly related to the size and reliability of available lexical corpora, but the general principle of *EHL* is to preserve and make use of every bit of information. Most of the databases are available for public viewing (some even for downloading) at <http://starling.rinet.ru>. At the present time they mainly focus on linguistic families of Eurasia and North Africa, where four ‘superfamilies’ can be identified with varying degrees of reliability: Sino-Caucasian, Eurasiatic, Afroasiatic, and Austric. A brief perspective on each of them is given below.

Sino-Caucasian

The existence of this ‘superfamily’ was first scientifically motivated by S. STAROSTIN [STAROSTIN 1984]. His evidence was based on three lower level language families: North Caucasian, Sino-Tibetan, and Yeniseian. Later two additional languages — Basque and Burushaski, usually treated as isolates, were added to the list.

The first comprehensive etymological dictionary of North Caucasian (B-type family) was published in 1994 [NIKOLAYEV & STAROSTIN 1994]. Descriptive sources for all languages and major dialects of the family were used in a step-by-step reconstruction of the proto-language that generally meets the criteria listed above. The reconstruction, with 2759 entries, is supported by 8 subordinate databases representing A-type daughter families (Nakh, Lezghian, Abkhaz-Adyghe, &c.). The main point of the work, tying together Nakh-Daghestanian (East Caucasian) and Abkhaz-Adyghe (West Caucasian) languages, is still far from being universally accepted in Western linguistics, mainly due to insufficient scientific debate (apart from one or two very brief critical reviews, not a single detailed discussion of the dictionary has appeared in Western press up to now), but enjoys much support among the majority of Russian experts on historical Caucasian linguistics.

Due to a lack of descriptive data for most languages, the only existing etymological dictionary of Sino-Tibetan languages [PEIROS & STAROSTIN 1996] was compiled through systematic comparison of only five available languages: Old Chinese, Tibetan, Burmese, Kachin (Jingpaw), and Lushai; however, even on the basis of these five it is possible to establish detailed phonological correspondences and arrive at a reconstruction. After publication the work was continued by adding new etymologies from published sources and improving the correspondences. An updated electronic version of the dictionary (2823 entries) also contains two subordinate databases: comparative Kiranti and a substantial historical database of Chinese characters (more than 4,000 entries), both prepared by S. STAROSTIN.

¹⁷ The network is supported by the STARLing software package developed by S. STAROSTIN.

The Yeniseian etymological dictionary was published in [STAROSTIN 1995], following the phonological reconstruction in [STAROSTIN 1982]. An updated version, which includes additional information from [WERNER 2002], contains 1059 entries.

There are several publications on Sino-Caucasian, most of them belonging to S. STAROSTIN. The family, like most C-type ones, has not yet found significant recognition. In this particular case, besides the usual reasons, there is also the complexity of correspondences (North Caucasian languages, in particular, are among the most phonetically complex in the world) and lack of full-length publications. Just before his premature passing away in September 2005, however, S. STAROSTIN finished a book-long manuscript on Proto-Sino-Caucasian, including detailed discussion of its comparative phonology and lexicon, which will be published in the near future (and is already available online). So far, the etymological database contains 1361 entries.

Glottochronology suggests the following rough dates of splitting for Sino-Caucasian and its descendants:

Sino-Caucasian	10 KYA	Sino-Tibetan	6 KYA	Basque	modern
North Caucasian	6 KYA	Yeniseian	2 KYA	Burushaski	modern

It is also highly likely that the Na-Dene family of North America belongs to the Sino-Caucasian superfamily. This has been postulated with the help of modern comparative data by S. NIKOLAYEV [NIKOLAYEV 1991], and supported on the basis of “mass comparison” by J. BENGTSON and M. RUHLEN; today, work in this direction is mainly advanced by J. BENGTSON. However, no EL is currently available for Na-Dene, and further research is needed to verify this hypothesis. Occasional attempts to broaden the scope of the family by including extinct languages of Eurasia such as Sumerian or Etruscan, as well as other language families in the Americas (e. g. Salishan), have not been successful so far.

Eurasiatic

In the 1960s V. M. ILLICH-SVITYCH presented a set of arguments in favour of a genetic relationship among several families of Eurasia and North Africa; the proposed ‘superfamily’ was called Nostratic ([ILLICH-SVITYCH 1971–1984]). Evidence in favour of Nostratic includes regular correspondences between daughter proto-languages, large etymological lists that rely heavily on the basic lexicon, and a small inventory of reconstructed grammatical morphemes. Several decades later S. STAROSTIN has proposed that it is better to talk about two coordinate superfamilies: Eurasiatic and Afroasiatic (for the former the old name ‘Nostratic’ is sometimes retained, to avoid confusion with GREENBERG’s ‘Eurasiatic’ — a similar classification proposal, although based mainly on the “mass comparison” method, that excludes Dravidian and Kartvelian). The *EHL* etymological network includes several databases of Eurasiatic languages:

- the Altaic comparative dictionary, published in 2003; its electronic version contains the proto-language database (2805 entries) supported by five subordinate reconstructions: Turkic, Mongolian, Tungus-Manchu, Korean, and Japanese;
- the Eskimo database (1774 entries); according to O. MUDRAK [MUDRAK 1984], Eskimo forms part of Eurasiatic¹⁸ and within it is particularly close to Altaic. It should be mentioned, however, that no database on Aleut is available so far;
- a detailed database on comparative Dravidian, prepared by G. STAROSTIN (2211 entries), based primarily on etymological data available in [BURROW & EMENEAU 1984], but incorporating a modified variant of the traditional reconstruction;
- the Indo-European database by S. NIKOLAYEV is arguably the largest collection of proto-morphemes of this well-studied family (3178 entries). Two subordinate databases — Germanic and Baltic — were also compiled by the same scholar;
- the Uralic database (1898 entries) is loosely based on K. RÉDEI’s etymological dictionary [RÉDEI 1988] with significant additions from a team of Uralic linguistics scholars working in Moscow. No subordinate databases are available so far, but the daughter families are undergoing intensive study;
- the Kartvelian database (1310 entries) is based on G. KLIMOV’s etymological dictionary ([KLIMOV 1998]) with some additions by S. STAROSTIN that reflect his study of the external contacts of this family.

¹⁸ The Eurasiatic origin of the Eskaleut languages was also proposed by J. GREENBERG.

No other B- or A-type families can be demonstrated to belong to Eurasiatic (one possible hypothesis suggests that the Chukchi-Kamchatkan languages also form parts of it, but it is also possible that in reality they belong to a different superfamily; the whole issue needs very serious work). The combined Eurasiatic database brings together reliable etymologies both from already published sources and those newly discovered by members of the *EHL* team (altogether 2077 entries of varying quality, but including ‘core’ ones).

The age of Eurasiatic and its major descendants may be crudely estimated as:

Eurasiatic	12 KYA				
Altaic	8 KYA	Dravidian	5 KYA	Uralic	6 KYA
Eskimo	2 KYA	Indo-European	7 KYA	Kartvelian	4–3 KYA

Afroasiatic

Afroasiatic is arguably the only ‘superfamily’ that is generally recognized by mainstream linguistics, despite (or, perhaps, due to) the fact that its languages are actually far less studied than those of Eurasiatic. It is formed by five daughter families (Semitic, Berber, Chadic, Cushitic, and Omotic) and one isolate (Egyptian). None of these families except for Semitic can boast a comprehensive etymological dictionary, and even for Semitic with its lengthy tradition of study, existing ELs are either outdated [COHEN 1974] or, as of now, still incomplete [MILITAREV & KOGAN 2000–2003]. The situation with other branches of Afroasiatic is even less fortunate.

As a result, daughter families are represented differently in the Afroasiatic network of *EHL*. The Semitic database (compiled by A. MILITAREV) contains 2852 etymologies, each found in at least two languages of the family. Next to it is the Egyptian one (also compiled by MILITAREV) with 1495 entries. Other databases have about several hundred etymologies each.

The main Afroasiatic database consists of 3212 etymologies of varying quality. Its compilers, A. MILITAREV and O. STOLBOVA, brought together acceptable published etymologies (from dictionaries such as [OREL & STOLBOVA 1995]) and then enriched it with hundreds of newly discovered ones. Further research on the family in general is tightly connected with improving existing reconstructions for Cushitic, Omotic, and Chadic.

The ages of Afroasiatic and its descendants are roughly estimated as:

Afroasiatic	12 KYA	Cushitic	9 KYA	Semitic	7 KYA
Omotic	7 KYA	Chadic	7 KYA	Berber	3 KYA

Austric

The existence of the Austric superfamily is less clearly shown than that of the others. In contrast to Eurasiatic, Sino-Caucasian, and Afroasiatic, no detailed proto-Austric glossaries or equally detailed tables of correspondences between the various daughter branches of Austric have been produced. The number of these proposed daughter branches is four, and they are grouped in pairs: Austronesian / Tai-Kadai and Austroasiatic / Miao-Yao¹⁹. Until recently none of these daughter families was represented by a comprehensive etymological dictionary. In terms of known etymologies Austronesian fares significantly better than the others, reflecting a fairly long tradition of scholarship. The work, however, is mostly on languages of Western Indonesia, the Philippines, and parts of Eastern Oceania. Other territories are underrepresented, which often makes it difficult to attribute an etymology to the appropriate chronological level. An *EHL* Austronesian database is still under construction.

The situation with Tai-Kadai is slightly different. A representative collection of etymologies for one of its principal branches, Zhuang-Tai, has been published [LI 1977] and is available (with a few modifications) on the *EHL* network (1329 entries). A limited number of common Tai-Kadai forms (taken mainly from [PEIROS 1998]) are also available electronically.

The Austroasiatic family is represented much better. A preliminary version of its comparative dictionary by I. PEIROS contains 2457 entries, identified mainly through comparison of ten lower level families that include about all of Austroasiatic²⁰.

¹⁹ See [PEIROS 1998].

²⁰ Etymologies from *Austroasiatic Etymological Dictionary* by H. SHORTO have not yet been included.

For Miao-Yao no etymological dictionaries or convincing phonological reconstructions are available; as a result, no corresponding database is found in the *EHL* network either. All the Miao-Yao proto-forms used in Austric comparisons are taken from [PEIROS 1998].

A systematic search for Austric etymologies has not yet been conducted. In 2004–2005 S. STAROSTIN and I. PEIROS collected some lexical similarities in support of Austric, included in the highly provisional ‘Austric’ database. The 900 proposed etymologies represent mainly lexical similarities between Austroasiatic and Austronesian, although they are actually supported by basic (relatively simple) phonological correspondences. Thus, we can say that our knowledge of Austric is much more limited than that of other superfamilies; however, the discovery of a number of comparanda that fit the ‘core etymology’ requirements makes the Austric hypothesis quite plausible.

The approximate age of Austric and its descendants may be estimated thus:

Austric	10 KYA	Austronesian	5 KYA ²¹	Austroasiatic	7 KYA
		Tai-Kadai	5 KYA	Miao-Yao	4 KYA

“*Borean*”?

The *EHL* network of databases presents what we consider extremely strong evidence in support of the four superfamilies discussed above. Two of them — Sino-Caucasian and Eurasiatic — are based on reconstructions performed in full accordance with the comparative method, with regular phonetic correspondences established between reconstructed intermediate proto-languages. Knowledge of the other two — Afroasiatic and Austric — has not so far reached the same level of confidence; however, since it is possible to discuss the two theories in terms of regular phonetic correspondences rather than mere similarities, both can be considered scientific, and further work on them is promising.

Since Afroasiatic and Austric are still to be reconstructed and the other two superfamilies are still in need of serious improvements, it is not yet possible to apply strict comparative methods of investigation to even deeper chronological levels. However, the obstacles here are technical rather than theoretical. The widespread idea that comparative research has an impassable threshold of about 10 KYA (such a period of time is claimed to be enough to make related languages lose all relevant similarity) does not take into account the fact that the main objects of research in this case are not modern languages, but reconstructed proto-languages which turn out to be more similar to one another than their modern day descendants. Thus, solid reconstructions for C-type families like Eurasiatic, Sino-Caucasian, and others should eventually help²².

At the present time it is possible to discuss such ‘ultra-deep’ relationships only on a very speculative level. Numerous morphemic similarities between various language families of Eurasia have already been spotted in the past as potential indication of such a relationship; many, if not most, of these similar forms (traced back to high level reconstructions) were compiled by S. STAROSTIN into a special database and later supplemented by some of his own findings. Since such morphemic comparisons are rather numerous (several hundred at least), chance resemblances are not very probable, and a ‘Borean’ super-superfamily hypothesis, open to ‘bona fide’ discussion, has been formulated²³. Statistical analysis of attested similarities shows that if such a taxon really exists, its initial division was as follows:

- (i) Eurasiatic and Afroasiatic (= ILLICH-SVITYCH’s ‘Nostratic’);
- (ij) Sino-Caucasian
- (iij) Austric.

The estimated age of ‘Borean’ would be around 15–17 KYA.

²¹ Our calculations exclude Austronesian languages of New Guinea and some surrounding islands.

²² It should be noted that the amount of information recoverable for protolanguage states always remains in direct proportion to the number of languages used for comparison. Since lexical loss and replacement, in most cases, occurs independently in daughter languages, the probability of any given morpheme disappearing without a trace in three of them is less than in only two, and so on. This ensures that, given a sufficient number of languages or language branches, the morphemic inventory of the reconstructed ancestral language will be just as large (sometimes even larger!) than that of any single one of its descendants. Such is the case for commonly accepted families like Indo-European, Turkic, Semitic, etc., and there is little reason to doubt that a different situation has to be proposed for Type C families.

²³ The term was originally used by H. FLEMING for a somewhat different linguistic entity [FLEMING 1991].

Other families

The ‘Borean’ hypothesis currently links together the 4 superfamilies described above. At the same time, the real scope of ‘Borean’ remains unknown, since we still lack deep level reconstructions of many families in the American, African, and Indo-Pacific, and Australia. This means that we have no certain means of verifying whether some of these families can also form part of the hypothetical ‘Borean’.

So far, preliminary research has been carried out for Khoisan in South Africa, resulting in a classification and a set of provisional reconstructions by G. STAROSTIN [2003, 2008], whose Khoisan databases are already incorporated into the *EHL* etymological network. Comparison with the ‘Borean’ data has not produced any conclusive results, suggesting that Khoisan, at least, cannot be included in ‘Borean’, although genetic connections on an even deeper level might be possible.

The preliminary lexicostatistical study of Sub-Saharan African languages identifies at least 4 other superfamilies: Niger-Congo (not quite identical to the Niger-Congo super-family proposed by J. GREENBERG), East Sudanic, Central Sudanic, and Kordofanian, plus a number of smaller branches whose position is so far very unclear (such as Songhay or Atlantic languages). We still need to investigate how these superfamilies are connected to each other and to Borean. This can be done, however, only through an extensive etymological study of available data.

The situation with native languages of the Americas is different, and still far from being resolved. According to J. GREENBERG, they can all be classified into two small stocks (Eskimo-Aleut, Na-Dene) and one huge super-family — Amerind. We have significant evidence that the Eskimo-Aleut family is part of Eurasiatic, while Na-Dene seems to be related to Sino-Caucasian languages (although in the latter case a completely convincing demonstration is still lacking).

The status of the Amerind proposal remains unclear. The main source for lexical lookalikes between these languages remains J. GREENBERG’s monograph. We have already mentioned that it was and still is heavily criticized, sometimes for good reasons; this, however, does not eliminate the problem itself — why exactly are there so many lexical similarities found not just between some of the proposed daughter families, but frequently between all the families of the hypothetical Amerind superfamily? Even if we rule out everything that is ‘non-core’ (i. e. forms with scarce distribution, far-flung semantics, &c.), explaining all the rest away as chance resemblances would simply be closing our eyes on the problem. This leaves us with two options: either the similarities are of common genetic origin, or they result from intensive language contact, which is a less probable option given the peculiarities of their dispersal.

A very preliminary proposal based on *EHL* exploratory studies suggests that in the Americas one can find at least three types of grouping:

(i) The Almosan superfamily (Algic, Salishan, Wakashan, and some other languages) might be related to Chukchee and Nivkh languages of North Asia, forming a so-called “Beringian” superfamily; connections with “Borean” have been noticed as well²⁴.

(ij) A number of families (Penutian, Hokan, Mayan, Mixe-Zoque, Maipuran, Pano-Tukanoan, &c.) presumably form a different superfamily, also with resemblances to “Borean”²⁵.

(iij) At the same time, we were not able to detect any external relations for such well-established families as Siouan, Gulf, or Otomanguean.

Distant relationships among the Papuan (= Non-Austronesian) languages of New Guinea and aboriginal languages of Australia remain to be investigated. It is possible that in that region we could distinguish up to 4–6 superfamilies (Trans-New-Guinea, Australian, East Papuan, &c.), none being properly reconstructed. Some lexical similarities have also been spotted between Trans-New-Guinea morphemes and some of the alleged ‘Borean’ roots, but these remain too scarce to establish a firm connection.

7. The ‘bottleneck’ scenario.

Anatomically modern human beings seem to have evolved in Africa around two hundred thousand years ago ([ALEMSEGED, COPPENS, & GERAADS 2002]). It is not known for certain when they acquired language of the modern type. The second wave of migration of anatomically modern human beings out of Africa, according to genetic and archaeological data, seems to have taken place around 60 to 50 KYA. Compared with the previous known wave (which occurred of the order of a hundred and thirty KYA and is known to have reached as far as Palestine) it was definitely more successful, populating Eurasia and the Indo-Pacific region, including Australia. In Western Europe it gave rise to the Aurignacian culture,

²⁴ Proposed by S. NIKOLAEV, who is now working on a detailed justification.

²⁵ Proposed by S. NIKOLAEV and I. PEIROS in their survey of existing linguistic classifications.

with its remarkable paintings, engravings, and sculpture in addition to the widespread Upper Paleolithic tools, suggesting that the human beings of the second wave were behaviorally as well as anatomically modern. That makes it likely that they already possessed a language or languages of the modern type.

For tens of thousands of years afterwards, the usual kinds of linguistic transformation presumably took place, producing daughter languages, which themselves gave rise to daughter languages, and so forth. In that way a considerable degree of linguistic diversity would have been achieved. However, that amount of diversity need not necessarily be reflected in the diversity of attested languages. Instead, we may be dealing here with a “bottleneck” effect, in which a great many languages (but not necessary all of them) descend from a single ancestor.

The climatic changes near the height of the last Ice Age some twenty thousand years ago shrank drastically the territories suitable for human habitation, with ice caps and deserts occupying a large fraction of the land mass. We may picture the human beings of that time confined to refugia often separated by hostile areas. Under those conditions linguistic diversity could have been greatly reduced and it may therefore be the case that all or most of the languages of subsequent times are descended from a single ancestor, the tongue of a particular refugium. If the similarities of attested languages are found to suggest a common origin for all or most of them, that origin could well be a speech that survived the height of the Ice Age when most others did not. With the improvement of climatic conditions, humans began to move out of their refugia, colonizing territories previously unsuitable for permanent occupation. This led to growth and subsequent division of their communities, resulting in the development of new languages.

This is a different story from that of “monogenesis”, according to which the latest common ancestor of all or most attested languages would be the earliest human language of the modern type. According to the bottleneck scheme, by contrast, all or most of the diversity of attested languages would have developed over some twenty thousand instead of fifty thousand years, making it somewhat more plausible that one might discover evidence of common descent if such common descent is actually correct. In addition, the bottleneck idea allows the age of modern language to be pushed back to any time between fifty thousand years ago and two hundred thousand years ago, when anatomically modern humans appeared.

Suppose it is true, as mentioned above, that an enlarged Borean taxon embraces most of the African superfamilies as well as the “Amerind” languages, then the age of such a “supersuperfamily” would agree very roughly with the 20 000 years we attribute to the bottleneck. If some or all of the ‘global roots’, found across most of the world’s super-families²⁶, are genuine, then they could relate to such a time horizon. Another feature of the bottleneck scheme with a protolanguage less than 20 000 years old is that the migrations of the speakers of the descendant languages (such as Eurasiatic, Afroasiatic, &c.) need have nothing to do with the ‘out of Africa’ notion which refers to much earlier times.

The phenomenon of linguistic bottleneck is encountered on smaller scales, as in the case of the Australian aboriginal languages. It is believed that humans reached the continent at least 40 KYA. Yet it is generally agreed that all or nearly all of the attested Australian languages form a single superfamily and the similarities among them make it extremely difficult to believe that the proto-language of that superfamily could be older than twelve thousand years or so²⁷. It seems that a single language or the descendants of a single language spread over all or nearly all of Australia at that comparatively recent date. It could have been a local language of Australia or of New Guinea; if we had to guess, we would probably choose the latter, on the basis of some lexical similarities. In any case, one can search for genetic or cultural traits that might have been introduced along with the language in question.

A very familiar example of a bottleneck is the domination of Europe by Indo-European languages. Here most of the currently spoken European languages, with but a few exceptions like Basque (whose native speakers constitute a small linguo-geographic “refugium”), are descended from a single language spoken, probably in Southern Russia, some six or seven thousand years ago. Since we are used to that idea, we should be able to entertain the possibility of bottlenecks having occurred on much wider scales.

²⁶ For more on ‘global roots’, see e. g. [RUHLEN 1994]. Many of the particular connections proposed by Ruhlen, as well as other researchers, have been heavily — and often justly — criticized (e. g. by D. RINGE, L. CAMPBELL, and others), yet a decisive statistical demonstration that would once and for all reject all of the accumulated evidence as non-evidence is still missing. We apply the same cautionary approach to the issue of ‘global roots’ as we do towards GREENBERG’S ‘Amerind’ and all the other theories arrived at through the ‘mass comparison’ method: all of the comparative data obtained that way are valid as material for further research, to be gradually accepted or discarded as the families in question are subjected to standard comparative analysis.

²⁷ I. PEIROS, in his unpublished lexicostatistical study of aboriginal Australian languages, estimates the age of Proto-Australian as constituting around 10,000 to 12,000 years.

8. Conclusion

The data accumulated within the *EHL* etymological network vary in quality and ‘convincing force’. The program, however, manages to collate the results of half a century’s research on distant language relationships (so far, mostly within Eurasia) with results recently achieved by the *EHL* team which has the benefits of better data access, methodology that is modernized (but still firmly rooted in tradition), and computer handling of data. Much remains to be done, but even now certain prehistoric scenarios of linguistic development within the last 20 000 years can be drawn up, with the most probable one looking as follows:

- (i) At the height of the Ice Age humans were forced to take refuge in one or several zones suitable for survival, causing a decrease of the linguistic diversity that presumably existed before.
- (ij) With the improvement of climatic conditions, humans began to move out of their refugia, colonizing territories previously unsuitable for permanent occupation. This led to growth and subsequent division of their communities, resulting in the development of new languages.
- (ijj) In the process of spreading various linguistic groups suffered different fates; some disappeared with or without any traces, while others expanded, spreading their languages over vast territories or shifting from one language to another²⁸.
- (iv) One of the most successful survivors of the Ice Age may be a hypothetical ‘Borean’ super-superfamily, whose age is estimated as 15–17 KYA. Inconclusive, but significant evidence for ‘Borean’ is provided by preliminary comparison of four super-families, the historical reality of which is no longer questioned by *EHL* members: Eurasiatic, Afroasiatic, Sino-Caucasian and Austric.
- (v) Preliminary data also indicate possible connections between Borean and some superfamilies of Africa, America, and the Indo-Pacific region, not included in the four superfamilies mentioned above. Further research into distant relationships of languages is needed to find out whether these additional superfamilies are related to ‘Borean’ on a higher level or are hitherto unidentified branches of ‘Borean’. The question of the original ‘Borean’ homeland also remains open.

Further research into distant relationships of languages is needed to find out whether there are other Ice Age survivors that are related to ‘Borean’ on a higher level or turn out to be potentially undiscovered members of the family itself. The question of the original ‘Borean’ homeland also remains open.

A c k n o w l e d g m e n t s

We would like to thank all participants of the *EHL* program for providing data and sharing their knowledge with us. We also gratefully acknowledge the generous support for the *EHL* project on the part of Mr. Jerry MURDOCK and Bryan J. and June B. ZWAN Foundation.

R e f e r e n c e s

- ALEMSEGED, COPPENS, & GERAADS 2002 — Z. ALEMSEGED, Y. COPPENS, and D. GERAADS. Hominid Cranium from Omo: Description and Taxonomy of Omo-323-1976-896 // *American Journal of Physical Anthropology*. 117 (2).
- BERGSLAND & VOGT 1962 — K. BERGSLAND and H. VOGT. On the Validity of Glottochronology // *Current Anthropology*. 3; pp. 115–153.
- BLUST 2000 — R. BLUST. Why Lexicostatistics Doesn’t Work: the ‘Universal Constant’ Hypothesis and the Austronesian Languages // *Time Depth in Historical Linguistics*. Ed. by Colin RENFREW, April McMAHON & Larry TRASK. McDONALD Institute for Archaeological Research, Cambridge; pp. 311–332.
- BURLAK & STAROSTIN 2005 — С. А. БУРЛАК, С. А. СТАРОСТИН. *Сравнительно-историческое языкознание [Comparative-Historical Linguistics]*. 2-е изд. М. [Second edition. Moscow]: “Academia”.
- BURROW & EMENEAU 1984 — T. BURROW and M. B. EMENEAU. *A Dravidian Etymological Dictionary*. Second Edition. Oxford Publishers.
- CALDWELL 1856 — R. CALDWELL. *A Comparative Grammar Of The Dravidian or South Indian Family of Languages*, 1st ed. London.
- CAMPBELL 1997 — L. CAMPBELL. *American Indian Languages: the Historical Linguistics of Native America*. Oxford University Press, New York.

²⁸ Technically, language shift is a kind of language loss.

- CAMPBELL 2004 — L. CAMPBELL. *Historical Linguistics: an Introduction*. 2nd edition. Edinburgh: Edinburgh University Press, Cambridge MA: MIT Press.
- COHEN 1974 — D. COHEN. *Dictionnaire des racines sémitiques ou attestées dans les langues sémitiques*. Paris, MOUTON, La Haye.
- CROWLEY & DIXON 1981 — T. CROWLEY and R. M. W. DIXON. *Tasmanian // Handbook of Australian Languages*, Vol. 2. Ed. by R. M. W. DIXON & B. J. BLAKE. Canberra, Australian National University Press.
- DIMMENDAAL 1995 — G. J. DIMMENDAAL. Do Some Languages Have a Multi-Genetic or Non-Genetic Origin? An Exercise in Taxonomy // *Cinquième Colloque de Linguistique Nilo-Saharienne*. Éd. par Robert NICOLAÏ et Franz ROTTLAND. Köln: Rüdiger KÖPPE Verlag; pp. 357–372.
- DYBO & STAROSTIN 2008 — A. V. DYBO and G. S. STAROSTIN. In Defense of the Comparative Method, or the End of the Vovin Controversy // *Аспекты компаративистики*. [т.] 3 [*Aspects of Comparative Linguistics*. [v.] 3 / *Orientalia et Classica: Труды Института восточных культур и античности*. Вып. XIX. М.: Российский государственный гуманитарный университет [Papers of the Institute of Oriental and Classical Studies. Issue XIX. Moscow: RSUH Publishers]; pp. 119–258.
- DYEN 1965 — I. DYEN. A Lexicostatistical Classification of the Austronesian Languages // *International Journal of American Linguistics*. Memoir 19. Bloomington, Indiana.
- FLEMING 1991 — H. C. FLEMING. A New Taxonomic Hypothesis: Borean or Boralean // *Mother Tongue Newsletter*. V. 14.
- FOLEY 2005 — W. A. FOLEY. Linguistic Prehistory in the Sepik-Ramu Basin // *Papuan Pasts: Cultural, Linguistic and Biological Histories of Papuan-Speaking Peoples*. Ed. by Andrew PAWLEY, Robert ATTENBOROUGH, Robin HIDE and Jack GOLSON. Canberra: Pacific Linguistics; pp. 109–144.
- GORDON 2005 — R. G. Jr. GORDON (ed.). *Ethnologue: Languages of the World*. Fifteenth edition. Dallas, Tex.: SIL International.
- GREENBERG 1966 — J. GREENBERG. *The Languages of Africa*. Indiana University, Bloomington. MOUTON & C^o., The Hague, Netherlands.
- GREENBERG 1987 — J. GREENBERG. *Language in the Americas*. Stanford University Press, Stanford.
- HOCK & JOSEPH 1996 — Hans H. HOCK and Brian D. JOSEPH. *Language History, Language Change, and Language Relationship: An Introduction to Historical and Comparative Linguistics*. Berlin & New York: MOUTON DE GRUYTER.
- ILICH-SVITYCH 1971–1984 — В. М. Иллич-Свитыч. *Опыт сравнения ностратических языков (семитохамитский, картвельский, индоевропейский, уральский, дравидийский, алтайский): Введение. Сравнительный словарь*. М.: «Наука», Главная редакция восточной литературы [An Attempt at Comparative Dictionary of the Nostratic Languages (Semitic-Hamitic, Kartvelian, Indo-European, Uralic, Dravidian, Altaic). Moscow: “Nauka” publishers]. V. 1 (b–K): 1971; V. 2 (l–j): 1976; V. 3 (p–q): 1984.
- KLIMOV 1998 — G. A. KLIMOV. *Etymological Dictionary of the Kartvelian Languages*. Berlin: MOUTON DE GRUYTER.
- LI 1977 — LI FANG-KUEI [李方桂, Lǐ Fāngguì]. *A handbook of Comparative Tai*. Honolulu: University Press of Hawai’i.
- MARCANTONIO 2002 — A. MARCANTONIO. *The Uralic Language Family: Facts, Myths and Statistics*. Publications of the Philological Society. Cambridge.
- MATTESON 1972 — E. MATTESON. Proto Arawakan // *Comparative Studies in Amerindian Languages*. Ed. by Esther MATTESON. *Janua Linguarum, series practica*, 127. The Hague: MOUTON; pp. 160–242.
- MILITAREV & KOGAN 2000–2003 — A. MILITAREV and L. KOGAN. *Semitic Etymological Dictionary*. Münster: Ugarit Verlag. Volume 1: 2000. Volume 2: 2003.
- MUDRAK 1984 — О. А. МУДРАК. К вопросу о внешних связях эскимосских языков [On the External Relations of Eskimo Languages] // *Лингвистическая реконструкция и древнейшая история Востока*. Сб. тез. М.: ИВ АН СССР [Linguistic Reconstruction and the Prehistory of the East. Moscow, Institute of Oriental Studies]; pp. 80–87.
- NIKOLAEYV 1991 — S. L. NIKOLAYEV. Sino-Caucasian Languages in America // *Dene-Sino-Caucasian Languages*. Ed. by V. SHEVOROSHKIN. Bochum; pp. 42–66.
- NIKOLAYEV & STAROSTIN 1994 — S. L. NIKOLAYEV and S. A. STAROSTIN. *A North Caucasian Etymological Dictionary*. Moscow, “Asterisk” publishers.
- OREL & STOLBOVA 1995 — V. OREL and O. STOLBOVA. *Hamito-Semitic Etymological Dictionary: Materials for a Reconstruction*. Leiden & New York.
- PAYNE 1991 — D. L. PAYNE. A Classification of Maipuran (Arawakan) Languages Based on Shared Lexical Retentions // *Handbook of Amazonian Languages*, Vol. 3. Ed. by Desmond C. DERBYSHIRE and Geoffrey K. PULLUM. Berlin: MOUTON DE GRUYTER; pp. 355–499.

- PEIROS 1997 — I. PEIROS. Macro Families: Can a Mistake Be Detected? // *Indo-European, Nostratic, and Beyond: Festschrift for Vitalij V. SHEVOROSHKIN* / Ed. by Irén HEGEDŰS, Peter A. MICHALOVE and Alexis MANASTER RAMER. JIES Monograph No. 22. Institute for the Study of Man, Washington D.C.; pp. 265–292.
- PEIROS 1998 — I. PEIROS. Comparative Linguistics in Southeast Asia // *Pacific Linguistics Series C-142*. Canberra, Australian National University.
- PEIROS 2000 — I. PEIROS. Family Diversity and Time Depth // *Time Depth in Historical Linguistics*. Ed. by Colin RENFREW, April McMAHON & Larry TRASK. McDONALD Institute for Archaeological Research, Cambridge; pp. 75–108.
- PEIROS & STAROSTIN 1996 — I. PEIROS and S. STAROSTIN. *A Comparative Vocabulary of Five Sino-Tibetan Languages* (6 vols.). Melbourne.
- POKORNY 1958 — J. POKORNY. *Indogermanisches etymologisches Wörterbuch*. 2 Bände. FRANCKE Verlag, Bern und München (Bern–Stuttgart 1989, 2. Aufl.; Tübingen–Basel 1994, 3. Aufl.).
- RÉDEI 1988 — K. RÉDEI. *Uralisches Etymologisches Wörterbuch*. Akadémiai Kiadó, Budapest.
- SANDS 1998 — B. SANDS. Eastern and Southern African Khoisan // *Evaluating Claims in Distant Linguistic Relationships*. Ed. by R. VOSSEN. *Quellen zur Khoisan-Forschung / Research in Khoisan Studies*, Bd 14. Hamburg: Rüdiger KÖPPE.
- STAROSTIN 1982 — С. А. СТАРОСТИН. Праенисейская реконструкция и внешние связи енисейских языков [Proto-Yeniseian Reconstruction and the External Relations of Yeniseian Languages] // *Studia Ketica*, vol. 3. Leningrad: “Nauka” publishers; pp. 144–237.
- STAROSTIN 1984 — С. А. СТАРОСТИН. Гипотеза о генетических связях сино-тибетских языков с енисейскими и северокавказскими языками [A Hypothesis about the Genetic Connections between Sino-Tibetan, Yeniseian, and North Caucasian Languages] // *Лингвистическая реконструкция и древнейшая история Востока*. Сб. тез. М.: ИВ АН СССР [*Linguistic Reconstruction and the Prehistory of the East*. Moscow: Institute of Oriental Studies]; pp. 19–38.
- STAROSTIN 1995 — С. А. СТАРОСТИН. Сравнительный словарь енисейских языков [A Comparative Dictionary of Yeniseian Languages] // *Studia Ketica*, vol. 4. М.: «Языки Русской Культуры» [Moscow: “Languages of Russian Culture”]; pp. 176–315.
- STAROSTIN 2000 — S. A. STAROSTIN. Comparative-Historical Linguistics and Lexicostatistics // *Time Depth in Historical Linguistics* / Ed. by Colin RENFREW, April McMAHON & Larry TRASK. McDONALD Institute for Archaeological Research, Cambridge; pp. 223–259.
- STAROSTIN 2007 — С. А. СТАРОСТИН. Определение устойчивости базисной лексики [Definition of the stability of the basic lexicon] // С. А. СТАРОСТИН. *Труды по языкознанию*. М.: «Языки Славянской Культуры» [*Works on Linguistics*. Moscow: “Languages of Slavic Culture”]; pp. 827–839.
- STAROSTIN 2003 — G. S. STAROSTIN. A Lexicostatistical Approach towards Reconstructing Proto-Khoisan // *Mother Tongue*. VIII; pp. 83–128 [“81–126” in the on-cover TOC].
- STAROSTIN 2008 — G. S. STAROSTIN. From Modern Khoisan Languages to Proto-Khoisan: the Value of Intermediate Reconstructions // *Аспекты компаративистики*. [Т.] 3 [*Aspects of Comparative Linguistics*. [v.] 3] (= *Orientalia et Classica*: Труды Института восточных культур и античности. Вып. XIX. М.: Российский государственный гуманитарный университет [Papers of the Institute of Oriental and Classical Studies. Issue XIX. Moscow: RSUH Publishers]); pp. 337–470.
- VASILYEV & MILITAREV 2008 — М. Е. ВАСИЛЬЕВ, А. Ю. МИЛИТАРЁВ. Глоттохронология в сравнительно-историческом языкознании. Модели дивергенции языков [Glottochronology in Comparative-Historical Linguistics and the Models of Linguistic Divergence] // *Аспекты компаративистики*. [Т.] 3 [*Aspects of Comparative Linguistics*. [v.] 3] (= *Orientalia et Classica*: Труды Института восточных культур и античности. Вып. XIX. М.: Российский государственный гуманитарный университет [Papers of the Institute of Oriental and Classical Studies. Issue XIX. Moscow: RSUH Publishers]); pp. 509–536.
- WERNER 2002 — H. WERNER. *Vergleichendes Wörterbuch der Jenissej-Sprachen*. Band 1: A–K, Band 2: L–S, Band 3: *Onomastik*. Wiesbaden, Germany: HARRASSOWITZ Verlag.

Р е з ю м е

Убеждение, что данные современных языков могут быть использованы для реконструкции праязыков с возрастом в 10 000 и более лет, сегодня разделяется далеко не всеми специалистами. Данная статья, частично основанная на результатах совместных исследований российских и американских участников проекта Института Санта Фе «Эволюция языков», посвящена обсуждению методологии макрокомпаративистики и различным способам преодоления связанных с ней теоретических препятствий. Во второй части статьи подводятся основные итоги деятельности проекта, а также предлагаются различные направления дальнейших исследований в области глубинного языкового родства и лингвистической предыстории человечества.

Аблаут в глагольных основах нахско-дагестанских языков

В статье проводится сопоставительный анализ рядов вокалических чередований (аблаута), обнаруживаемых в глагольных системах языковых групп нахско-дагестанской семьи. Показано, что разнообразие этих рядов, в той или иной степени присущее всем этим группам, не позволяет, как это часто принято, трактовать нахско-дагестанские глагольные корни как моноконсонантные. Также показано, что вокалические чередования такого рода, как правило, не удастся объяснить исходя из внутренней истории той или иной подгруппы, и что, следовательно, в какой-то форме глагольный аблаут должен реконструироваться и для пра-нахско-дагестанского, хотя конкретная реконструкция пока остается делом будущего.

Традиция изучения дагестанских языков характеризуется явным пренебрежением к вокалической составляющей корневой морфемы. Как известно, здесь долгое время господствовала и продолжает существовать и по сей день теория моноконсонантизма, принимаемая исследователями в качестве аксиомы. Так, в одной из последних по времени выхода работ на сравнительно-историческую тематику читаем: «теория корня в дагестанских языках является одной из сложных, что обусловлено моноконсонантным характером корневой морфемы. В восточнокавказских языках увеличение числа моноконсонантных реконструкций корня происходит по мере продвижения от современного состояния к более отдаленной перспективе» [Сулейманов 2000: 27].

Хотя основные положения этой теории увязываются с работами А. С. Чикобава, главный ее тезис восходит, по-видимому, к Н. С. Трубецкому, которому принадлежит высказывание о том, что «правосточнокавказские глагольные корни состояли в большинстве своем из одного-единственного шумного согласного» [Трубецкой 1987: 341]. Между тем трудно не обратить внимания на то, каким образом интерпретировал исследователь роль гласных в составе глагольной основы: по его мнению, с помощью специфических для разных групп глаголов чередований они передавали видо-временные значения. В соответствии с такой позицией гласные не должны исключаться ни из сравнительно-исторических, ни из синхронных описаний.

Глагольный аблаут представлен в современных нахско-дагестанских языках довольно неравномерно. Достаточно ярко он проявляется, например, в нахских языках. Здесь, как известно, выделяется несколько разновидностей спряжения в зависимости от способов противопоставления глагольных основ совершенного (однократного) и несовершенного (многократного) вида. В чеченском языке выделяется десять таких разновидностей, противопоставляющих три ступени чередования [Дешериева 1979: 83–84], ср. чеч. (формы инфинитива, настоящего и близкого прошедшего времени):

1-й тип:	$\bar{a} — \bar{o} — \bar{e}$:	$\bar{a}ла(н)$ ‘сказать’ — $\bar{o}лу$ — $\bar{e}ли(н)$;
2-й тип:	$a — o — e$:	$таса(н)$ ‘бросить’ — $тосу$ — $теси(н)$;
3-й тип:	$\bar{e} — \bar{o}ь // уоь — \bar{y}й // e$:	$д\bar{e}ша(н)$ ‘читать’ — $доьшу$ — $д\bar{y}йши(н)$;
4-й тип:	$e — оь — и$:	$верза(н)$ ‘выздоровливать’ — $воьрзу$ — $вирзи(н)$;
5-й тип:	$ий — уьй — ий$:	$лийча(н)$ ‘купаться’ — $луьйчзу$ — $лийчи(н)$;
6-й тип:	$и — уь — и$:	$тилла(н)$ ‘класть’ — $туьллу$ — $тилли(н)$;
7-й тип:	$o — y — уьй // оь$:	$тоха(н)$ ‘бить’ — $туху$ — $туьйхи(н)$;
8-й тип:	$o — y — уь$:	$кхолла(н)$ ‘творить’ — $кхуллу$ — $кхуьлли(н)$;
9-й тип:	$y — y — уьй // оь$:	$уза(н)$ ‘тянуть’ — $узу$ — $оьзна$;
10-й тип:	$y — y — уь$:	$дуза(н)$ ‘наполнять’ — $дузу$ — $дуьзи(н)$.

Хотя очевидно, что некоторые разновидности легко объединяются (например, с одной стороны, ряды первый и второй и, с другой стороны, третий и четвертый, отличающиеся лишь долготой гласного, или седьмой и девятый ряды, а также восьмой и десятый, отличающиеся качеством базового гласного), разнообразие чередований налицо (отметим также не вошедшие в представленные выше типы неправильные глаголы).

В ингушском языке, по мнению Р. И. Долаковой [1967: 219], имеется девять типов аблаута, ср.:

1-й тип:	$\bar{a} — оа — аь$:	$\bar{a}ха$ ‘пахать’ — $оах$ — $аьхар$;
2-й тип:	$\bar{a} — а — аь$:	$к\bar{a}га$ ‘блеснуть’ — $к\bar{a}г$ — $к\bar{a}ьгар$;

3-й тип:	<i>a — o — e:</i>	<i>мага</i> 'мочь' — <i>мог</i> — <i>мегар</i> ;
4-й тип:	<i>a — a — e:</i>	<i>дарста</i> 'полнеть' — <i>дарст</i> — <i>дерстар</i> ;
5-й тип:	<i>ē — e — ий:</i>	<i>дѣкъа</i> 'делить' — <i>дѣкъ</i> — <i>дийкъар</i> ;
6-й тип:	<i>o — o — e:</i>	<i>тоха</i> 'бить' — <i>тох</i> — <i>техар</i> ;
7-й тип:	<i>ов — ов — ай:</i>	<i>ловза</i> 'бить' — <i>ловз</i> — <i>лайзар</i> ;
8-й тип:	<i>ӯв — ув — ий:</i>	<i>лӯвча</i> 'блеснуть' — <i>лувч</i> — <i>лийчар</i> ;
9-й тип:	<i>и — у — и:</i>	<i>дита</i> 'оставлять' — <i>дут</i> — <i>дитар</i> .

В бацбийском насчитывается четыре типа аблаута *a — u*, *a — e*, *o — e* и *u — e* [ЧРЕЛАШВИЛИ 1999: 199], ср.: *вахѣ* 'пошел' — *ихѣ* 'шел', *хатлмлї* 'спросил' — *хетлмлї* 'спрашивал', *дохкї* 'продал' — *дехкї* 'продавал', *тумлѣ* 'срезал' — *темлѣ* 'резал' и др.

В специальной литературе была высказана мысль о вторичности аблаутных чередований в нахском глаголе: «... для образования дюративного аспекта от моментного был использован префикс *i*, а современная система образования аспекта — результат слияния означенного гласного с гласным моментного аспекта глагола. Следует добавить, что в дальнейшем к оформленным по аспекту формам присоединились так называемые обстоятельственные префиксы (*q-*, *y-*, *h²-*, *s-*) и экспоненты грамматических классов, образующие новые глагольные основы. Таким образом, образование аспекта в нахских языках усложнилось... Заключаем, в образовании аспекта в нахских языках аблаут не имел места» [ИМНАЙШВИЛИ 1977: 134]. Тем не менее мнение исследователей склоняется в пользу первичности аблаута в этой языковой группе. В «Северокавказском этимологическом словаре» [НИКОЛАЕВ & STAROSTIN 1994: 101] была предложена следующая схема типов общенахского аблаута:

Сов. вид	Несов. вид	Сов. вид	Несов. вид
* <i>e:</i>	* <i>a:</i>	* <i>a:</i>	* <i>e:</i>
* <i>a</i>	* <i>e:</i>	* <i>o:</i>	* <i>e:</i>

Рассматривая взаимоотношения нахских и дагестанских языков, Б. К. Гигинейшвили выдвигал глагольный аблаут в качестве одного из показателей специфики нахских языков: «Это характерное для морфологии нахского глагола чередование корневых гласных совершенно чуждо дагестанским языкам как на нынешнем этапе их развития, так и исторически» [Гигинейшвили 1977: 24].

Между тем данные дагестанских языков позволяют усомниться в подобном суждении. Наиболее интересны в этом отношении данные даргинского языка, где также имеется противопоставление однократного и неоднократного видов, выражаемое чередованием гласных, ср.:

арсес 'полететь' — *урсес*, *бѣбклес* 'умереть' — *бубклес*;
абхъес 'открыть' — *ибхъес*, *бухес* 'отнести' — *бихес*, *бехъес* 'разрушить' — *бихъес*;
бяхъес 'ударить' — *бирхъес*;
беркес 'съесть' — *букес*, *бергес* 'съесть' — *бугес*;
бердес 'износить', 'сносить' — *удес*, *берцес* 'спастись' — *уцес*;
абердес 'вырвать' — *авдес*;
белсес 'сплести' — *лусес*, *белклес* 'написать' — *луклес*,
айес 'дать' — *лугес*, *хъес* 'привести' — *бихъес*.

Как видно из примеров, в некоторых типах аблаутное чередование дополняется оппозицией наличия/отсутствия консонантных элементов. Исследователи высказывались об исконности аблаута в даргинском: «Сравнение способов образования коррелятивных видовых форм... дает основания для вывода о том, что категория вида была унаследована... от даргинского языка-основы почти в таком виде, в каком она продолжает функционировать и поныне. Поэтому сходства в образовании видовых форм... идут гораздо глубже, чем различия. Однако имеющиеся между ними различия ни в коем случае нельзя отнести к категории несущественных, малозначительных, поскольку они в какой-то мере отражают состояние после праязыковой эпохи в развитии отдельных идиомов и их категорий, в том числе и категории вида» [Хайдаков 1985: 128]. Реконструкцию протодаргинского аблаута находим в «Северокавказском этимологическом словаре» [НИКОЛАЕВ & STAROSTIN 1994: 101], где была предложена схема, включающая пять типов:

Сов. вид	Несов. вид	Сов. вид	Несов. вид	Сов. вид	Несов. вид
* <i>a</i>	* <i>u</i>	* <i>e</i>	* <i>u</i>	* <i>y</i>	* <i>u</i>
* <i>a</i>	* <i>y</i>	* <i>e</i>	* <i>y</i>		

Как показывают материалы современных лезгинских языков, категория вида (аспекта) уже утратила здесь свой первоначальный облик и сохраняется в форме противопоставления так называемых видовых основ: основы совершенного, или недлительного, вида (терминатива) и основы несовершенного, или длительного, вида (дуратива), от которых образуются соответствующие временные формы [АЛЕКСЕЕВ 1985: 67]. Основным показателем видовых противопоставлений здесь является наличие оппозиции \emptyset (терминатив) — *-p-/-л-* (дуратив). Наряду с этим здесь выявляются и рефлексы общелезгинского глагольного аблаута, реконструированного С. А. Старостиным (см. [NIKOLAYEV & STAROSTIN 1994: 166–168]). Схему чередований здесь можно представить в виде следующей таблицы:

Основа	терминатива	дуратива	потенциалиса
Первый гласный корня	<i>и</i> <i>а</i> <i>е</i> <i>о</i>	<i>аь</i> <i>и</i> <i>е</i> <i>о</i>	<i>ы</i> <i>а</i> <i>е</i> <i>о</i>

Как видно из таблицы, основы с гласными *е* и *о* не имели чередований. Что касается аблаутных чередований, то в современных языках в значительной степени произошла нейтрализация некоторых членов оппозиции, что можно проследить на следующих примерах:

1. **и/*аь/*ы*

В арчинском языке лучше сохранилась оппозиция ступеней чередования **и/*ы*, реализуемая как в пределах одной глагольной парадигмы (а), так и лексическими противопоставлениями (б), в то время как третья ступень представлена ныне самостоятельными глагольными лексемами (в), ср.:

- (а) *кар-ас* ‘вести, брать с собой’ (**и*) — дур. *оркир* — терм. *ока* (**ы*);
- (б) *ло-с* ‘давать’ (**и*) — *ола-с* ‘продавать’ (**ы*);
- (в) *шел-с* ‘побежать’ (**и*) — *гьелриалс* ‘бежать’ (**аь*);
- (а, в) *хьле-с* ‘идти’ (**и*) — дур. *орхьлур*, терм. *охьла* (**ы*) — *гьерхьла-с* ‘ходить’ (**аь*).

В рутульском языке налицо реликтовые примеры оппозиции ступеней чередования **и/*аь*, представленной близкими по значению лексемами (а), а также ступеней **и/*ы* в рамках глагольной парадигмы, ср.:

- (а) *к-учьва-с* ‘начинать’ (**-ичьве-*) — *к-аьчьва-с* ‘начинать; входить’ (**-аьчьва-*)¹;
- (б) терм. ед.ч. *йикьи-р* (**и*) ‘умер’ — мн. ч. *л-ыркьи-р* (**ы*) ‘умерли’.

В цахурском сохранился общелезгинский аблаут **и/*аь*, противопоставляющий основы терминатива (потенциалиса) и дуратива, у глаголов, характеризующихся отсутствием в основе конечного сонанта (к ним относятся, например, *гиклас* ‘убивать’, *гихьас* ‘попасть, оказаться’, *алишес* ‘покупать’, *гичлес* ‘входить’, *гихвас* ‘убегать’, *хьшес* ‘свежевать’, см. [АЛЕКСЕЕВ 1985: 80]). Ранее [ibid.] мы высказывали мнение, что сохранение в ряде цахурских глаголов чередования *и/е* могло быть обусловлено определенными процессами с участием дуративного *-й-* (< **-р-*), поскольку видовое противопоставление \emptyset — *-й-* имеет место в глаголах с конечными сонантами, ср.:

	Целевая форма	Дуратив
‘умереть’	<i>хьиклас</i>	<i>хьекла</i> (<i>*хьи-й-кла</i>)
‘заставить’	<i>аликас</i>	<i>илека</i> (<i>*али-й-ка</i>)
‘спать’	<i>кьалихьас</i>	<i>кьилехьа</i> (<i>*кьали-й-хьа</i>), но
‘оставлять’	<i>гьа̄сарас</i>	<i>гьа-й-с̄ар</i>
‘бросать’	<i>гьувохьарас</i>	<i>гьуво-й-хьар</i>
‘звать’	<i>хьотлалас</i>	<i>хьо-й-тлал</i>
‘бежать’	<i>кьадахьванас</i>	<i>кьада-й-хьван</i>

Сопоставление данных цахурского языка с аналогичными явлениями в других лезгинских языках дает основания отказаться от этой гипотезы.

В крызском языке реликты общелезгинского аблаута **и/*аь* выявляются в формах повелительного нескольких глаголов: *хьий-идж* ‘быть’ — пов. *с-аьхь*; *кьаьй-идж* ‘умирать’ — пов. *с-аькь*; *кур-идж* ‘зарезать’ — пов. *с-аькр*; *джир-идж* ‘жарить’ — пов. *с-аьджир* и т. п.

¹ В цахурском данная основа сохранила видовую оппозицию этих ступеней чередования **и/*аь*: *ичь-ес* ‘войти’ — дур. *ечьле* и т. п., а в арчинском — лексическое противопоставление основ **и/*аь*: *члу-бус* ‘входить’ — *ачьла-с* ‘прятаться’.

2. *a / *u

Этот тип аблаута сохранил свою продуктивность в северном диалекте табасаранского языка, ср.: дубек. *a-в-хь-ус* 'упасть' — дур. *и-в-хь-ур-*, *алд-а-в-тл-ус* 'срубить' — дур. *илд-и-в-тл-ур-* и т. п.

Исконные грамматические функции сохранились у данного типа аблаута и в цахурском, ср. *с-аклал-ес* 'вернуться' — дур. *с-йклар* и т. п.

Кроме того, налицо межъязыковые изоглоссы, отражающие различные ступени аблаута, ср.: арч. *айлар* 'болеть' (*a) — таб. дур. *ийру хьуз* 'болеть', аг. *итар хьас* 'болеть', лезг. *тла-з* 'болеть', крыз. *тит-аьдж* 'ныть' (*u).

Последний тип аблаута имеет прямую аналогию в андийских языках, где чередование *u / *o противопоставляет формы единственного и множественного числа, ср. анд. ед. *вухиду* — мн. *вохиду* 'получать, покупать', ед. *вугьиду* — мн. *вогьиду* 'стоять; быть', ед. *вучьиду* — мн. *вочьиду* 'встать', ед. *вусинну* — мн. *восинну* 'находить', ед. *вучинну* — мн. *вачинну* 'упасть', *вукьиду* — мн. *вокьиду* 'приходить', *вукьуду* — мн. *вокьуду* 'расти', *вучьиду* — мн. *вочьиду* 'резать', *вуцуду* — мн. *воцуду* 'наполнить', *вухьиду* — мн. *вохьиду* 'зарезать', *вутинну* — мн. *вотинну* 'мочить', *вухьиду* — мн. *вахьиду* 'выявить', *вучьиду* — мн. *вачьиду* 'платить', *вучьиду* — мн. *вочьиду* 'умереть', *вукьуду* — мн. *вокьуду* 'быть' [ЦЕРЦВАДЗЕ 1954: 437–440]. Общий принцип аблаута в андийском глаголе следующий: «более закрытые гласные меняются на более открытые, при этом более открытые гласные выражают понятие множественного числа» [ibid.: 440]. Соответственно, исходные *-a-*, *-o-*, *-e-* чередованиям не подвергаются.

Эти же принципы прослеживаются, хотя и менее отчетливо, в других андийских языках, ср., например: год. ед. *вуклал* — мн. *баклал* 'положить', *вухи* — мн. *баху* 'восходить, показаться', *вухи* — мн. *баху* 'попасть, упасть', *вукли* — мн. *бакли* 'быть', *вун* — мн. *мани* 'уходить, идти', *вухи* — мн. *баху* 'вытаскивать, вынимать', *вукь* — мн. *бакь* 'гнать' [САИДОВА 1973: 114–115]. В каратинском языке «в глаголах, в которых классный показатель присоединяется к основе посредством гласного *и* (< *i*), *о*, во множественном числе этот гласный заменяется посредством *а*» [МАГОМЕДБЕКОВА 1971: 116], ср. III ед. *буча* — I–II мн. *бачва* 'мыть', III ед. *боиле* — I–II мн. *баиле* 'пойти', III ед. *бихь* — I–II мн. *бахь* 'зарезать', III ед. *му'у* — I–II мн. *ма'у* 'расти'. В ахвахском языке чередование гласных наблюдается в глаголах с корневым *-и-*: III ед. *бикьурулла* — I–II мн. *бакьурулла* 'делить', III ед. *битурулла* — I–II мн. *батурулла* 'проиграть', III ед. *бихьурулла* — I–II мн. *бахьурулла* 'держат' и др. Как отмечает З. М. Магомедбекова, «иногда, во множественном числе для класса человека, вместо ожидаемого сопровождающего гласного *i* встречается *a*. Возможно, это вызвано стремлением разграничить мн. число класса человека от единственного числа третьего класса» [МАГОМЕДБЕКОВА 1967: 36].

В отличие от ахвахского и других андийских языков, но так же, как в андийском, гласный *-a-* в чамалинском появляется в обеих формах мн. числа: III ед. *буклва* — мн. I–II *баклва*, III–IV *йаклва* 'быть', III ед. *муна* — мн. I–II *мана*, III–IV *йана* 'идти', III ед. *бицла* — мн. I–II *бацла*, III–IV *йацла* 'схватить', III ед. *бихла* — мн. I–II *бахла*, III–IV *да-хла* 'брат, покупать' [БОКАРЕВ 1940: 54–55; 1949: 76–77]. Имеются также глаголы с корневым *-и-*, не дающие чередования, ср.: III ед. *бичла* — I–II мн. *бичла* 'умереть', III ед. *бигьла* — I–II мн. *бигьла* 'победить', III ед. *билла* — I–II мн. *билла* 'проснуться', видимо по семантическим причинам.

Из приведенного краткого обзора видно, что аблаут *и/а*, а также *у/а* имеют общеандийский характер и наблюдаются в генетически общих лексемах, ср., например: анд. *бикл-/боклв-*, год., чам., тинд., кар. *букл-/баклв-*, багв. *букл-/бакл-*, ахв. *бикл-/баклв-* 'быть'. Этот и ему подобные примеры Т. Е. Гудава [1959: 13] считал подтверждением функциональной членимости «сопровождающего» гласного. Как было показано, они ограничиваются корневыми *-и-* и *-у-* (< **-и-*) и не выходят за пределы оппозиции числовых форм, так что говорить о других функциях глагольного аблаута в андийских языках не приходится.

В плане общедагестанских корреляций семантическая связь категорий вида и числа не представляется нам неестественной. В работах по кавказоведению эта связь подчеркивалась неоднократно. Так, в специальной литературе было замечено, что количественные характеристики реализуются в глагольной парадигме с помощью не только синтаксической категории числа, но и собственно глагольной категории вида: «В пределах глагола функционирование категории ко-

личества носит специфический характер. Здесь мерилom является не количество предметов, а частотность действий, которая может быть выражена через соответствующие формы категории — это ее мерилa. Одни формы в характеризуемой категории выражают однократное, одноразовое действие, а другие — многократное, многократное действие» [Хайдаков 1985а: 4].

Заметим, что охарактеризованная выше категория числа андийского глагола, по-видимому, находит реликтовое отражение в лексических оппозициях аварского языка, ср.: *буклине* 'быть' — *бакл-ар-изе* 'собираться' < *'быть (о многих)', где *-ар-* — суффикс учащательности; *щевезе* 'достигать' — *барцизе* 'зреть' < *'достигать (о многом, многих)' и нек. др. Что же касается известного в лингвистической литературе аварского глагольного аблаута типа *пирхизе* 'блеснуть, вспыхнуть' — *перхэзе* 'сверкать, блестеть' и т. п., то его вторичность, обусловленность регрессивной ассимиляцией достаточно хорошо показана специалистами [Чикобава, Церцвадзе 1962: 32–33]; [Микаилов 1958: 132–136].

Иные функции имеет аблаут в цезских языках. Так, в ряде глаголов восточноцезских языков классные оппозиции выражаются чередованием корневых гласных, ср.: гунз. *гэла* I, *гила* II–III, *гула* IV–V 'класть', *наа* I, *ниа* II–III, *нуа* IV–V 'приходить', *млэгла* I, *млумла* II–III, *млымла* IV–V 'бросать'. Возведение этих чередований к реально засвидетельствованным или же реконструируемым сегментным классным показателям не удается, за исключением форм II класса, где можно полагать исконным **й* (*гила* < **гэ-й-ла*, *ниа* < **нэ-й-а*, *млумла* < **млэ-й-мла*). Не исключено, что в данной группе глаголов произошло переосмысление исторического показателя I класса **в* (*гула* < **гэ-в-ла*, *нуа* < **нэ-в-а*, *млумла* < **млэ-в-мла*), связанное с формированием новой, нулевой формы I класса. Несколько труднее представить себе процесс **нэ-б-а* > ? > *нуа* и т. п.

Между тем в структуре западноцезских языков имеются факты, в значительной степени сходные с обрисованными в других подгруппах нахско-дагестанских языков. В данном случае мы имеем в виду форму общего времени в инхокаринском диалекте хваршинского языка, которая образуется удлинением корневого главного, вставкой *ā/ō* в скопление корневых согласных или дифтонгизацией корневых *у, ы, и*, ср.: *ахъа* 'спать' — *āхъ*, *леза* 'взять, купить' — *лэз*, *кюлла* 'прыгать' — *кюл*, *лулха* 'сыпать' — *лулōх*, *шварда* 'прыгать' — *шварад*, *угъа* 'умереть' — *увогъ*, *илла* 'сказать' — *ийол* и т. п. (см. [Бокарев 1959: 166–167]). Сходным образом формируется общее время в собственно хваршинском диалекте, ср.: *леса* — *лэса* 'спать', *тилла* — *тйлла* 'отдать', *луха* — *луха* 'копать', *ишана* — *ишōна* 'жарить', *шварда* — *шварōда* 'прыгать' и т. п. (см. [Шарафутдинова, Левина 1961: 116]).

При некоторых различиях в говорах в целом цезский язык использует аналогичный принцип образования общего времени, при котором происходят следующие преобразования корневых гласных: шайтл. *а* — *ā*, *о* — *ō*; *и, е* — *ē*; кидер., мокок. *а, у, о* — *ā*; *и, е* — *āь*; асах. *а, у, е, и* — *ā*; *у* (во втором слогe) — *ва*, ср. *аула* 'есть' — *āул*, *кюлла* 'прыгать' — *кюллу* // *кюллу*, *икла* 'идти' — *ēклу* / *āьклу* / *āьклу*, *-ехура* 'убивать' — *-ехвār* и др. [Имнайшвили 1956: 421–423]. Как полагал Д. С. Имнайшвили [ibid.: 423], «формантом настоящего общего в дидойском языке являлся гласный *а*, а нынешний способ — выражение этой категории изменением гласных основы глагола — результат слияния форманта *а* с гласными основы глагола». В хваршинском он предполагал [ibid.: 425] в качестве исходного аффа. **-о-* (*-ой-*). Как видим, его гипотеза аналогична той, что была предложена им для нахских языков, с той лишь разницей, что в качестве исконного аффикса выступает иной согласный.

Как видно из проведенного выше обзора, в той или иной степени глагольный аблаут представлен во всех подгруппах нахско-дагестанских языков (по-видимому, это явление не находит отражения в лакском и хиналугском языках, а также в некоторых представителях лезгинской группы). Несмотря на то что кавказоведение ныне располагает фундаментальным этимологическим словарем северокавказских языков, в котором впервые осуществлена реконструкция общевосточных глагольных основ, проблема реконструкции общевосточнокавказского глагольного аблаута сохраняется, заключаясь в недостаточном лексическом материале (притом что аблауту, как можно полагать, были подвержены не все основы, а лишь их часть), на котором можно было бы выявить регулярные соответствия в этой области. Проблема осложняется тем, что помимо первого гласного основы в систему глагольного аблаута входили также чередования второго гласного основы, а в систему видовых противопоставлений в целом — оппозиция \emptyset — *-р-(-л-)* и редупликация основы.

- АЛЕКСЕЕВ 1985 — М. Е. АЛЕКСЕЕВ. *Вопросы сравнительно-исторической грамматики лезгинских языков. Морфология. Синтаксис*. М.: «Наука».
- БОКАРЕВ 1940 — А. А. БОКАРЕВ. О классных показателях в аваро-андо-цезских языках // *Язык и мышление*. Т. X. М.
- БОКАРЕВ 1949 — А. А. БОКАРЕВ. *Очерк грамматики чамалинского языка*. М.—Л.
- БОКАРЕВ 1959 — Е. А. БОКАРЕВ. *Цезские (дидойские) языки Дагестана*. М.
- ГИГИНЕЙШВИЛИ 1977 — Б. К. ГИГИНЕЙШВИЛИ [ბ. გიგინეიშვილი]. *Сравнительная фонетика дагестанских языков [დაღესტნური ენების შედარებითი ფონეტიკა]*. Тбилиси: Издательство Тбилисского университета.
- ГУДАВА 1959 — Т. Е. ГУДАВА. *Сравнительный анализ глагольных основ в аварском и андийских языках*. Махачкала.
- ДЕШЕРИЕВА 1979 — Т. И. ДЕШЕРИЕВА. *Исследование видо-временной системы в нахских языках (С привлечением материала иносистемных языков)*. М.: «Наука».
- ДОЛАКОВА 1967 — Р. И. ДОЛАКОВА. Ингушский язык // *Языки народов СССР*. Т. IV. *Иберийско-кавказские языки*. М.: «Наука»; стр. 210–227.
- ИМНАЙШВИЛИ 1956 — Д. С. ИМНАЙШВИЛИ [დ. იმნაიშვილი]. К спряжению глагола в дидойском языке [ზმნის უღვლილებისათვის დიდოურ ენაში] // *იბერიულ-კავკასიური ენათმეცნიერება / Иберийско-кавказское языкознание*. Т. 8. თბილისი / Тбилиси; стр. 413–457.
- ИМНАЙШВИЛИ 1977 — Д. С. ИМНАЙШВИЛИ [დ. იმნაიშვილი]. *Историко-сравнительный анализ фонетики нахских языков [ნახური ჯგუფის ენათა ფონეტიკის ისტორიულ-შედარებითი ანალიზი]*. თბილისი: “მეცნიერება” / Тбилиси: «Мецниереба».
- МАГОМЕДБЕКОВА 1967 — З. М. МАГОМЕДБЕКОВА [ზ. მაჰომედიბეკოვა]. *Ахвахский язык: Грамматический анализ, тексты, словарь [ახვახური ენა. გრამატიკული ანალიზი, ტექსტები, ლექსიკონი]*. თბილისი: “მეცნიერება” / Тбилиси: «Мецниереба».
- МАГОМЕДБЕКОВА 1971 — З. М. МАГОМЕДБЕКОВА [ზ. მაჰომედიბეკოვა]. *Каратинский язык: Грамматический анализ, тексты, словарь [კარატული ენა. გრამატიკული ანალიზი, ტექსტები, ლექსიკონი]*. თბილისი: “მეცნიერება” / Тбилиси: «Мецниереба».
- МИКАЙЛОВ 1958 — Ш. И. МИКАЙЛОВ. *Сравнительно-историческая фонетика аварских диалектов*. Махачкала.
- САИДОВА 1973 — П. А. САИДОВА. *Годоберинский язык*. Махачкала.
- СУЛЕЙМАНОВ 2000 — Н. Д. СУЛЕЙМАНОВ. *Словообразование и структура слова в восточно-лезгинских языках*. Махачкала.
- ТРУБЕЦКОЙ 1987 — Н. С. ТРУБЕЦКОЙ. *Избранные труды по филологии*. М.: «Прогресс».
- ХАЙДАКОВ 1985 — С. М. ХАЙДАКОВ. *Даргинский и мегебский языки (принципы словоизменения)*. М.: «Наука».
- ХАЙДАКОВ 1985a — С. М. ХАЙДАКОВ. Способы выражения множественного числа в дагестанских языках // *Категория числа в дагестанских языках*. Махачкала.
- ЧИКОБАВА, ЦЕРЦВАДЗЕ 1962 — არნ. ჩიქობავა, ილ. ცერცვაძე [А. С. ЧИКОБАВА, И. И. ЦЕРЦВАДЗЕ]. *ხუნძური ენა / კავკასიურ ენათა კათედრის შრომები*, 2 [Аварский язык / Труды кафедры кавказских языков, 2]. თბილისი / Тбилиси.
- ЧРЕЛАШВИЛИ 1999 — К. Т. ЧРЕЛАШВИЛИ [კ. ჭრელაშვილი]. *Бацбийский язык // Языки мира. Кавказские языки*. М.: «Наука»; стр. 197–203.
- ШАРАФУТДИНОВА, ЛЕВИНА 1961 — Р. ШАРАФУТДИНОВА, Р. ЛЕВИНА. *Хваршинский язык: Предварительное сообщение // Вопросы изучения иберийско-кавказских языков*. М.: Изд. АН СССР.
- NIKOLAYEV & STAROSTIN 1994 — S. L. NIKOLAYEV and S. A. STAROSTIN. *A North Caucasian Etymological Dictionary*. Moscow, “Asterisk” publishers.

S u m m a r y

The main goal of the article is to briefly analyze the series of vocalic gradation (Ablaut) observed in the verbal systems of all the subgroups of the Nakh-Daghestanian language family. It is shown that the diversity of vocalic gradation, attested in most of these subgroups, does not justify the frequent treatment of Nakh-Daghestanian verbal roots as primarily monoconsonantal. It is also shown that such vocalic gradations usually cannot be explained through the inner history of the subgroups themselves, meaning that a certain variety of Ablaut must have existed already in Proto-Nakh-Daghestanian. However, the exact reconstruction of this variety is still a task for the future.

Once Again on the Comparison of Personal Pronouns in Proto-Languages

The article discusses the importance of comparison of personal pronouns for studies dealing with long-range linguistic relationship. It is recognized that individual similarities between pronouns reconstructed for various low-level protolanguages may be ascribed to chance resemblance or borrowing, but the same is much less probable for comparisons between entire paradigmatic systems of pronouns. Based on a brief survey of the principal pronominal forms and markers in most of the protolanguages reconstructed for the world's language families, it is asserted that many of the current hypotheses on macrofamilies, such as Nostratic, Sino-Caucasian, Austric, and others, are indeed supported by such paradigmatic similarities.

For those scholars who engage in long-range comparison of the world's languages, personal pronouns have been a viable source of genetic data. Indeed, as was already noticed by the first Indo-Europeanists nearly two hundred years ago, personal markers, both syntactically independent and bound with verbs or nouns, are among the most stable elements of the language's morphology. The Ukrainian pronominal form *мене* 'me' and the Irish form *tú* 'thou' both sound very much the same as they must have sounded some 6000 years ago in the speech of Indo-Europeans, and bear the same meanings. Three personal pronominal forms are deservedly placed in the Swadesh list of the one hundred most stable lexical items in human vocabulary.

Moreover, personal pronouns are relatively rarely borrowed between languages, especially from a genetically different language stocks. It would be an extreme thing to say that they cannot be borrowed at all — since they *can* and *are*, as shown by a number of cases — but this happens much more rarely than with nominal or verbal lexical items. Instances of the first person singular pronoun 'I' being a loanword in a language are extremely scarce, probably no more than ten reliable examples known so far, and almost all of them can be explained by some peculiar sociolinguistic (or grammatical) conditions.

These two unique characteristics of personal pronouns — internal stability and resistance to external borrowing — make the analysis of personal pronouns, both lexical and morphological, extremely important for comparative linguistic studies.

Can personal pronouns serve as solid proof or, at least, a considerable heuristic argument in proposing distant genetic relationship between languages? This question has long been a subject of discussion in various fields of linguistics. Personal markers were a strong argument in favor of the Altaic theory advocated by numerous authors starting from J.-F. STRAHLENBERG ([1730]). They were one of the keystones in postulating the Nostratic macrofamily by H. PEDERSEN, and the Amerind family in the Americas by J. GREENBERG ([2000]). In Papuan linguistics, personal markers arguably remain as the most important factor in positing linguistic relationship and elaborating classifications, most of which are not yet supported by conventional lexical comparative analysis ([FOLEY 1986: 3]). Pronouns are widely drawn upon elsewhere to prove genetic relationship, and, therefore, methodologically a positive answer to the above question would help us to validate a number of macrofamily hypotheses still under discussion.

Unsurprisingly, opponents of long-distance relationship claims have their own reaction. A number of respectable linguists have challenged the argument that personal pronouns may be essential for long-range genetic comparison. The most widespread and easily understandable objections are quoted as follows:

1. Personal pronouns *can* be borrowed. Well known examples of this kind are drawn mainly from closely related languages, like English, where *they* is a Scandinavian loanword, or Dravidian, where Kolami is claimed to have borrowed the 2nd person singular *niv* from Old Telugu ([KRISHNAMURTI 2001: 91]). More distant languages also demonstrate sporadic examples: thus, Micronesian Chamorro borrowed Spanish *yo* 'I' ([TOPPING 1973: 107]), while Ambonese Malay adopted *ose* 'thou' from Portuguese *Você* 'you (Sg.)'. Languages of New Guinea tend to "exchange" certain personal pronouns quite often ([LAYCOCK-Z'GRAGGEN 1975: 732, 737]), though this phenomenon has not yet been sufficiently studied.

2. Personal pronouns are, in fact, mostly very short lexical items usually consisting of CV roots which are hard to compare, because the chance of coincidence is very high ([CAMPBELL 2003]). Moreover, it has been tentatively shown that pronominal patterns tend to crystallize an opposition of a nasal to a stop consonant as the basic phoneme of the pronominal root ([NICHOLS 2003]). Even if this hypothesis is not entirely correct, this pattern works in a great number of the world's languages.

In our opinion, the most responsible answer to both of these points is not to try and prove the opposite; rather it is to use personal pronouns for language comparison not just as isolated lexical items, but as a paradigmatic subsystem of morphology.

In the majority of the world's languages, pronouns constitute a balanced paradigm. The types of such paradigms are very well presented by M. CYSOUW ([CYSOUW 2003]), the most basic being the opposition of the two pronouns of the 1st and 2nd persons singular: the pronominal system is limited by this binary opposition in a number of languages, e. g. Golin and Salt-Yui in New Guinea ([FOLEY 1986: 70]). Other structural samples include the number opposition, with four basic members like *I — we, thou — you*. Still other systems are augmented even further, including 3rd person pronouns, usually adopted from the demonstrative domain; the introduction of the category of clusivity (inclusive / exclusive pronouns of the 1st person plural); more number categories (dual / trial / paucal &c.). Finally, personal pronouns can be declined for case, as in most Indo-European languages, adding even more forms to the pronominal paradigm. Nominative and oblique forms are frequently derived from suppletive roots, e. g., the Indo-European 1st person singular nominative *eǵHo(m) vs. oblique *me-.

It seems clear that in order to prove that the languages are genetically related, it is the entire pronominal paradigm that should be compared rather than isolated personal pronouns as disconnected lexical items. Adopting the latter approach, we may certainly encounter loanwords or mere coincidences, like Indo-European *me 'me' = Sumerian (Eme-sal) *me* 'id.'. However, comparison of entire paradigms drastically decreases the probability of either chance similarity or borrowing.

In those languages where personal pronouns constitute paradigms, there are no known cases of borrowing the *whole paradigm* from another language, provided exceptional situations with pidgin and mixed languages are excluded¹ ([THOMASON–EVERETT 2005]).

One of the most recent claims ([CAMPBELL–POSER 2008: 214]) is that the paradigmatic approach to comparison is equally invalid for long-range comparison purposes. The main reason here is that, as the authors claim, personal pronouns are essentially iconic “phonosymbols”, an assertion that is as insufficiently grounded as it is simply untrue; if it were so, pronominal comparison would be useless altogether, since pronouns would tend to share the same phonological features in all language families — which is clearly not the case, as can be easily seen from the chart below. CAMPBELL & POSER also claim that pronouns are subject to analogical reformations, which is obviously correct (we state the same in [BABAЕV 2008: 209]). However, analogical reformations usually have internal reasons and should not be expected to make one language's paradigm closer to that of another.

Paradigm comparison is also important for historical linguistics, because in long-range studies any comparison of materially short cognates with a CV or VCV structure, where the consonant is the only reliable sound to be compared, is dubious by definition. Therefore, comparison between proto-languages, aimed at reconstructing a 10 000-year-old or older linguistic situation, should be made between reconstructed cognate paradigms of personal pronouns rather than between personal pronouns as lexical roots.

Such attempts have, in fact, been made, some of them quite successful. It was noticed long ago that the pattern exhibited by personal pronoun paradigms in a number of language families in Northern Eurasia is remarkably similar; this pattern can be described as *mV for the 1st person vs. *tV for the 2nd person. The Nostratic/Eurasiatic theory proposes a genetic link between languages that present this pattern, namely, Indo-European, Altaic, Uralic, Yukaghir and, most probably, Chukchi-Kamchatkan. Indeed, (a) the linguistic probability of borrowing this pattern is practically zero; (b) the statistical probability of a chance similarity in such a geographical cluster is also close to zero; and c) a pattern like this is not attested for any other proto-language of any other family on Earth. This makes the “pronominal factor” one of the strongest bits of grammatical evidence for the Nostratic hypothesis.

Another example is the American pronominal pattern *na- (1st pers.) vs. *mi (2nd pers.), first noted by A. TROMBETTI ([1905]) in a great number of the languages of both Americas and later de-

¹ One well known case is that of the Thai language, which uses Chinese and English personal pronouns in specific sociolinguistic contexts ([COOKE 1968: 11 – 12]). However, they do not replace the common Thai personal pronouns.

veloped by J. GREENBERG into one of the key arguments of his Amerind theory. Again, the spreading of the pattern is extremely wide, and the chance of borrowing or coincidence is lower than low.

The present article is inspired by a tentative list of reconstructed pronominal forms for the first two persons in the world's languages presented by M. RUHLEN ([1994])². However, we find it necessary to continue RUHLEN's effort by substantially adding to the sketch pencilled by his pioneering work. Our task will be to present an updated and accurate comparative chart of personal pronoun paradigms, reconstructed for proto-languages of most of the world's language families, to be subject to further analysis.

To make our results as precise as possible, we will rely on the most recent data collected by scholars, especially concerning languages of Africa and the Pacific region. Although many families still lack thorough proto-language reconstructions, most of their pronominal systems have been at least tentatively sketched out since RUHLEN's work was published fourteen years ago. Apart from independent pronominal forms, we will try to include bound verbal and nominal affixes denoting person as well, provided they differ from the former.

Finally, it is necessary to give the appropriate reference links to etymological dictionaries and/or comparative grammars for each reconstructed form that we quote for a proto-language. Unfortunately, no such references are given in RUHLEN's list, which makes it harder to resolve all the problems that may emerge during one's analysis of it. We have attempted to find as many links as possible for the reconstructions, bearing in mind that for a number of language families (African for example) no comparative work exist at all, and for these cases we make our own tentative reconstruction based on the available dictionaries and word lists (these reconstructions are marked by a double asterisk).

The chart below will provide sufficient material for such an analysis. Along with the usual pronominal forms reconstructed for the family's proto-language, we also mark a number of peculiar forms from separate languages (or language groups) of the family in case they differ from the main family pattern — e. g., Svan forms are placed alongside the reconstructed Proto-Kartvelian ones, because Svan is considered the most morphologically archaic dialect of the family, and a number of judgments on Proto-Kartvelian pronouns has been made based exclusively on very peculiar Svan forms.

We have attempted to make our chart below as taxonomically correct as possible. The majority of the families listed go back to proto-languages which should have existed at roughly the same chronological period: 6 to 8 thousand years ago, considered by many of today's mainstream linguistics as the limit for comparative reconstructions. And, although a number of macro-families have also been recognized in the last decades, for some of them it is not always easy to reconstruct a single system of personal pronouns. Thus, the Afrasian pronominal reconstructions are still under serious discussion. Niger-Congo and Austro-Asiatic may be taken as even better examples: these two macrofamilies seem to be well established and widely accepted, but the personal pronoun systems for their proto-languages have not yet been verified. Therefore, we will only demonstrate pronominal forms for their subfamily proto-languages: Mande, Gur, Dogon, &c. and Munda and Mon-Khmer respectively. The Australian macrofamily is so far an unverified hypothesis (although widely accepted), and the New Guinean 'phyla' are even less so ([FOLEY 1986]). The most complicated situation lies in the Americas: the classification of solidly reconstructed families includes more than 60 taxonomic units ([CAMPBELL–MITHUN 1979: 39–46]), and, although it has been reduced to around a dozen macrofamilies by GREENBERG and RUHLEN ([2007: 6]), their theory is not universally accepted and still needs further verification.

The above chart, with its schematic information, cannot even pretend to be perfect, since each form in this matrix is subject to long and deep analysis per se. But even this sketchy picture still allows to make a number of interesting conclusions, based on the patterns of personal pronoun paradigms.

1. The languages of northern Eurasia show a pattern “*M—T*” which is found nowhere else in the world on the family proto-language level. Namely, these are Indo-European, Uralic, Yukaghir, Altaic and Chukchi-Kamchatkan languages. This phenomenon has been identified long ago and constitutes a part of the morphological evidence for the Nostratic hypothesis.

2. A similar pattern “*M—S*” is present in Indo-European, Kartvelian, Altaic and, arguably, Chukchi-Kamchatkan, where the form **γǝš* ‘thou’ is of an unclear origin. In Indo-European, the distribution of

² *Editors' note:* A similar list of 1st and 2nd person pronouns has been earlier produced by S. L. NIKOLAEV during his talk on the Nostratic Seminar in Moscow in April of 1985. The list was, however, much shorter, including only those major families across Eurasia, Africa, and North America for which at least some sort of protolanguage reconstruction had already been established or sketched.

“*M—T*” vs. “*M—S*” is morphological (the former is found in the pronominal system, the latter in the verb system), while in Altaic the variability is dialectal (Mongolic vs. Turkic and Tungus-Manchu). In both Indo-European and Altaic, however, the two systems are projected onto the proto-language level.

It is necessary to add that Nostratic languages also share the case paradigm for the 1st person pronoun: the form **mVn(V)-* may be reconstructed for the Nostratic proto-language on the basis of Indo-European, Altaic, Uralic and Kartvelian languages ([ИЛЛИЧ-СВИТЫЧ 1971: 6]; [БАБАЕВ 2008]).

Chart 1. Proto-language Personal Pronouns

Language Family	‘I’	‘Thou’	‘We’	‘You (Pl.)’
Indo-European ³	n <i>*eg’ho(m)</i> obl <i>*me(ne)</i> v st <i>*-H-</i>	<i>*tū, t(u)e</i> v <i>*-s-</i>	n <i>*me(s), *we(i)-</i> obl <i>*ne/o-</i>	n <i>*iū-</i> obl <i>*we-</i> v <i>*-te</i>
Uralic ⁴	n <i>*mE</i> obl <i>*mEnV</i> v st <i>*-k</i>	n <i>*tE</i> obl <i>*tEnV</i>	<i>*mE</i>	<i>*tE</i>
Yukaghir ⁵	<i>mit</i>	<i>tit</i>	<i>met</i>	<i>tet</i>
Altaic ⁶	n <i>*bi</i> obl <i>*mi-n-, ? *ŋa</i>	n <i>*si, *t’i</i> obl <i>*si-n-, ? *na</i>	<i>*ba / *bu</i> obl. <i>*ma-n- ~</i> <i>*mu-n-</i>	<i>*t’a, *su</i> obl. <i>*su-n-</i>
Turkic	n <i>*bi</i> obl <i>*män</i>	n <i>*si</i> obl <i>*sän</i> v prt <i>*-ŋ</i>	<i>*bi-ř</i> v prt <i>*-k</i>	<i>*si-ř</i>
Mongolian	obl <i>*nama-</i>	<i>čī < *t’i</i>	obl <i>man-</i>	obl <i>tan-</i>
Korean	<i>na < *ŋa, čě</i>	<i>ne < *na</i>		<i>ne-huy</i>
Japanese	<i>*a < *ŋa, wa(re)</i>	<i>na(re)</i>		
Kartvelian ⁷	<i>*me(n)</i> v subj <i>*(x)w-</i>	<i>*si</i> v subj <i>*x-</i> v obj <i>*g-</i>	<i>*čwe(n)</i>	<i>*tkve(n)</i>
Svan ⁸			excl <i>naj / nāj</i> v incl subj <i>l-</i> v incl obj <i>gw-</i>	
Dravidian ⁹	n <i>*nyān</i> obl <i>*yan-</i> v abs <i>*-kV</i>	n <i>*nīn</i> obl <i>*nin-</i> v abs <i>*-ti</i>	n excl <i>*nyām</i> obl excl <i>*yam-</i> n incl <i>*nām</i> obl incl <i>*nam-</i>	n <i>*nīm</i> obl <i>*nim-</i>
Chukchi-Kamchatkan ¹⁰	<i>*yə-m</i>	<i>*yə-š</i> obl <i>*yə-n</i>	<i>muri</i>	<i>turi</i>
Eskimo (W) ¹¹	rel <i>-ma</i> abs <i>-qa</i>	rel <i>-pək</i> abs <i>-n</i>	rel <i>-m-ta</i> abs <i>-(x)put/-wut</i> v du <i>-(x)puk</i>	v rel <i>-pəsi</i> v abs <i>-(x)si</i> v du <i>-(x)tək</i>
Elamite ¹²	v <i>(h)u</i> obl <i>un</i> abs <i>-k</i>	<i>nu, ni</i> v <i>-(a)t(i)</i>	incl <i>ela, elu</i> excl <i>nika, niku</i>	<i>num</i>
Chadic ¹³	I <i>*ʔan-i</i> II <i>*ʔi(ya)</i>	masc I <i>*ka, II</i> <i>*ku</i> fem <i>*ki(m)</i>	incl <i>*muni</i> excl <i>*na, *ʔyina</i>	<i>*kuni, *kuna</i>

³ [BEEKES 1995: 207–211].

⁴ [РЕДЕЙ 1988]; [ХАЙДУ 1985: 225, 234]; [НОНТИ 1993: 124]; [ХЕЛИМСКИЙ 1982].

⁵ [НИКОЛАЕВА–ХЕЛИМСКИЙ 1997: 166].

⁶ [EDAL 225, 1237]; [ИТАВАСНИ 1998: 130–135].

⁷ [КЛИМОВ 1998], [ШМИДТ 1984].

⁸ [ТУИТЕ 1997: 18–26].

⁹ [DED 5160]; [СТАРОСТИН 2006].

¹⁰ [МУДРАК 2000: 39–40, 97, 145–146].

¹¹ [BERGSLAND 1986]; [МЕНОВЩИКОВ 1997].

¹² [ДЬЯКОНОВ 1979].

¹³ [BLAŽEK 1991: 37].

Language Family	'I'	'Thou'	'We'	'You (Pl.)'
Berber ¹⁴	subj * <i>ənakkw</i> obj * <i>īy</i> , * <i>ū(w)</i>	subj masc * <i>kayy</i> , fem * <i>kamm</i> obj masc * <i>ək</i> , <i>ik</i> , fem * <i>am</i>	* <i>na-</i>	* <i>-tum-</i>
Cushitic ¹³	subj * <i>ʔani</i> , obj * <i>yi</i> / * <i>yu</i>	subj *(<i>ʔan</i>)- <i>ti</i> obj * <i>ku</i> / * <i>ki</i>	subj * <i>muni</i> obj * <i>na</i>	*(<i>ʔan</i>)- <i>tinV</i>
OmotiC ¹³	N * <i>ta</i> , * <i>na-</i> , S * <i>ʔi</i>	N * <i>ni</i> , S * <i>ya</i>	N incl * <i>ni</i> , excl * <i>nu</i> S * <i>wV</i>	N * <i>ʔantu</i> , S * <i>yV</i>
Semitic ¹³	*(<i>ʔan</i>)- <i>āku</i> , *(<i>ʔan</i>)- <i>ā</i>	masc *(<i>ʔan</i>)- <i>ta</i> fem *(<i>ʔan</i>)- <i>ti</i>	* <i>hina</i> , * <i>naḥnu/a</i> , * <i>na/i-</i>	masc *(<i>ʔan</i>)- <i>tumu</i> fem *(<i>ʔan</i>)- <i>tina</i>
Egyptian ¹³	(<i>in</i>)- <i>k</i> , - <i>j</i>	masc - <i>k</i> , <i>kw</i> fem <i>tm(t)</i> , - <i>t</i>	(<i>in</i>)- <i>n</i>	<i>tn</i>
Sino-Tibetan ¹⁵	* <i>ŋā</i>	* <i>naŋ</i> , * <i>k^wV-</i>	—	—
North Caucasian ¹⁶	I (n) * <i>zo</i> (E * <i>zō-</i> , W * <i>sa</i>) II (obl) * <i>nV</i> / * <i>mV</i>	I * <i>wo</i> (E * <i>wō-</i> / * <i>g^wə</i> , WC * <i>wa</i>) WC fem * <i>ba</i> II * <i>yu-</i> , * <i>ɣwV-</i>	EC incl * <i>Lä</i> : EC excl *(<i>nV</i>) <i>ži</i> WC * <i>ḥa</i>	EC * <i>z^wě</i> WC * <i>s^wa</i>
Hurro-Urartian ¹⁷	*(<i>j</i>) <i>es-</i> / * <i>so-</i> v - <i>f-</i>	* <i>we-</i> v - <i>u/o-</i>	*(<i>j</i>) <i>es-</i> / * <i>so-</i>	* <i>we-</i>
Yeniseian ¹⁸	I * <i>ʔaʒ</i> , II * <i>b-</i>	I * <i>ʔaw</i> , II * <i>kV-</i> / <i>ʔVk-</i>	* <i>ʔaʒəŋ</i>	* <i>ʔawoŋ</i>
Na-Dene ¹⁹	* <i>šwī</i> / * <i>xwī</i>	* <i>ni</i> , * <i>wī</i> (Tlingit)	* <i>na-</i> , * <i>nu-</i> * <i>ta-</i> , * <i>tu-</i>	* <i>ʔuX-</i>
Burushaski ²⁰	<i>ʒa</i> poss <i>a-</i>	<i>un</i> obl <i>gu-</i> / <i>gō</i>	<i>mi</i>	<i>ma</i>
Basque ²¹	<i>ni</i> v trans * <i>-da-</i>	(<i>h</i>) <i>i</i> v trans masc * <i>-ga-</i>	<i>gu</i>	<i>zu</i>
Sumerian ²²	<i>ŋae</i> poss - <i>ŋu</i> * <i>mV</i> (?)	<i>zae</i> poss - <i>zu</i>	<i>mēdē</i> poss - <i>me</i>	<i>mēzē</i> , - <i>ze</i>
Nahali ²³	<i>ʒo</i> obl <i>eŋ-</i>	<i>nē</i>	<i>ʒo</i> , <i>tye-ko</i>	<i>nē(ko)</i>
Nivkh ²⁴	<i>ni</i>	<i>či</i>	excl <i>nin</i> du <i>me-</i> incl <i>mer</i> , <i>mir</i>	<i>čin</i>
Ainu ²⁵	<i>ku-</i> obj ' <i>en-</i> / -' <i>en-</i>	' <i>e-</i>	excl abs <i>ci-</i> , - <i>as-</i> excl obj ' <i>un-</i> incl subj <i>a-</i> incl obj ' <i>i-</i>	' <i>eci-</i> , <i>es-</i>
Thai-Kadai ²⁶	* <i>k(ā)u</i>	* <i>m(a)u</i>	<i>rau</i> , * <i>tu</i>	<i>miŋ</i> , <i>suu</i>

¹⁴ [ДЪЯКОНОВ 1967: 218–227]; [BLAŽEK 1991].¹⁵ [BENEDICT 1972: 93]; [STAROSTIN 1984].¹⁶ [STAROSTIN 1984]; [NICHOLS 2003: 293].¹⁷ [DIAKONOFF–STAROSTIN 1986: 81–82].¹⁸ [STAROSTIN 1984].¹⁹ [BENGTSON 1992].²⁰ [BERGER 1974: 24].²¹ [TRASK 1997: 218].²² [ДЪЯКОНОВ 1967: 62]; [КАНЕВА 2006: 59–61].²³ [SHAFFER 1941: 363]; [PINNOW 1966: 189].²⁴ [GRUZDEVA 1998].²⁵ [VOVIN 1993: 79, 84, 106, 144]; [SHIBATANI 1990: 25–31, 45–50].²⁶ [BENEDICT 1942: 586]; [BENEDICT 1975: 203–208]; [LI FANG-KUEI 1977: 361–363].

Language Family	'I'	'Thou'	'We'	'You (Pl.)'
Austronesian ²⁷	*(w) 'aku / *ku	*kaw, *(ka)su	excl *(k)ami incl *(k)ita	*(ka)mu
Mon-Khmer ²⁸	*(V)ko / *ki, *ʔVn / *ʔVy	*miʔ / *mVj	*hiʔ incl *bən du *ʔaʔ	*(b)naʔ, *piʔ
Munda ²⁹	*iŋ	*me	excl *le incl *bi(n)	*pe
Miao-Yao ³⁰	*ko(u), *wanj Yao *ʔja	*koŋ, *kam Yao *mway	*(m)pV	Miao du *mu Yao *nV-
Andaman ³¹	do(l)- / tu- Onge mi	poss ŋu(l)- / ŋo-	mu(l)- et-	ŋu(l)- / ŋo-
Kusunda ³²	tsi	tok	nu	nok
Papuan (general types) ³³	I nV II kV / gV / ŋV III tV / dV	I kV / gV / ŋV II ma III na	I ni II mV / pV III ki / ti	I ki II mi III nik
West Papuan ³⁴	*tV	*nV		
Toricelli ³⁵	*k-	*yV, *t-, *k-	*m- / *p-	*y-
Trans- N. Guinean ³⁶	*na	*ka	*ni	*ki
Australian ³⁷	*ŋa-	*ŋin	incl *ŋali excl *ŋay	*nhurra
Algic ^{38 39}	*nV / *-Vn, v *-Vk v obj *-Vy	*-Vm(?), *-Vt v obj *-Vš	*-nān	*-wāw
Keresiouan	*(V)nā-			
Hokan	*n ^y a / *n ^y i	*mi / *ma		
Penutian	*n-	*mV		
Aztec-Tano ⁴⁰	*neʔ	*ʔeme	*ta-	*ʔeme
Chibcha	*na-sV	*mue-ya(nV)		
Quechua	*nu-qa	*qa-m		
Panoan		*mi	*no	*mato
Nilo-Saharan ⁴¹	*āgā *āh	*īni	incl *ānà excl *ām	*wém
Kordofanian ⁴²	*ji	*ŋa	excl *-r(i)- incl *-n(a)-	*-j(a)-
Atlantic ⁴³	*mi obj *ne	*mo obj *'i, *mV	*tV (сев.) *sV ?	*nV
Mande (S) ⁴⁴	*N̄, *mā	*ī / *ē, *bV	excl *yī/ō incl *koa	*kā

²⁷ [ZORC 1995: 1169, 1191]; [BENEDICT 1975: 203–209].

²⁸ [CABATON 1905: 336–337]; [SHORTO 2006: 69]; [PINNOW 1966: 184]; [HAYES 1992: 161, 168]; [PEIROS 1998].

²⁹ [PINNOW 1966: 166–167].

³⁰ [BENEDICT 1975: 203, 205]; [PEIROS 1998].

³¹ [PORTMAN 1887: 5]; [BLEVINS 2007: 167]; [BURENHULT 1996].

³² [WATTERS 2006: 44].

³³ [WURM 1975: 191].

³⁴ [CAPELL 1975: 678].

³⁵ [LAYCOCK 1975: 768–770].

³⁶ [FOLEY 2000].

³⁷ [BLAKE 1979: 31–69]; [BLAKE 1988: 7]; [EVANS 1988: 103]; [EVANS 2003: 19]; [HARVEY 2003].

³⁸ [PROULX 1985: 84–87].

³⁹ American Indian language forms below are based on [GREENBERG–RUHLEN 2007: 121–123, 229–230], unless footnoted.

⁴⁰ [STEELE 1979: 447].

⁴¹ [EHRET 2001: 225–231].

⁴² [SCHADEBERG 1981: 182–184].

⁴³ [POZDNIAKOV–SEGERER 2004].

⁴⁴ [ВЫДРИН 2006].

Language Family	'I'	'Thou'	'We'	'You (Pl.)'
Ijo ⁴⁵	*a, *i/ε	*i	*wɜ	*o
Dogon ⁴⁶	<i>mi</i> poss <i>n</i>	<i>u-</i> / <i>uwV</i>	<i>be, eme</i>	<i>e</i> poss <i>ebV</i>
Gur-Adamaua ⁴⁷	*mi / *ma	*mo	*-a(n)	*-u(n) / *-ui(n)
**Gur ⁴⁸	*ma	*u / *bV	*tV	*i, *nV
Ubangi ⁴⁹	*mí	*mε	*ʔélé	*ʔéné
**Kwa ⁵⁰	subj *N obj *mε	*o, *wo	*ne ?	*mo / *mu
**Benue-Congo	subj *N obj *mi	*o, *u	excl *tV	
**Bantoid	*N obj *mε	*a, (obj?) *wo	*se	*ne, *we
Bantu ⁵¹	subj *Ni- obj *mε	subj *u / *we obj *ku	*tu *sV	*nV *mu
Khoisan ⁵²	N *mV, S *ηV C *ti- / *ta-	*a	excl *e incl (du) *mV	*-u

3. Another common paradigm, shared by the languages grouped within the Nostratic/Eurasianic hypothesis, is the “*K—M*” pattern for the 1st person stative (intransitive) vs. relative (transitive). The Uralic intransitive opposition of the 1st person markers *-*k* (intransitive) and *-*m* (transitive) ([ХЕЛИМСКИЙ 1982]) is usually compared with the Turkic preterite (1st pl.) *-*k* and present *-*m*, and the Eskimo absolutive *-*qa* together with relative *-*ma*. The Hittite “series” of verb conjugation in *-*hi* and *-*mi*, usually regarded as the ancient Indo-European opposition of stative and transitive, might also be derived from the Nostratic type, since there are clues for the velar (or postvelar) origin of the Indo-European laryngeal phoneme(s). The Nostratic postvelar form was proposed by ИЛИЧ-СВИТЫЧ ([ИЛИЧ-СВИТЫЧ 1971: 149]).

The syntactic function-related “*K—M*” opposition of the 1st person forms finds a similar correlation in the 2nd person, where the stative counterpart of “*K*” is usually “*T*”. The pattern “*K—T*” in its stative aspect is found in Dravidian appellatives ([ZVELEBIL 1990: 35–36]), in Indo-European (where *-*t-* is the 2nd person perfective affix in the same paradigm as the 1st person **H*), as well as in Elamite and Afrasian families, which are usually considered as the closest relatives of Nostratic or Eurasianic by long-range comparativists.

Again, it should be noted that this particular system is never reconstructed for any other proto-language around the world. This uniqueness, hardly representing the result of cross-borrowing, cannot be interpreted as anything other than a trace of genetic kinship between the languages of Eurasia.

4. Further evidence for another reconstructed macrofamily, Sino-Caucasian, or even broader, Dene-Caucasian, said to include North Caucasian, Sino-Tibetan, Yeniseian, Na-Dene languages and possibly some ancient tongues of Europe and the Near East (Hurrian, Urartian, Hatti, Basque) as well, is also found in the chart above. The pattern “*S/Z—W*” is easily traced in North Caucasian, Yeniseian, Na-Dene, and Hurro-Urartian languages. The primary consonant phoneme in the 1st person is not that clear, but it is definitely a coronal sibilant/affricate. Nahali data might be considered here as well.

It is interesting that the Sino-Tibetan proto-language paradigm does not fit this pattern. This discord might, however, be explained by the hypothesis that Proto-Sino-Caucasian, in fact, had two suppletive pronominal paradigms. STAROSTIN ([1984]) explicitly showed that the second set of pronouns in North Caucasian should be reconstructed as **nV* / **mV* (1st person sg.) – **yu* / **ywV* (2nd person sg.), which correlates well with such Sino-Tibetan forms as, respectively, **ηā* and **kwV*, as

⁴⁵ [WILLIAMSON 2004].

⁴⁶ [BENDOR-SAMUEL & al. 1989: 176]; [SEGERER 2002–2007].

⁴⁷ [BOYD 1989: 207].

⁴⁸ [BEACHAM 1968: 93]; [CASIMIR 1988: 57]; [PROST 1974: 646]; [RENNISON 2000].

⁴⁹ [MONIÑO 1995].

⁵⁰ [SEGERER 2002–2007].

⁵¹ [MEEUSSEN 1965]; [GUTHRIE 1967–1971, IV: 226–240]; [АКЦЕHOBA–ТОПOPOBA 1990: 146-150]; [NURSE–PHILIPPSON 2003].

⁵² [GREENBERG 1963: 70–74]; [STAROSTIN 2003]; [GÜLDEMANN 2005].

well as Yeniseian prefixed pronouns **b-* and **kV-*. The Burushaski forms might also fit here, as proposed by BENGTON ([1997]) in his cognate list for a theory of Burushaski as a Sino-Caucasian language. It should be noted that, typologically, both Yeniseian and Burushaski share the two-paradigm system of personal pronouns with North Caucasian.

5. Languages of South-Eastern Asia and the Pacific seem to share the pattern “*K—M*” (for the opposition between the 1st and 2nd singular persons), as shown by Tai-Kadai, Miao-Yao, Mon-Khmer, and Austronesian data. These are the families grouped by P. BENEDICT into the hypothetical Austric macrofamily. Additional similarities can be seen between single forms of pronouns in Austric languages (like **Vŋ* ‘I’); some of them are shared by Ainu, which is also considered to be related to Austric ([BENGTSON—BLAŽEK 2000]). The protoforms of pronouns for the Austro-Asiatic protolanguage are still foggy.

It is curious that the same “*K—M*” pattern also characterizes a number of languages in New Guinea: this is the type identified by WURM ([1975]) as “Set II” of the Papuan personal pronouns. It is emphasized that set II is supposed to mark the most archaic strata of human language in New Guinea; nowadays, it is used in languages of the Torricelli phylum and those languages of the Trans-New-Guinea phylum which show a strong substratum influence. Although genetic relations between Austric languages of Asia and the Pacific have not yet been thoroughly studied, this coincidence might become a hint for any further studies in this regard.

6. Languages viewed under the Niger-Congo macrofamily umbrella show distinct similarities in the pronominal system. A pattern “*M—W*” for the 1st and 2nd person singular pronouns is displayed by most of the language families of Niger-Congo. Moreover, at least one 1st person plural pronoun can probably be reconstructed as **tV* based on Bantu, Atlantic, and Gur languages. It seems to us that the Niger-Congo system actually consisted of two series of pronouns, one subject, one object (oblique), just as it looks like in numerous languages of practically all the families. The other 1st person singular pronoun **N* (representing a variable syllabic nasal) most probably denoted subject forms opposed to the object (oblique) **mi*.

Niger-Congo forms tend to show some basic similarity in the pattern (*N*[any nasal] — *W*), in the line of the hypothesis of their common genetic roots ([BLENCH 2007]). Kordofanian and Ijoid languages in this regard lie closer to Nilo-Saharan than to common Niger-Congo patterns. But there is still obvious lack in solid basis for the proto-forms, so these clues should be used carefully before reliable reconstructions are made for all sub-families of Niger-Congo.

7. Another striking similarity (“*T—N*”, basically 1st pers. **tV* vs. 2nd pers. **nV*, with plural forms in **-k*) lies between Kusunda, a language isolate of a forest gatherers’ tribe of central Nepal, and the West Papuan languages whose speakers are thought to represent the latest wave of Papuan migration to the island of New Guinea ([WURM & al. 1975: 947]). More anthropological and cultural similarities can be found between Papuans and Kusunda, so the two languages do deserve deeper linguistic comparison ([WHITEHOUSE & al. 2004]).

It goes without saying that the more we delve into language prehistory, the more observations like these will be formulated. Non-evident cognates may well be hidden behind phonetic changes, impossible to uncover without a thorough analysis.

However, some more general comments might be even more valuable from the point of view of linguistic typology.

It would be incorrect to call the choice of phonemes for personal pronouns a linguistic universal. Neither in proto-languages nor in living ones does there seem to be any inclination towards a “resonant — stop” phonematic distribution like the one proposed by J. NICHOLS. From a statistical point of view, it is obvious that there are plenty of proto-languages with both the 1st and the 2nd person pronouns based on nasals (even such a phenomenon as sharing the same nasal sounds for both persons is quite common in living languages, e. g. Korean), or both formed by stop consonants (Miao, Afrasian, Austronesian, Kusunda, &c.). There are also clear cases when both basic pronouns are built on the basis of sibilants. Although this statistical survey is relatively small, from a certain point of view it is more precise than one that would include all the living languages of the world, since the disproportionality in absolute language numbers among the various families would spoil the picture in the latter case. Thus, several hundred Austronesian languages will show phonetic values that will decisively overwhelm those of the single Basque language, but only because they are closely interrelated, while Basque is an isolate.

Subsequently, it is also incorrect to claim that “chance resemblance” can play an important part in pronominal comparison between languages of different families. There are absolutely no coinci-

dences in paradigm patterns between the languages which are not thought to be genetically related by modern long-range comparativists. One can certainly find similarities that may be hard to explain (like the above-mentioned Basque and Sumerian, or Papuan and Nivkh which share the “N—K” pattern). However, we are dealing here not just with single isolates instead of family proto-languages, but also with vastly diverse chronological periods: Basque and Sumerian data are separated by a time gap of at least 3000 years. It is only the comparison between chronologically close proto-languages of large families that may show us some clues in how to reconstruct deep linguistic prehistory.

References

- BEACHAM 1968 — C. G. BEACHAM. *The Phonology and Morphology of Yom*. A Thesis Submitted to the Faculty of the Hartford Seminary Foundation. Hartford.
- BEEKES 1995 — R. BEEKES. *Comparative Indo-European Linguistics: an Introduction*. Amsterdam—Philadelphia: John BENJAMINS.
- BENDOR-SAMUEL & al. 1989 — J. BENDOR-SAMUEL, E. OLSEN, A. WHITE. Dogon / J. BENDOR-SAMUEL (ed.) // *Niger-Congo Languages*. NY—London: LANHAM; pp. 169–177.
- BENEDICT 1942 — P. BENEDICT. Thai, Kadai and Indonesian: a New Alignment in Southeastern Asia // *American Anthropologist*, vol. 44, N^o. 4; pp. 576–601.
- BENEDICT 1972 — P. BENEDICT. *Sino-Tibetan: a Conspectus*. Cambridge.
- BENEDICT 1975 — P. BENEDICT. *Austro-Tai Language and Culture*. The Hague.
- BENGTSON 1992 — J. BENGTSON. Notes on Sino-Caucasian / V. SHEVOROSHKIN (ed.) // *Dene-Sino-Caucasian Languages*. Bochum: Brockmeyer; pp. 67–129.
- BENGTSON 1997 — J. BENGTSON. Ein Vergleich von Buruschaski und Nordkaukasisch // *Georgica*, 20; pp. 88–94.
- BENGTSON–BLAŽEK 2000 — J. BENGTSON, V. BLAŽEK. Lexical parallels between Ainu and Austric and their implications // *Archiv Orientální*, 68.2; pp. 237–258.
- BERGER 1974 — H. BERGER. *Das Yasin-Burushaski (Werchikwar)*. Wiesbaden.
- BERGSLAND 1986 — K. BERGSLAND. Comparative Eskimo-Aleut phonology and lexicon // *Journal de la Société Finno-Ougrienne*, vol. 80; pp. 63–137.
- BLAKE 1979 — B. BLAKE. Australian Case Systems: Some Typological and Historical Observations / S. A. WURM (ed.) // *Australian Linguistic Studies*. Canberra; pp. 323–394.
- BLAKE 1988 — B. BLAKE. Redefining Pama-Nyungan // *Aboriginal Linguistics*, vol. 1; pp. 1–90.
- BLAŽEK 1991 — V. BLAŽEK. The Microsystems of Personal Pronouns in Chadic, Compared with Afroasiatic // *Studia Chadica et Hamitosemitica*. Aktes den internazionalen Symposions zur Tschadsprachenforschung, Frankfurt am Main, 6–8 Mai 1991. Köln; S. 36–57.
- BLENCH 2007 — R. BLENCH. Further Evidence for a Niger-Saharan Macrophyllum // *Advances in Nilo-Saharan Linguistics*. Proceedings of the 8th Nilo-Saharan Linguistics Colloquium. Köln: Rüdiger KÖPPE Verlag; pp. 12–24.
- BLEVINS 2007 — J. BLEVINS. A Long Lost Sister of Proto-Austronesian? Proto-Ongan, Mother of Jarawa and Onge of the Andaman Islands // *Oceanic Linguistics*, vol. 46, N^o. 1; pp. 155–198.
- BOYD 1989 — R. BOYD. Adamawa-Ubangi / J. BENDOR-SAMUEL (ed.) // *Niger-Congo Languages*. NY—London: LANHAM; pp. 178–215.
- BURENHULT 1996 — N. BURENHULT. Deep Linguistic Prehistory with Particular Reference to Andamanese // *Lund University Working Papers*, 45; pp. 5–24.
- CABATON 1905 — A. CABATON. Dix dialectes indochinois recueillis par Prosper ODEND’HAL // *Journal Asiatique*, mars-avril 1905; pp. 265–344.
- CAMPBELL 2003 — L. CAMPBELL. How to Show Languages Are Related: Methods for Distant Genetic Relationship / B. JOSEPH, R. JANDA (eds.) // *The Handbook of Historical Linguistics*. Oxford: BLACKWELL; pp. 262–282.
- CAMPBELL–MITHUN 1979 — Introduction: North American Indian Historical Linguistics in Current Perspective / L. CAMPBELL, M. MITHUN (eds.) // *The Languages of Native America: Historical and Comparative Assessment*. Austin & London: University of Texas Press; pp. 3–69.
- CAMPBELL–POSER 2008 — L. CAMPBELL, W. POSER. *Language Classification. History and Method*. Cambridge: Cambridge University Press.

- CAPELL 1975 — A. CAPELL. The West Papuan Phylum / S. A. WURM (ed.) // *Papuan Languages and the New Guinea Linguistic Scene*. (= *Pacific Linguistics*, Series C, N° 38). Canberra; pp. 667–716.
- CASIMIR 1988 — K. K. CASIMIR. *Lexicon der Tagbana-Sprache*. Dissertazion zur Erlangung des Grades eines Doktors der Philosophie an der Fakultät für Linguistik und Literaturwissenschaft an der Universität Bielefeld. Bielefeld.
- COOKE 1968 — J. R. COOKE. Pronominal Reference in Thai, Burmese and Vietnamese // *University of California Publications in Linguistics*, 52.
- CYSOUW 2003 — M. CYSOUW. *The Paradigmatic Structure of Person Marking*. Oxford: Oxford University Press.
- DED — T. BURROW, M. B. EMENEAU. *A Dravidian Etymological Dictionary*. Oxford, 1984.
- DIXON 1980 — R. DIXON. *The Languages of Australia*. Cambridge: Cambridge University Press.
- EDAL — S. STAROSTIN, A. DYBO, O. MUDRAK. *An Etymological Dictionary of Altaic Languages*. HdO Vol. 8-1, 8-2, 8-3. Leiden: BRILL, 2003.
- EHRET 2001 — C. EHRET. *A Historical Comparative Reconstruction of Nilo-Saharan*. Köln: KÖPPE.
- EVANS 1988 — N. EVANS. Arguments for Pama-Nyungan as a Genetic Subgroup, with Particular Reference to Initial Laminalization // *Aboriginal Linguistics*, 1; pp. 91–110.
- EVANS 2003 — *The Non-Pama-Nyungan languages of northern Australia: Comparative Studies of the Continent's Most Linguistically Complex Region* / N. EVANS (ed.). Canberra: "Pacific Linguistics".
- FOLEY 1986 — W. FOLEY. *The Papuan Languages of New Guinea*. Cambridge: Cambridge University Press.
- FOLEY 2000 — W. FOLEY. The Languages of New Guinea // *Annual Review of Anthropology*, October 2000, Vol. 29; pp. 357–404.
- GREENBERG 1966 — J. GREENBERG. *The Languages of Africa*. Indiana University, Bloomington. MOUTON & C°, The Hague, Netherlands.
- GREENBERG 2000 — J. GREENBERG. *Indo-European and Its Closest Relatives: The Eurasiatic Language Family*. Vol. 1: *Grammar*. Stanford: Stanford University Press.
- GRUZDEVA 1998 — E. GRUZDEVA. *Nivkh*. München.
- GUTHRIE 1967–1970 — M. GUTHRIE. *Comparative Bantu. An Introduction to the Comparative Linguistics and Pehisotry of the Bantu Languages*. Vol. I (1967), II (1971), III (1970), IV (1970). London.
- GÜLDEMANN, ELDERKIN 2005 — T. GÜLDEMANN, E. ELDERKIN. On External Genealogical Relationships of the Khoe Family / M. BREZINGER, C. KÖNIG (eds.) // *Khoisan languages and linguistics: the Riezlern symposium 2003*. Quellen zur Khoisan-Forschung 17. Köln: Rüdiger KÖPPE.
- HARVEY 2003 — M. HARVEY. Reconstruction of Pronominals among the Non-Pama-Nyungan Languages / N. EVANS (ed.) // *The Non-Pama-Nyungan languages of northern Australia: Comparative Studies of the Continent's Most Linguistically Complex Region*. Canberra: "Pacific Linguistics"; pp. 475–513.
- HAYES 1992 — L. V. HAYES. On the track of Austric. Part I: *Introduction* // *MKS XXI*; pp. 143–177.
- HONTI 1993 — László HONTI. Zur Morphologie ugrischer Personalpronomina // *MSFOU* 215, 121–127, Helsinki.
- ITABASHI 1998 — ITABASHI Yoshizo [板橋 義三]. The Old Japanese Personal Pronouns as an Etymological Problem // *UAJ*, Vol. 70.
- KLIMOV 1998 — G. A. KLIMOV. *Etymological Dictionary of the Kartvelian Languages*. Berlin: MOUTON DE GRUYTER.
- KRISHNAMURTI 2001 — Bh. KRISHNAMURTI. *Comparative Dravidian Linguistics*. Oxford University Press.
- LAYCOCK 1975 — D. C. LAYCOCK. The Torricelli Phylum / S. A. Wurm (ed.) // *Papuan Languages and the New Guinea Linguistic Scene* (= *Pacific Linguistics*, Series C, N° 38). Canberra; pp. 767–780.
- LAYCOCK–Z'GRAGGEN 1975 — D. C. LAYCOCK, J. Z'GRAGGEN. The Sepik-Ramu Phylum / S. A. WURM (ed.) // *Papuan Languages and the New Guinea Linguistic Scene* (= *Pacific Linguistics*, Series C, N° 38). Canberra; pp. 729–763.
- LI FANG-KUEI 1977 — LI FANG-KUEI [李方桂, Lǐ Fāngguì]. *A handbook of Comparative Tai*. Honolulu: University Press of Hawaii.
- MEEUSSEN 1965 — A. E. MEEUSSEN. *Reconstructions grammaticales du bantou*. Tervuren: Musée Royal de l'Afrique Centrale (MRAC).

- MONIÑO 1995 — Y. MONIÑO. *Le Proto-Gbaya : Essai de linguistique comparative historique sur vingt-et-une langues d'Afrique centrale*. Louvain-Paris : PEETERS.
- NICHOLS 2003 — J. NICHOLS. Diversity and Stability in Language / B. JOSEPH, R. JANDA (eds.) // *The Handbook of Historical Linguistics*. Oxford: BLACKWELL; pp. 283–310.
- NURSE–PHILIPPSON 2003 — D. NURSE, G. PHILIPPSON (eds.) / *The Bantu Languages*. London – NY: ROUTLEDGE *Language Family Series*.
- PEIROS 1998 — I. PEIROS. *Comparative Linguistics in Southeast Asia*. Canberra: Pacific Linguistics.
- PINNOW 1966 — H.-J. PINNOW. A Comparative Study of the Verb in the Munda Languages / N. ZIDE (ed.) // *Studies in Comparative Austroasiatic Linguistics*. The Hague: MOUTON & C^o.; pp. 96–193.
- PORTMAN 1887 — M. W. PORTMAN. *Manual of the Andamanese Languages*. Delhi.
- POZDNIAKOV–SEGERER 2004 — K. POZDNIAKOV, G. SEGERER. Reconstruction des pronoms atlantiques et typologie des systèmes pronominaux // *Systèmes de marques personnelles en Afrique*. (= *Collection «Afrique et Langage»*, 8); pp. 151–162.
- PROST 1974 — A. PROST. Le Nõtrea, langue des Nõtrea dits Bulba // *Les langues de l'Atakora*. Bulletin de l'IFAN, T. XXXVI, sér. B, n^o. 3; pp. 628–659.
- PROULX 1985 — P. PROULX. Proto-Algic II: Verbs // *International Journal of American Linguistics*, Vol. 51, N^o. 1. (Jan., 1985); pp. 59–93.
- RÉDEI 1988 — K. RÉDEI. *Uralisches Etymologisches Wörterbuch*. Akadémiai Kiado, Budapest.
- RENNISON 2000 — R. RENNISON. *Koromfe–English/French/German Dictionary* [Forthcoming].
- RUHLEN 1994 — M. RUHLEN. First- and Second-Person Pronouns in the World's Languages // *On the Origin of Languages: Studies in Linguistic Taxonomy*. Stanford: Stanford University Press.
- SCHADEBERG 1981 — T. SCHADEBERG. *The Heiban Group*. Vol. 1: *The Survey of Kordofanian*. Hamburg: Helmut BUSKE Verlag.
- SEGERER 2002–2007 — G. SEGERER. *Les marques personnelles dans les langues africaines [base de données en ligne du LLACAN: <http://sumale.vjf.cnrs.fr/pronoms/>]*.
- SHAFFER 1941 — R. SHAFER. Nahali: A Linguistic Study in Paleoethnography // *Harvard Journal of Asiatic Studies*, v. 34.
- SHIBATANI 1990 — SHIBATANI Masayoshi [柴谷 方良]. *The Languages of Japan*. Cambridge: Cambridge University Press.
- SHORTO 2006 — H. L. SHORTO. *A Mon-Khmer Comparative Dictionary*. Canberra.
- STAROSTIN 2003 — G. S. STAROSTIN. A Lexicostatistical Approach towards Reconstructing Proto-Khoisan // *Mother Tongue*. VIII; pp. 83–128 [“81–126” in the on-cover TOC].
- STEELE 1979 — S. STEELE. Uto-Aztecan: An Assessment for Historical and Comparative Linguistics / L. CAMPBELL, M. MITHUN (eds.) // *The Languages of Native America: Historical and Comparative Assessment*. Austin & London: University of Texas Press: pp. 444–544.
- STRAHLENBERG 1730 — J.-F. STRAHLENBERG. *Das Nord- und Östliche Theil von Europa und Asia, in so weit solches das gantze Russische Reich mit Siberien und der grossen Tatarey in sich be-greiffet in einer historisch-geographischen Beschreibung der alten und neuern Zeiten*. Lübeck.
- THOMASON–EVERETT 2005 — S. THOMASON, G. EVERETT. Pronoun Borrowing // *Berkeley Linguistics Society*, 27; 301–315.
- TOPPING 1973 — D. M. TOPPING. *Chamorro Reference Grammar*. Honolulu: Hawaii University Press.
- TRASK 1997 — L. TRASK. *The History of Basque*. London–NY: Routledge.
- TROMBETTI 1905 — A. TROMBETTI. *L'unità d'origine del linguaggio*. Bologna.
- TUITE 1997 — K. TUITE. Svan // *Languages of the World / Materials*, 139. München–Newcastle: LINCOM Europa.
- VOVIN 1993 — A. VOVIN. *A Reconstruction of Proto-Ainu*. Leiden: BRILL.
- WATTERS 2006 — D. WATTERS. Notes on Kusunda Grammar: a Language Isolate of Nepal // *Himalayan Linguistics Archive*, 3 (2006); pp. 1–182.
- WHITEHOUSE & al. — P. WHITEHOUSE, T. USHER, M. RUHLEN, W. S.-Y. WANG. Kusunda: an Indo-Pacific Language in Nepal // *PNAS*, Vol. 101, N^o. 15, April 2004. <http://www.pnas.org/content/101/15/5692.full>.
- WILLIAMSON 2004 — K. WILLIAMSON. *Proto-Ijoid Reconstructions* [Forthcoming].

- WURM 1975 — Personal Pronouns / S. A. WURM (ed.) // *Papuan Languages and the New Guinea Linguistic Scene*. (= *Pacific Linguistics*, Series C, N° 38). Canberra; pp. 191–217.
- WURM & al. — S. A. WURM, D. C. LAYCOCK, C. L. VOORHOEVE, T. E. DUTTON. Papuan Linguistic Prehistory, and Past Language Migrations in the New Guinea Area / S. A. WURM (ed.) // *Papuan Languages and the New Guinea Linguistic Scene*. (= *Pacific Linguistics*, Series C, N° 38). Canberra, 1975; pp. 935–960.
- ZORC 1995 — D. ZORC. A Glossary of Austronesian Reconstructions / D. T. TRYON (ed.) // *Comparative Austronesian Dictionary, an Introduction to Austronesian Studies*. Part I, fasc. 2. Berlin; pp. 1105–1197.
- ZVELEBIL 1990 — K. ZVELEBIL. *Dravidian Linguistics : An Introduction*. Pondicherry.
- АКСЁНОВА–ТОПОРОВА 1990 — И. С. АКСЁНОВА, И. Н. ТОПОРОВА. *Введение в бантуистику (Имя. Глагол.)* [*An Introduction to Bantu Linguistics: Noun, Verb*]. М.: «Наука».
- БАБАЕВ 2008 — К. В. БАБАЕВ. *Происхождение индоевропейских показателей лица* [*The Origins of the Indo-European Personal Markers*]. М.–Калуга: «Эйдос».
- ВЫДРИН 2006 — В. Ф. ВЫДРИН. Личные местоимения в южных языках манде [Personal Pronouns in Southern Mande] // *Труды Института лингвистических исследований*. Т. 2, Ч. 2. СПб: «Наука»; стр. 327–413.
- ДЪЯКОНОВ 1979 — И. М. ДЪЯКОНОВ. Эламский язык [The Elamite Language] // *Языки Азии и Африки*. Т. III. М.; стр. 37–49.
- ДЪЯКОНОВ 1967 — И. М. ДЪЯКОНОВ. *Языки древней Передней Азии* [*Languages of the Ancient Southwest Asia*]. М.
- ИЛЛИЧ-СВИТЫЧ 1971 — В. М. ИЛЛИЧ-СВИТЫЧ. *Опыт сравнения ностратических языков (семитохамитский, картвельский, индоевропейский, уральский, дравидийский, алтайский): Введение. Сравнительный словарь* [*A Tentative Comparative Dictionary of the Nostratic Languages (Semitic-Hamitic, Kartvelian, Indo-European, Uralic, Dravidian, Altaic)*]. Т. I (b–Қ). М.: «Наука», Главная редакция восточной литературы.
- КАНЕВА 2006 — И. Т. КАНЕВА. *Шумерский язык* [*The Sumerian Languages*]. СПб.: «Петербургское Востоковедение».
- МЕНОВЩИКОВ 1997 — Г. А. МЕНОВЩИКОВ. Азиатских эскимосов язык [The Language of Asian Eskimos] // *Языки мира. Палеоазиатские языки*. М.: «Индрик»; стр. 75–80.
- МУДРАК 2000 — О. А. МУДРАК. *Этимологический словарь чукотско-камчатских языков* [*An Etymological Dictionary of Chukchi-Kamchatkan Languages*]. М.: «Языки русской культуры».
- НИКОЛАЕВА–ХЕЛИМСКИЙ 1997 — И. А. НИКОЛАЕВА, Е. А. ХЕЛИМСКИЙ. Юкагирский язык [The Yukaghir Language] // *Языки мира. Палеоазиатские языки*. М.: «Индрик»; стр. 75–80.
- СТАРОСТИН 2006 — Г. С. СТАРОСТИН. Еще раз к вопросу о личных местоимениях в дравидийских языках [Once again on personal pronouns in Dravidian languages] // *Аспекты компаративистики*. Вып. 2 [*Aspects of Comparative Linguistics. Issue 2*] (= *Orientalia et Classica. Труды Института восточных культур и античности*. Вып. XI [*Papers of the Institute of Oriental and Classical Studies. Issue XI*]). М.
- ХАЙДУ 1985 — П. ХАЙДУ. *Уральские языки и народы* [*The Uralic languages and peoples*]. М.: «Прогресс».
- ХЕЛИМСКИЙ 1982 — Е. А. ХЕЛИМСКИЙ. *Древнейшие венгерско-самодийские языковые параллели: Лингвистическая и этногенетическая интерпретация* [*The Earliest Hungarian-Samoyed Language Parallels: Linguistic and Ethnogenetic Interpretation*]. М.: «Наука».
- ШМИДТ 1984 — К. Х. ШМИДТ. Типологическое сопоставление систем картвельского и индоевропейского глагола [Typological Comparison of Kartvelian and Indo-European Verbal Systems] // *Вопросы языкознания*, 33; стр. 48–57.

Р е з ю м е

Статья посвящена вопросу о том, насколько значимым является сопоставление данных личных местоимений и местоименных показателей для исследований в области дальнего родства языков. Утверждается, что, хотя отдельные, бессистемные сходства между местоимениями в праязыках неглубокого уровня могут быть объяснены случайным сходством или как заимствования, такая ситуация гораздо менее вероятна при сопоставлении целых местоименных парадигм. Основываясь на кратком обзоре основных форм личных местоимений и местоименных показателей, реконструируемых специалистами для подавляющего большинства праязыков, автор делает вывод, что многие из ныне существующих гипотез о дальнем родстве (таких, как ностратическая, сино-кавказская, австрическая и т. п.) подтверждаются наличием как раз такого рода парадигматических сходств.

Языковые контакты и генеалогическая классификация

Статья посвящена вопросу об элементах языковой конвергенции в процессе становления новых языков, по мнению автора, все более актуальному по мере накопления сведений о «смешанных» языках (пиджинах и креолах), несмотря на постоянное игнорирование или недооценку его со стороны представителей традиционной компаративистики. Показано, что существующие модели классификации, основанные на морфологических и лексических (лексикостатистических) критериях, зачастую оказываются бессильными из-за невозможности детального учета таких элементов. При этом проблема оказывается актуальной не только для изначально неродственных языков, подвергшихся пиджинизации, но и для языков одной семьи, дивергенция которых одновременно сопровождалась интенсивными взаимными контактами (т. н. «проблема генеалогического пня», согласно термину А. Б. Долгопольского).

Статья «Языки мира» в Лингвистическом энциклопедическом словаре — как это и принято — представляет собой обзор родственных связей языков. Обзору предшествует пояснение: каждая языковая семья «происходит из группы близких друг к другу диалектов, которые в древности были диалектами одного языка или входили в один языковой союз (группу территориально близких языков, обладающих совокупностью общих черт)» [ИВАНОВ 1990₂: 609] (курсив мой. — В. Б.). Такое определение языковой семьи внутренне противоречиво. В самом деле, представим, что в глубокой древности имелось два в любом смысле неродственных между собой языка Я и L, каждый из которых не имел диалектных различий. Со временем язык Я эволюционировал в группу близких диалектов А, Б, В, Г, Д, Е, Ж, а язык L — в группу диалектов А, В, С, D, E, F. Носители диалектов Б, Г, Ж и С, F довольно рано отделились от своих родственников, а носители диалектных цепей А—В—Д—Е и А—В—D—E мигрировали в направлении друг друга:

$$\text{Б Г Ж} \quad \Rightarrow \text{А—В—Д—Е} \Rightarrow \quad \leftarrow \text{А—В—D—E} \leftarrow \quad \text{С F}$$

При этом диалекты Д и D остались на периферии нового ареала, а А, В, Е, А, В и Е заключили языковой союз на несколько тысячелетий:

$$\begin{array}{ccccccc} \text{Б} & \text{Г} & \text{Ж} & & \text{Д} & \text{—————} & \text{А} & \text{————} & \text{В} & \text{————} & \text{Е} \\ & & & & & & | & & & & | \\ & & & & & & \text{А} & \text{————} & \text{В} & \text{————} & \text{Е} & \text{—————} & \text{D} & & \text{С} & \text{F} \end{array}$$

Языковой союз позднее распался, и входившие в него языки эволюционировали в течение нескольких тысячелетий самостоятельно.

Как мы будем классифицировать современные языки А, Б, В, Г, Д, Е, Ж, А, В, С, D, E, F, восходящие к соответствующим древним диалектам? Может показаться, что теоретически существуют разные решения:

С двумя семьями. Языки А, Б, В, Г, Д, Е и Ж восходят к праязыку *Я, а языки А, В, С, D, E и F — к праязыку *L. В этом случае предки языков, входящих в каждую из постулируемых семей, в отдаленной древности «были диалектами одного языка».

Или с тремя. Языки Б, Г и Ж составляют одну семью, языки С и F — другую; принцип объединения тот же: предки современных языков «в древности были диалектами одного языка». Третью семью составляют языки А, В, Е, А, В и E, объединяющиеся на том основании, что их предки «в древности входили в один языковой союз». Положение языков Д и D оказывается спорным, поскольку наследство языкового союза в них выражено слабо. Для трех семей будет реконструировано три праязыка: *Я₁, *L₁ и *Ω.

Однако оба решения «неверны». В первом случае мы целиком проигнорировали вторую половину определения (относительно языкового союза), во втором — проигнорировали его первую часть для половины языков. При буквальном следовании определению мы, при наличии достаточной информации, неизбежно должны принять третье решение и констатировать

наличие трех семей: двух независимых, в каждой из которых предки современных языков «в древности были диалектами одного языка» (1. А, Б, В, Г, Д, Е, Ж и 2. А, В, С, D, E, F), и пересекающейся с ними третьей семьи, образуемой языками, чьи предки «в древности входили в один языковой союз» (А, В, Е, А, В и Е).

На практике такое решение никогда не принимается. Но на практике общность языков, чьи предки всего лишь находились в интенсивном контакте, компаративист никогда и не называет языковой семьей. Происхождение реверанса в сторону языкового союза понятно: это всего лишь дань уважения имени Н. С. Трубецкого, именно *имени*, а не взглядам, выраженным в известной статье 1939 г. [Трубецкой 1987]. Но в этой работе Трубецкой предстает кем угодно, но никак не компаративистом, что убедительно показал В. А. Дыбо [2003: 69–73].

Важный постулат компаративистики гласит: «между родственными языками ⟨...⟩ имеются лишь регулярные соответствия, закономерность которых и устанавливается в результате введения системы-посредника, т. е. в результате реконструкции праязыка» [В. А. Дыбо 2003: 71]. Вопрос о том, как соотносится праязык, выявленный с помощью определенных технических приемов (назовем его *п р а я з ы к о м*₁), с реальными коммуникативными системами прошлого, далеко не прост. В период романтической юности сравнительно-исторического языкознания считалось, что реконструированный праязык₁ и реальный язык давних предков (назовем его *п р а я з ы к о м*₂) идентичны; опытный компаративист может научиться говорить на нем и даже создавать литературные произведения, которые могли бы быть восприняты пра-родителями с той же легкостью, с какою сочинялись их потомком-компаративистом.

Позднее стараниями младограмматиков выяснилось, что воссоздать реальные праязыки₂ вряд ли возможно, а социологическая школа вывела эту проблематику за пределы науки: «лингвисту нет вообще другого дела кроме как истолковывать системы соответствий, устанавливаемые между разными языками ⟨...⟩ которые одни представляют осязаемую реальность и, следовательно, единственный предмет сравнительной грамматики индоевропейских языков» [Мейе 1938: 107].

В рамках «истолковывания системы соответствий» продолжали уточняться реконструкции архетипических праформ, но они рассматривались лишь как удобные этикетки для соответимых элементов современных языков, являющихся потомками... Чьими потомками? Разумеется, какого-то праязыка₂, в исторической реальности которого не сомневался, пожалуй, никто, кроме Н. С. Трубецкого.

Итак, за демонстративным устранением праязыка₂ из предметной области компаративистики скрывалось известное лукавство, поскольку постулирование его существования и дивергентного развития по-прежнему оставалось фундаментом научной компаративистики.

Что же касается реконструированных праязыков₁, основного результата сравнительно-исторических исследований, то большинство компаративистов, видимо, никогда не считали их отвлеченными абстрактными конструктами, предназначенными лишь для удобства описания синхронных соотношений между языками. За ними всегда виделась историческая реальность: это были по возможности точные приближения к *п р а я з ы к а м*₂ — реальным коммуникативным средствам древних народов; степень возможного приближения праязыка₁ к соответствующему праязыку₂ зависит от полноты данных. В идеале — впрочем, недостижимом — конструкт и реальность должны быть идентичны. Неоднозначность реконструкций элементов праязыка₁ часто объясняется диалектными различиями в праязыке₂, но согласно классической компаративистской модели дивергировавший на диалекты праязык — это поздний этап развития ранее единообразного общего языка, не имевшего диалектного членения; именно эту стадию, единый общий язык, реальное существование которого постулируется компаративистской моделью языковой эволюции, и логично называть праязыком₂.

Графической интерпретацией этой модели служит генеалогическое древо — имеющий одну исходную вершину ориентированный «граф, в котором два и более узлов могут иметь единого предка, но никакой узел не может иметь более чем одного предка» [Бурлак, Старостин 2001: 47]. Каждый нетерминальный узел такого графа соответствует праязыку определенного уровня глубины, а все восходящие к нему терминальные узлы символизируют дочерние по отношению к этому праязыку языки, образующие некую общность (семью, ветвь, группу и т. п.).

Теоретически процесс реконструкции описывается следующим образом: «Уральский праязык — это результат компаративистской процедуры, проведенной над финно-угорским и самодийским реконструированными праязыками, которые сами являются результатом той же процедуры, проведенной над реконструкциями же: над праприбалтийско-финским, прасаамским, прамордовским, прамарийским, прапермским, праугорским и над прасеверносамодийским, праселькупским, прасаяносамодийским» [В. А. Дыбо 2003: 77]. Эта картина, с одной стороны, идеализированна (на практике компаративисты не дожидались и даже не испытывали нужды — по крайней мере поначалу — в реконструкции прамордовского или праселькупского, чтобы приступить к реконструкции вышестоящих праязыков), с другой — в качестве идеальной неполна; при классификации финно-угорских языков обычно упоминаются и промежуточные таксоны: *финно-пермские языки* рядоположены угорским, а *обско-угорские* — венгерскому (см., например, соответствующие статьи в ЛЭСе или генеалогические деревья этой семьи в «*Основах финно-угорского языкознания*» [1974: 38–39]).

Все упомянутые в предыдущем абзаце праязыки — это научные конструкторы, праязыки₁. Им должны соответствовать постепенно все далее расходящиеся между собой цепочки непрерывных языковых традиций, представленные на синхронных срезах реальными праязыками₂, и ведущие от общеуральского₂ к, скажем, общехантыйскому₂ и общеэнецкому₂, — языкам реально не засвидетельствованным, однако породившим, в частности, современные хантыйские и энецкие диалекты. Существенно отметить, что скудость информации практически никогда не позволяет выяснить в деталях, что и как происходило с единой языковой традицией в ходе ее продвижения к более позднему таксономическому этапу, когда мы констатируем наличие двух (или нескольких) заведомо разошедшихся традиций.

Ранняя компаративистика исходила из того, что эволюция от праязыка₂ любой исторической глубины к современным языкам — процесс сугубо дивергентный. «Потом выяснилось, что такой путь сравнительно редок. Обычно диалекты, разграничиваясь в один период, снова сближаются впоследствии» [Шайкевич 1995: 206]. Древодвидная модель эволюции языков представляет нам ее как дискретный процесс от одного узла к другому: отделившись от родственного, некий идиом выглядит как монолитный вплоть до дивергенции в следующем узле. Но говорить о подлинном единстве языковой традиции можно лишь в применении к моменту, который соотносится с самым началом ветви, исходящей из узла дивергенции. До следующего узла, когда языковой традиции предстоит «бесповоротно» разделиться на две или несколько самостоятельных, проходят многие сотни лет эволюции, в ходе которой языковая традиция уже не монолитна, а ее близкородственные варианты находятся в достаточно сложном дивергентно-конвергентном взаимодействии. Такое развитие представляется естественным не только для диалектов одного языка, но и для диалектов недавно разошедшихся языков. Диалект одного языка, оказавшийся волею судеб в сфере влияния другого (но близкородственного) языка, в ходе эволюционного процесса может постепенно стать диалектом этого последнего. Так, сотню лет назад восточнославянские говоры всей Витебской и запада Смоленской губ. обоснованно считались белорусскими, но те из них, что оказались на территории РСФСР, столь же обоснованно признаются сейчас (велико)русскими¹, точно так же нидерландские говоры района Эммерих—Везель—Гельдерн, окончательно оказавшегося с 1815 г. в составе Пруссии, постепенно перешли на положение немецких говоров, притом что в течение всего XIX в. литературный нидерландский кое-где продолжал использоваться как язык школы и церкви [Plank 1988].

Если мы будем принимать во внимание лишь языковые ситуации, разделенные большими промежутками времени, мы, вроде бы, вправе сказать, что некоторая популяция сменила один родной язык на другой: жители Велижского уезда, став подданными Екатерины II, говорили по-белорусски, а их отдаленные потомки, родившиеся при Б. Н. Ельцине, говорят по-русски, аналогичные перемены произошли и с языком крестьян из-под Гельдерна от времен Фридриха Вильгельма III до Г. Коля. Но при более подробной временной шкале совершенно

¹ БССР была образована 01.01.1919 из части Минской губ.; в 1924–1926 гг. ее территория поэтапно увеличивалась, однако вся бывшая Смоленская, три уезда Витебской (Себежский, Невельский, Велижский) и небольшая часть Могилевской губ. остались в составе РСФСР.

очевидна непрерывность языковой эволюции: не носители меняют язык, а сам язык меняется, причем во многих случаях столь медленно, что изменения не осознаются носителями.

Сходные процессы заведомо могли иметь место и в прошлом, в связи с чем факт существования некоторых праязыков₂ оказывается под сомнением. Вернемся к уральскому примеру. Вряд ли можно усмотреть основания для постулирования существования единого, диалектно не расчлененного обскоугорского праязыка₂ после миграции языковых предков современных венгров на запад. Реальность же обскоугорского единства связана с продолжением постоянных контактов территориально смежных этнических групп, говоривших на близкородственных идиомах; благодаря контактам здесь удерживались определенные ретенции и распространялись специфические инновации, которые и составляют суть «обскоугорских» характеристик современных языков. Неменьшие сомнения возникают и в существовании финноугорского праязыка₂ как единого идиома: еще в общеуральский период праязык₂ представлял собой диалектный континуум, причем «его „досаамодийские“ и „доугорские“ диалекты обладали рядом общих черт, которые отличали их от „допермских“ и особенно от „дофинно-волжских“ диалектов» [Хелимский 1991: 87], подробнее см. [Хелимский 1982].

Такое положение кажется вполне естественным, когда языковая эволюция происходит на едином пространстве, пусть и обширном, но, так сказать, «континуальном», не имеющем серьезных внутренних границ. Но конвергентные процессы, не позволяющие расклассифицировать родственные языки строго древовидным образом, характерны и для «дискретно» распространенных языков, например для островной Океании, см. [Беликов 1989].

Подобные факты, конечно, не новость для компаративистики; вот какое теоретическое осмысление получают они у одного из ведущих современных компаративистов. При исследовании близкородственных идиомов, реконструированные исторические взаимоотношения которых суммарно отражены на некоем генеалогическом древе, могут выявиться «следы архаического явления», в результате тщательного исследования его распределения по рассматриваемым идиомам производится их классификационная перегруппировка, и «мы получаем другое, причем, очевидно, более старое дерево с другими узлами» [А. В. Дыбо 2003: 62]. В этом контексте под *старым деревом* следует понимать такое, где исконные дивергентные процессы не затемнены последующими конвергенциями; именно оно должно отвечать идеалу компаративистской теории. А под *новым* — вероятно, такое дерево, на конфигурации которого отразилось взаимовлияние начавших расходиться близкородственных идиомов. Компаративистика всегда начинает с построения «новых»² деревьев. Позднее в отдельных (редких) случаях удается вскрыть детали исторических взаимоотношений между предками современных идиомов. Казалось бы, можно приступить к построению «старого» дерева, но эта задача в полном объеме оказывается неосуществимой.

Идея о разделении славянских языков на три группы в массовом сознании приобрела характер общего места. Однако специалисты с недавних пор считают положение более сложным: «родословное древо по отношению к славянским, как и применительно ко многим другим языкам, является удобным схематическим упрощением, но оно в очень малой степени отражает реальные исторические процессы развития диалектов» [Иванов 1990: 96]. Как альтернативу «новому» дереву с тремя подгруппами славянских языков Вяч. Вс. Иванов приводит «схему связей между диалектами праславянской языковой области», которая «тоже является условной, но для определенного периода (ок. сер. 1-го тыс. и несколько ранее) она могла отвечать определенной исторической реальности» [ibid.]. Взаимоотношение семи узлов этой схемы с реально существующими через полторы тысячи лет их потомками таково, что попытка построения «старого» генеалогического дерева, отражающего «следы архаических явлений», обречена на неудачу. Остановлюсь лишь на одном пункте. Как было убедительно показано, севернокривичский диалект тысячу лет назад ни фонетически, ни морфологически не обладал особой близостью к другим диалектам восточнославянской территории, а «неосуществление второй палатализации»³ отличает древний севернокривичский диалект не только от восточнославянских, но и от всех прочих славянских вообще» [За-

² Чтобы избежать путаницы, термины «новое» и «старое» дерево, по А. В. Дыбо, ниже будут ставиться в кавычки.

³ Переход заднеязычных *k, *g, *x в свистящие, при отсутствии которого современным русским *целый* и *серый* соответствовали бы слова с начальными k и x. — В. Б.

лизняк 1995: 37]. Этот диалект не вымер, его традицию продолжают современные говоры значительной части Псковской и Новгородской областей, постепенно конвергировавшие с другими восточнославянскими говорами и в ходе формирования русского языка ставшие его частью. Появление севернокривичского узла на «старом» славянском дереве исторически оправдано, однако трудно надеяться, что все псковичи легко поверят в генеалогию, согласно которой их язык от языка тверитян и смолян отстоит столь же далеко (если не далее), как от языка жителей Праги и Сараева.

Сходной оказывается ситуация с любой разветвленной группировкой языков, если удастся провести ее детальный диахронический анализ. Вот еще один пример. По отражению различных черт арийского консонантизма древние иранские диалекты, представлявшие собой языковой континуум (и продолжающие их современные языки), образуют два типа группировок: по одному признаку «вычленяется центральный ареал и несколько маргинальных», по другому — «западная и восточная группы диалектов». При этом «лексический уровень показывает иное членение континуума» [Эдельман 2003: 181, 182].

Когда-то в начале 1970-х А. Б. Долгопольский на одной университетской лекции так образно охарактеризовал невозможность построения дивергентного дерева для разветвленной группы языков, современное состояние которых отражает многочисленные следы их взаимовлияний: «...у всех языков имеются генеалогические деревья, а у тюркских — генеалогический пень». Перефразируя, можно сказать: попытка преобразования «нового» генеалогического дерева в «старое» без обращения к конвергенции приводит к созданию генеалогического пня. По этой причине компаративисты избегают строить «старые» деревья, даже понимая условность многих узлов «новых»: «промежуточные праязыки являются некоторой схематизацией, полезной при формулировке выявленных генеалогической классификацией языков соотношений, но не обязательно отвечающей некоторой исторической реальности» [Иванов 1990₁: 96]. Последнее утверждение нуждается в уточнении: за такого рода «искусственными» праязыками₁ не стоят реальные праязыки₂, но историческая реальность не ограничивается фактически функционировавшими в прошлом языковыми системами; взаимодействие синхронно существовавших языков, приводившее к разного рода конвергентным явлениям, не менее реально. Именно подобными процессами и обусловлена специфика реконструируемых праязыков₁. Впрочем, такую «контактную реальность» понимать надо достаточно абстрактно: при реконструкции языковой эволюции нечасто удастся разнести разновременные явления, поэтому на классификациях современных языков, особенно исторически не документированных, отражается их некая суммарная панхронная контактная история, и именно ее выражением являются классификационные таксоны типа *обскоугорские языки* или *восточнославянские языки*. Сопоставим две приведенные выше цитаты из Вяч. Вс. Иванова: с одной стороны, «родословное древо (...) в очень малой степени отражает реальные исторические процессы», а с другой — оно полезно «при формулировке выявленных генеалогической классификацией языков соотношений». Второе справедливо именно потому, что классификационная дендрограмма очень хорошо суммарно отражает всякого рода реальные исторические процессы.

Правда, классификационная дендрограмма — это «новое» дерево, заметно отличное от идеального с точки зрения компаративистики «старого» родословного древа. Это разные объекты, причем первый для компаративистики — продукт побочный. В. А. Виноградов, проводя аналогию между эволюционными таксономиями в биологии и лингвистике, указывает, что «понятия „генетический“ и „генеалогический“, часто употребляемые в лингвистике как синонимы, в биологии различаются: первый предполагает изучение структуры организма в терминах генов (минимальных единиц функции), второй — изучение собственно филогенетической истории вида. По-видимому, разграничение этих понятий было бы полезно и в лингвистике, где генетический анализ означал бы тот аспект сравнительно-исторического изучения языков, который связан с установлением регулярных соответствий, а генеалогия — общую историю языков» [Виноградов 1982: 260]. В этих терминах родословное (оно же «старое») древо — продукт генетического анализа, а генеалогическая классификация — отражение «общей истории языков».

Пока речь шла о языках, ретроспектива которых мыслится как непрерывное развитие некоей единой языковой традиции. Считается, что если бы мы могли совершить путешествие по времени от праиндоевропейского по каждой из его дочерних ветвей до современности, то ни-

где не обнаружили бы проблем во взаимопонимании между дедом и внуком или между ровесниками, говорящими на идиомах, срок разделения которых исчисляется десятилетиями.

Марк Твен и его современник Р. Киплинг вряд ли испытывали серьезные затруднения при чтении произведений друг друга или английских фольклорных текстов, рождавшихся по обе стороны Атлантики. Однако едва ли кто-то из них без перевода или серьезной предварительной подготовки полностью понял бы смысл такого образчика народной лирики начала XX в. [Laycock 1977: 61]:

Yu meri wantok,
Yu giamanim mi tasol.
Yu raitim nem bilong mi
Ananit long lek bilong yu⁴.

Самые ранние следы зарождения «ответвления английского», на котором написаны эти поэтичные строки, появились при жизни родителей Марка Твена и Р. Киплинга, а как самостоятельный язык он сформировался уже после того, как младший из двух современников стал публиковаться; к концу жизни классиков этот язык — ток-писин стал языком повседневного бытового общения части населения Новой Гвинеи, языком церкви и администрирования.

О непрерывности языковой традиции тут говорить не приходится.

Ток-писин относится к числу креольских языков⁵ — типу «смешанных» языков, понаслышке известному всем лингвистам. Креольские языки, как и смешанные языки других типов, обычно выпадают из общих обзоров типа «классификация языков мира», поскольку в таких текстах всегда излагаются генеалогические схемы языковых семей, а размещение этих языков по генеалогическим деревьям сопряжено с очевидными натяжками. Игнорирование компаративистикой этих языков лишь отчасти связано с тем, что им трудно подобрать место в классификации, главное то, что они равным счетом ничего не могут дать для реконструкции праязыков разного уровня, поэтому компаративистам совсем неинтересны⁶. Креолисты отвечают компаративистике взаимностью: рассуждая о родственных связях креольских языков, они абсолютно игнорируют парадигму сравнительно-исторического языкознания и имеют свои (и разные) взгляды на то, что следует называть родством.

Дело в том, что креольские и другие контактные языки (о которых речь пойдет далее) есть результат слияния исторически не связанных языковых традиций, причем в случае креольских языков эти традиции нельзя считать непрерывными.

При эволюции единой языковой традиции для повторного слияния двух начавших расходиться идиомов требуется их взаимопонимание. Но взаимопонимание обусловлено не столько сходством двух идиомов, сколько языковой ситуацией, в которой оказались их носители. Монолингвам для достижения взаимопонимания необходимо, чтобы единственно известные им идиомы действительно были очень близки, чуть ли не тождественны, минимальный опыт межъязыковой коммуникации быстро «портит» монолингва, у него появляются навыки восприятия иного идиома, то есть навыки билингвизма. Билингвы же достигают взаимопонимания независимо от материального и типологического сходства используемых ими языков. Взаимопонимание, если оно массовое, часто ведет к взаимосближению. При конвергенции генетически далеких или условно неродственных языковых традиций (различных *языков*, не *диалектов*) происходит их типологическое и материальное схождение. Оно может быть односторонним или в различной степени взаимным (тут уместно вспомнить такие термины, как *заимствование*, *языковой союз*, *субстрат*, *суперстрат* и т. п.). Из периода такого взаимодействия языко-

⁴ Вот перевод, близкий к буквальному: *Ты, женщина моего народа, ты обманула меня, вот и все. [А ведь] ты написала мое имя на внутренней поверхности своих бедер.*

⁵ Строго говоря, основания считать ток-писин креольским языком появляются лишь во второй половине XX в., но при обсуждении генетической квалификации контактных языков противопоставление *креолов* и *расширенных пиджинов* оказывается несущественным, так же как и противопоставление путей приобретения «креольского статуса» — *постепенная креолизация vs. резкая креолизация*. Содержание этих терминов будет уточнено ниже.

⁶ Точнее, они бесполезны для «частных компаративистик»: креолы, обязанные своим происхождением индоевропейским языкам, не помогают индоевропеистике и т. п. Но сами креольские языки в ряде случаев оказываются сгруппированными в семейства, все члены которых восходят к единому пиджину; исследование взаимоотношений членов таких семейств для теоретической компаративистики могло бы быть небезынтересным.

вые традиции выходят с разной степенью «чистоты» и жизнеспособности, вплоть до прекращения существования одной из них. Но в отдельных случаях сближение различных языковых традиций приводит к их слиянию. Наиболее распространенный тип такого слияния — образование функционально ограниченных языков межэтнической коммуникации, пиджинов.

Пиджины не результат полноценного двуязычия, а альтернатива ему. Они возникают для решения ограниченных коммуникативных задач при отсутствии готового инструмента, пригодного для этой цели. Каждая из контактирующих сторон пытается использовать по возможности упрощенную форму известного ей языка, при этом готовность идти на уступки в пользу варианта, предлагаемого контрагентом, зависит от заинтересованности в эффективности коммуникации. В результирующем смешанном языке преобладает компонент, связанный с языком менее заинтересованной стороны. Однако «выстраивает» новую языковую систему именно наиболее заинтересованная («побежденная» в языковом отношении) сторона, чей вклад в новую языковую систему оказывается также значительным. Дальнейшие судьбы пиджинов могут быть различны, нас интересует тот путь, на котором они превращаются в функционально полноценные языки. Это происходит, когда для некоторого социума пиджин становится единственным языком внутригруппового общения; в его рамках такой *расширенный* (*extended*, или *expanded*) пиджин начинает обслуживать все коммуникативные потребности, обычно *креолизуясь*, то есть становясь родным (усваиваемым в младенчестве) языком⁷. Креолизации может подвергнуться не только расширенный пиджин, но и более ранние стадии его развития, в том числе и начальная («жаргонная»)⁸.

До начала расширения/креолизации пиджин располагает максимум тремя-четырьмя сотнями корней, большая их часть восходит к одному языку — *лексификатору*; этот бедный «пракреольский» словарь включает не более 70–80 единиц из списка Сводеша. Дальнейшее наращивание лексикона идет за счет внутренних ресурсов, «естественных» ономастопоэтических образований и заимствований [Беликов 1991₂]. Кроме престижного суперстратного языка-лексификатора источником заимствований (не только лексических) становятся также родные для носителей нового языка субстратные языки. Последних, как правило, много, вклад каждого не особенно заметен⁹, но обычно они оказываются родственны и/или типологически близки, поэтому осмысленно говорить об их суммарном влиянии на структуру контактного языка. Если контакт бывшего пиджина с лексификатором продолжается, то в процессе заимствования происходит регуляризация звуковых корреспонденций и грамматическая «подстройка» под престижный язык. В конце концов конвергентное взаимодействие креола с языком-лексификатором может перевести его на положение чего-то близкого к диалекту этого последнего. Креолы часто попадают в поле зрения лингвистов только на этой стадии. Но есть радикальные креолы, которые разрывают связь с лексификатором на время или навсегда. В приложении приведены 100-словные списки двух таких языков.

Сарамакка креолизовался в джунглях Суринама на рубеже XVII–XVIII вв. среди беглых рабов, язык общения которых подвергся «двойному смешению»: большинство из них прибыло в Суринам с опытом общения на английском пиджине, а на плантациях осваивало местный португальский креол, который был родным языком меньшинства «лесных негров». Результатом конвергенции стал новый креол, где единицы португальского и английского происхождения численно соотносятся как 1 : 2. Языком административного управления Суринама тогда и позднее был голландский. Ток-писин восходит к сложившемуся к середине XIX в. в южной части Тихого океана английскому пиджину бичламар, процесс его медленной креолизации начался также при доминировании английского языка, но до этого в начале XX в. при германской администрации на Новой Гвинее ток-писин уже перешел на стадию расширенного пи-

⁷ В ряде языковых ситуаций первые носители креольской разновидности такого контактного языка становятся дву- и многоязычными, часто в детстве, но это не меняет сути: креольский язык остается их родным языком, пусть и наряду с другими.

⁸ Последний случай называется *резкой креолизацией* (*early creolization* у Д. Биккертона, *abrupt creolization* у С. Томасона), которой противопоставлена *постепенная* (*gradual*) креолизация.

⁹ Редким исключением является креол бербис (на р. Бербис, восток Гайяны), где в 100-словном списке Сводеша 57 единиц этимологизируются из нидерландского и 38 — из восточного иджо (Нигерия) [KOWENBERG 1996:1351]. Подчеркну, что такие подсчеты креолистов не дают оснований для лексико-статистических оценок; «этимологизируются из» вовсе не означает, что эти единицы совпадают с таковыми в языке-источнике: как увидим ниже, семантические отличия могут быть разительными.

джина. В позднем ток-писине началась фонетическая и семантическая «подстройка» под английский, однако на базисной лексике она мало отразилась.

Фонетические соответствия между языком-лексификатором и радикальным креолом далеки от регулярности, ср. передачу английского дифтонга в ток-писине и сарамакка то как *o* (*ó*), то как *u* (*ú*):

английский	ток-писин	сарамакка	значение	№
<i>bone</i> ['bəʊn]	<i>bun</i>	<i>bónu</i>	‘кость’	(10) ¹⁰
<i>nose</i> ['nəʊz]	<i>nus</i>	<i>núsu</i>	‘нос’	(61)
<i>smoke</i> ['sməʊk]	<i>smok</i>	<i>sumúku</i>	‘дым’	(78)
<i>stone</i> ['stəʊn]	<i>ston</i>	<i>sitónu</i>	‘камень’	(81)

В стословном списке немало свидетельств того, что создатели пиджина, весьма приблизительно восприняв форму некоторой частотной единицы целевого языка, наполняли ее столь же приблизительно с точки зрения лексификатора семантикой. В родовом значении ‘птица’ (6) в обоих языках используются единицы, имеющие в лексификаторе более конкретную семантику. Глаголы ‘сидеть’ (74) и ‘стоять’ (79) восходят к приказам *sit down!*, *stand up!*. Угроза *убить* со стороны носителя языка-лексификатора обычно заканчивалась лишь нанесением побоев, поэтому неудивительно, что глагол *kilim* приобрел в ток-писине значение ‘бить’, в значение же ‘убить’ (43) используется *mekim i dai* (букв. «сделать мертвым») или же *kilim i dai* (‘избить до смерти’). Этимологом для слова со значением ‘полный’ (32) послужило выражение *full up* ‘переполненный, битком набитый’. Вопросительные местоимения восходят к готовым фразам: *wanem* (< *what name?*) ‘что’ (96) и *husat* (< *who's that?*) ‘кто’ (98). В ситуации складывания пиджина коммуникативная потребность в понятии ‘готовить пищу’ появляется раньше, чем в глаголе ‘жечь’, неудивительно, что способ выражения второго понятия (12) восходит к первому (ср. *Paya i kukim haus* ‘пожар сжег дом’). Обозначение единичного объекта может восходить к форме множественного числа, если таковое более употребимо (т.-п. ‘зуб’ (89) — *tit* < *teeth*, сарам. ‘ухо’ (21) — *jési* < *ears*). Специфика контактной ситуации проявляется в том, что «эталон», доступный первым носителям пиджина, часто оказывается далеким от нормативного языка-лексификатора. Носители последнего в процессе коммуникации широко использовали метафору, часто пейоративную, создатели же пиджина «вычисляли» семантику, исходя из общечеловеческих допущений и понятийной структуры собственных языков, не осознавая ни метафоричности, ни пейоративности. К результатам этого процесса в стословном списке относится наполнение внешней оболочки стандартного английского имени *Mary* семантикой «женщина вообще» (99)¹¹. Сходных единиц немало и за пределами стословного списка, ср. *manki* ‘мальчик, не прошедший обряда инициации’ (из *monkey* ‘обезьяна’), *bagarapim* ‘сломать, испортить, вывести из строя’ (из слэнгового *bugger up* ‘загубить, испортить’ от *bugger* ‘заниматься содомией’).

Стословный список хорошо иллюстрирует способ, каким развивающийся язык «набирает» лексикон. Здесь широко представлены единицы со значительным расширением семантики, в обоих языках одинаково выражаются единицы ‘кусать’ (7) и ‘есть’ (23), ‘кора’ (3) и ‘кожа’ (75), при необходимости вносится конкретизация: так, собственно ‘кора’ может «уточняться» как *skin bilong diwai* и *páu kákisa* (ср. №90, ‘дерево’: т.-п. *diwai*, сарам. *páu*). Такого рода «уточнения» возможны и для других единиц стословного списка, но их параллели находятся за его пределами. Так, в ток-писине *gras*, *sit* и *skru* «сохраняют» значения этимологов (‘трава’, ‘экскременты’ и ‘винт, винтовое соединение’, откуда ‘сустав’), а для значений ‘перо’ (27), ‘волосы’ (36), ‘пепел’ (2), ‘колени’ (44) возможны «уточненные» номинации *gras bilong pisin* (букв. ‘«трава» птицы’), *gras bilong het* (букв. ‘«трава» головы’), *sit bilong paia* (букв. ‘«экскременты» огня’), *skru bilong lek* (букв. ‘сустав ноги’)¹². Очевидно, первоначально у этих и подобных единиц было лишь то узкое значение, в котором имела коммуникативная потребность на стадии раннего пиджина (‘есть’, ‘кожа’ ‘трава’ и т. п.).

¹⁰ Здесь и ниже в круглых скобках приводится номер единицы стословного списка.

¹¹ Слово *meri* обозначает также любую особь женского пола, ср. *bulmakau meri* ‘корова’, *pukpuk meri* ‘самка крокодила’; исключение сделано лишь для европейской женщины, которая всегда именуется *misis*.

¹² Ср. также *gras bilong solwara* ‘водоросль’ (*solwara* — ‘море’, из *salt* + *water*), *i gat gras* ‘быть волосатым, покрытым шерстью’ (*gat* — ‘иметь’), *sit bilong lam* ‘сажа’ (*lam* — лампа), *skru bilong han* ‘локоть’ (*han* — ‘рука’), *mama bilong skru* ‘гайка’ (*mama* — ‘мать, бабка, тетка по отцу’).

В развивающемся пиджине нужда в обозначении культурных артефактов могла возникнуть раньше, чем появились стандартные способы именованья многого из того, что компаративисты вполне обоснованно относят к базисной лексике. В результате в ток-писине этимонами для представленных в стословном списке единиц 'сердце' (40) и 'рог' (41) послужили слова со значением 'часы' и 'гребенка', ср. еще *banis* (из англ. *fence*) 'забор', а также 'грудная клетка, ребро'. Такая неожиданная метафорика не чужда и языкам с непрерывной традицией, где она характерна в первую очередь для эмоционально окрашенного просторечия и жаргонов, ср.: *кумпол* и *котелок* 'голова', *грабки/грабли* 'руки', *шары* 'глаза' и т. п., но тут они имеют достаточно жесткую контекстуальную привязку типа фразеологической: *по кумполу получают* или *бьют*, *котелок не варит*, *грабки тянут*, *шары вылупляют*, но, скажем, сочетания [?]*прикрыть шары*, [?]*бить по котелку* или [?]*почесать кумпол левой грабкой* вряд ли допустимы (*чешут* обычно *репу*, причем без заполнения инструментальной валентности). Нормативными стандартными обозначениями соответствующих частей тела подобные слова становятся исключительно редко¹³, неслучайно в откровенно прескриптивном первом издании словаря ОЖЕГОВА¹⁴ [1949] нет ни одного из упомянутых слов, в менее ригористичном МАСе *котелок* в этом значении имеет помету *прост.*, *шары* — *груб. прост.*, но *грабки/грабли*, *кумпол*, *репа* отсутствуют¹⁵.

Следует иметь в виду, что многие случаи экзотической с точки зрения языка-лексификатора компоновки сем в одну лексему обязаны своим существованием не столько метафорическому расширению значения, сколько устройству понятийной структуры субстратных языков. Например, объединение сем *кора* и *кожа* в одной лексической единице — вполне стандартное явление для многих языков Африки и Океании, так что слова *skin* и *kákisa*, начав на ранней стадии формирования соответствующих языков употребляться в значении 'кожа', в сознании носителей складывающихся языков уже с самого начала обозначали также и кору дерева. В других случаях семантика единого понятия языка-лексификатора с точки зрения структуры субстратных языков оказывалась недопустимо обобщенной в сравнении с понятийной структурой субстратных языков, и, несмотря на предельную словарную бедность раннего пиджина, в нем появлялись излишние на взгляд европейца «дубликаты». Так, характерное для ток-писина противопоставление инклюзивного местоимения *yumi* эксклюзивному *mipela* фиксируется уже в раннем бичламаре [Беликов 1998: 66–69], при этом «строительный материал» для материализации субстратной семантики берется из языка-лексификатора (*you, me, fellow*). Ярче всего субстратный характер семантики креольских языков проявляется в хорошо структурированных лексических подсистемах. Вот один пример: создатели ток-писина усвоили внешние оболочки английских слов *brother* и *sister*, но наполнили их более привычным содержанием: *brata* переводится как 'брат (мужчины), сестра (женщины)', а *susa* — как 'брат (женщины), сестра (мужчины)'¹⁶. Сходным образом дело обстоит и с грамматическими элементами, формально восходящими к лексификатору, но содержательно отражающими структуру субстратных языков, например, суффикс транзитивности *-im* в ток-писине (19, 33, 39, 72) восходит к английскому местоимению *him* 'его', но содержательно соответствует рефлексам праокеанийского **-aki* в родных языках большинства создателей контактного языка. Выбор между транзитивной и нетранзитивной формой глагола (да и вообще его употребление) обусловлен концептуальной структурой мира океанийцев и часто далек от того, как употребляется в лексификаторе формальный этимон. Например, гла-

¹³ Ср. русск. *глаз*, первоначально значившее как раз 'шарик'.

¹⁴ В разделе «От редакции» словарь назван «общедоступным пособием, которое призвано содействовать повышению культуры речи широких народных масс» [Ожегов 1949: III].

¹⁵ В созданном на основе МАСа «Большом толковом словаре», претендующем на включение «лексикой всех стилистических пластов и литературных жанров, в которых функционирует русский язык на протяжении всего 20 века», в том числе и единиц, для которых характерно «фактическое использование (...) в устной речи» [БТС 1998: 3], стилистическая норма менее строга: *прост.* превращается в *разг.*, *груб. прост.* — в *разг.-сниж.*; к слову *репа* без помет появляется пояснение «о глупой, плохо соображающей голове» (и семантическое ограничение, и отсутствие пометы вряд ли можно считать лексикографической удачей), слов *кумпол* и *грабки* (а также менее маркированного *грабли*) нет по-прежнему.

¹⁶ Глагол *переводится* использован неслучайно: с европейской точки зрения слова *brata* и *susa* многозначны, однако в действительности это не так, *brata* в с е г д а обозначает сиблинга того же пола, а *susa* — сиблинга противоположного пола.

гол *dring(-im)* 'пить' в словосочетании *dring susu* 'сосать грудь' оказывается непреходным и выступает в бессуффиксальной форме, поскольку *susu* 'грудь', не поглощаясь в процессе «питья», не является прямым объектом. В других контекстах это переходный глагол: *dringim sup* 'есть суп', *dringim suga* 'жевать сахарный тростник', *dringim smok* 'курить'.

Итак, для обоих рассматриваемых языков связь с английским бесспорна, но какова она? В рамках собственно компаративистики проблема родственных связей креолов практически не обсуждается, зато среди креолистов начиная с 1950-х годов идет непрерывная дискуссия на эту тему. Д. Тейлор возродил идеи Г. Шухардта о двойном родстве креольских языков, утверждая, что от одного «родителя» они получают в наследство словарь, а от другого — грамматику [Тейлор 1972: 492]. Родство первого типа он предлагал называть генетическим, а родство второго типа — глубинным (*basic*). Р. Холл [1972] и У. Вайнрайх [1972] не видели достаточных оснований для противопоставления креолов языкам иного генезиса и настаивали на индоевропейском характере вест-индских креолов. Структурное, а иногда и материальное сходство в грамматике англо- и франкокреольских языков Антильских островов Р. Холл объяснял в первую очередь схожим африканским субстратом, а У. Вайнрайх — конвергентными процессами. Р. Томпсон [1972] даже выдвинул гипотезу о родстве всех креольских языков между собой и возможности возвести общекреольскую грамматику к раннему португальскому морскому пиджину, а через него — к средневековому средиземноморскому пиджину, известному как сабир, или лингва-франка. Происходившая при этом многократная смена основной части словаря при сохранявшейся в основном неизменной грамматической структуре получила наименование релексификации. Противники теории релексификации считали креолы специфическим результатом эволюции соответствующих романских и германских языков; для объяснения структурных отличий от португальского, французского, голландского и других участвовавших в этом процессе языков был введен термин *реграмматикализация*, под которой понималась смена грамматики при сохранении словаря. Дискуссия о родственных связях креольских языков продолжается по сей день, ее участники находят все новые аргументы и контраргументы. Начиная с конца 1970-х акцент сдвинулся на обсуждения генезиса грамматических систем; традиционным точкам зрения была противопоставлена гипотеза биопрограммы Д. Биккертона [BICKERTON 1977; 1981; 1988], согласно которой важнейшие структурные характеристики «подлинных» креолов, восходящих к ранним нестабильным пиджинам, имеют своим истоком генетически обусловленные принципы устройства человеческого языка¹⁷.

Конечно, ни «глубинное родство» по Тейлору, ни «грамматическое родство» языков, прошедших релексификацию, с точки зрения сравнительно-исторического языкознания родством не называется. Является ли родством связь креольского языка с его лексификатором — вопрос непростой. Для выявления родства компаративистика требует регулярности фонетических соответствий, в первую очередь на основном словарном фонде, причем отсеиваются любые заимствования, в том числе и из родственных языков; сходство грамматических формантов желательно, но совсем необязательно; степень родства между языками принято оценивать по числу закономерных совпадений в списке Сводеша (опять же без учета заимствований).

Сложности начинаются с отделением исконной лексики от заимствований; чтобы это сделать, исследователь должен располагать детальными сведениями о составе лексикона контактного языка на тот момент, с которого он склонен отсчитывать его историю как самостоятельной языковой системы. Ясно, что креолизовавшийся или расширенный пиджин являются таковыми, но их диахроническая ретроспектива континуальна вплоть до начальной стадии возникающего пиджина, когда его фонетика и лексика не сложились еще в устойчивую систему. Достаточно полные сведения о словарях стабильных пиджинов немногочисленны, но имеющиеся данные однозначно указывают: эти языки не располагают всеми единицами списка Сводеша. На достаточ-

¹⁷ Гипотеза Биккертона пережила пик популярности; по мнению С. Томасон, сейчас она «считается жизнеспособной теорией в основном среди формально-ориентированных лингвистов, не занимающихся креолистикой» [THOMASON 2001: 178]. Для Биккертона вопрос о родстве креольских языков не стоит: онтогенез креольской грамматики повторяет филогенез языка человечества, и перед значительностью этого события вопрос о происхождении лексических наполнителей генетически заданных универсальных грамматических структур оказывается попросту нерелевантным.

но позднее появление ряда базисных слов указывает и типологический анализ лексикона ряда независимо возникших креолов [Беликов 1987]¹⁸. На рассматриваемом материале весьма вероятно позднее появление единиц 'колено' (45) и 'кровь' (9); первое в ток-писине обозначается перифрастически, а в сарамакка заимствовано из нидерландского *knie*; второе в обоих языках заимствовано, соответственно, из нем. *Blut* и нидерл. *bloed* (если слова креолов возводить к англ. *blood* ['bld], придется постулировать неожиданный и нигде более не отмеченный переход [ʌ] > [u]).

Другой вопрос: насколько правомерно сопоставлять контактные языки с нормативными вариантами лексификаторов? Эталоном при создании пиджинов были не они, а их диалектно и жаргонно окрашенные «родственники», стословные списки которых могли иметь свою специфику. Соотношение слов *pluma* и *pena* в нормативном португальском известно: преобладает второе, в двуязычных словарях оно всегда оказывается первым, изредка единственным переводом слов типа *feather*, *перо* и т. п., однако мы не знаем, каково было положение в территориальных и социальных вариантах португальского языка XVI–XVII вв., служивших лексификаторами для контактных языков. Между тем, ставим ли мы «плюс» при сопоставлении срм. *puíma* (27) с его португальским эквивалентом, зависит именно от этого¹⁹.

При анализе единиц *save* (45) и *pikinini* (73) в ток-писине велик соблазн не считать их восходящими к языку-лексификатору. Соответствующие слова находятся на периферии общеанглийского словаря, о чем говорят и колебания в орфографии: *savvy* ~ *savvey* 'понимать' и *pickaninny* ~ *picaninny* ~ *piccaninny* 'негритенок' (часто воспринимается как «неполиткорректное»), оба имеют иберороманскую этимологию; по ВЕБСТЕРУ, первое фиксируется с 1785, второе — с 1653 г. Однако не исключено, что в жаргонах английских моряков Нового времени оба слова входили в основной словарный фонд и именно из этого регистра проникли в «общий» словарь²⁰. Соответствующие единицы имеются во всех англокреольских языках Атлантики в значении 'знать' и 'маленький' и/или 'ребенок'; их первоисточником явно служил португальский пиджин, но дальнейшее направление воздействия установить здесь невозможно (в английский морской жаргон через английские пиджины или наоборот). Что касается тихоокеанских креолов, здесь все ясно: оба слова в значениях 'знать' и 'ребенок' отмечены в прибрежном тихоокеанском пиджине середины XIX в. и представлены во всех его потомках (ток-писин, бислама Новых Гебрид, пиджин Соломоновых островов²¹, брокен Торрессова пролива, североавстралийский криол) [CLARK 1979]; [Беликов 1998: 66–68]. Таким образом, *save* (45) ток-писина соответствует *savvy* лексификатора, как и *pikinini* в значении 'ребенок', однако то же слово в значении 'семя' (73) — инновация ток-писина, призванная заполнить лексическую лауну раннего пиджина²².

Вышесказанное показывает известную условность лексико-статистических сопоставлений контактных языков и их лексификаторов, но если по возможности аккуратно такие подсчеты произвести, то английский окажется в равной степени лексически близким ток-писину и другим германским (около 70% совпадений), а уровень близости ток-писина с датским или немецким на 20% ниже. Попарные сопоставления между английским, русским и таджикским

¹⁸ Любопытно, что словарные лакуны в зарождающемся и в вымирающем языке оказываются сходны: первыми забываются как раз те единицы, которые отсутствуют в ранних пиджинах [POLINSKY 1996].

¹⁹ Ср. сложности при выборе эквивалента единице *belly* (4) для восточнославянских диалектов: *живот* ~ *брюхо* ~ *пузо* ([С. Л. Николаев: устное сообщение]).

²⁰ Над речью относительно замкнутого профессионального сообщества не особенно довлела необходимость придерживаться стандарта, и здесь действовал обычный для жаргона закон: семантически стертые частотные слова легко вытеснялись эмоционально притягательными заимствованиями. П. Мюльхойслер приводит наблюдение очевидца, хотя и относящееся к ситуации на Тихом океане в XIX в., однако общезначимое: «Капитан и команда забавляются смешением языков. Они слышат то или иное чужое слово и начинают его использовать вместо соответствующего английского. Так, на нашем корабле вместо 'быстро' всегда говорят *quillequille*, вместо 'есть' — *kaikai*, вместо 'мясо' — *bulmakau* и так далее» [MÜHLHÄUSLER 1986: 98]; напомним, что единицы 'есть' и 'мясо' входят в 100-словный список.

²¹ Это креолизовавшийся расширенный пиджин с самоназванием *Pijin*.

²² Различить значения *pikinini* помогает контекст, ср.: *pikinini bilong sayor* 'овощные семена', *pikinini paul* 'цыпленок', *no save karim pikinini* 'быть бездетным' (букв. 'не уметь приносить (= рожать) детей'); как видно из последнего примера, глагол *save* также многозначный, ср. еще значения 'понимать', 'иметь обыкновение': *Mi save pinis* 'Я понял', *Mi no save giaman* 'Я не имею обыкновения врать', *Yu save smok?* 'ты куришь?'

дают 27–32 %, а сравнение ток-писина с русским и таджикским — 20–21 %. «Не зная истории формирования неомеланезийских языков [англокреольских языков юго-запада бассейна Тихого океана], подобную картину можно проинтерпретировать единственным образом: пранеомеланезийский отделился от общеиндоевропейского ранее расхождения между собой остальных групп (...) Существенная близость неомеланезийского словаря к германскому — результат сильного и длительного английского влияния» [Беликов 1991: 103–104].

Существенно отметить, что «английский» («германский», «индоевропейский») характер знаменательных и служебных морфем радикальных креолов касается в первую очередь плана выражения. Структуры значений отдельных лексических и грамматических единиц, как и структуры семантических полей, повторяют в контактных языках не соответствующие структуры языков-лексификаторов, а структуры субстратных языков. В таких условиях связь креольских языков с лексификаторами не хочется называть генетической.

Выше говорилось, что исходным пунктом для конвергенции, приводящей к возникновению креольских языков, служат начальные стадии формирования ситуации двуязычия, когда более заинтересованная в коммуникации сторона не стала продвигаться в освоении языка партнеров далее первого шага, поскольку минимальные коммуникативные задачи были решены. Другие типы контактных языков являются продуктом полноценного двуязычия. Относительно недавно они получили общее наименование «переплетенные» (*intertwined*) языки [ARENDS & al. 1994]²³. Все эти языки представляют собой результат конвергенции неродственных языковых традиций, причем каждая сохранилась в результирующем смешанном языке в отчетливо опознаваемом виде.

Ранее других из числа «переплетенных» языков в научный оборот были вовлечены некоторые цыганские криптолекты, сохранившие собственный основной словарный фонд, но использующие грамматику языка окружающего населения, «греческого, английского, датского, шведского, норвежского, окситанского, французского, испанского, бразильского португальского, каталанского, баскского, турецкого и западноармянского» [ВАККЕР, МУЙСКЕН 1994: 47]. Функционально эти «языки» вроде бы соответствуют искусственным тайным языкам типа офенского, близка и их структура, ср. пример из текста на **англоромани**, записанного в конце XIX в. [ВАККЕР, МУЙСКЕН 1994: 41]:

Palla bish besh-es apopli the Beng wel'd and pen'd: Av with man-di.
Через 20 лет-PL снова DEF Дьявол пришел и сказал: пойдём с я-DAT

Генезис таких цыганских «языков» не вполне ясен; по-видимому, их все же можно считать искусственными образованиями в том смысле, что неизвестная окружающим лексика намеренно сохранялась с конспиративными целями. Только что приведенный пример, будучи фрагментом сказки, может использоваться как контраргумент. Но, во-первых, об аутентичности текста можно говорить, лишь зная мелкие детали, сопровождавшие его порождение, поскольку в число полевых записей не так уж редко попадают тексты, специально сконструированные для исследователя и естественным образом никогда не порождаемые. Ср. запись несколько надуманного содержания, впервые опубликованную в 1820 г. и с тех пор часто встречающуюся в работах по офенскому языку: *Масу зетил ёный ховряк, в хлябом костре Москве мастырится клёвая оклюга. На мастырку эбетой биряют скень юс. Поерчим на масовском остряке и повершаем да пулим шивару* 'Мне сказывал один господин, что в столичном городе Москве строится чудесная церковь, на устройство которой делаются чрезвычайные пожертвования. Поедем и мы (туда) на моей лошади и посмотрим да купим товару' (цит. по: [БОНДАЛЕТОВ 1974: 77]). Во-вторых, собственный фольклор может существовать и в таких языках, искусственный характер которых не вызывает сомнений, ср. приводимые В. Д. Бондалетовым образцы песенного творчества на офенском «языке» владимирских коробейников, масовском «языке» одоевских торговцев, жгонском «языке» костромских шерстобитов [ивид.: 78–81]; «на спор» на эти «языки» могли переводиться и общерусские песни [ивид.: 19].

Символическая ценность языков типа англоромани как средства выражения групповой идентичности и их использование в этнически однородной среде также мало о чем говорит, поскольку то же самое характерно и для языков заведомо искусственных. Мало того, в определенных стереотипных ситуациях употребление искусственного языка было, по-видимому,

²³ П. Муйскен называет автором термина Н. Смита [MUYSKEN 2001:481].

обязательным. Ср. описание ситуации проводов офеней в дорогу, которую этнограф назвал бы совершением обряда перехода из мира *своих* во внешний мир *чужих*: «Хозяин ⟨...⟩ распоряжаясь угощением ⟨...⟩ выражается уже на искусственном офенском языке: „Поханка, привандай масыгам гомыры похлябе ⟨...⟩“» [Гольшев 1880].

В пользу искусственности языков типа англоромани у нас есть лишь один довод — существенная, возможно, ведущая роль функции сокрытия информации в их функциональном спектре. Что же касается генезиса неанглийской (неиспанской и т. д.) части их лексиконов, то тут важно помнить, что ее передача от поколения к поколению шла естественным путем и цыганская ее подсистема представляет собой результат эволюции непрерывной языковой традиции, в конечном счете восходящей к праиндоевропейскому.

Среди языков сходной структуры есть такие, которые не имеют «засекречивающей» функции. Примером может служить язык **эйну**, использующийся во внутригрупповой коммуникации небольшой этнической общности, живущей на востоке Китая в ряде пунктов вдоль южной ветви Великого шелкового пути от Кашгара до Хотана (на значительном расстоянии не только от персидско-таджикского региона, но и от ираноязычного ареала в целом). Все эйну двуязычны, еще в детстве усваивают уйгурский язык, которым они пользуются при любых внешних контактах. С фонологической и грамматической точек зрения язык эйну вполне укладывается в рамки уйгурского, диалектом которого он одно время считался. Характерно, например, сингармонически обусловленное распределение заимствованных из уйгурского показателей: *-lar/-ler* (мн. ч.), *-da/-de/-tal/-te* (локатив), *-raq/-rek* (сравнительная степень) и т. д. [LEE-SMITH 1996, 856], ср.:

эйну	уйгурский	значение
<i>xana-da</i>	<i>öj-de</i>	‘в доме’
<i>šeb-de</i>	<i>keč-te</i>	‘ночью’
<i>xurd-raq</i>	<i>kičik-rek</i>	‘меньше’
<i>kemte-rek</i>	<i>az-raq</i>	‘моложе’

Хотя практически все служебные морфемы эйну функционально и материально соответствуют уйгурским, взаимопонимание между двумя языками исключено в силу доминирования в эйну персидской лексики. Из 24 единиц стословного списка, встретившихся в процитированной работе, 15 персидских (*uštur* ‘живот’, *kalaŋ* ‘большой’, *xor-* ‘есть’, *ateš* ‘огонь’, *ser* ‘голова’, *bisjar* ‘много’, *kox* ‘гора’, *šeb* ‘ночь’, *nis* ‘нет’, *ječ* ‘один’, *kes* ‘человек’, *xurd* ‘маленький’, *saŋ* ‘камень’, *du* ‘два’, *ab* ‘вода’) и 6 уйгурских (*mən* ‘я’, *awu ~ ewu / ašu ~ ešu* ‘тот’, *bu* ‘этот’, *sen* ‘ты’, *nime* ‘что’, *kim* ‘кто’). Несмотря на ярко выраженную агглютинацию и сингармонизм, эйну придется признать индоевропейским (а именно иранским) языком. Если воспользоваться 35-элементным списком Яхонтова, соотношение иранских и тюркских элементов будет 6 : 6 (к подчеркнутым иранским добавится отсутствующее у Сводеша *bad* ‘ветер’).

Генезис эйну вполне ясен: персидский идиом, на протяжении многих столетий испытывавший сильную тюркскую интерференцию, в конце концов почти лишился исходной грамматики и половины базовой лексики. Происхождение тех или иных элементов языковой структуры эйну практически всегда очевидно, однако способ формирования из них единой системы характерен только для этого языка, и вопрос о его генетической принадлежности остается неопределенным; имея полный стословный список, формально его решить будет просто, но любое решение вызовет и у иранистов, и у тюркологов известное неудовольствие.

Еще более последовательно разведены словарь и грамматика в языке **медиаленгва** (центральный Эквадор). Совершенно очевидно, что лексически медиаленгва — ответвление испанского²⁴, но, даже не вникая в суть его агглютинативной морфологии и синтаксиса, еще более очевидно, что это всего лишь язык кечуа с испанским наполнением, ср. следующий пример [MUYSKEN 1997: 365], AC — аккузатив, NM — номинализатор, BN — бенефактив, PR — прогрессив, I — 1-е лицо:

²⁴ Приводимый П. Мюйскеном фрагмент базового словаря медиаленгва [1997: 421–423] содержит 51 единицу стословного списка Сводеша, среди них лишь 7 совпадают с местным кечуа (из которых 3 — испанские заимствования). Слов испанского происхождения — 47, но в 3 случаях имеются расхождения с нормативным испанским (*sirru* resp. *montaña* ‘гора’, *chikitu* resp. *pequeño* ‘маленький’, *bos* resp. *tú* ‘ты’).

Медиаленгва:	<i>Unu</i>	<i>fabur-ta</i>	<i>pidi-nga-bu</i>	<i>bini-xu-ni</i>	
Кечуа:	<i>Shuk</i>	<i>fabur-da</i>	<i>maña-nga-bu</i>	<i>shamu-xu-ni</i>	
	одно	одолжение-АС	просить-NM-BN	приходить-PR-I	
Испанский:	<i>Vengo</i>	<i>para</i>	<i>pedir</i>	<i>un</i>	<i>favor</i>
Русский:	<i>Я пришел</i>	<i>чтобы</i>	<i>попросить</i>	<i>об одном</i>	<i>одолжении</i>

Хотя родство представленных здесь индоевропейских языков (испанского и русского) в этом примере не прослеживается, их структурная противопоставленность другой паре языков видна вполне наглядно. Будучи языком исключительно внутрисемейного общения, медиаленгва функционально схожа с айну, но при внешних контактах его носители пользуются обоими «исходными» языками.

Этот язык возник в начале XX в. после постройки железной дороги Кито–Гуаякиль (1909); молодежь, получившая возможность работать по найму, активно осваивала престижный испанский язык, до того мало распространенный среди сельского индейского населения. «Прототип» медиа-ленгва, вероятно, представлял собой своеобразную манеру общения, ставшую символом групповой идентичности определенной части молодежи, воспринявшей элементы городской культуры и противопоставившей себя традиционному кечуанскому обществу. Типологические параллели хорошо известны²⁵, однако обычно членство в такого рода социальных группах (и использование их атрибутики) имеет возрастные ограничения. Здесь же часть «новых кечуа» мигрировала в города, где ассимилировалась в языковом и культурном отношении с испаноговорящими эквадорцами, а те, кто остался в традиционных местах обитания, обрели (по крайней мере в собственных глазах) более высокий социальный статус. Новый язык, символ этого статуса, стал главным средством внутригруппового общения. В рамках этого социума «сейчас для части детей первым выученным языком становится испанский, родным языком некоторых стариков является кечуа, но родной язык среднего поколения — медиа-ленгва. На медиа-ленгва говорят все, но только в пределах своего сообщества» [МУYSKEN 1997: 374].

Итак, «переплетенные» языки могут не только сложиться «сами собой», но и быть в известном смысле продуктом сознательного смешения независимых языковых традиций. Из числа таких языков наиболее экзотически выглядит **мичиф**, распространенный по обе стороны американо-канадской границы, в основном в Северной Дакоте и Саскачеване. Он возник в начале XIX в. среди потомков от браков индейских женщин с французскими торговцами мехами, называвшими себя *voyageurs*. Для нескольких сотен своих носителей мичиф — родной язык и часто единственный, который был им известен в детстве. Сейчас все говорящие на мичифе двуязычны и владеют английским, заимствования из которого проникают в современный мичиф. Вклад в мичиф алгонкинского языка кри, французского и английского иллюстрируется следующим примером [ВАККЕР, РАПЕН 1996: 1174]:

Un vieux opahikê-t ê-nôhchikê-t, êkwa un matin ê-waniskâ-t ahkosi-w, but kêyapit ana wî-nitawi wâpaht-am ses pièges. Sipwêhtê-w. Mêkwat êkotê tasîhkê-t, une tempête. Maci-kîsikâ-w. Pas moyen si-misk-ahk son shack. Wanisi-n.

‘Один старик был-охотником, охотился и однажды утром проснулся, заболел, но все же он хотел-идти-смотреть свои ловушки. Он пошел. Тем-временем там, [пока-]он-был-занят, [случилась] буря. Плохо-распогодилось. Нет способа найти-ему его хижину. Он-потерялся’.

Носители языка не могут выразиться на нем иначе, поскольку ни французские глаголы, ни существительные кри большинству из них неизвестны. «Коротко говоря, мичиф — язык с французскими существительными, числительными, артиклями и прилагательными, в то время как глаголы, указательные местоимения, послелогии, вопросительные слова происходят из кри» [ВАККЕР, РАПЕН 1996: 1172]. В разных общинах доля французских существительных колеблется от 83 до 94%, а глаголов кри — от 88 до 99% [ВАККЕР, МУYSKEN 1995: 45]. В действительности к кри восходит вся сложная глагольная система языка, а грамматически малонагруженные в тексте имена — французские. Регулярность звуковых соответствий между мичифом и французским,

²⁵ Ср. возникшие в среде русскоязычных хиппи тексты типа *Кабы я была кингица, спичит фёрстая герлица...*

равно как и между мичифом и кри, сомнению не подлежит. Остается выяснить, чей вклад в базисную лексику «основнее». Тут мы совершенно неожиданно оказываемся заложниками частеречной структуры стословного списка. Хотя известно, что глаголы заимствуются куда реже имен, семантические сдвиги у глаголов — явление вполне обычное. Вероятно, в связи с этим в стословном списке преобладают имена: на 54 существительных приходится 19 глаголов и 15 прилагательных²⁶, значит, французский компонент в мичифе автоматически становится более приоритетным. И в самом деле, подсчеты по словарю мичифа [LAVARDURE, ALLARD 1983] (без отсутствующего там слова №11 *breast*) дают 69 совпадений с французским стословным списком (для 16 из них имеются дублиеты из кри)²⁷; по 35-словному списку Яхонтова — 26 совпадений (6 дублиетов).

В ряде «переплетенных» языков граница между гетерогенными структурными компонентами проходит в пределах грамматики. Формально вопрос об их генетической принадлежности решается достаточно легко, но представленные в этих языках способы смешения языковых традиций в контексте настоящей статьи кажутся любопытными, поэтому остановлюсь на двух примерах.

В языке хэчжоу (河州, китайская пров. Ганьсу) большинство знаменательных морфем китайские, противопоставлены три тона (и фонетические соответствия с китайским, кажется, достаточно систематичны). Что касается служебных морфем, то в плане выражения они также в основном китайские, но функционируют вполне по-алтайски. «В концептуальном отношении хэчжоу — уйгурский язык. (...) Это выражается в принципах, управляющих порядком слов, и в том, как именное и глагольное словоизменение отражает уйгурское» ([LEE-SMITH 1996c: 872]). Ср. перевод фразы *Чем ты их будешь угощать?* на три языка, где структурные различия между хэчжоу и уйгурским сводятся к наличию в последнем личного показателя при глаголе, дублирующего подлежащее ([ibid.: 868]):

китайский:	你	用	什么	招待	他们
	<i>Nǐ</i>	<i>yòng</i>	<i>shénme</i> (> <i>shémmě</i>)	<i>zhāodài</i> (~ устар. <i>zhāodǎi</i>)	<i>tāmēn?</i>
	[ni˥˩]	[jʊŋ˥˩] ~ [ɥʊŋ˥˩]	[ʃzɥm˥˩.mɿ˥˩]	[ʃʒɑŋ˥˩.tʰaj˥˩] (~ [ʃʒɑŋ˥˩.tʰaj˥˩])	[tʰɑ˥˩.mən˥˩]
	ты	использовать	что	угощать	они
хэчжоу:	<i>Ni</i>	<i>tʰam-xa</i>	<i>šima-la</i>	<i>kʰuēte-li?</i>	
уйгурский:	<i>Sen</i>	<i>ular-ni</i>	<i>nime bilen</i>	<i>küt-mekçi-sen?</i>	
	ты	они-OBJ	что INSTR	ждать-INTENTION-2.ед.	
	(букв. 'Чем ты их ждешь?')				

Сходным образом устроены языки ўтунь (или ўтуньхуа, 五屯话, пров. Цинхай) и тэнвэнь (唐汪, пров. Ганьсу). «Вполне вероятно, что на территории Китая существует еще несколько таких языков» ([LEE-SMITH 1996a: 846]).

Формально говоря, хэчжоу — сино-тибетский язык, причем близкий к китайскому. Еще более очевидна генетическая отнесенность медновского языка (о. Медный, Командорские острова). Несмотря на значительное количество русских заимствований, нет сомнений в его принадлежности к алеутской ветви эскалеутской семьи: он сохранил исконную базовую лексику и значительную часть грамматики. Однако глагольная морфология целиком «заимствована» из русского (притом что сами глагольные корни обычно алеутские), ср. [GOLOVKO, VAKHTIN 1990: 118–119], «русский компонент» записан кириллицей:

Тысяча девятьсот двадцать четвёрты года я аг-ал Mīdnam ila. Aba-qalī-ла-я када вуяна-х тин [= себя] ayugn-ил. И вот я man akītax abā-ю. Пенсия-м kuga ū-ю, и всё равно я abā-ю. Ну, anīyup у меня тоже qalagīit. Ну, huzūŋi они aḡunā-ют щас. Ígluŋ у меня ū-ют.

‘Я родилась в 1924 году на острове Медном. Работать-начала я, когда война началась. И вот я с того-времени работаю. Пенсию получаю, и всё равно я работаю. Ну, детей у меня тоже много. Ну, все они большими-становятся сейчас. Внуки у меня есть’.

²⁶ В списке Яхонтова противопоставление имен и глаголов еще ярче: на 23 существительных приходится 3 глагола и 2 прилагательных!

²⁷ Среди лексики французского происхождения есть также слова, не совпадающие с единицами французского списка, ср. ‘камень’ (81): фр. *pierre* — мч. *rosh* (из фр. *roche* ‘скала’); ‘ночь’ (60): фр. *nuit* — мч. *kawundipishkaw* ~ *swayr* (из фр. *soir* ‘вечер’).

Медновский язык является очевидным продуктом достаточно полноценного русско-алеутского двуязычия. Скорее всего, он возник в среде *креолов* (так в Русской Америке называли потомков от браков русских промышленников с алеутками) как символ новой этничности, отличной и от алеутской, и от русской.

* * *

Подведем итоги.

Тип языковой эволюции определяется языковой ситуацией, в которую вовлечены носители языка: числом используемых языков, наличием/отсутствием генетических связей между этими языками, а также их социальной иерархией в глазах носителей.

При решении ограниченных коммуникативных задач в ситуации зачаточного двуязычия достаточно часто возникают функционально ущербные новые языки — пиджины, складывающиеся из фрагментов различных языковых традиций. В определенных обстоятельствах пиджины могут дать начало и полноценным новым языкам — креолам; в таком случае есть все основания говорить о возникновении новой языковой традиции.

В ситуации устойчивого развитого двуязычия может возникать новый социум, дистанцирующийся от существовавших ранее. В качестве символа складывающейся новой этничности в рамках нового социума в значительной степени сознательно конструируется новый язык как намеренное смешение различных языковых традиций. В том случае, если исходные традиции в результирующем языке представлены достаточно равномерно, этот язык вряд ли можно привязать к одной из «старых» языковых семей и следует говорить о возникновении новой, «скрещенной» языковой традиции (как в медиа-ленгва, мичифе и, возможно, в хэчжоу). Если же вклад одной из традиций «относительно мал» (как в случае с медновским языком), от него можно абстрагироваться при соотнесении нового языка с существующими генеалогическими деревьями.

Еще один класс «смешанных» языков представляют собой те, которые в ходе эволюции меняют свою генетическую принадлежность. Как говорилось, такой процесс негласно узаконивается в традиционной компаративистике для близкородственных идиомов. На мой взгляд, принципиальных отличий между перемещением по генеалогическому древу «восточнославянских» диалектов и языка эйну нет. На протяжении своей истории эйну неторопливо «растрачивал» иранские элементы и постепенно заменял их тюркскими. Вероятно, здесь можно говорить о непрерывности языковой традиции, но генетическая квалификация самой этой традиции постепенно изменяется. При достаточной информации последовательный компаративист мог бы даже вести своеобразный мониторинг и выявить момент «перехода количества в качество»: как только в словном списке число этимологически «новых» элементов превысит число «старых», происходит «диалектический прыжок» языка с одной ветви генеалогического древа на другую²⁸.

Экзотика эйну в том, что лингвисты застали его в «критической точке», когда оснований предпочесть ту или иную филиацию нет. Случай этот несчастный, но и не исключительный. Филиацию таких находящихся на перепутье языков компаративисты обсуждают редко. Любопытно, что при этом они не всегда «идут на принцип» и могут пренебречь важнейшим постулатом о примате базисной лексики над другими элементами языковой структуры. Если вся грамматическая система и значительная часть лексикона указывают на связи языка в одном, причем совершенно определенном, и «чужом» для специалиста в области частного языкознания направлении, идея признать язык «своим» всего лишь на основании данных части лексикона, хотя бы и базовой, вызывает невольное отторжение. Наиболее ярким представляется случай с языком **ма'а**, или мбугу (северо-восток Танзании). «Этот язык обладает системой согласовательных именных классов, а также глагольной системой, характерной для языков банту, причем особая близость обнаруживается с языками асу и шамбала. В то же время основной словарный фонд мбугу, а также личные и указательные местоимения не связаны с банту, но обнаруживают це-

²⁸ Точно так же и те креольские языки, которые, начав свое существование на заново возникшем генеалогическом древе, остались под воздействием своего языка-лексификатора, естественно, передвигаются на соответствующую «новую» ветку. Так, идиом, которым повседневно пользуются афроамериканцы, давно стал разновидностью американского английского.

лый ряд сходжений с кушитскими языками»²⁹ [Порхомовский 1982: 215]. В. Я. Порхомовский отмечает, что исследователи по-разному решают вопрос о генетической принадлежности этого языка, при этом ему самому как афразисту такой «неправильный» язык трудно счесть кушитским; «представляется более обоснованным отнесение языка мбугу к банту» [ibid.: 254].

Говоря о смене генетической филиации, мы, естественно, имеем в виду тот тип генетического древа, который выше был назван «новым». В этом смысле язык эйну вот-вот окажется на «новом» тюркском древе. Пока же он сидит на иранской ветви индоевропейского древа. Как долго мы будем готовы считать, что его «старая» филиация объективнее новой?

И как обстоит дело с англоромани? Можно ли считать, что его более ранние варианты (пока английский компонент списка Сводеша уступал индийскому) находились на том же индоиранском «старом» стволе, что и эйну? И верно ли, что, не собрав стословный список англоромани начала XXI в., мы ничего не можем сказать о его генетической принадлежности?

Приложение. Стословные списки двух радикальных креолов

№	метаязык	ток-писин	сарамакка	№	метаязык	ток-писин	сарамакка
1	all	<i>olgeta</i>	<i>hii, tíu</i>	27	feather	<i>gras</i>	<i>puúma</i>
2	ashes	<i>sit</i>	<i>síndja</i>	28	fire	<i>paia</i>	<i>fája, vèvé</i>
3	bark	<i>skin</i>	<i>kákisa</i>	29	fish	<i>pis</i>	<i>físi</i>
4	belly	<i>bel</i>	<i>bèè, baika</i> (уст.)	30	fly (v.)	<i>flai</i>	<i>buwá</i>
5	big	<i>bikpela</i>	<i>bigi, gaán</i>	31	foot	<i>fút, lek</i>	<i>fútu</i>
6	bird	<i>pisin</i>	<i>fóu</i>	32	full	<i>pulap</i>	<i>fíu</i>
7	bite	<i>kaikai</i>	<i>njan</i>	33	give	<i>givim</i>	<i>dá</i>
8	black	<i>blakpela</i>	<i>baáka</i>	34	good	<i>gutpela</i>	<i>búnu</i>
9	blood	<i>blut</i>	<i>buúu, sangá</i>	35	green	<i>grinpela</i>	<i>guúun, giín</i>
10	bone	<i>bun</i>	<i>bónu</i>	36	hair	<i>gras</i>	<i>uwíi ~ wiwíi</i>
11	breast	<i>susu</i>	<i>bóbi</i>	37	hand	<i>han</i>	<i>máu</i>
12	burn (tr.)	<i>kukim*</i>	<i>tjumá</i>	38	head	<i>het</i>	<i>hédi</i>
13	claw	<i>kapa</i>	<i>húnja,</i>	39	hear	<i>harim</i>	<i>jéi</i>
14	cloud	<i>klaut</i>	<i>bundjí, liba(hoos), wóluku</i>	40	heart	<i>klok</i>	<i>háti</i>
15	cold	<i>kolpela</i>	<i>kótò</i>	41	horn	<i>kom</i>	<i>tutú</i>
16	come	<i>kam</i>	<i>kó</i>	42	I	<i>mi</i>	<i>mi</i>
17	die	<i>dai, indai</i>	<i>dédè</i>	43	kill	<i>mekim i dai, kilim i dai</i>	<i>kíi</i>
18	dog	<i>dok</i>	<i>dágu</i>	44	knee	<i>skru</i>	<i>kiní</i>
19	drink	<i>dringim</i>	<i>bebé, diíngi</i>	45	know	<i>save</i>	<i>sábi</i>
20	dry	<i>drai</i>	<i>dèè</i>	46	leaf	<i>lip</i>	<i>uwíi ~ wiwíi</i>
21	ear	<i>yau</i>	<i>jési,</i>	47	lie	<i>slip</i>	<i>kándi</i>
22	earth	<i>graun</i>	<i>goón, dóti</i>	48	liver	<i>leva</i>	<i>lèbèn</i>
23	eat	<i>kaikai</i>	<i>nján, kómé</i>	49	long	<i>longpela</i>	<i>lánga</i>
24	egg	<i>kiau</i>	<i>wóbo</i>	50	louse	<i>laus</i>	<i>lòsu, piójo</i>
25	eye	<i>hai</i>	<i>wójo</i>	51	man	<i>man</i>	<i>wómi</i>
26	fat (n.)	<i>gris</i>	<i>fátu</i>	52	many	<i>planti</i>	<i>hía ~ hila</i>
				53	meat	<i>abus</i>	<i>gbamba, mbéti</i>
				54	moon	<i>mun</i>	<i>líba</i>
				55	mountain	<i>maunten</i>	<i>kúnunu</i>

²⁹ В. Я. Порхомовский не случайно так осторожно квалифицирует кушитский компонент этого языка, поскольку его небантуский лексикон неоднороден и соотносится с разными группами кушитской ветви. Кроме того, заметная часть словаря восходит к нилотскому языку масаи, в частности не менее семи единиц стословного списка: 'кусать', 'приходить', 'собака', 'сердце', 'знать', 'дождь', 'маленький' [THOMASON 1997: 475—476].

№	метаязык	ТОК-ПИСИН	САРАМАККА	№	метаязык	ТОК-ПИСИН	САРАМАККА
56	mouth	<i>maus</i>	<i>búka</i>	79	stand	<i>sanap</i>	<i>taámpu, tán-a-pê</i>
57	name	<i>nem</i>	<i>nê</i>	80	star	<i>sta</i>	<i>teéa</i>
58	neck	<i>nek</i>	<i>gangáa</i>	81	stone	<i>ston</i>	<i>sitónu</i>
59	new	<i>nupela</i>	<i>njú-njú</i>	82	sun	<i>san</i>	<i>sónu</i>
60	night	<i>nait</i>	<i>ndéti</i>	83	swim	<i>swim</i>	<i>sún</i>
61	nose	<i>nus</i>	<i>núsu</i>	84	tail	<i>tel</i>	<i>lábu</i>
62	not	<i>no</i>	<i>ná,</i>	85	that	<i>dispela</i>	<i>dí, dídé</i>
63	one	<i>wanpela</i>	<i>wán</i>	86	this	<i>dispela</i>	<i>dísi</i>
64	person	<i>man</i>	<i>sòmbè ~ sèmbè</i>	87	thou	<i>yu</i>	<i>ju ~ i</i>
65	rain	<i>ren</i>	<i>tjúba, vulá</i>	88	tongue	<i>tang</i>	<i>tóngò</i>
66	red	<i>retpela</i>	<i>bè</i>	89	tooth	<i>tit</i>	<i>tánda</i>
67	road	<i>rot</i>	<i>pási</i>	90	tree	<i>diwai</i>	<i>páu</i>
68	root	<i>rop</i>	<i>lútu</i>	91	two	<i>tupela</i>	<i>tú</i>
69	round	<i>raun</i>	<i>lóntu</i>	92	walk	<i>go</i>	<i>wáka</i>
70	sand	<i>waisan</i>	<i>sáнду</i>	93	warm	<i>hatpela</i>	<i>gbòlò, kéndi,</i>
71	say	<i>spik</i>	<i>táa, táki, fan</i>	94	water	<i>wara</i>	<i>wáta</i>
72	see	<i>lukim</i>	<i>sí</i>	95	we	<i>yumi [incl.], mipela [excl.]</i>	<i>u, wi</i>
73	seed	<i>pikinini</i>	<i>síi</i>	96	what	<i>wanem</i>	<i>andí</i>
74	sit	<i>sindaun</i>	<i>sindó</i>	97	white	<i>waitpela</i>	<i>wéti</i>
75	skin	<i>skin</i>	<i>kákisa, sinkí(n)i</i>	98	who	<i>husat</i>	<i>ambè ~ ámmè</i>
76	sleep	<i>slip</i>	<i>duumí</i>	99	woman	<i>meri</i>	<i>koósuma, mujée</i>
77	small	<i>liklik, smolpela</i>	<i>pikí</i>	100	yellow	<i>yelopela</i>	<i>koóko, donú, fòkò</i>
78	smoke	<i>smok</i>	<i>sumúku</i>				

Примечание. Акут (´) и его отсутствие над гласными *a, e, i, o, u* в сарамакка обозначают высокий и низкий тоны соответственно, гравис над *è* и *ò* — открытую артикуляцию, циркумфлекс над *ê* и *ô* — высокий тон у *è* и *ò*.

Л и т е р а т у р а

- БЕЛИКОВ 1987 — В. И. БЕЛИКОВ. Лексические замены в креольских языках: анализ стословного списка // *Возникновение и функционирование контактных языков*: Материалы рабочего совещания. М.: «Наука».
- БЕЛИКОВ 1989 — В. И. БЕЛИКОВ. Древнейшая история и реальность лингвистических дендрограмм // *Лингвистическая реконструкция и древнейшая история Востока*: Материалы к дискуссиям на международной конференции. Ч. 1. М.: «Наука».
- БЕЛИКОВ 1991₁ — В. И. БЕЛИКОВ. Компаративистика и креольские языки // *Сравнительно-историческое изучение языков разных семей. Лексическая реконструкция. Реконструкция исчезнувших языков*. М.: «Наука».
- БЕЛИКОВ 1991₂ — В. И. БЕЛИКОВ. Роль заимствований в становлении неомеланезийских языков // *Лексические заимствования в языках зарубежного Востока: Социолингвистический аспект*. М.: «Наука».
- БЕЛИКОВ 1998 — В. И. БЕЛИКОВ. *Пиджины и креольские языки Океании: Социолингвистический очерк*. М.: «Наука».
- БОНДАЛЕТОВ 1974 — В. Д. БОНДАЛЕТОВ. *Условные языки русских ремесленников и торговцев. Вып. I. Условные языки как особый тип социальных диалектов*. Рязань: Ряз. гос. пед. ин-т.
- БТС 1998 — *Большой толковый словарь*. СПб.: «Норинт».
- БУРЛАК, СТАРОСТИН 2001 — С. А. БУРЛАК, С. А. СТАРОСТИН. *Лингвистическая компаративистика*. М.: «УРСС».

- ВАЙНРАЙХ 1972 — У. ВАЙНРАЙХ. О совместимости генеалогического родства и конвергентного развития // *Новое в лингвистике*. Вып. VI. М.: «Наука».
- ВИНОГРАДОВ 1982 — В. А. ВИНОГРАДОВ. Функционально-типологические критерии и генеалогическая классификация языков // *Теоретические основы классификации языков мира: Проблемы родства*. М.: «Наука».
- ГОЛЫШЕВ 1880 — И. ГОЛЫШЕВЪ. Проводы офеней въ дорогу изъ дому для торговли и разговоръ ихъ на своемъ искусственномъ языкѣ // *Ежегодникъ Владимирскаго губернскаго статистическаго комитета*. Владимиръ. Т. 3. (Цит. по: [БОНДАЛЕТОВ 1974: 77]).
- А. В. ДЫБО 2003 — А. В. ДЫБО. Семантическая реконструкция в сравнительно-историческом языкознании // *Сравнительно-историческое исследование языков: Современное состояние и перспективы*. Тезисы докладов международной научной конференции. Москва, 22–24 января 2003 года. М.: Филол. ф-т МГУ.
- В. А. ДЫБО 2003 — В. А. ДЫБО. Язык — этнос — археологическая культура (несколько мыслей по поводу индоевропейской проблемы) // *Сравнительно-историческое исследование языков: Современное состояние и перспективы*. Тезисы докладов международной научной конференции. Москва, 22–24 января 2003 года. М.: Филол. ф-т МГУ.
- ЗАЛИЗНЯК 1995 — А. А. ЗАЛИЗНЯК. *Древненовгородский диалект*. М.: «Школа “Языки русской культуры”».
- ИВАНОВ 1990₁ — Вяч. Вс. ИВАНОВ. Генеалогическая классификация языков // *Лингвистический энциклопедический словарь*. М.: «Советская энциклопедия».
- ИВАНОВ 1990₂ — Вяч. Вс. ИВАНОВ. Языки мира // *Лингвистический энциклопедический словарь*. М.: «Советская энциклопедия».
- ЛЭС — *Лингвистический энциклопедический словарь*. М.: «Советская энциклопедия». 1990.
- МАС — *Словарь русского языка*. 2-е изд. М.: «Русский язык». 1981–1984.
- МЕЙЕ 1938 — А. МЕЙЕ. *Введение в сравнительное изучение индоевропейских языков*. М.–Л.: Соцэкгиз.
- ОЖЕГОВ 1949 — С. И. ОЖЕГОВ. *Словарь русского языка*. М.: ОГИЗ–ГИС.
- Основы финно-угорского языкознания (вопросы происхождения и развития финно-угорских языков)*. М.: «Наука». 1974.
- ПОРХОМОВСКИЙ 1982 — В. Я. ПОРХОМОВСКИЙ. Проблемы генетической классификация языков Африки // *Теоретические основы классификации языков мира: Проблемы родства*. М.: «Наука».
- ТЕЙЛОР 1972 — Д. ТЕЙЛОР. О классификации креолизованных языков // *Новое в лингвистике*. М.: «Наука». Вып. VI.
- ТОМПСОН — Р. ТОМПСОН. Заметка о некоторых чертах, сближающих креолизованные диалекты Старого и Нового Света // *Новое в лингвистике*. М.: «Наука», 1972. Вып. VI.
- ТРУБЕЦКОЙ 1987 — Н. С. ТРУБЕЦКОЙ. Мысли об индоевропейской проблеме // *Избранные труды по филологии*. М. [Первоначально: N. S. TRUBETZKOY. Gedanken über das Indogermanenproblem // *Acta linguistica*, 1939, Vol. 3].
- ХЕЛИМСКИЙ 1982 — Е. А. ХЕЛИМСКИЙ. *Древнейшие венгерско-самодийские языковые параллели: Лингвистическая и этногенетическая интерпретация*. М.: «Наука».
- ХЕЛИМСКИЙ 1991 — Е. А. ХЕЛИМСКИЙ. Самодийская лингвистическая реконструкция и праистория самодийцев // *Сравнительно-историческое изучение языков разных семей: Лексическая реконструкция. Реконструкция исчезнувших языков*. М.: «Наука».
- ХОЛЛ 1972 — Р. А. ХОЛЛ. Креолизованные языки и «генеалогическое родство» // *Новое в лингвистике*. М.: «Наука». Вып. VI.
- ШАЙКЕВИЧ 1995 — А. Я. ШАЙКЕВИЧ. *Введение в лингвистику*. М.: Российский открытый ун-т.
- ЭДЕЛЬМАН 2003 — Д. И. ЭДЕЛЬМАН. Еще раз об ареальной и темпоральной стратификации общеиранского праязыкового континуума // *Сравнительно-историческое исследование языков: современное состояние и перспективы*. Тезисы докладов международной научной конференции. Москва, 22–24 января 2003 года. М.: Филол. ф-т МГУ.
- ARENDS & al. — J. ARENDS, P. MUYSKEN, N. SMITH (eds.). *Pidgins and Creoles: An Introduction*. Amsterdam–Philadelphia: John BENJAMINS, 1995.

- BAKKER, MUYSKEN 1995 — P. BAKKER, P. MUYSKEN. Mixed Languages and Language Intertwining // [ARENDS et al. 1995: 41–52].
- BAKKER, PAPEN — Peter BAKKER, Robert A. PAPEN. Michif and Other Languages of the Canadian Métis // [WURM & al. 1996: 120; pp. 1171–1183].
- BICKERTON 1977 — D. BICKERTON. Pidginization and Creolization: Language Acquisition and Language Universals // A. VALDMAN (ed.) / *Pidgin and Creole Linguistics*. Bloomington.
- BICKERTON 1981 — D. BICKERTON. *Roots of Language*. Ann Arbor: Karoma Publishers.
- BICKERTON 1988 — D. BICKERTON. Creole Languages and the Bioprogram // F. J. NEWMAYER (ed.) / *Linguistics: The Cambridge survey*. Vol. 2. *Linguistic Theory: Extensions and Implications*. Cambridge: CUP; pp. 268–284.
- CLARK 1979 — R. CLARK. In Search of Beach-la-Mar: Towards a History of Pacific Pidgin English // *Te Reo*,. Vol. 22/23.
- GOLOVKO, VAKHTIN — E. V. GOLOVKO, N. B. VAKHTIN. Aleut in Contact: The CIA enigma // *Acta Linguistica Hafniensia*, 1990. Vol. 22.
- KOUWENBERG — Silvia KOUWENBERG. Berbice Dutch Creole // [WURM & al. 1996: 136; pp. 1147–1156]; on-line: http://www.caribbeanlanguages.org.jm/Kouwenberg_in_Atlas_Wurm_1996.pdf.
- LAVERDURE — P. LAVERDURE, I. R. ALLARD. *The Michif Dictionary*. Winnipeg: Pemmican Publications, 1983.
- LAYCOCK — D. S. LAYCOCK. Creative Writings in New Guinea Pidgin // Wurm S. A. (ed.). *New Guinea Area Languages and Language Study*. Canberra, 1977. Vol. 3.
- LEE-SMITH 1996_a — Mei W. LEE-SMITH. Some Hybrid Languages in China // [WURM & al. 1996: 081; pp. 845–850].
- LEE-SMITH 1996_b — Mei W. LEE-SMITH. The Ejnu Language // [WURM & al. 1996: 082; pp. 851–864].
- LEE-SMITH 1996_c — Mei W. LEE-SMITH. The Hezhou Language // [WURM & al. 1996: 083; pp. 865–874].
- MÜHLHÄUSLER 1986 — P. MÜHLHÄUSLER. *Pidgin and Creole Linguistics*. Oxford: BLACKWELL.
- MUYSKEN 1997 — P. MUYSKEN. Media Lengua // S. THOMASON (ed.) / *Contact Languages: A Wider Perspective*. Amsterdam: John BENJAMINS.
- MUYSKEN 2001 — P. MUYSKEN. Intertwined Languages // *Concise Encyclopedia of Sociolinguistics*. Amsterdam &c.: ELSEVIER Science; p. 481.
- PLANK 1988 — P. H. VAN DER PLANK. Growth and Decline of the Dutch Standard Language Across the State Borders // *IJSL*, № 74.
- POLINSKY 1996 — Maria POLINSKY [М. С. ПОЛИНСКАЯ]. *American Russian: An endangered language?* [Manuscript.] USC–UCSD; on-line: <http://ling.ucsd.edu/~polinsky/pubs/american-russian.pdf>.
- THOMASON 1997 — S. THOMASON. Ma'a (Mbugu) // S. THOMASON (ed.) / *Contact Languages: A Wider Perspective*. Amsterdam: John BENJAMINS.
- THOMASON 2001 — S. THOMASON. *Language Contact: An Introduction*. Washington (D.C.): Georgetown Univ. Press.
- WURM & al. 1996 — Stephen A. WURM, Peter MÜHLHÄUSLER, Darrell T. TRYON. (eds.). *Atlas of the Languages of Intercultural Communication in the Pacific, Asia, and the Americas*. Berlin–New York: MOUTON–DE GRUYTER. Vol. II.2. *Texts*.

Summary

The article discusses the issue of elements of linguistic convergence in the process of arisal of new languages. In the author's opinion, the importance of this issue grows in direct proportion to the amount of data we have on “mixed” languages (pidgins and creoles), despite the fact that it is frequently ignored or underestimated by traditionally-oriented scholars of comparative-historical linguistics. It is shown that the existing models of genetic classification, based on morphological and lexical (lexicostatistical) criteria, frequently do not work due to the impossibility of fully taking this issue into account. In addition, this problem is actual not only for originally non-related languages undergoing pidginization, but also for languages that belong to the same family whose divergence was at the same time accompanied by intense mutual contacts (this problem, after A. B. DOLGOPOLSKY’s term, is named “the issue of the genealogical stump”).

Dravidian Numerals

Although the most probable affiliation of the Dravidian language family on the macro-comparative level lies with the Nostratic superfamily, potential ties between Proto-Dravidian numerals and known Nostratic words for numerals remain scarce. The article summarizes most of the existing hypotheses on the origins of Dravidian numerals, both internal and external, and adds a few others based on the author's theory of a possible Australian substratum in Proto-Dravidian. It is concluded that the latter theory may shed some much needed light on this complicated issue

Numerals belong to the relatively stable part of lexicon of almost all language families, although they are certainly not immune against borrowing. This lack of immunity can be convincingly demonstrated on data from certain Dravidian languages: thus, Brahui has borrowed all the numerals above "3" from Persian, while other Dravidian languages have introduced them from Indo-Aryan languages: Malto (everything starting with "1", although inherited forms for "1" and "2" are also in parallel use), Kurukh (above "3"), Kuwi and Kui (above "2"; in Kui inherited forms for "3"–"7" are in parallel use), Pengo (above "2"), Kolami (above "4", besides parallel inherited forms), Gondi (above "7", besides parallel inherited forms). On the other hand, some non-Dravidian languages, e. g. Nihali, borrowed their numerals from their Dravidian neighbours: e. g. Nihali *irar* "2", *mōḥ(o)* "3", *nāḷku, nālo* "4" ([KUIPER 1966: 74–75]).

The main purposes of the present study are to describe the inherited Dravidian numerals, to try to understand their structure through internal etymologization, and, where possible, to discuss their external parallels. External comparison is understood here as drawing upon data from hypothetically related languages, hypothetical old substrata and neighbouring language families and language isolates. As far as external genetic relationship is concerned, the situation is approached here from the point of view of the Nostratic hypothesis, proposing a common origin for Afroasiatic, Kartvelian, Indo-European, Uralic, Altaic, and Dravidian languages. (It should be mentioned that Robert CALDWELL, the author of the first comparative grammar of Dravidian languages, was one of the first scholars who had already speculated about a distant relationship among these language families). Special attention will be paid to the data of the Elamite language, sometimes assumed to be closely related to Dravidian (McAlpin 1981).

We know next to nothing about pre-Dravidian substrata in India, with the potential exception of Nihali, although for this language the Austric affiliation seems most probable. In respect to this, the old hypothesis about a possible Dravidian-Australian relation ([MÜLLER 1882: 95–98]; [CALDWELL 1913: 75–77, 395]; recently again [DIXON 1980: 236, 488–489]) is reinterpreted here from this point of view — namely, that the pre-Dravidian substratum could be related to Australian languages ([BLAŽEK 1992: 421–431]). Finally, out of all the neighbouring languages it is most reasonable to take into particular account the Munda family, as one of the most common sources of borrowing into Dravidian.

Dravidian cardinal numerals have been reconstructed and etymologized as follows:

"one"

1.1. **oru*-(C)/**ōr*-(V) ([DEDR: №990 (a)]; [Zvelebil 1977: 34]) = **or*- (G. STAROSTIN).

Etymology:

1.1.1. ANDRONOV ([1994: 169]) thinks that *-r* in **oru*-/**ōr*- is secondary relative to the root **ol*-, discussed below. It is reasonable to think that the inlaut *-r* of the following numeral **iru*-/**īr*- could have influenced the transition **ol*- > **or*-. Since the sequence **wo*- is not attested in Dravidian, it is possible to speculate about a protoform **wol*-, which would be compatible with East Cushitic **wal*-/**wil*- > Saho *wili* "one", Somali *wal* "all", Elmolo *wol* "together", Oromo *wol(-i)* "together, with", Sidamo *wole* "other" ([SASSE 1982: 188–189]).

1.1.2. There is a hypothetical possibility to identify here a substratum influence of the Australian type; cf. examples from various groups of Pama-Nyungan: Karanya *uru* ([CURR II: №104]), Pitta-Pitta *ururu* ([KLUGE 1938: 68 after ROTH]) (two languages of the Pitta-Pitta group); Karuwali (Karna subgroup of the Dieri group) *orru* (CURR) = *uru*, Wongkumara (Ngura subgroup of the Dieri group) *warra* "1" ([CURR II, №106, 52]; [SCHMIDT, *Anthropos* 7, 1912: 492]).

1.2. **onru* ([DEDR: №990 (d)]) = **on-tu* ([KRISHNAMURTI 2001: 255]) = **ond*- (G. STAROSTIN).

Etymology:

1.2.1. ANDRONOV ([1994: 168–169]) reconstructs an original **oŋ-tu* < **ol-tu* on the basis of Tamil *ol-*, Malayalam *ollu-* “to unite”, cf. also Tamil *ol* “end”. Regularity of the development *-l + t- > -nr-* may be demonstrated, for instance, on the example of the Tamil verb *al-* “to be not so-and-so” in the 3rd p. sg. ntr. *antū* ([DEDR: №234]).

1.3. **okk-* “one, single, alone” ([DEDR: №990 (b)]) = **ok(k)-* (G. STAROSTIN).

Etymology:

1.3.1. ANDRONOV ([1994: 169]) derives it from **ol- + -kk-*, via **ork-*, cf. Konḍa *uRku uRku* “one each”.

1.3.2. KRISHNAMURTI ([2001: 255]) prefers analyzing this and the other forms as a combination of the bare root **o* with suffixes *-r-*, *-n-*, *-k-*; the root **o* here is the same as in Old Tamil *o* “to unite”. This solution is perhaps compatible with the idea of ANDRONOV ([1978: 240]), who speculates about the same primitive root reflected in Malayalam *o-* “to be similar”.

1.3.3. ANDRONOV ([1978: 240]) also admits a possible connection with IE **oǵ-no/ko/yo-* “one”.

1.4. **oŋti* “alone, single” ([DEDR: №990 (c)]).

“two”

2.1. **iru-(C) / *īr-(V)* ([DEDR: №474]; [ZVELEBIL 1977: 34]) = **ir-* (G. STAROSTIN).

Etymology:

2.1.1. CALDWELL ([1913: 331]) derived the numeral **īr-* “2” from the verb **īr-*, attested in Tamil *īr* (*-pp-*, *-tt-*) “to drag along, pull, attract, carry, flay, draw, paint, write”, Malayalam *īr* “splitting, sawing”, *īruka* “to saw, split”, Kannada *īr* “to pull, draw”, Parji *īrp-* “to pull”, Gadba *īr-* “to pull, drag”, Gondi *ric-* “to cut with saw”. However, Pengo *nir-* “to pull”, if related, probably excludes this etymology ([DEDR: №542]).

2.1.2. CALDWELL himself also tried to find parallels outside the Dravidian family. His comparison with the potential Kartvelian counterparts is undoubtedly remarkable: Georgian *or-*, *vor-*, Megrelian *žir-*, *žar-*, Laz *žu(r)-*, *žu(r)-*, *jur-*, Svan *jōri*, *jori*, *jerbi* < **jerwi* < **jewri* < **jori*, derivable from proto-Kartvelian **jor-* “2” ([KLIMOV 1998: 144–145]). Yet the fact that it is the only common Kartvelian lexeme reconstructed with an initial **j-* is rather suspicious.

2.1.3. Again, there is a substratum alternative indicated by Australian forms: Tiwi (one of the non-Pama-Nyungan languages spoken at the Melville Island in north from the Arnhem Land) *yur-rara* “2” ([BLAKE 1981: 112]); Wailpi *yierlina* “2”, Kaurna *illa* “2”, *yerrábbola* “4” (both Yura subgroup of the South-Western group of the Pama-Nyungan family), Meyu (dialect of Kaurna) *yerra* “mutually, both”, *yerrabula* “4” = ‘dual of’ *yerra* ([KLUGE 1938, 59–61]).

Note: The North Munda numeral **iral-* “8”, attested in Santali, Birhor *irəl*, Mundari *iralia*, Ho *irilia*, Kurku *ilar(ia)* ([PINNOW 1959: 86]), stands isolated within both Munda and Austro-Asiatic. The numeral can be explained as having a Dravidian origin, provided it reflects a compound consisting of the Dravidian numeral **ir-* and the verb **al-* “to be not so-and-so” > Tamil *al-* id., Malayalam *alla* “is not that”, not thus”, Kota *ala-* “to be not so-and-so”, Toda *alogy* “except”, Kannada *alla* “to be not so-and-so”, Kodagu *alla* id., Gondi *hal* “not”, Malto *-l-* ‘negative morpheme’, Brahui *all-* ‘base of past negative tenses of *anning* “to be”’ ([DEDR: №234]); thus, **ir-al* = “two-is not”. It is possible to imagine a (North) Dravidian origin for this specific Northern Munda isogloss.

2.2. **utri* “pair” > Tulu *udri* “a match, pair”, Telugu *uddi* “a match, an equal, a rival; equal”, *uddincu* “to pair, match, couple” ([DEDR: №623]).

2.2.1. Within the framework of the hypothesis of an Australian-type substratum influence it is tempting to wonder if this word could not be related to Pama-Nyungan **kutyarra* “2” ([BLAKE 1988: 43]: formed from **kutyá* by the non-singular suffix *-rra*; cf. the reconstruction of the 2nd person pronoun: sg. **NHu-* : du. **NHunpala* : pl. **NHurra*, by [EVANS 1988: 103]). Loss of the expected initial **k-* is not unprecedented in Dravidian, especially before back vowels, cf., e.g., Tamil *konay/ōnāy*, Malayalam *konayi/ōnāyi* “wolf” ([ANDRONOV 1994: 85]).

“three”

3.1. **muv-(C) / *mū-(V)* ([DEDR: №5052]; [ZVELEBIL 1977: 34–35]) = **mū-* (G. STAROSTIN) = **muH-* ([KRISHNAMURTI 2001: 330]: plus the neuter marker **-ntu*).

Etymology:

3.1.1. ANDRONOV ([1994: 169–170]) assumes segmentation into **mūn-* and the neuter marker **-tu*. In the first edition of his *Comparative Grammar of Dravidian Languages* ([1978: 242]) ANDRONOV speculates about a derivation from Dravidian **muŋ-* > Tamil *muŋ* “in front”, *muŋai* “front, face, eminence, point, edge”, Malayalam *mun* “priority in space and time”, Kota *mun-* “front, fore”, *mon* “point”, Toda *mun* “in front”, *min* “sharp point”, Kannada *mun* “that which is before, in front of”, *muntu* “the front

part or side, front”, Kodagu *miñña* “in front, further”, *mone* “sharp point”, *muṅ gay* “forearm”, Tulu *munè*, *monè* “point, end”, Telugu *muni* “first, former, previous, front”, Kolami *muni* “sting of scorpion”, *mut* “before” = Naiki *mund* id., Parji *muna vanda* “forefinger”, *mundi* “in front”, *mūni* “tip, point”, Gadba *mundel* “the front”, Gondi *munne* “in front of”, *mūne* “ahead”, Konḍa *mundala* “in front”, Kuwi *munu* “point of needle”, Kurukh *munddh*, *mund* “first, ahead of, previous to”, Malto *mundi* “formerly, in ancient times”, Brahui *mōn* ([DEDR: №5020]). He suggests the development “protruding finger” > “middle finger” > “third finger”. The root **muṅ-* is indeed used to designate one of the fingers, namely, “forefinger” in Parji *muna vanda*, probably in the sense of “first finger” (if the “thumb” is not included).

3.1.2. Alternatively, accepting a substratum origin, it is necessary to take into account some of the Australian forms for the numeral “3” (all from the South-Western group of the Pama-Nyungan family): Natingero dialect of Kalamai *mow* (Mirniny subgroup), Yungar dialect of Wadjuk *moa*, Wardand *mow*, Warrango *mowe*, Ngokgurring *mow*, Nyakinyaki *mow* (all from the Nyunga subgroup). There are also longer forms in various languages of the South-Western group: Natingero (see above) *monga*, Luritja or Kukatja *munngorra*, Bedengo *murrngul*, Jumu (‘Lake Amadeus’) *mun-kuripa*, and Malgana (‘Gascoyne River’) *manguraba* or (‘Sharkes Bay’) *mangaranu* “3” ([KLUGE 1938: 54–55]), where the second component can be identified with Nawu (Yura subgroup of the South-Western group) *karbu* “3” ([KLUGE 1938: 56]).

“four”

4.1. **nāl* ([DEDR: №3655]; [ZVELEBIL 1977: 34]; (G. STAROSTIN).

Etymology:

4.1.1. Following KITTEL ([1908]), CALDWELL ([1913: 335]) speculated about a relation between the numeral **nāl* “4” and the adjective **nal* “good” > Tamil *nal* (*nar-*) “good”, *nalla* “good, fine, excellent, abundant”, Malayalam *nal* “good, fine”, *nalla* “good, right, fine, handsome, real, true”, Toda *naś* “beauty”, Kannada *nal* “goodness, fairness, fineness”, *nala(vu)*, *nalivu* “pleasure, delight”, Kodagu *nallē* “good”, Tulu *nala*, *nalū* “good, cheap”, Telugu *naluvu* “beauty, ability, beautiful”, Gondi *nelā* “good” ([DEDR: №3610]). But CALDWELL himself admitted that the semantic development remains unexplained.

4.1.2. ANDRONOV ([1994: 170]) mentions that Tamil *nālu* means both “4” and “several”, similarly Malayalam *nālu* and Telugu *nālugu*. But the primary meaning “several” would be understandable only if “four” were the highest numeral. It seems more probable to suppose a secondary development “four” → “several”, not vice versa.

4.1.3. On the other hand, there are some remarkable external parallels. CALDWELL ([1913: 335]) concluded: “The resemblance between the Finnish tongues and the Dravidian, with respect to the numeral “four”, amounts almost to identity, and can scarcely have been accidental.” The Fenno-Ugric data look as follows: **neljä* (> **neljä*) > Finnish *neljä*, Estonian *neli*, gen. *nelja* | Lappish North *njälljē*, Lule *nielja*, Kildin *nielj*, Akkala *niel’* | Mordvin Erzya *ni’le*, Moksha *ni’lä* | Mari *nəl* | Udmurt Kazan *ni’l’*, Sarapul *ni’l’*; Komi Permyak *ni’ol’*, East *nul’* | Khanty Vach *ni’lə*, Obdorsk *ni’l*; Mansi Tavda *ni’l’r*, Pelymka *ni’lə*, Sosva *ni’la*; Hungarian *négy* id., *negyven* “40” ([UEW: 315–316]) || ? Altaic: Middle Korean *nəyh* id. || Tungus **hö[l]gün* “6” ([BLAŽEK 1999: 130]).

4.1.4. Alternatively, Dravidian “4” could be of substratum origin if we accept its Australian connections: cf. the forms for “4” in some Australian languages, all from the vast Pama-Nyungan family: Nawu (Yura subgroup of South-Western group) *nulla* (№63: ‘Gawler Range’ by [CURR 1886]), unidentified language *nalira* (№155 from ‘Tambo, Barcoo River’), Koa (Maric subgroup of Pama-Maric group) *nadera* (№140: ‘Diamantina River, Middleton Creek’), Maraura (Narrinyeri group) *nailko* (№84: ‘Murray River’). These forms were collected by E. M. CURR in his monumental collection of aboriginal language vocabularies *The Australian Race*, I–III, published in Melbourne 1886; here quoted according to [TROMBETTI 1923: 83] who was the first to compare Dravidian & Australian forms).

4.2. **kirt-a* “one-fourth” > Kannada *gidda*, *girda*, *gira* “a fourth part”; *giddana*, *gidna* “the fourth part of a solige”, Telugu *gidda*, *gidde* “one-fourth of the *sōla* measure” ([DEDR: №1553]).

4.2.1. It is possible to speculate about a relation with Pama-Nyungan **kutyarra* “two” (see 2.2). In some Australian languages the numeral “4” is formed from the numeral “2” by means of the dual suffix **-pa/ula*, cf. Potaruwutj (Kulin group of the Pama-Nyungan family) *kirtpan* & *kurtpun* ([CURR III: 492, 494]: ‘Hopkins River’; quoted after [KLUGE 1938: 74]).

“five”

5.1. **cayN_(C) / *cay_(V)* ([DEDR: №2826]; [ZVELEBIL 1977: 34–35]) = **śai-* (G. STAROSTIN).

Etymology:

5.1.1. ANDRONOV ([1994: 171–172]) connects the numeral “5” with Dravidian **kay-/ *key-* “hand” > Tamil *kai* “hand, arm; elephant’s trunk; handle”, Malayalam *kai*, *kayyi* id., *kayyu* “the hand”, Kota *kay* “hand, arm”, Toda *koy* id., Kannada *kay*, *kay(y)i*, *key* “hand, forearm; handle; trunk of elephant”, Kodagu *kay* “hand, arm”, Tulu *kai* “hand; handle”, Telugu *cēyi*, *cey(y)i* “hand, arm; elephant’s trunk”, *kai* “the hand”, Kolami *ki*·, *kīy*, *kiyu*, *key* “hand, arm”, Naiki *kī* “hand”, Parji *key* id., Gadba *ki*, *kiy*, *kiyyu* id., Gondi *kay*, *kai* id., Konḍa *kiyu*, *kivu* id., Pengo *key*, Manda *kiy* id., Kui *kaju*, *kagu* “hand, arm; elephant’s trunk”, pl. *kaska*, Kuwi *kēyū*, *keyyu*, *keyu*, pl. *keska*, Kurukh *xekkha* “hand, arm”, Malto *qege* “hand” ([DEDR: №2023]). ANDRONOV’s solution is undoubtedly acceptable from a semantic point of view, but it is difficult to understand why the initial consonant of the numeral became palatalized, in contrast with “hand” (this is the main objection of [KRISHNAMURTI 2001: 255]). It should be stressed that in the case of “hand”, only the Telugu form *cēyi*, *cey(y)i* undergoes (regular) palatalization (but not *kai* “the hand”, possibly a re-borrowing from Kannada). On the other hand, it is legitimate to suppose the influence of the following numeral **caṛu/ *cār* “6”.

It is also possible to speculate about the relationship between “5” and Konḍa *sēna*, Pengo *hēni* “many” ([DEDR: №2824]).

5.1.2. Altaic: Middle Korean *tàsās* “5” = *tā* “all” + *son* “hand”; *suyn* “50” ([BLAŽEK 1999: 130]).

5.1.3. There are interesting parallels in Austro-Asiatic languages: geographically closest is Khasi *san* “5” (Nagaraja); further cf. proto-Mon **(m-)suun* id. ([DIFFLOTH 1976]), &c.

“six”

6.1. **caṛu*-(C)/**cār*-(V) ([DEDR: №2485]; [ZVELEBIL 1977: 35]) = **śād-* (G. STAROSTIN).

6.1.1. ANDRONOV ([1994: 172]) rejects the proto-Dravidian status of the alternation **caṛu*-(C)/**cār*-(V), which is found only in South Dravidian. As the starting point, he reconstructs **cāl-* + *-tu* with the neuter marker **-tu*, an almost universal formative of numerals; in this case, the root can be identified with **cāl-* > Tamil *cāl* “to be abundant, full, sufficient, great”, Malayalam *cāla* “richly, fully”, Kannada *sāl(u)* “to be sufficient or enough, suffice”, Telugu *cālu* “to be able, capable, bear, endure, be enough, sufficient”, *cāla* “abundant(ly)”, Kolami *sāl* “to be able, can”, Gondi *hālna* “completely”, *āl-* “to be able”, Konḍa *sāl-* “to be capable of, be suitable”, Kuwi *hāl-* “to suffice, be enough to” ([DEDR: №2470]). The primary meaning of the numeral “six” = “the abundant [one]” is quite natural, cf. the most probable etymology of the Indo-European numeral **(k)seks-* ~ **(s)ueks-* “6” based on the root **ueks-* “to grow” > Lithuanian *vešétiz* “to grow vigorously; flourish” (see [BLAŽEK 1999: 239–241] with references). For the change **-l + t- > *-r-* Andronov (l.c.) finds an analogy in certain Tamil non-case sandhi patterns, such as *āṛidu* “banyan [is] bad” < *āl tīdu* (Tamil *āl*, *ālam* “banyan”, &c.: [DEDR: №382], & Tamil *tītu* “evil, fault, defect”: [DEDR: №3267]).

6.1.2. It is possible to speculate about a compound of the type **cay* “5” + **orū-* “1”.

“seven”

7.1. **eṛu*-(C)/**ēṛ*-(V) ([DEDR: №910]; [ZVELEBIL 1977: 35]) = **eṛ-u-* / **ēṛ-* ([KRISHNAMURTI 2001: 63]) = **jōṛ-* (G. STAROSTIN: vocalization after Gondwan **jōṛ-* while **-ē-* in other branches could have been influenced by the following numeral **eṇ-* “8”; in his dissertation [STAROSTIN 2000: №350] reconstructs **ē*, i. e. **jēṛ-*, in his transcription **jēr-*).

7.1.1. CALDWELL ([1913: 342]) explained this numeral as a verbal noun, derived from **eṛ(u)* “to rise” > Tamil *eṛu* “to rise, ascend (as heavenly body), rise by one’s own power, originate, be excited, aroused, increase, grow, begin”, *eṛucci* “rising, ascent, elevation, starting, origin, beginning, birth”, *eṛumai*, Malayalam *eṛu* “height, prominence”, Kota *eṛ* “weight”, Kannada *eṛ*, *ēṛ(u)* “to stand up, rise, awake, spring up, be produced, be obtained”, Kodagu *ē-ḷ-* “to get up”, Tulu *erkuni* “to rise or collect”, *erka* “full; fullness”, Brahui *harfing* “to raise, support, carry (off)”, &c. ([DEDR: №851]). Although the semantic motivation is not transparent (< **“full”*?; cf. Written Mongolian *doluyan* “7” : Turkic **dōl-* “to fill” or Turkic **jätti* “7” : **jät-* “be enough”, Turkish dial. *yetiz* “all, whole, full”; see Blažek 1999, 116), this etymology remains most promising.

7.1.2. CALDWELL ([1913: 342]) also speculated about a relation with Turkic **jätti* “7” (on the etymology, see 7.1.1), but Dravidian **-ṛ-* and Turkic **-t(t)-* has not been recognized by anybody as a valid correspondence.

7.1.3. There are other possible cognates in Altaic: Turkic **jōṛ* “100” | Written Mongolian *yerü* “the most of”, *yerüdügen* “for the greatest part, generally”, *yerüñkei* “common” | Middle Korean *jār* “10”, *jārāh* “a big quantity, number” | Old Japanese *jōrò-du* “10.000” (see Blažek 1999, 117). The pri-

mary semantics is in good agreement with the Dravidian verb **eṛ(u)* “to rise” &c., discussed in §7.1.1. For both Altaic and Dravidian a common Nostratic denominator **jeřU* may be established.

Note: The North Munda numeral **eya* “7”, attested in Santali *eae*, Mundari *eja*, *ea(e)*, Ho *iya*, *aia*, Birhor *eae*, *aea*, Kurku (*y*)*eya* ([PINNOW 1959: 269]), stands isolated within Munda and Austro-Asiatic in general. It could be borrowed from a Dravidian donor language in which Proto-Dravidian **-ɣ-* changed into *-y-*. Such a change is described, e.g., for Irula & Kota (Irula *kuyi*, Kota *kōy* “pit, hole” : Tamil *kuṛi*: [DEDR: №1818]); Toda (*tōy* “to be lowered” : Tamil *tāṛ* id.: [DEDR: №3178]); Manda (*nuy-* “to wash” < **noṛ-* / **noṭ-*: [DEDR: №3783]); and Kurukh (*kiyya* “under”, also Brahui *kī* “below, down” < **kīṛ-*: [DEDR: №1619]), cf. [ZVELEBIL 1970: 150, § 1.34.3–4; 151, § 1.34.5; 153, §§ 1.34.11.3 & 1.34.14]; [KRISHNAMURTI 2001, 42–75]. The best candidate for a donor language for North Munda would be Proto-North Dravidian or, rather, its earlier ancestor that still preserved the higher numerals.

“eight”

8.1. **eṭtu* / **eṇ* ([DEDR: №784]; [ZVELEBIL 1977: 35]) = **eṇṭtu* / **eṇ* ([ANDRONOV 1994: 173]) = **eṇ-* (*-nṭ-*) (G. STAROSTIN).

8.1.1. Already CALDWELL ([1913, 345–346]) identified this numeral with the stem **eṇ* “number”, **eṇ-* “to count, calculate”, joined by the neuter suffix *-(t)tu*; cf. the data: Tamil *eṇ* “thought”, calculation, number”, *ēṇi* “number”, *eṇṇu* “to think, consider, determine, esteem, conjecture, count, reckon”, Malayalam *eṇ* “number, thought”, Toda *ōṇ-* “to count”, *ōṇm* “counting, numbers”, Kannada *eṇike* “counting, number, thinking, observation”, *eṇṇu* “to count, think”, Kodagu *ēṇṇ-* “to say, tell”, Tulu *eṇṇuni* “to count, think”, *eṇṇige* “calculation, estimation”, Telugu *enṇu* “to count, reckon”, *encu* “to count, reckon, enumerate, think, consider, believe”, Parji *eja* “number”, *eja cāj-* “to count”, Pengo *eja ki-* id., Manda *eji ki-* id., Kuwi *eji kinai* id. ([DEDR: №793]).

8.2. **eṇ(u)-pa(k)ti* is reconstructible for Tulu, Telugu, Kolami and Gondi.

8.2.1. Although TYLER ([1986: 10]) accepts the identification of the word **eṇ* “number” with the numeral “8”, he alternatively offers to interpret the compound **eṇ(u)-pa(C)ti* as “a pair [subtracted from] ten” or “rest of ten”, where the first component can be identified with Tamil *ēṇai* “other; the rest” ([DEDR: №919]) or **iṇay* > Tamil *iṇai* “pair, couple, likeness, union”, Malayalam *iṇa* “pair, couple, union, companion”, Kannada *eṇe*, *eṇa* “a couple, pair, connexion, equality, similarity, a match”, Kodagu *ēṇe* “double”, *ēṇe makka* “twins”, Tulu *inè*, *iṇè* “a couple, pair, companion, mate”, Telugu *ena* “equal, equality, a match” ([DEDR: №457]). Although such constructions are not altogether improbable, the phonetic problems of **i* vs. **e* and **ṇ* vs. **ṇ* remain unresolved.

8.2.2. ANDRONOV ([1994: 173]) decides that the second component here is secondary, caused by analogy with the numeral “9”.

“nine”

9.1. **oṇ-* / **oṛ-paktu* ([DEDR: №1025]).

9.1.1. Apparently “one [subtracted from] ten” (so [CALDWELL 1913: 347], and independently GUNBERT).

9.2. **toḷ-*(*pak-*)*tu* ([DEDR: №3532]) = **toḷ-pad-* (G. STAROSTIN).

9.2.1. CALDWELL ([1913: 348]) suggested that **toḷ-* be identified with Dravidian **toḷ-* / **toḷ-* “before”, directly attested in Kuwi *toll(i)e*, *toli* “before, formerly”, cf. further Gondi *tollē* “previous”, Telugu *toli* “beginning; first, former, previous, old”, Tamil *tol* “old, ancient”, *tollai*, *toḷḷai* “antiquity, ancientness” ([DEDR: №3516]). The semantic motivation “before [ten]” is quite acceptable.

9.2.2. The first component is almost universally identified with the root **toḷ-*, attested in Tamil *toḷ* “to perforate, bore with an instrument”, *toḷḷai* “hole, perforation, pit, anything tubular, fault, defect”, Malayalam *toḷḷa* “hole, cavity”, Kota *toyḷ-* “to pierce”, *toḷ* “hole, vagina”, Kannada *toḷe* “hole, cavity”, Tulu *toluvè* “hole”, *toḷu* “hole; empty”, Telugu *toli* “hole”, Gondi *tullana* “to be pierced” ([DEDR: №3528]; cf. also [ANDRONOV 1994: 174]). However, the semantic motivation “defect [ten]” or “empty [ten]” seems rather vague.

“ten”

10.1. **pak-tu* ([DEDR: №3918]) = **paH-* & neuter suffix **-tu* ([KRISHNAMURTI 2001: 328]) = **paT-* (G. STAROSTIN).

10.1.1. CALDWELL ([1913: 351]) accepted Gundert’s idea that Dravidian “10” may be borrowed from Sanskrit *pañkti-* “set of five”, cf. also the ordinal numeral *pakthá-* “fifth”. The divergence of South Dravidian, Telugu, and the branches represented by Kolami and Gondi is dated to the middle of the 2nd millennium BC by ANDRONOV ([1994: 13]), or even later, to c. 1200–1000 BC by ([STAROSTIN

2000: 217–218]). In any case, contacts between Indo-Aryans and the ancestors of those Dravidians who preserved this numeral were possible.

10.1.2. ANDRONOV ([1994: 176]) proposes a long string of changes (**pax-* < **pav-* < **pam-* < **paŋ-* < **pal-*) so that the numeral can be identified with the root **pal-*, attested in such forms as Tamil *pala* “many, several, diverse”, *paŋmai* “plurality, multitude”, Malayalam *pala* “many, several, various”, Toda *peś* “a large number, many”, Kannada *pala*, *palavu* “much, many, several, various”, *palavar* “several persons”, Telugu *palu* “many, several, various, different”, *paluvuru*, *paluvuṇḍru* “many persons”, Malto *palware* “to be multiplied, be bred” ([DEDR: №3987]). From a semantic point of view this idea is acceptable, but the series of the proposed transformations is rather long and unconvincing.

10.1.3. CALDWELL ([1913: 353]) offered another possibility of internal etymologization, namely, based on the idea of “duplicity” (“10” = “2×5”?) or “share” (“¹/₁₀”?), cf. Tamil *pakku* “fracture, duplicity”, *pāttu* “dividing, sharing, share, half”, *pātti* “division, section, part, share”, *paṅkam* “portion, division”, Toda *paxy* “division”, Kannada & Tulu *pagadi* “tribute, tax”, &c., all from the verb **pak-* “to divide/be divided” ([DEDR: №3808]).

“hundred”

11.1. **nūr(-tu)* ([DEDR: №3729]) = **nūd-* (G. STAROSTIN).

11.1.1. CALDWELL ([1913: 354]) argued for a starting-point in Tamil *nūru* “to crush, pulverize, reduce to powder; powder, dust, flour”, Malayalam *nūru* “powder”, &c. ([DEDR: №3728]).

11.1.2. MENGES ([1968: 97]) sought a cognate in Turkic **jūr* “100”, but, according to “classic” consonant correspondences among Nostratic languages, Dravidian **n-* is only compatible with Turkic **j-* when the latter reflects Altaic **ń-* (> Mongolian **n(i)-*, Tungusic **ń-*). For Turkic **-r-*, on the other hand, the regular counterpart in Dravidian has been identified as Proto-Dravidian **-r-* (ILLIC-SVITYČ 1971: 150; 170]). The cognates in other Altaic languages exclude the genetic relationship of these numerals, cf. Written Mongolian *yerü* “the most of”, *yerüdügen* “for the greatest part, generally”, *yerünkei* “common” | Middle Korean *jār* “10”, *jārāh* “a big quantity, number” | Old Japanese *jōrō-du* “10.000” (see [BLAŽEK 1999: 117]).

“thousand”

12.1. Telugu *vēyi*, *veyi*, *veyyi*, pl. *vēlu* “1000”, *vēna-vēlu* “thousands by thousands”.

12.1.1. Derived from a Proto-Dravidian root, attested in such forms as Tamil *viyam* “extensiveness, height”, *viyal* “greatness, width, expansion”, *viyaṅ* “greatness, vastness, excellence”, Malayalam *viyam* “extension”, Gondi *weeya* “high” ([DEDR: №5404]).

Conclusions.

1) For the higher numerals from “6” to “10”, as well as “100” and “1000”, internal Dravidian etymologies seem most probable. The same could be said about the numeral “5”, but only if its phonetic relation with the word “hand” can be made clearer.

2) Some of these higher numerals (“7”, “8”) may have been lost in North Dravidian languages, but their former presence in that subgroup can be reconstructed on the basis of their (allegedly) being borrowed into North Munda languages.

3) The most surprising results concern the numerals from “1” to “4”. In between the potential cognates in other branches of Nostratic and the potential Australian parallels, which should indicate a substratum origin for the Dravidian stems, it is the latter comparanda that seem more convincing. On the other hand, it is so far unclear how well distributed the quoted forms are among Australian languages; if they can be shown to represent recent innovations, which is not impossible, the value of these connections will be lessened.

The results of the present study are summarized in the following table:

Form	Internal etymology (§§)	External parallels (§§)
<i>*oru</i> / <i>*ōr</i> “1”		A: PN <i>*urr-</i> / <i>*warr-</i> 1 (1.1.2.) N: EC <i>*wal-</i> / <i>*wil-</i> together (1.1.1.)
<i>*iru</i> / <i>*īr</i> “2”		A: Kaurua <i>yerra</i> both, Tiwi <i>yurarra</i> 2 (2.1.3.) or N: K <i>*jor-</i> 2 (2.1.2.)
<i>*utr-i</i> “pair”		A: PN <i>*kutyarra</i> 2 (2.2.1.)
<i>*muv</i> / <i>*mū</i> or <i>*mūn-</i> “3”	<i>*muṅ</i> front, point (3.1.1.)	A: PN (SW) <i>*mow</i> 3 (3.1.2.) A: PN (SW) <i>*muṅ-kurV(ba)</i> 3 (3.1.2.)

Form	Internal etymology (§§)	External parallels (§§)
* <i>nāl</i> “4”		A: PN: Nuwa <i>nulla</i> ; Maraura <i>naikō</i> 4 (4.1.4.) N: FU * <i>neljä</i> 4 AL: MK <i>nəyh</i> 4 (4.1.3.)
* <i>kirt-i</i> “ ¹ / ₄ ”		A: PN * <i>kutyarra-pula</i> 4 = dual of 2 (4.2.1.)
* <i>cayN</i> /* <i>cay</i> “5”	* <i>kay</i> /* <i>key</i> hand (5.1.1.) or CD * <i>cēn</i> - many (5.1.2.)	N: AL: MK <i>tasas</i> 5, <i>suyn</i> 50, <i>son</i> hand (5.1.2.) AU: Khasi <i>san</i> 5 (5.1.3.)
* <i>caru</i> /* <i>cār</i> “6”	< * <i>cāl-tu</i> : * <i>cāl</i> be abundant (6.1.1.)	
(<i>j</i>) <i>erū</i> /(<i>j</i>) <i>ēr</i> “7”	* <i>er(u)</i> rise (7.1.1.)	
* <i>eṇ</i> “8”	* <i>eṇ</i> number; calculate (8.1.1.)	
* <i>on-</i> /* <i>or-paktu</i> “9”	one [subtracted from] ten (9.1.1.)	
* <i>toḷ-(pak-)tu</i> “9”	[one] before ten (9.2.2.)	
* <i>pak-(tu)</i> “10”	Tamil <i>pakku</i> duplicity (10.1.3.)	
* <i>nūr-(tu)</i> “100”	* <i>nūru</i> powder (11.1.1.)	
Telugu <i>veyyi</i> “1000”	Tamil <i>viyan</i> greatness (12.1.1.)	

Appendix 1: What is known about Elamite numerals?

The hypothesis of a particularly close relationship between Elamite and Dravidian was first formulated by CALDWELL ([1856]; cf. [1913: 57; 65]); later, additional attempts at demonstrating this relationship were undertaken by A. TROMBETTI, F. BORK, I. DIAKONOFF and, particularly, J. MCALPIN. If this relationship were really close, one would expect some cognates among numerals. Below we list what is known about the Elamite numerals according to the interpretation of W. HINZ & H. KOCH, the authors of the monumental *Elamisches Wörterbuch* ([1987]):

ki “1” (also, with various extensions, *kik*, *kikki*, *kikqa*; with the personal marker, *kir*) — see [EW: 459, 465, 467, 468–469].

mar “2” (also syllabically written variants *ma-ir*, *mar-ra*) — see [EW: 876, 860].

ziti “3” — see [EW: 1305].

tuku “5”? — see [EW: 356].

barba “80”? — see [EW: 147].

It is apparent that among the known Elamite numerals there are no cognates in the system of Dravidian numerals (with the possible partial exception of *mar* “2”, cf. Dravidian **maṛu* “other, next, again”: [DEDR: №4766]). It can also be mentioned that F. KÖNIG ([1965: 42, fn. 15]) offered to interpret the Middle Elamite word *nulkippi* as “4 pairs”, i. e. “8”. If his solution were correct, the hypothetical root **nul-* could be a cognate of Dravidian **nāl* “4”. However, HINZ & KOCH ([EW: 1016]) interpret this word quite differently, namely, as a plural form of the noun ‘fertility-maker’.

Appendix 2: A survey of the inherited Dravidian cardinal numerals

=(C)/=(V)	* <i>oru</i> /* <i>ōr</i> “1”	* <i>onru</i> “1”	* <i>oṅṭi</i> “alone”	* <i>okk-</i> “1”	* <i>iru</i> / * <i>īr</i> “2”	* <i>muv</i> / * <i>mū</i> “3”
DEDR	[№990 (a)]	[№990 (d)]	[№990 (c)]	[№990 (b)]	[№474]	[№5052]
Tamil	<i>oru</i> (C)/ <i>ōr</i> (V)	<i>onru</i> nr. one; n. <i>orrai</i> one of a pair, odd number	<i>oṅṭi</i> that which is single, one who is alone; <i>oṅṭai</i> one of a pair	<i>okka</i> together	<i>iraṅṭu</i> ; adj. <i>iru</i> (C)/ <i>īr</i> (V) <i>iru-p/ṽatu</i> 20 <i>iruvar</i> 2 men <i>iraṅṭai</i> pair	<i>mūnru</i> & <i>mūn u</i> adj. <i>mu</i> (CC) / <i>mūv</i> (V) <i>mu-ppatu</i> 30 <i>muvar</i> 3 men
Mala-yalam	<i>oru</i> / <i>ōr</i>	<i>onnu</i> n., <i>orra</i> one, single, odd		<i>okka</i> together	<i>raṅṭu</i> ; adj. <i>iru</i> (C)/ <i>īr</i> (V) <i>iru-patu</i> 20 <i>iruvar</i> 2 men <i>iraṅṭa</i> double	<i>mūnnu</i> <i>mu-/nu-ppatu</i> 30 <i>muvar</i> 3 men
Kasaba		<i>o(y)ndu</i> / <i>onnu</i>			<i>reṅṭu</i> ; <i>irppattu</i> 20	<i>mūyndu</i> <i>muvattu</i> 30

=(C)/=(V)	* <i>oru</i> / * <i>ōr</i> “1”	* <i>onru</i> “1”	* <i>oṅṭi</i> “alone”	* <i>okk-</i> “1”	* <i>iru</i> / * <i>īr</i> “2”	* <i>muv</i> / * <i>mū</i> “3”
DEDR	[N ^o 990 (a)]	[N ^o 990 (d)]	[N ^o 990 (c)]	[N ^o 990 (b)]	[N ^o 474]	[N ^o 5052]
Kurru		<i>oṅḍu</i> / <i>uṅḍu</i>			<i>reṅḍu</i> / <i>ronḍu</i> <i>iravay</i> / <i>iridi</i> 20	<i>mūru</i>
Irula					<i>raṅḍu</i> / <i>reṅḍu</i>	
Kota	<i>or, o-r, o</i>	<i>oḍ</i> 1 <i>pan oṅḍ</i> 11			<i>eyḍ</i> <i>ir va-d</i> 20	<i>mu-ṅḍ</i> <i>muat</i> 30
Toda	<i>wī-r, oṣ, o</i>	<i>wiḍ</i>	<i>waty</i> single, odd < Ta <i>oṭṭai</i>		<i>e-d;</i> <i>i-r o-r</i> 2 years <i>ī foḥ</i> 20 <i>īm</i> double	<i>mu-d</i> <i>mu poḥ</i> 30
Kan- naḍa	<i>or</i> (C) / <i>ōr</i> (V)	<i>ondu</i> 1 thing	<i>oṅṭi</i> 1, alone		<i>er(a)ḍu, erar</i> adj. <i>ir(u), ik, ic</i> <i>irpattu</i> 20 <i>irbar</i> 2 men	<i>mūru</i> adj. <i>mu(k), mū</i> <i>mū-vattu</i> <i>muvar</i> 3 men
Koḍagu	<i>ori</i>	<i>ondi</i> 1 thing <i>pannandi</i> 11		<i>okka</i> to- gether	<i>daṅḍi</i> n. <i>iru-vadi</i> 20 <i>ibba</i> 2 men	<i>mu-ndi</i> <i>mu-vē</i> 3 men <i>nuppadi</i> “30”
B. Ku- rumba		<i>-oṅḍe</i>			<i>-aḍḍu</i> , adj. <i>aḍ</i>	<i>-mu:ru</i>
Šolega		<i>oṅḍu</i> / <i>o</i>			<i>eruḍu</i> / <i>o</i> <i>ippattu</i> 20	<i>mu:r^u</i> / <i>o</i> <i>mu:vattu</i> 30
Tulu	<i>or</i> adj., <i>oru</i> single	<i>oṅḷi</i> 1 thing <i>ondikē</i> joining	<i>oṅṭi</i> alone, single		<i>raddu</i> n. adj. <i>iru, ir-</i> <i>irva</i> 20 <i>irveru</i> 2 men	<i>mūji</i> <i>muppa</i> 30 <i>mūv(v)eru</i> 3 men
Koragu		<i>oṅḷi</i> / <i>uṅḷi</i>			<i>eyḍi</i>	<i>mūji</i>
Telugu	<i>orumu</i> be united	<i>oṅḍu</i> 1 thing	<i>oṅṭi</i> singleness	<i>oka</i> 1, single	<i>reṅḍu</i> n. <i>iru-, ir-</i> comp. <i>iruva(d)i</i> 20 <i>ibbaru</i> 2 men	<i>mūḍu</i> , inscr. <i>mūnru</i> adj. <i>mu, mū</i> <i>muppa(d)i</i> 30 <i>muvuru</i> 3 men
Kolami				<i>okkon</i> m. <i>okkod</i> f. <i>ok maṭ</i> once	<i>indiy</i> 2 things <i>iddar</i> 2 men <i>i-ral</i> 2 women <i>irve</i> 20	<i>mu-ndiy</i> 3 things <i>muggur</i> 3 men
Naiki				<i>okko(n)</i> m. <i>okko(d)</i> f. / n. <i>okoda</i> once	<i>inding</i> 2 things = Ch. <i>erndi</i> <i>iddar</i> 2 men = Ch. <i>iroṭer</i> <i>iral</i> 2 women = Ch. <i>ira</i>	<i>mūndiy</i> 3 things = Ch. <i>mūndi</i> <i>muggur</i> 3 men = Ch. <i>mug(g)ur</i> <i>muyal</i> 3 women = Ch. <i>muy(y)a</i>
Parji				<i>ok</i> adj., <i>okur</i> m., <i>okal</i> f.	<i>ir</i> adj. <i>irḍu</i> 2 things <i>irul</i> 2 men <i>iral</i> 2 women	<i>muy</i> adj. <i>mū:duk</i> 3 things <i>muvir</i> 3 men <i>muyal</i> 3 women
Gadba				<i>ukur</i> m., <i>okuṭ</i> f., <i>ukuṭ</i> n.	<i>iḍḍig</i> n., <i>ir</i> adj. <i>iṅḍi</i> 2 things <i>irul</i> 2 men <i>iral</i> 2 women	<i>mūnu</i> <i>mūṅḍ</i> n. <i>muvur</i> m. <i>muyal</i> f.
Gondi	<i>oror, ore</i> m.	<i>unḍi, unḍī,</i> <i>undi</i>	<i>unṭhal</i> n.	<i>ōkō</i> one each	<i>rand(u)</i> non-m. <i>irvu^r/r</i> m. <i>ranṭe</i> pair	<i>mūṅḍ</i> non-m. <i>muvir</i> m. <i>muhk</i> 3 each
Konḍa	<i>or-</i>	<i>unri</i> f. / n.			<i>ri^(p)</i> - <i>ri²er</i> 2 men <i>runḍi</i> f./n.	<i>mūnri</i> n. <i>mu²er</i> 3 men <i>mūRku mūRku</i> 3 each
Pengo	<i>ro, ronje</i>				<i>ri; rindaṅ</i> n. <i>rikar</i> 2 men <i>rindek</i> 2 women	

=(C)/=(V)	* <i>oru</i> / * <i>ōr</i> “1”	* <i>onru</i> “1”	* <i>oṅṭi</i> “alone”	* <i>okk-</i> “1”	* <i>iru</i> / * <i>īr</i> “2”	* <i>muv</i> / * <i>mū</i> “3”
DEDR	[N ^o 990 (a)]	[N ^o 990 (d)]	[N ^o 990 (c)]	[N ^o 990 (b)]	[N ^o 474]	[N ^o 5052]
Manda	<i>ru, runḍi</i>				<i>ri; rikar</i> 2 men <i>rikehiṅ</i> 2 wome	
Kui	<i>ro, ronḍi</i> n. <i>roanju</i> m.				<i>rī; rīṅḍi</i> f./ n. <i>ri’ari</i> 2 men <i>ri kōṛi</i> 40	<i>mū.</i> ; <i>mūṅḍi</i> f./ n. <i>mu’ar</i> 3 men <i>mū kōṛi</i> 60
Kuwi	<i>rō; ro’esi</i> 1 man <i>ronḍi</i> 1 woman				<i>rī; ri’ari</i> 2 men <i>riṅḍi</i> f./ n.	
Kurukh	<i>or^ot</i> m., <i>ol^ox</i> f. <i>ormā</i> all	<i>ōn(d)</i> one whole	<i>oṅṭā</i> 1 thing	<i>oṅghon</i> once	<i>ē:r, ēṅḍ</i> 2 thing <i>irb</i> 2 men	<i>mūnd</i> 3 things <i>nubb</i> 3 men
Malto	<i>ort, -e</i> m., <i>-i</i> f. <i>orme</i> all	<i>-ond</i> 1 thing			<i>-is</i> 2 things <i>iwr</i> 2 men	
Brahui	<i>asiṭ</i> 1 (entity) <i>asi</i> adj. <i>asīke</i> once				<i>irat</i> 2 (entities) <i>irā</i> adj.	<i>musiṭ</i> 3 (entities) <i>musi</i> adj. <i>musika</i> thrice
=(C)/=(V)	* <i>nāl</i> “4”	* <i>cayN</i> / * <i>cay</i> “5”	* <i>caru</i> / * <i>cār</i> “6”	* <i>eṛu</i> / * <i>ēr</i> “7”		
DEDR	[N ^o 3655]	[N ^o 2826]	[N ^o 2485]	[N ^o 910]		
Tamil	<i>nāl(u), nālku, nāṅku</i> 4 <i>nāṛpatu, nappatu</i> 40 <i>nalvar</i> 4 men	<i>aintu, aṅcu</i> ; adj. <i>ai</i> 5 <i>aim-patu</i> 50 <i>aivar</i> 5 men	<i>āru</i> 6 <i>aṛu-patu</i> 60 <i>aṛuvar</i> 6 men	<i>ēru</i> 7 <i>eṛu-patu</i> 70 <i>eṛuvar</i> 7 men		
Mala- yalam	<i>nāl(u), nāṅku, nān</i> 4 <i>nalpatu</i> 40 <i>nāl(uv)ar</i> 4 men	<i>aṅcu</i> ; adj. <i>ai</i> 5 <i>am-patu</i> 50 <i>aivar</i> 5 men	<i>āru</i> 6 <i>aṛu-patu</i> 60 <i>aṛuvar</i> 6 men	<i>ēru</i> 7 <i>eṛu-patu</i> 70 <i>eṛuvar</i> 7 men		
Kasaba	<i>nālu; nālattu</i> 40	<i>aṅju; ayvattu</i> 50	<i>āru; aṛavattu</i> 60	<i>ēlu; eḷuvattu</i> 70		
Kurru	<i>nālu; nālabbay</i> 40	<i>aṅju; yābay</i> 50	<i>āru; aravay</i> 60	<i>ōgu; ḍabbay</i> 70		
Kota	<i>na·ng; na·n digl</i> 4 months; <i>na·lvat</i> 40 <i>na·r pa·d</i> 40 days	<i>aṅj</i> 5 <i>ay vat</i> 50	<i>a·r</i> <i>ar vat</i> 60	<i>e·y, dial. e·g</i> <i>eḷ vat</i> 70		
Toda	<i>no·ng; pa·ng</i> 14 <i>naḷ poḥ</i> 40	<i>üz, dial. üj; pu·j</i> 15 <i>e boḥ</i> 50	<i>o·r; pa·r</i> 16 <i>aṛoḥ</i> 60	<i>öw; pa·w</i> 17		
Kan- naḍa	<i>nāl(u), nāl(u)ku, nā(ku)</i> <i>nalvattu</i> 40 <i>nalvar</i> 4 men	<i>ay(i)du</i> ; adj. <i>ai</i> <i>ay-vattu</i> 50 <i>aybaru</i> 5 persons	<i>āru</i> <i>aṛu/a-vattu</i> 60 <i>aṛuvar,</i> <i>arvaru</i> 6 men	<i>ēru</i> <i>ēr-pattu</i> 70 <i>ēṛvaru</i> 7 men		
Koḍagu	<i>na·lī; na· padī</i> 40 <i>na·vē</i> 4 men	<i>aṅji; aim-badī</i> 50 <i>ayvē</i> 5 men	<i>a·rī; aru-vadī</i> 60	<i>ē·lī; eḷu-vadī</i> 70 <i>ēvvē</i> 7 men		
B. Ku- rumba	<i>na:kṛ</i>	<i>əyidu</i>	<i>a:rṛ</i>	<i>ō:lṛ</i>		
Šōlega	<i>na:kṛ/ō; na:lvattu</i> 40	<i>ayidu/ō; aivattu</i> 50	<i>a:rṛ/ō; aṛuvattu</i> 60	<i>ō:lṛ/ō; yappattu</i> 70		
Tulu	<i>nālṛ; nālpa</i> 40 <i>nālverṛ</i> 4 men	<i>ainṛ</i> 5 things; <i>aiva</i> 50 <i>aiverṛ, aiyerṛ</i> 5 men	<i>āji; āj(i)pa, ājip(p)a</i> 60	<i>ēḷṛ; ēḷpa, erpa</i> 70 <i>ēḷverṛ</i> 7 men		
Koragu	<i>nālu</i>	<i>ayṅṅu</i>				
Telugu	<i>nālugu, nāluvu</i> <i>naluvadi, nalubadi</i> 40 <i>naluguru, naluvuru</i> 4 men	<i>ayidu, ēnu</i> 5 (things) <i>ē:badi, yabhai</i> 50 <i>ayidu-guru, ēguru,</i> <i>ēvuru</i> 5 men	<i>āru</i> <i>aṛu-vadi, aruvai</i> 60 <i>āruguru, āruvuru</i> 6 men	<i>ēḍu, inscr. (7th c.) ēṛu</i> <i>ḍebbadi, ḍebbhai</i> 70 <i>ēḍuguru</i> 7 men		
Kolami	<i>na·liṅ</i> 4 things <i>nalgur</i> 4 men <i>nallav</i> 4 women	<i>aid & ayd</i> <i>segur</i> 5 men	<i>ār</i> <i>argur</i>	<i>eḍ, ēṛ</i> <i>ergur, eḍgur</i> 7 men		
Naiki	<i>nālīṅ</i> 4 things Ch. <i>nāli</i> non-m. <i>nalgur</i> 4 men <i>nallaḷ</i> 4 women	Ch. <i>sēndi</i>	Ch. <i>sādi</i>			

(C)/(V)	*nāl “4”	*cayN / *cay “5”	*caru / *cār “6”	*erū / *ēr “7”	
DEDR	[N ^o 3655]	[N ^o 2826]	[N ^o 2485]	[N ^o 910]	
Parji	<i>nālu(k)</i> 4 things adj. <i>nel</i> ; <i>nelvir</i> 4 men <i>nelal</i> 4 women	<i>cē:du(k)</i> 5 things adj. <i>cem/n</i> ; <i>cevir</i> 5 men <i>ceyal</i> 5 women			
Gadba	<i>nalgur</i> , <i>nalvur</i> m. <i>nandal</i> , <i>naṇḍal</i> f. <i>nālig</i> , <i>nālug</i> n.				
Gondi	<i>nālung</i> ; <i>nālk</i> 4 each <i>nāl-vir/-vur/-gur</i> 4 men <i>nāluṅ</i> , <i>lālū</i> non-m.	<i>saiyung</i> , <i>sīyuy</i> , <i>hayuy</i> , <i>ayṅ</i> non-m., <i>sey(y)ur</i> , <i>sīvir</i> , (<i>h</i>) <i>ayvur</i> m., <i>saik saik</i> 5 each	<i>sārung</i> , <i>harung</i> (M.), <i>ārū sar^{vir}/-vur</i> , (<i>h</i>) <i>arvur</i> m., <i>sarne</i> sixth day	(<i>y</i>) <i>ērung</i> , <i>yedung</i> , <i>ērū</i> : non-m., <i>ēr-vur/-vir</i> , <i>ēr(v)ur</i> m. <i>yērḱ</i> 7 each	
Koṇḍa	<i>nāl^{er}</i> m., <i>nālgī</i> non-m.			<i>ērū</i>	
Kui	<i>nal</i> ; <i>nālgī</i> non-m. <i>nālur</i> 4 men	<i>siṅgi</i> 5 things, <i>sēṅgi</i> f./n., <i>sēṅ gōri</i> 100	<i>saj</i> , <i>sajgi</i> 6 things dial. <i>hāja</i> 6	<i>oṛi</i> ; <i>oḍgi</i> 7 things <i>oṛi gōri</i> 140	
Kurukh	<i>nāx</i> 4 things <i>naib</i> m./ f.				
(C)/(V)	*eṭṭu / *eṅ “8”	*on- / *or-paktu “9”	*toḷ-(paC-)tu	*pak-tu “10”	*nūr-tu “100”
DEDR	[N ^o 784]	[N ^o 1025]	[N ^o 3532]	[N ^o 3918]	[N ^o 3729]
Tamil	<i>eṭṭu</i> ; <i>eṅ-patu</i> 80 <i>eṅmar</i> , <i>eṅvar</i> 8 men	<i>oṅpatu</i> , <i>ompotu</i>	<i>toṅṭu</i> 9 <i>toṅ-nūru</i> 90 <i>toḷḷ-ayiram</i> 900	<i>pattu</i> , <i>paktu</i> ; <i>paḱ-pattu</i> 100 = 10 x 10 <i>patin-</i> 10 in teens <i>-patu</i> 10 in tens <i>patinmar</i> 10 men	<i>nūru</i> , obl. <i>nūrru-</i> <i>nūrruvar</i> 100 men
Mala- yalam	<i>eṭṭu</i> ; <i>eṅ-patu</i> 80 <i>eṅmar</i> 8 men	<i>ompatu</i>	<i>toṅ-nūru</i> 90 <i>toḷḷ-ayiram</i> 900	<i>pattu</i> , obl. <i>patin-</i> <i>-patu</i> in tens <i>patinmar</i> 10 men	<i>nūru</i> , obl. <i>nūrru-</i> <i>nūrru^u/avar</i> 100 men
Kasaba	<i>eṭṭu</i> ; <i>embattu</i> 80	<i>ombattu</i> 9	<i>tombattu</i> 90	<i>pattu</i>	
Kurru	<i>oṭṭu</i> ; <i>eṅ(a)bay</i> 80	<i>ombedu</i> 9		<i>pot(t)u</i>	
Kota	<i>eṭ</i> ; <i>em bat</i> 80	<i>orba-d</i> , dial. <i>onba</i> 9 <i>patrmba-d</i> 19	<i>tombat</i> 90	<i>pat</i> ; <i>paṅ</i> 11, <i>pad-</i> in 13, 14, <i>padn-</i> in 15–18 <i>-vat/-bat</i> in some tens	<i>nu·r</i>
Toda	<i>ōṭ</i> ; <i>pu-ṭ</i> 18	<i>winboṭ</i> ; <i>pu-nboṭ</i> 19 <i>e-ṅ boṭ</i> 90		<i>pot</i> ; <i>pon-</i> in 11–13; <i>-(f)oth/-poṭ/-boṭ</i> in tens	<i>nu·r</i>
Kana- ṇḍa	<i>eṅṭu</i> ; <i>eṅ-pattu</i> 80 <i>eṅbar</i> 8 men	<i>omb(h)attu</i>	<i>tom-b(h)attu</i> 90	<i>pattu</i> ; <i>padin-</i> teens in 15–18; <i>padi-</i> in 13, 14; <i>pann-</i> in 11, 12; <i>-vattu/</i> <i>-vattu/-battu</i> in tens <i>padimbar</i> 10 men	<i>nūru</i> <i>nūr(v)ar</i> 100 men
Koḍagu	<i>ēṭṭi</i> ; <i>ēm-badi</i> 80	<i>oymbadi</i>	<i>tom-badi</i> 90	<i>patti</i> ; <i>padin-</i> in 14–18, <i>padi-</i> in 13, <i>pann-</i> in <i>pannandi</i> 11, <i>pannerandi</i> 12 <i>-vadi/-padi/-badi</i> in tens	<i>nu·ri</i> , obl. <i>nu-iṭ-</i>
B. Ku- rumba	<i>-ōṭṭu</i>	<i>-embadu</i>		<i>-pattu</i>	
Šōlega	<i>eṅṭ^u/o</i> 8; <i>yambattu</i> 80	<i>ombattu</i> 9; <i>tombattu</i> 90		<i>attu</i> 10	<i>nu:ru</i>
Tulu	<i>eṅ(u)ma</i> 8; <i>eṅpa</i> 80	<i>ormba</i>	<i>soṅpa</i> 90	<i>pattu</i> 10; <i>patt-</i> in 11, <i>padu-</i> in 12–14, <i>paduṅ-</i> in 15–19, <i>-va/-pa</i> , obl. <i>oṭu</i> in tens	<i>nūdu</i> , obl. <i>nūta-</i> 100

(C)/(V)	*eṭṭu/*eṇ “8”	*oṇ/*or-paktu “9”	*toḷ-(paC-)tu	*pak-tu “10”	*nūr-tu “100”
DEDR	[N ^o 784]	[N ^o 1025]	[N ^o 3532]	[N ^o 3918]	[N ^o 3729]
Telugu	<i>enimidi</i> 8; <i>enu-badi</i> , <i>enabhai</i> 80 = inscr. <i>eṇ(um)bodi</i> ; <i>enamandru</i> 8 persons		<i>tommidi</i> 9 <i>tombadi</i> , <i>tombhai</i> 90 <i>tomma(n)nūru</i> 900 <i>tommaṇḍuguru</i> , <i>tommaṇḍru</i> 9 persons	<i>padi</i> ; <i>padun-</i> in 11, 15–18 <i>padu-</i> in 13, 14, <i>pan-</i> in 19; <i>-vadi/-padi/-badi</i> in tens inscr. <i>padunru</i> 10 persons	<i>nūru</i> , obl. <i>nūṭa-</i> <i>nūruḡ/vuru</i> 100 men
Kolami	<i>enumadi</i> , <i>enumidi</i> 8 <i>enmāṭar</i> 8 persons		<i>tomdī</i> 9 <i>tomater</i> 9 men	<i>padī</i> , <i>padi</i>	
Gondi	<i>armu^r/l</i> , <i>yermud</i> , <i>enmidi</i> , <i>tenmidi</i> 8 <i>armuhk</i> 8 each	<i>eḍmu</i> , <i>unmā</i> 9 <i>unmak</i> 9 each	<i>tomidi</i> , <i>tumidi</i> 9 <i>tombai</i> 90	<i>pad(i)</i> , pl. <i>patk</i>	<i>nūr</i> , pl. <i>nuhk</i>

Abbreviations of languages and language periods

Australian; **AL**taic; **AU**stroasiatic; **CD** — Central Dravidian; **EC** — Eastern Cushitic; **FU** — Fenno-Ugric; **K**artvelian; **MK** — Middle Korean; **No**stratic; **PN** — Pama-Nyungan; **SW** — South-Western.

References

- ANDRONOV 1978 — Михаил С. АНДРОНОВ [Mikhail S. ANDRONOV]. *Сравнительная грамматика дравидийских языков*. М.: «Наука», Главная редакция восточной литературы [A Comparative Grammar of the Dravidian Languages. Moscow: “Nauka”].
- ANDRONOV 1980 — Михаил С. АНДРОНОВ [Mikhail S. ANDRONOV]. *Язык брауи*. М.: «Наука», Главная редакция восточной литературы [The Brahui Language. Moscow: “Nauka”].
- ANDRONOV 1994 — Михаил С. АНДРОНОВ [Mikhail S. ANDRONOV]. *Сравнительная грамматика дравидийских языков*. 2-е изд. Москва: Российская Академия наук — Институт востоковедения [A Comparative Grammar of the Dravidian Languages. 2nd Ed. Moscow: Russian Academy of Sciences — Institute for Oriental Studies].
- BLAKE 1981 — Barry J. BLAKE. *Australian Aboriginal Languages*. London–Sydney–Melbourne–Singapore–Manila: ANGUS & ROBERTSON Publishers.
- BLAKE 1988 — Barry J. BLAKE. Redefining Pama-Nyungan: Towards the Prehistory of Australian Languages // *Aboriginal Linguistics*, 1; pp. 1–90.
- BLAŽEK 1992 — Václav BLAŽEK. Australian Elements in Dravidian Lexicon? // *Nostratic, Dene-Caucasian, Austric and Amerind* / Ed. Vitaly SHEVOROSHKIN. Bochum: BROCKMEYER; pp. 421–431.
- BLAŽEK 1999 — Václav BLAŽEK. *Numerals: Comparative-Etymological Analyses and Their Implications*. Brno: MASARYKOVA univerzita.
- BRAY 1934 — Sir Denys BRAY. *The Brahui Language*. Part III: *Etymological Vocabulary*. Delhi: Manager of Publications.
- CALDWELL 1913 — Robert CALDWELL. *A Comparative Grammar of the Dravidian or South Indian Family of Languages*. New Delhi: Munshiram MANOHARLAL. 1913 [1974].
- CURR 1886 — Edward M. CURR. *The Australian Race: Its Origin, Languages, Customs, Place of Landing in Australia, and the Routes by Which It Spread in Itself over That Continent*. I – III. Melbourne: John FERRES, Government Printer.
- DEDR — *A Dravidian Etymological Dictionary*. Second Edition. By Thomas BURROW & Murray B. EMENEAU. Oxford: Clarendon Press 1984.
- DIFFLOTH 1976 — Gérard DIFFLOTH. Mon-Khmer numerals in Aslian languages // Gérard DIFFLOTH & Norman H. ZIDE. (eds.) / Special issue: *Austroasiatic Number Systems* (= *Linguistics* 174: Special publication). The Hague: MOUTON & C^o B.V.; pp. 31–37.
- DIXON 1980 — Robert DIXON. *The Languages of Australia*. Cambridge: University Press.

- EVANS 1988 — Nicholas EVANS. Arguments for Pama-Nyungan as a Genetic Subgroup, with Particular Reference to Initial Laminalization // *Aboriginal Linguistics*, 1; pp. 91–110.
- EW — *Elamisches Wörterbuch*. by Walther HINZ & Heidemarie KOCH. Berlin: REIMER. 1987.
- ILLIČ-SVITYČ 1971 — В. М. Иллич-Свитыч [Vladislav M. ILLIČ-SVITYČ]. *Опыт сравнения ностратических языков (семитохамитский, картвельский, индоевропейский, уральский, дравидийский, алтайский): Введение. Сравнительный словарь*. М.: «Наука» [A Tentative Comparative Dictionary of the Nostratic Languages (Semitic-Hamitic, Kartvelian, Indo-European, Uralic, Dravidian, Altaic)]. Т. I (b–K)..
- KITTEL 1908 — Ferdinand KITTEL. *Grammar of the Kannada Language*. Mangalore.
- KLIMOV 1998 — Georgij A. KLIMOV. *Etymological Dictionary of the Kartvelian Languages*. Berlin–New York: MOUTON–DE GRUYTER.
- KLUGE 1938 — Theodor KLUGE. *Die Zahlbegriffe der Australier, Papua und Bantuneger nebst einer Einleitung ueber die Zahl: Ein Beitrag zur Geistesgeschichte des Menschen*. Berlin–Steglitz: Selbstverlag.
- KÖNIG 1965 — Friedrich W. KÖNIG. *Die elamischen Königsinschriften*. Graz: *Archiv für Orientforschung*, Beiheft 16.
- KRISHNAMURTI 2001 — Bhadriraju KRISHNAMURTI. *Comparative Dravidian Linguistics*. Oxford: University Press.
- KUIPER 1966 — Franciscus B. J. KUIPER. The Sources of the Nahali Vocabulary // *Studies in Comparative Austroasiatic Linguistics* / Ed. by Norman H. ZIDE. London–The Hague–Paris: MOUTON; 57–81.
- MCALPIN 1981 — D. W. MCALPIN. Proto-Elamo-Dravidian: The Evidence and Its Implications // *Transactions of American Philosophical Society* 71/3. Philadelphia.
- MENGES 1968 — Karl H. MENGES. *The Turkic Languages and Peoples: An Introduction to Turkic Studies*. Wiesbaden: HARRASSOWITZ.
- MÜLLER 1886 — Friedrich MÜLLER. *Grundriss der Sprachwissenschaft*, II.1. Wien; 1–98.
- PINNOW 1959 — Heinz-Jürgen PINNOW. *Versuch einer Historischen Lautlehre der Kharia-Sprache*. Wiesbaden: HARRASSOWITZ.
- SASSE 1982 — Hans-Jürgen SASSE. *An Etymological Dictionary of Burji*. Hamburg: BUSKE.
- SCHMIDT — Wilhelm SCHMIDT. Die Gliederung der Australischen Sprachen // *Anthropos*, Bd. 7 (1912), Bd. 8 (1913), Bd. 9 (1914), Bd. 12–13 (1917–1918).
- STAROSTIN 2000 — Георгий С. СТАРОСТИН [Georgij S. STAROSTIN]. *Сравнительная фонетика дравидийских языков*. Диссертация на соискание ученой степени кандидата филологических наук. М. [A Comparative Phonology of the Dravidian Languages. PhD. Thesis. Moscow].
- TROMBETTI 1923 — Alfredo TROMBETTI. *Elementi di glottologia*. Bologna: ZANICHELLI.
- TYLER 1986 — Stephen TYLER. *Proto-Dravido-Uralian*. Austin (Rice University) [Ms].
- TYLER 1990 — Stephen TYLER. Summary of Noun and Verb Inflectional Correspondences in Proto-Dravidian and Proto-Uralian // *Proto-Languages and Proto-Cultures* / Ed. by Vitaly SHEVOROSHKIN. Bochum: BROCKMEYER, 68–76.
- UEW — *Uralisches etymologisches Wörterbuch* / Ed. by Károly RÉDEI. Akadémiai Kiadó. 1986–1988.
- ZIDE 1978 — Norman H. ZIDE. *Studies in the Munda Numerals*. Mysore: Central Institute of Indian Languages.
- ZVELEBIL 1970 — Kamil ZVELEBIL. *Comparative Dravidian Phonology*. The Hague–Paris: MOUTON.
- ZVELEBIL 1977 — Kamil ZVELEBIL. *A Sketch of Comparative Dravidian Morphology*. The Hague–Paris–New York: MOUTON.
- ZVELEBIL 2001 — Kamil ZVELEBIL. *Nilgiri Areal Studies* / Ed. by Jaroslav VACEK & Jan DVOŘÁK. Prague: The Karolinum Press.

Р е з ю м е

Автор разделяет мнение, согласно которому дравидийская языковая семья является одной из ветвей ностратической макросемьи; вместе с тем отмечается, что в системе числительных, восстанавливаемой для прадрavidийского, надежных ностратических параллелей практически не обнаружено. В статье перечислены все основные гипотезы о происхождении дравидийских числительных, включая как внешние, так и внутренние этимологии; к ним автор добавляет ряд собственных, предлагаемых им в рамках разрабатываемой им теории о наличии австралийского субстрата в прадрavidийском. Согласно выводам автора, австралийская этимологизация ряда дравидийских числительных в целом оказывается предпочтительнее, чем известные сопоставления их с ностратическим материалом.

Некоторые проблемы сравнительно-исторического иранского языкознания¹

Языки иранской семьи, носители которых расселились по разным регионам Евразии, контактировали с носителями языков других семей и между собой, наслаивались на разные субстраты и подверглись значительной типологической перестройке. Это представляет немало трудностей для их сравнительно-генетического изучения. В статье рассматриваются: 1) ареальная и темпоральная стратификация праиранского языка, представлявшего диалектный континуум; 2) этапы филиации арийской «ветви» индоевропейской семьи, приведшие к выделению праиранского; 3) генетическое членение иранской семьи; 4) вопросы методики сравнительно-исторической реконструкции.

Иранская языковая семья, входящая (через определенные промежуточные этапы) в большую индоевропейскую в виде ее «ветви», при всей, казалось бы, прозрачности генетических позиций как отдельных составляющих ее языков, так и семьи в целом, создает еще немало проблем в изучении ее истории. Это относится к исследованию иранских языков в сравнительно-историческом, историко-типологическом и ареальном аспектах. В какой-то мере трудности изучения истории этих языков связаны с трудностями разграничения генетического, типологического и ареального компонентов, стимулирующих диахронические изменения. Последнее обстоятельство диктуется внешней историей языков, составляющих данную семью: иранские языки — живые и зафиксированные в памятниках, — будучи индоевропейскими, оказались разбросаны по огромным территориям Евразии, причем ни в одной из областей их нынешнего распространения они не являются аборигенными. Отсюда — языковые контакты различного типа, воздействие различных субстратов, адстратов, суперстратов, не исключая, впрочем, спонтанных процессов, общих генетически и типологически с другими индоевропейскими языками Европы и Азии.

Кроме того, определение хода истории каждого конкретного языка затруднено тем, что основная часть вымерших языков не имеет современных продолжений, а современные языки не имеют древних памятников письменности. В качестве образцов «прямой преемственности» можно назвать лишь две линии: 1) древнеперсидский — среднеперсидский — современные персидский, таджикский, дари; 2) согдийский — ягнобский. Однако и здесь мы имеем дело с продолжениями диалектов, не отраженных памятниками письменности более раннего состояния.

В настоящей статье предполагается обратить внимание читателей на некоторые проблемы истории иранских языков, существенные, как представляется, не только для иранистики, но и для диахронического изучения других индоевропейских языков Европы и особенно Азии. Большая часть этих проблем была рассмотрена в ирановедческой литературе разных лет, однако теперь, когда подводятся итоги лингвистических исследований второй половины XX века и намечаются перспективы сравнительного изучения языков XXI века, имеет смысл заново обратиться внимание на определенный комплекс проблем, связанных с историей иранской языковой семьи в целом и в отдельных ее компонентах.

Известно, что сравнительно-историческое иранское языкознание, которое с самого зарождения в начале XIX века формировалось и развивалось в русле индоевропейского языкознания [АБАЕВ 1962: 3], продолжалось в традициях историко-генетических исследований и в XX веке. Следует подчеркнуть, что в нашей стране эти традиции не исчезали даже в годы запрета на индоевропеистику и на все, что с нею было связано. (В тот период оперирование понятиями генетического родства языков требовало немалого мужества, которое проявлял, например, в конце 40-х — начале 50-х годов проф. Б. В. Миллер при чтении курса «Введение в иранскую филологию» на Филологическом факультете МГУ.)

При этом сохранение традиций не препятствовало прогрессу сравнительно-исторической иранистики, впитывавшей то новое, что давала лингвистическая теория и практика XX в. Так, бурное развитие описательной грамматики, постоянно вводящей в научный обиход неизученные или слабо изученные прежде иранские языки и диалекты, помимо накопления огром-

¹ Этимологическая база данной статьи создавалась в ходе разработки корпуса *Этимологического словаря иранских языков*, работа над которым финансируется РГНФ, проекты №№ 02-04-00071, 05-04-04170, 08-04-00198.

ного нового языкового материала, способствовало выработке подхода к языку в его синхронном состоянии как к определенной системе. Это сказалоь и на исторических исследованиях, в частности, на представлении реконструируемых единиц и фрагментов предшествующего состояния (одного языка или праязыка для группы языков, либо семьи в целом) как элементов некоей системы, связанных между собой определенными отношениями. Были учтены, в частности, достижения фонологии, что позволило впоследствии внести определенные коррективы в представление о сути таких необходимых понятий сравнительно-исторического исследования, как «фонетический закон» и «звукосоответствия». Были более четко сформулированы различия между фонетическими изменениями и/или тенденциями в прошлых состояниях языка и последующей фонологизацией (или ре-фонологизацией) фонетических элементов и их отношений, образовавшихся в результате этих процессов; были уточнены различия между соответствиями фонем и частичными соответствиями позиционных вариантов и т. д.

Большую помощь компаративистическому изучению иранских языков оказал осознанный учет результатов их исследования в типологическом (включая синхронную и диахроническую типологию, как структурную, так и контенсивную, см. [ОИТИИЯ 1975]) и ареальном аспектах [РАСТОРГУЕВА 1964]; [ЭДЕЛЬМАН 1968]. Тем самым применение в сравнительно-исторических штудиях результатов исследования языков другими методами, а также учет достижений теоретической лингвистики XX в., способствовали тому, что в методическом плане сравнительное иранское языкознание конца XX в. не оставалось на уровне XIX в. и даже первой половины XX века.

Помимо сравнительно-исторического изучения языков иранской семьи в целом или языков в пределах отдельных «ветвей», развивалось также сравнительно-историческое описание каждого иранского языка по отдельности, результатом чего явилось многотомное издание «Основы иранского языкознания» [ОИЯ 1979–2008], в котором такое освещение получил практически каждый из известных иранских языков — живых и вымерших. (Следует заметить, что в плане системности синхронного анализа и глубины исследования истории большинства языков, включая «малые» бесписьменные, этот труд намного четче и современнее — в хорошем смысле этого слова, — чем сопоставимый по времени обобщающий компендиум [СЛИ 1989]).

Все это способствовало более глубокому и реалистичному подходу к сравнительно-историческим штудиям в иранистике и сказалоь на интерпретации и реинтерпретации ряда диахронических законов и частных трансформаций.

Рассмотрим основные из них.

I. Реконструкция общеиранской праязыковой системы.

Как известно, классическая иранистика XIX (и это продолжилось в начале XX века), как и классическая индоевропеистика того периода, в реконструкциях отдельных элементов праязыковой системы и целостной прасистемы базировалась в основном на материале языков древних памятников. Филологический метод был тогда решающим, и сама методика компаративистики зарождалась и совершенствовалась на благодатном для этого материале древних индоевропейских языков. Данные живых языков использовались в этих целях в значительно меньшей степени, причем привлекались в основном «большие» языки, имевшие письменную традицию и лучше изученные. Из многочисленных живых иранских языков, особенно «малых» бесписьменных, только часть имела тогда первые описания и публикации материалов, иногда в очень несовершенной записи; часть языков была практически известна только по названию, а некоторые еще не были даже открыты. Не были известны тогда и письменные памятники некоторых вымерших языков (согдийского, хорезмийского, бактрийского, сакского, парфянского).

В этот период реконструкция древней фонетической системы (и отчасти фонологической, хотя этот аспект еще не был сформулирован), морфологии, элементов лексики (включая праязыковые корни, основы или слова) базировалась на свидетельствах древних же иранских языков — авестийского и древнеперсидского — и считалась наиболее архаичной, общеиранской, лежащей в основе истории всех иранских языков. Известный к тому времени материал более поздних языков, вымерших и живых, не противоречил в целом этой системе, а случаи отражения в этих языках древних элементов, не вписывающихся в традиционную реконструкцию, еще не привлекали к себе достаточного внимания (хотя и констатировались иногда как интересные факты), а иногда рассматривались как инновации. Единственным исключением конца XIX в., результаты которого частично сказались на иранистике начала XX в., была праиранская реконструкция Хр. Бартоломе [ВАРТОЛОМАЕ 1895–1901], который кроме древних иранских языков

(с учетом также древнеиндийских данных), принимал во внимание и показания поздних языков, включая живые, хотя установленные им закономерности также относятся больше к древним иранским языкам, чем к более поздним. В методическом плане — четким разграничением двух систем: 1) праязыковой системы и 2) той древнеиранской, которая зафиксирована в языках древних памятников, — Хр. Бартоломе намного опередил не только современников, но и многих историков языка последующих периодов. (В трудах отдельных исследователей даже середины XX в. в качестве прототипа того или иного слова современного языка дается авестийское слово.). Подробнее об этом см. [РАСТОРГУЕВА, ЭДЕЛЬМАН 1981, 123–124]).

Дальнейшее приращение материала вымерших языков не-древнего периода и живых языков, особенно бесписьменных «малых» языков и диалектов, усилившееся в 20-е — 30-е годы XX в. и продолжающееся поныне, с одной стороны, и историко-лингвистическое осмысление этого материала, — с другой (см., например, труды Э. Бенвениста, В. Хеннинга по согдийскому языку, труды В. И. Абаева по осетинскому [АБАЕВ ОЯФ]; [ИЭСОЯ], Г. МОРГЕНСТЕРНЕ по бесписьменным языкам Памиро-Гиндукушского региона [MORGENSTIERNE PFL], [EVSh], [ID] и многие другие), — обусловили новую ступень развития компаративистических исследований языков иранской семьи. Это сказалось не только на выявлении места этого материала в общей схеме развития иранской и шире — арийской языковой семьи, но и на коррекции самой схемы филиации арийской языковой общности, а также на коррекции реконструируемых протосистем праарийского и праиранского состояния (см. об этом ниже).

Обнаружилось, что некоторые «отклонения» в живых и вымерших языках (даже имеющих большую традицию изучения) от «классической» схемы развития обусловлены не инновациями, а сохранением архаизмов, либо развитием особых рефлексов архаичного состояния, отличных от традиционно признаваемых.

Так, например, в числительном «тринадцать» в среднеперсидском языке *sēzdah* {*syčdh*} и в его продолжениях — классическом персидском и таджикском: кл. перс. *sēzdah*, тадж. *sezdah* «тринадцать», — появление *-z-* традиционно считается результатом аналогии в среднеперсидском с числительным ср.-перс. *pānzdah* {*p'nčdh*}, ман. {*p'nzdh*}), «пятнадцать». Это выглядит несколько странным, поскольку в промежуточном числительном «четырнадцать» (ср.-перс. *čahārdah* {*čh'ldh*}, кл. перс. *čahārdah* и тадж. *čahordah*, *čordah* из праиран. **čatvar-daša-* > др.-перс. **čathvar-datha-*) вставной *-z-* не появляется. Кроме того в самом ср.-перс. *pānzdah* «пятнадцать» и его продолжениях: кл. перс. *pānzdah*, тадж. *ponzdah*, — развитие консонантной группы *-nz-* является исключением из историко-фонетического правила отражения здесь древней группы **-nč-* в виде *-nj-*, см. [РАСТОРГУЕВА, МОЛЧАНОВА 1981, 78]). Скорее можно предположить, что ср.-перс. *pānzdah* «пятнадцать» продолжает сочетание **panča-ča daša*, букв. «пять-и десять» (как и большинство других числительных этого разряда, продолжающих сочетания с энклитическим союзом **-ča* «и») при последующей утрате первого **-ča-* (исхода основы) в результате гаплогонии.

В таком случае ср.-перс. *sēzdah* (> кл. перс. *sēzdah*, тадж. *sezdah*) «тринадцать» может либо иметь срединное *-z-* из того же соединительного союза **-ča* (аналогично некоторым другим числительным), либо продолжать еще более раннее сочетание. В праарийских прототипах двусловных числительных, кроме композитов и сочетаний с соединительным союзом, были и сочетания полноформенных слов типа арийск. **trajas daća* (ср. ведич. *tráyo-daša-*). В праиранском они могли продолжиться в виде сочетания **trajah daša*, отразишегося в части языков как **traj(a)-daša*, или — при раннем сращении общеарийского сочетания **trajas + daća* — в виде праиранского диалектного композита **trajaz-daša* (с озвончением **s > *z* перед звонким **d*). Продолжением последнего и могут являться ср.-перс. *sēzdah*, кл. перс. *sēzdah*, тадж. *sezdah* «тринадцать».

Было замечено также, что иранские языки, особенно бытующие (либо бытовавшие) в изолированных ареалах, включая «малые» бесписьменные живые языки, хранят иногда рефлексы более архаичного состояния, чем то, которое зафиксировано в древних памятниках, либо выявляют свои собственные инновации, не прошедшие через те этапы, через которые прошли инновации древних языков. А это дало возможность существенно откорректировать общеиранскую праязыковую фонетическую и фонологическую системы (и определить различия между ними), которые оказались более архаичными, чем системы языков древних памятников. Это обстоятельство помогло уточнить действительно праиранский облик реконструируемых словоформ, основ, корней и некорневых морфем, то есть фрагменты праязыковых морфолого-синтаксической и лексической подсистем.

Так, например, традиционно в иранистике считалось (и это положение попало в некоторые индоевропейские труды), что праиранскими рефлексам позднейшеевропейских палатальных

согласных $*\hat{k}$, $*\hat{g}$, $*\hat{gh}$ являются «свистящие» $*s$, $*z$, что объединяет иранские языки со славянскими. Однако материал иранских языков показал, что это не совсем так — свистящие $*s$, $*z$ наследовали палатальным не во всех фонетических позициях: яркое исключение — отражение их в виде шипящих в соседстве с согласными — свидетельствует о более раннем шипящем характере рефлексов, ср., например, и.-е. $*oktō(u)$ «восемь» > др.-инд. $aṣṭā$, $aṣṭāu$, праиран. $*ašṭā-$, $*ašṭāu$. Кроме того ассимиляция рефлексов палатальных не была синхронной во всех диалектах древнеиранского континуума. Свистящими они станут (но не в составе консонантных групп) в большинстве языков позднее, причем этот процесс, характерный для центрального ареала, не в полной мере охватит окраинные ареалы иранского континуума. В наиболее раннем общеиранском состоянии это были мягкие шипящие — глухая $*ś$, сходная с древнеиндийской, и звонкая $*ž$. В дальнейшем в большинстве языков и диалектов центрального ареала они переходят в $*s$, $*z$, но в окраинных ареалах представляющие их звукотипы дольше сохраняют иное, мягкое шипящее звучание. Именно этим можно объяснить их переход (кстати, давно установленный) в древнеперсидском соответственно в θ и $*\delta$ (с отражением последнего в клинописных памятниках в виде d). Ср., например, отражение $*\hat{k}$ в и.-е. $*dekṃ-t$ «десять»: праиран. $*daśa-$ > авест. $dasa-$, шугн. $ḍīs$, пушту las , осет. $dæx$ и т. д., но др.-перс. $*daθa-$ > совр. перс. $dāh$. В другом окраинном ареале — в скифо-осетинской группе диалектов — произошло совпадение $*ś$ с рефлексами $*š$ и затем с $*s$ в единую фонему, о чем свидетельствует сегодняшний фонологический статус s в осетинском (как единственной глухой фонеме-субиланта, не противопоставленной $š$, но реализующейся по диалектам в виде свистящих, свистяще-шипящих и шипящих звукотипов). Этим же объясняется сохранение шипящей артикуляции рефлексов $*ś$ в группе $*śu$ в третьем окраинном ареале — в сакско-ваханской ветви восточноиранских языков, ср., например, отражение и.-е. $*ekwo-s$ «конь, лошадь» > праарийск. $*ácya-$ > праиран. $*áśya-$, откуда хот.-сак. $aśsa-$, вах. $yaš$, при авест. $aspa-$ [MORGENSTIERNE PFL II: 469]; [BAILEY DKS: II]; [СТЕБЛИН-КАМЕНСКИЙ ЭСВЯ: 422]; [ЭСИЯ I: 243–246]. Характерно также отражение праиранского $*śu$ > др.-перс. s (например, др.-перс. $asa-$ «конь, лошадь» в составе композита $asa-bāra-$ «всадник»), в отличие от праиранского $*sp$, отражающегося всегда в виде др.-перс. sp .

Традиционно считалось (на основании показаний древнеперсидского и авестийского языков), что в число основных генетических черт, отличающих иранские языки от индоарийских, входит отражение праарийских глухих придыхательных согласных $*ph$, $*th$, $*kh$ в виде праиранских глухих щелевых $*f$, $*θ$, $*x$, в отличие от древнеиндийского, где они сохраняются как смычные придыхательные. Более пристальное внимание к иранским языкам разных эпох и разных ареалов, а также учет различий между фонетическими тенденциями и фонологизацией их результатов (либо отсутствием такой фонологизации) показали, что, даже отвлекаясь от того известного ныне факта, что серия глухих придыхательных не была фонологизирована в общеарийском состоянии, случаи артикуляции глухих смычных с придыханием далеко не во всех иранских языках приводили к их спирализации.

Ряд языков сохранил их смычное произнесение и соответствующее последующее развитие. К тому же в тех языках, где эта спирализация произошла (главным образом, в языках западноиранской ветви, но и спорадически в некоторых восточноиранских), она выявляет иногда лексически и фонетически (т. е. позиционно) закреплённые варианты со спиранным или со смычным согласным в прототипе. Во-первых, характерно отсутствие спирализации глухих в позиции после $*s$ и перед сонорными после носовых, забиравших себе часть воздушной струи (что свидетельствует кстати о слабости элемента аспирации в праарийских диалектах, легших в основу иранского континуума). Во-вторых, часть языков указывает на отсутствие спирализации в прототипах и в других позициях. Ср., например, авест. $kaofa-$ «гора, горб животного», др.-перс. $kaufa-$ > перс. kuh «гора», мундж. $kifa$ и др., где $f < *p(h)$, но пушту kup «сгорбившийся», $kubay$ «горбатый (о животном)», вах. $kəp$ «горб», язг. $kəp$, ст.-вандж. kup , kub «камень; гора, скала» и др., восходящие в конечном счете к и.-е. $*keu-p$. Особенно богата исключениями анлаутная позиция, где в разных языках сказывались разные тенденции к анлаутной интенсивности артикуляции (включая тенденцию к смычности или к аспирации глухих согласных), то есть отражение элементов фонетической и фонологической синтагматики, в частности, пограничных сигналов слова.

В числе общеиранских праязыковых черт традиционно называлась также спирализация праарийских глухих смычных в позиции перед согласными и неслоговыми вариантами сонантов, то есть переход в этой позиции (в условиях отсутствия следов аспирации) $*p$, $*t$, $*k$ в щелевые $*f$, $*θ$, $*x$ (типа $*pr > *fr$; $*tr > *θr$; $*kr > *xr$), однако и здесь мы можем говорить скорее о фонетической тенденции, чем об общеиранском фонологическом законе: многие языки

(особенно восточноиранские) выявляют следы сохранения смычности глухими согласными в этой позиции в древних диалектах. Эти следы выявляются в виде сохранения либо самого смычного согласного, либо элемента смычки в рефлексах консонантной группы. Ср. авест. *puθra-* «сын», др.-перс. *puça-* < **puθra-*, йидга *pūl, pūr*, сак. *pūrā* < **puθra-*, но вах. *pətr*, шугн. *pus* «сын» < **putra-*; то же в рефлексах числительного «три»: ягн. *tiray, si/aray* (где диалектные варианты с *t — s* продолжают праягнобское **θ*) < **θrai-*, шугн. *aray* «три» < **θrai-*, но в близкородственном шугнанскому язг. *cūy* < **trai-* (с сохранением элемента смычки), вах. *truy* < **trai-*. Это указывает на относительно поздние случаи спирализации, разделяемые даже не всеми языками одной и той же генетической группы. (Подробнее о статусе глухих щелевых в праиранском см. [ЭДЕЛЬМАН СГВЯ-Ф: 54–56, материал 139–142, 166–168, 186–187]).

В определенном уточнении нуждается также тезис об совпадающем отражении двух серий арийских звонких — придыхательных и непридыхательных — в единую праиранскую серию звонких смычных непридыхательных, т.е. **bh, *b > *b; *dh, *d > *d; *gh, *g > *g* с сохранением их впоследствии в начале слова в западноиранских языках в виде смычных **b-, *d-, *g-* и спирализации в восточноиранских в щелевые **v-, *δ-, *γ-*. Как показывает материал, разделение будущего иранского континуума на ареальные группы происходило в очень ранний период, когда аспирация еще была актуальна, а фонологическая оппозиция «смычная — щелевая» для звонкой серии еще не выработалась. В восточной части континуума (дольше контактировавшей с индоарийскими языками) аспирированные **bh, *dh, *gh* сохранялись дольше и могли уподобить себе неаспирированные. При этом арийские звонкие согласные аспирированной серии могли реализоваться как спиранты [**v-, *δ-, *γ-*] (подробнее о спирантной реализации аспирированных еще в позднем индоевропейском состоянии см., например, [КНОВЛОСН 1965: 153]; [СЕМЕРЕНЬИ 1980: 159], в индоарийских языках [ЕЛИЗАРЕНКОВА 1974: 181]). В западной части праиранского континуума аспирация звонких утрачивалась ранее, и направление уподобления было обратным, имея результатом [**b-, *d-, *g-*].

Тем самым, в целом справедливый тезис о совпадении двух серий праарийских звонких смычных согласных — придыхательных и непридыхательных, то есть **bh* и **b; *dh* и **d; *gh* и **g* в праиранскую единую серию звонких согласных, обозначаемую обычно **b, *d, *g*, — носит фонологический характер: при отсутствии в тот период оппозиции «смычные — щелевые» это была единая «совпавшая» серия звонких согласных, которая могла реализоваться и реализовалась в анлауте на фонетическом уровне в разных ареалах неодинаково. В западном ареале преобладала смычная артикуляция (наследующая неаспирированным и усиленная тенденцией к анлаутной интенсивности в виде смычности звонких), в восточном — щелевая (наследующая аспирированным). И только позднее, после становления оппозиции «смычные — щелевые» — в уже разделившихся языковых группах и даже отдельных языках, оказалось, что в языках западной и восточной групп начальные звонкие согласные включены в противоположные члены этой оппозиции. Срединная позиция, особенно интервокальная, способствовала фонетической спирализации звонких (с последующим соответствующим фонологическим отождествлением), позиция после **n* — их смычному отражению, и это в большинстве языков обеих групп.

Преимущественная связь звонких аспирированных и неаспирированных с их рефлексами (при разных направлениях их унификации) в разных диалектах праиранского континуума отразилась также в срединных консонантных группах, где просматриваются дальние следы действия «закона Бартоломе», актуального для общеарийского состояния. Здесь также совпадение двух серий арийских звонких согласных в западных диалектах вызывало их унификацию по типу неаспирированных (то есть в виде смычных), в восточных — по типу аспирированных (то есть в виде щелевых), что сформировало впоследствии один из генетических признаков противопоставления западных и восточных иранских языков (см. ниже).

Имеются и другие уточнения и исключения из историко-фонетических (практически — историко-фонологических) признаков, традиционно считавшихся общеиранскими.

Таким образом, уже на основании историко-фонетических и историко-фонологических данных ясно, что праязыковое состояние предстает в виде не единого «ствола» генеалогического дерева, а в виде континуума диалектов, с относительно «продвинутым» центральным ареалом и до некоторой степени архаичными (или развивавшимися несколько иными путями) маргинальными, окраинными. Существенными были и различия между древними диалектами западной и восточной зон. Отдельные историко-фонологические изоглоссы (причем достаточно древние) отделяли также древние диалекты юго-западной зоны от всех остальных,

диалекты юго-восточной зоны от других и т. д. (Эту «немонолитность», в которой возможны еще расхождения и схождения мелких диалектов, я постаралась отразить в схеме 1, см. ниже).

Отмечаются и другие признаки раннего и постепенного членения праиранского континуума, выявляющиеся на более высоких уровнях языковых систем иранских языков, включая морфологию, лексику, семантику. При этом и здесь материал живых языков и диалектов, особенно окраинных (или, во всяком случае, изолированных в силу природных или социальных условий), предоставляет иногда сведения о более архаичном праиранском уровне, чем древние языки. Это касается как определенных подсистем, так и частных фактов, например, наличия или отсутствия определенных лексем в праязыковую эпоху.

Так, живые «малые» языки Памиро-Гиндукушского региона (большинство памирских, мунджанский) донесли до нас трехсерийную бицентрическую систему указательных местоимений типа латинской (то есть с ориентацией как на пространственную сферу говорящего, так и на сферу собеседника), реконструированную для индоевропейского состояния еще К. Бругманом (*Ich-Deixis*, *Du-Deixis*, *Jener-Deixis*, с выполнением последним также функций *Der-Deixis*) и утраченную затем в подавляющем большинстве индоевропейских языков с трансформацией ее в двухсерийную моноцентрическую (т. е. с отсчетом только от сферы говорящего), типа «этот — тот». Характерно, что в указанных живых иранских языках сохраняются не только материальные рефлексy древних местоимений всех трех серий (они имеются и в других иранских языках, например, в пушту, согдийском), но и их основные функциональные нагрузки (подробнее см. [Беликов 1972]; [Эдельман 1976]; [СГВЯ-М: 220–225]; [Юсуфбеков 1998]).

Особую роль в истории иранских языков сыграло отсутствие в общеиранском состоянии глагола с абстрактной семантикой «иметь». Выработка его в последующие периоды из глагола **dar-*, обозначавшего прежде «держатъ, схватывать», охватила не все иранские языки. Часть живых языков сохранила за этим глаголом только древний круг значений, и благодаря этому мы можем проследить механизм развития не только лексических, но и морфолого-синтаксических последствий такого семантического архаизма (например, создание эргативообразных конструкций с переходными глаголами в прошедших временах, которым посвящена большая литература; развитие связочных оборотов разного типа и др.).

В реконструкции определенных слов, словоформ, словоизменительных и словообразовательных моделей, установлении относительной хронологии их продуктивности и степени их распространенности по ареалам также наблюдаются интересные свидетельства живых языков (включая «малые» бесписьменные) и диалектов. Например, реконструируемое на основании письменных памятников и современных «больших» языков обозначение «пальца» (руки или ноги), находящее аналогию в древнеиндийском, является индоиранским вторичным построением: оно соответствует модели превосходной степени и первоначально обозначало «большой палец» — праиран. **angušta-*, **angušti-*. При этом некоторые живые иранские языки, не очень близкие друг другу генетически и бытующие в не соприкасающихся ареалах, хранят рефлексy более раннего — индоевропейского образования, имеющего параллели (с различной семантикой) в других индоевропейских языках, — праиран. **anguli-*, **anguri-*. Из этого обстоятельства можно заключить, что в общеиранском состоянии сосуществовали (по диалектам?) два обозначения «пальца» — более архаичное **anguli-* и относительно новое (но тоже доиранское) **angušta-*. Имеются и другие, более частные примеры сохранения в «малых» окраинных или изолированных языках и диалектах архаизмов, не зафиксированных в древнеиранских памятниках или утраченных ко времени их создания (примеры см. [Эдельман 1996]).

Для характеристики общеиранской лексической системы существенны и иные, «отрицательные» в этом плане свидетельства. К ним относятся, например, слова, построенные в ряде языков (даже не очень близко родственных), бытующих в едином ареале, по единой (иногда субстратной) словообразовательной модели из единого же исконного материала, но не восходящие к общеиранской лексеме. Выявление таких слов не всегда просто, однако оно необходимо во избежание неоправданной проекции их прототипа на общеиранский экран. Таковы, например, личные местоимения 2 л. мн. ч. в ряде языков восточноиранского ареала (особенно входящих в Центральноазиатский языковой союз), построенные из исконного материала по субстратной модели.

Характерны прослеживаемые в лексике элементы диглоссии, которые могли быть в праязыковой период (о них см. ниже).

Все эти факты отсутствия монолитности общеиранского состояния помогают уточнить не только языковую систему диалектов, образывавших общеиранский праязыковой континуум, но

и этапы филиации арийской общности и затем иранской языковой семьи, о чем уже неоднократно приходилось писать [Эдельман 1982; 1992; 1996; 1999; 2002: 21 и сл.] и о чем будет сказано также ниже. Они же играют существенную роль в анализе хода истории конкретных иранских языков.

Таким образом процесс «распада» общеиранского праязыка и выделения из него определенных языков, диалектов, групп и подгрупп не был единовременным, «одноактным»: он начался, возможно, во времена выделения иранских диалектов из более ранних состояний — общеарийского и индоиранского (о них см. ниже) и не был единообразным и одновременным по всему ареалу праиранского континуума (см. об этом также [Оранский 1979: 119–121]).

От реконструкции общеиранской праязыковой системы зависит реконструкция других этапов развития иранских языков: а) ретроспективная реконструкция более ранних этапов, освещающая генетические отношения иранской семьи с другими арийскими и далее — с другими индоевропейскими языками, то есть этапы филиации более ранних общностей, приведшие к вычленению иранских языков, и б) проспективная — поэтапная филиация праиранского континуума и вычленение генетических групп, подгрупп и единичных языков.

II. Реконструкция этапов вычленения иранской языковой семьи из более ранних.

Наиболее сложным явилось выявление этапов вычленения иранской семьи из общеарийской, поскольку в традиционной схеме ставился знак равенства между общеарийским и индоиранским состоянием. К тому же общеарийской системе приписывался ряд черт, присущих на деле только древнеиндийскому, в частности, наличие упоминавшейся выше серии глухих придыхательных согласных фонем и некоторых других. Материалы открытых в первой половине XX века ряда живых и вымерших языков иранской семьи, а также живых языков, входящих в арийскую семью, но не относящихся генетически ни к иранским, ни к индоарийским, дали возможность более тщательной проработки тех инноваций, которые свидетельствовали о вычленении иранских языков из более ранних общностей.

Так, выявилось, что вопреки традиционной схеме, арийская общность распадалась не на две «подветви» — иранскую и индоарийскую (или «индийскую»), а как минимум на три или четыре: индоарийскую, иранскую, нуристанскую (традиционное название языков последней группы в начале XX века — «кафирские») и дардскую. Возможно были и другие группы и отдельные языки. Нуристанские языки, рассматривавшиеся раньше в ряде работ как «промежуточные» между индоарийскими и иранскими, оказались самостоятельной группой, рано отделившейся от других арийских, а дардские языки, традиционно считавшиеся генетическим дальним ответвлением от праиндоарийского, могли (все или часть из них) на раннем этапе составлять самостоятельную группу (или группы), не представлявшую ответвление ни одной из этих «подветвей», но входившую в древности в языковой союз с диалектами праиндоарийского [Коган 2000: 26–27; 2003: 99; 2005].

Основными свидетельствами этапов распада арийской общности являются (на данном этапе изученности) историко-фонетические (точнее, историко-фонологические) законы. Они указывают, что из известных нам языков первыми из общеарийского состояния выделились нуристанские, образовавшие генетически единую группу с общими материальными инновациями, и только позднее оставшийся континуум диалектов, который можно условно назвать собственно индоиранским, просуществовав какое-то время и выработав общие инновации, разделился на индоарийскую (или индийскую), дардскую и иранскую группы (либо на большее число групп, если учесть возможность бытования в прошлом не дошедших до нас языков). Об этом говорят следующие процессы:

Отражение поздних индоевропейских палатальных $*\hat{k}$, $*\hat{g}$, $*\hat{g}h$ различается в языках разных групп, причем их различия особенно отчетливо указывают на постепенность распада арийской семьи: 1) и.-е. $*\hat{k}$ через этапы раннеарийск. $*\acute{c}$ позднеарийск. $*\acute{c}$ отразился в виде пранурист. c (с отражением затем в виде s , z , то есть с продвижением палатальной аффрикаты $*\acute{c}$ вперед при сохранении аффрикатного характера), в то время как в собственно индоиранских языках позднеарийская $*\acute{c}$ утрачивала аффрикатность, продолжая сохранять палатальную реализацию в виде мягкой шипящей щелевой $*\acute{s}$: позднеарийск. $*\acute{c}$ > др.-инд. $*\acute{s}$, дард. $*\acute{s}$, праиран. $*\acute{s}$ (см. выше). Ср., например, и.-е. $*dek̑m-t-$ «десять»: нуристанские — язык кати duc , ашкун dus , прасун $lāze$ (с закономерным отражением $*d > l$) «десять». То же в преконсонантной позиции: ср. и.-е. $*ok̑tō(u)$ «восемь» > раннеарийск. $*a\acute{c}tā(u)$ > позднеарийск. $*a\acute{c}ta(u)$ - > пранурист. $*acta(u)$ - (отразившееся в языке прасун в виде $\bar{a}stē$ «восемь» [Kuiper 1978: 101]), но индоиран. $*a\acute{s}ta(u)$ > др.-инд. $a\acute{s}tā$,

aštāu, праиран. **aštā-*, **aštāu*. Аналогичным образом отражаются звонкие палатальные: и.-е. **ǵ* > раннеарийск. **j* > позднеарийск. **j* > пранурист. **ǰ* > *ǰ*, *z*, но др.-инд. *ǰ*, праиран. **z*; и.-е. **ǵh* > раннеарийск. **jh* > позднеарийск. **jh* > пранурист. **ǰh* > *ǰ*, *z*, но др.-инд. *h*, праиран. **z*. Подробнее об отражении палатальных в арийских языках см. [MORGENSTIERNE 1945: 228]; [BUDDRUSS 1977: 23 и сл.]; [ЭДЕЛЬМАН 1992: 56, 60–61]; [КОГАН 2005: 46 и сл.].

Имеются также следы других процессов в нуристанских языках, свидетельствующие об их относительно раннем отделении от арийской общности, когда индийские и иранские языки еще входили в относительное «индоиранское» единство. Например, для исторической фонетики и фонологии нуристанских языков характерны: отсутствие следов аспирации звонких и глухих согласных, ограничения действия правила *RUKI* и нек. др., для лексики — отсутствие ареальной инновационной модели построения местоимений 2 л. мн. ч. и ряд других.

В результате схема этапов филиации арийской общности, изображенная в [ЭДЕЛЬМАН 1992; 1999: 13] выглядит следующим образом (см. *Схему 1*). Схема с уточнением относительной хронологии отделения дардской группы, но с упрощением — отсутствием графического обозначения диалектной неоднородности праязыковых континуумов — в [КОГАН 2005: 202] имеет несколько иной вид (см. *Схему 2*):

Схема 1

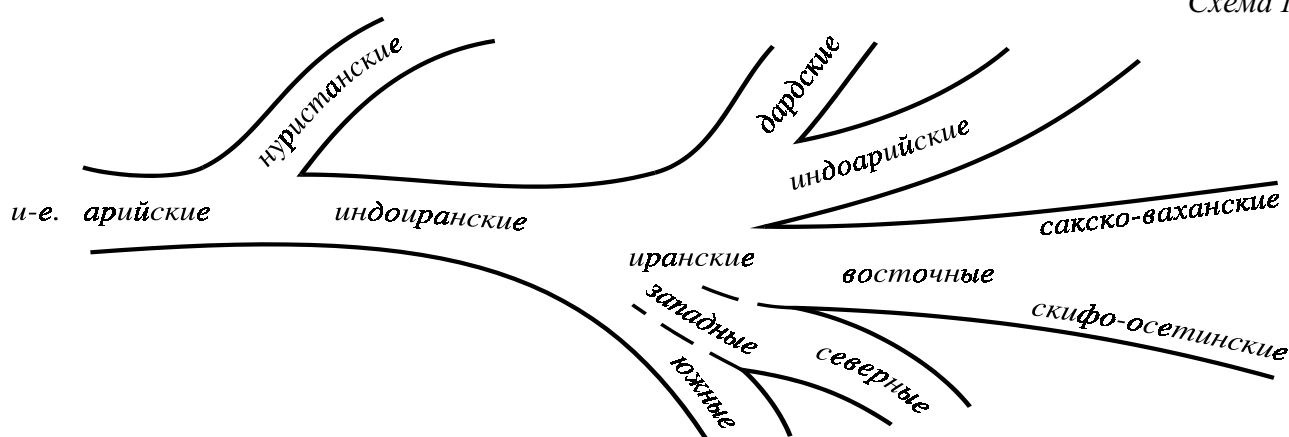
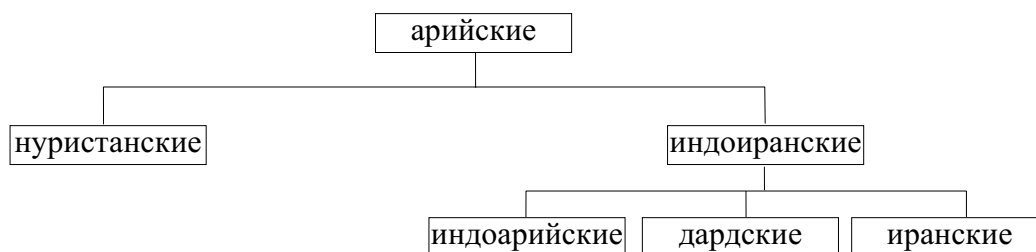


Схема 2



Отдельную проблему составляют генетические отношения иранских языков со славянскими, о чем существует большая литература. Рассмотрение сходств и различий между иранскими и славянскими языками на разных языковых уровнях — в исторической фонетике и фонологии, морфологии и синтаксисе, некоторых лексических элементах — показывают, что общих материальных инноваций между иранскими и славянскими языками, эксклюзивных для этих языков, минуя общеарийское и индоиранское состояние в истории иранских языков, не прослеживается. Следовательно нет и доказательств более тесного генетического родства, чем родство, восходящее к состоянию, общему для всей группы «сатэм» индоевропейской языковой семьи. Нет и доказательств вторичного конвергентного «средства» в рамках языкового союза в праязыковом (во всяком случае для иранских языков) периоде, хотя имеются свидетельства существования в прошлом ареальных конвергентных языковых и этнических союзов, в который входили скифо-сарматские племена и часть славян. Однако это относится к значительно более позднему периоду и охватывает лишь относительно небольшую часть ареалов распространения иранских и славянских языков. Прослеживаются также общие или сходные типологические инновации в иранских и славянских языках, интересные сами по себе, часть из них носит ареальный, контактный характер. Однако для определения степени генетического родства они не показательны (подробнее [ЭДЕЛЬМАН 2002]).

III. Генетическая классификация иранских языков.

Иранские языки — живые и вымершие — делятся на основные генетические группы, которые называются условно «западноиранской» и «восточноиранской», согласно первичному расселению ираноязычных племен (по отношению к соляным пустыням Иранского нагорья), существенно отличному от их исторического и современного распространения. Классифицирующими признаками принадлежности языков к той или иной группе традиционно считаются отражения определенных общеиранских согласных в обеих группах. Принято считать, что: 1) праиранские звонкие согласные **b*, **d*, **g* в начальной позиции сохраняются в виде зап.-иран. **b-*, **d-*, **g-*, но переходят в вост.-иран. щелевые **v-*, **δ-*, **γ-*; 2) срединные консонантные группы **-ft-*, **-xt-* сохраняются в виде зап.-иран. *-ft-*, *-xt-*, но озвончаются в вост.-иран. *-vd-*, *-γd-*; 3) праиранская **č* в анлауте сохраняется как зап.-иран. *č-*, но дентализуется в вост.-иран. *c-*; 4) праиранская **č* в инлауте сохраняется как зап.-иран. *-č-* (> *-ž-*, *-z-*), но дентализуется в вост.-иран. *-c-* > *-j-*, *-z-*; 5) праиранская **h-* сохраняется в западноиранских языках, но утрачивается в восточноиранских, см. [ОРАНСКИЙ 1979: 131 и сл.].

Исследования языков обеих групп в последние десятилетия внесли, однако, некоторые коррективы в это положение. О причинах и относительной хронологии появления признака (1) уже говорилось выше. В связи с этим признак (2) предстает не как позднее озвончение древних «глухих» консонантных групп в восточноиранских языках при сохранении их глухости в западноиранских, а как дальнейший результат построения значительно более ранних арийских консонантных групп, сформировавшихся согласно «закону Бартоломе»: в восточных языках эти группы отражаются как содержавшие арийскую звонкую аспирированную (по типу **gh + t > *gdh > gd > γd*), в западных — как содержавшие звонкую неаспирированную или глухую согласную (по типу **g + t*, **k + t > *kt > xt*), то есть продолжают в отражении звонких согласных ту же тенденцию, которая наблюдается в признаке (1), подробнее см. [ЭДЕЛЬМАН СГВЯ-Ф: 57–59].

Иная относительная хронология и иные причины способствовали появлению признаков (3) и (4). Как показывают материалы языков, изучение которых относится к XX веку, рефлексам праиранской **č* в начальной и срединной позициях далеко не во всех восточноиранских языках служит *c* и наследующие ему переднеязычные согласные: большая часть языков этой группы сохраняет палатальный фокус, выявляя рефлекс в виде *č*, *j*, *ž*, *ç*, причем различия в этом плане наблюдаются даже между близкородственными языками (например, севернопамирскими), а в части западных языков отмечается отражение **č* в срединной позиции в виде *-z-*, а в начальной позиции в единичных языках — в виде *c-*. Тем самым дентализация **č* затронула лишь часть восточноиранских языков, а также единичные западные, и следовательно была относительно поздней, развившейся в разных языках независимо, под влиянием определенных типологических процессов², в некоторых случаях поддержанных ареальными артикуляционными тенденциями. Поэтому она не может служить генетическим признаком, разделяющим восточные и западные языки. То же относится к признаку (5) — сохранению или утрате **h*. Подробнее о фонологических генетических чертах этих языковых групп и о поздних явлениях см. [ЭДЕЛЬМАН СГВЯ-Ф]; [РАСТОРГУЕВА 1990].

В результате основными историко-фонологическими признаками, разделяющими две самые большие генетические группы иранских языков, являются в действительности только признаки (1) и (2), связанные с дальними отражениями общеарийских звонких согласных — аспирированных и неаспирированных, о которых говорилось выше.

Что касается более дробного членения — разделения западноиранских языков на северную и южную группы — оно базируется на другом историко-фонологическом признаке — различии в отражении праиранских **ś*, **ž* — общеиранских рефлексов позднеиндоевропейских палатальных **k̂*, **ĝ*, **ĝh*, и, следовательно, могло быть более поздним, чем разделение иранских языков на восточные и западные. Дивергенция восточноиранских диалектов была изначально более дробной, о чем свидетельствуют рефлексы тех же праиранских **ś*, **ž* и содержавших их групп согласных (подробнее см. [ЭДЕЛЬМАН СГВЯ-Ф: 40–47]).

² Фонетическая тенденция к дентализации была стимулирована самим происхождением праиранской **č*, восходящей — через общеарийскую **č* к общесатэмной **k* в позиции перед гласными переднего ряда и сонантом **i* (в слоговом и неслоговом вариантах). Рефлекс праиранской **č* в основной массе слов продолжают пребывать в аналогичных фонетических позициях, что в одних языках способствует сдвигу реализующих их звукотипов в дентальную зону, в других — сохранению ими мягкого, палатального характера.

Интересна в этом плане генетическая позиция языка Авесты — собрания священных гимнов зороастрийцев, то есть языка, который был живым во II тысячелетии до н. э. (разные традиции относят проповеди Заратуштры к 1700 или к 1200 годам до н. э.) и священные тексты на котором передавались тысячелетиями устно. По фонетическим признакам он не может быть отнесен однозначно ни к западной, ни к восточной ветви, поскольку отражение в нем праиранских звонких согласных в анлаутной позиции по «западному» типу могло быть результатом передачи текста западноиранскими жрецами, а в отражении срединных консонантных групп наблюдаются различия, зависящие от хронологии их происхождения: древние (в застывших формах, в лексикализованных причастиях и т. д.) отражают действие «закона Бартоломе», поздние (в относительно продуктивных или во всяком случае прозрачных формах) трансформированы по «западному» типу, возможно, также под влиянием произношения западноиранских жрецов. При этом в данном языке имеются собственные инновации, не разделяемые другими иранскими языками. Не исключено продолжение здесь третьей ветви, не дошедшей до нас в виде более поздних представителей.

Методически существенно, что все рассмотренные генетические признаки базируются на историко-фонологических данных.

Признавая общую методическую ценность морфологических изоглосс в генетической классификации языков, приходится констатировать, что при данном уровне наших знаний о живых и вымерших иранских языках и об их истории можно установить лишь относительно частотные рефлексии тех или иных древних форм в языках одной или другой группы, при этом ни одна из морфологических изоглосс не охватывает ту или иную группу языков целиком. Те морфологические (точнее, морфолого-синтаксические) изоглоссы, которые считались классифицирующими в начале XX века (например, порядок компонентов определительной конструкции имен: препозиция определения определяемому в восточноиранских языках и постпозиция его в большинстве западноиранских), уже тогда не охватывали все языки одной или другой группы. После введения в научный обиход нового материала многих не известных ранее языков, они оказались еще менее показательными.

«Чистые» генетические изоглоссы морфологического уровня, которые однозначно делили бы иранские языки на группы и подгруппы, не прослеживаются. В определенной мере это связано со спецификой истории иранских языков: раннее «упрощение» древних парадигм и существенная структурно-типологическая перестройка практически всех иранских языков после древнего периода способствовали утрате многих древних морфологических элементов, которые были свойственны праиранскому периоду и еще сохранялись в языке Авесты. К тому же отсутствие древних письменных памятников восточноиранских языков и скудость материала западноиранских — древнеперсидских — памятников не дают возможности проследить полные парадигмы древних диалектов — ни западных, ни восточных, и тем самым не дают возможности выявить соотношение древних морфологических форм и их функций. Подробнее см. [Эдельман СГВЯ-М: 257–267].

То же относится к лексическим изоглоссам: фиксировавшиеся в начале XX века, до введения в обиход материала большей части живых и вымерших языков, лексические различия между восточными и западными иранскими языками также, как выяснилось, не охватывают все языки той или иной группы, или даже большинство из них. Так, например, предполагавшиеся различия в названиях собаки: др.-иран. **spaka-* для западноиранских языков и **kuti-* — для восточноиранских, — оказались «размыты» последующими исследованиями: первое слово на деле представляет собой отражение праиранского **śya-ka-* от **śyan-* : **śun-* < и.-е. **k̑on-* : *kun-* «собака», а его собственные рефлексии наблюдаются не только в западноиранских, но и в восточноиранских языках: пушту, ваханском, скифских диалектах, хотаносакском. Часть таких различий может относиться не к диалектным, а к жанровым вариантам и представлять собой элементы диглоссии [КАЗАМА 1986], подробнее об этом говорится в другой работе [Эдельман СГВЯ-Л].

Таким образом, признавая принципиальную ценность всеуровневых или разноуровневых изоглосс для анализа генетической дивергенции языковой семьи, в частном случае иранских языков мы вынуждены, в силу специфики материала, ограничиться историко-фонетическими, точнее — историко-фонологическими изоглоссами.

IV. Методика сравнительно-исторической реконструкции в иранистике

Как известно, сравнительно-историческая реконструкция призвана воссоздать более раннее состояние языка, группы языков, праязыка для языковой группы и/или семьи. При этом праязыковая

реконструкция опирается на закономерные соответствия спонтанно развившихся рефлексов пра-системы в разных языках-наследниках. Это — в «чистом» виде. В случае иранской языковой семьи (как, наверное, и всякой другой — при углубленном исследовании) приходится постоянно иметь в виду различные свидетельства, как облегчающие, так и затрудняющие такую реконструкцию.

Как уже говорилось, учет в праязыковой реконструкции всех языков, а не только древних, позволил значительно «отодвинуть» в хронологическом плане общеиранское праязыковое состояние. Это относится как к относительной хронологии (общеиранское состояние стало системно ближе к общеарийскому), так и к абсолютной (предположительно, не позднее рубежа II–III тысячелетия до н.э., учитывая приблизительное время бытования языка Авесты с его уже сформировавшимися диалектными отклонениями от общеиранского, см. выше).

При этом подход к реконструированному состоянию как к некоторой системе потребовал уже осознанного привлечения приемов, относящихся к типологическому и ареальному методам. Практически эти приемы применялись в сравнительно-исторической реконструкции и прежде, еще в XIX веке, но без четкого осознания их как относящихся к другим лингвистическим методам и без учета иерархии методов в этом процессе.

Для изучения истории конкретных иранских языков и реконструкции более раннего состояния каждого из них важным оказался учет положений как структурной, так и контенсивной типологии (о содержании данных понятий см. [Климов 1973; 1977]), что позволило опираться, с одной стороны, на возможность наследования иранскими языками ряда категорий позднеиндоевропейского состояния, включая рефлекс элементов активного строя (или оппозиции «живое — неживое»), а с другой, — на развитие в большинстве из них эргативных и эргативообразных построений, а также элементов поздней активности, включая субстратные. В этом плане большую роль сыграла разработка методики внутренней реконструкции предшествующего состояния языка, базирующейся на правилах и «закономерных исключениях» языка в его синхронном состоянии, а также разработка «промежуточной» (иногда ее называют также «встречной») реконструкции более раннего состояния конкретного языка или языковой группы, опирающейся на материал синхронного состояния языка и реконструируемого древнего (вплоть до праязыкового) состояния.

Осознанное использование приемов, относящихся к типологическому методу, в сравнительно-исторической реконструкции оказалось чрезвычайно важным для изучения истории практически всех иранских языков, но в особенности бесписьменных и/или не имеющих письменной традиции, поскольку исходными данными для изучения истории таких языков могли служить только, с одной стороны, их синхронное состояние, с другой, — реконструированное праиранское, отстоящее от него на четыре тысячелетия. В таких условиях реконструкция промежуточных этапов изменения языка стала просто необходимой, а без учета данных типологии определение степени вероятности той или иной предшествующей системы или того или иного изменения едва ли было бы возможным. Это относится к изучению и изменения языковой системы в целом, и изменения отдельных элементов или функций унаследованных элементов, аранжировки их роли в системных изменениях и др. (что отражено, в частности, в коллективных монографиях [ОИТИИЯ 1975]; [ОИЯ 1979–2008]).

Немалую роль в изучении истории иранских языков сыграл также учет ареального фактора, что важно в условиях их разбросанности на обширном пространстве, где они неавтохтонны, различных конвергентных процессов в разных регионах, наслоении разных иранских языков на разносистемные субстраты, адстратное влияние других языков, включая влияние одних иранских языков на другие, особенно со стороны языков, наиболее престижных в данном месте в данную историческую эпоху, и т. д. Этот фактор отражается практически на всех языковых уровнях, что особенно заметно в языковых союзах, в которые входит часть иранских языков или входила в разные эпохи. В некоторых случаях ареальный фактор влияет и на историко-фонологические изменения, вызывая к жизни некоторые изоглоссы, затеняющие сравнительно-историческую картину филиации языковой группы и отдельных подгрупп.

Так, например, для древней эпохи в западном ареале были характерны северо-западные заимствования в древнеперсидском языке (преимущественно из мидийского языка, престижного в определенный исторический период), но для последующих периодов, начиная со средневековья, важным фактором стали заимствования из персидского языка (ставшего наиболее престижным) в северо-западные языки и диалекты, что — при фонетическом сходстве слов — затемняет картину

спонтанного изменения фонетики этих языков. Обилие заимствованной лексики и влияние ареальной фонетики (и фонологии) долгое время вызывали дискуссию о генетической принадлежности «малых» островных иранских языков Центральной Азии: парачи и ормури. Их причисляли то к северо-западной, то к восточной группе, к последней — благодаря обилию восточноиранской и дардской лексики и ареальной фонологии. И только глубокий всесторонний анализ с учетом их разновременных контактов с разными (не только иранскими) языками и с разделением изменений «домиграционного» и «послемиграционного» периодов в их истории позволил на основании историко-фонетического анализа отнести их к реликтовым северо-западным языкам-пришельцам (подробнее см. [Ефимов 1986, 14–17, материал: 27 и сл.; 1997: 420–423, материал: 424 и сл.]).

Другой пример: ареальная группа, состоящая из части восточноиранских языков и единичных западноиранских и входящая в Центральноазиатский языковой союз, вырабатывает фонологический ряд церебральных фонем, инвентарь которого отличен от индоарийского аналогичного ряда, но совпадает в целом с субстратным — языка бурушаски. При этом часть церебральных фонем в исконном материале закономерно продолжает определенные праиранские консонантные группы.

Интересен в методическом плане историко-фонетический пример перехода праиранской $*d > l$ в ряде иранских языков региона Гиндукуша (в пушту, мунджанском, йидга, бактрийском). Этот переход традиционно рассматривался в ряде трудов как генетически общий, объединяющий эти языки в генетическую подгруппу, отличную от других иранских языков. Были также попытки объяснить аналогичный переход $*d > l$ в иранских языках других ареалов влиянием бактрийского языка или их особой генетической близостью с бактрийским. При этом не учитывалось, что особое близкое родство языков, входящих, в свою очередь, в одну семью, определяется, как известно, не артикуляционным сходством рефлексов единичного древнего прототипа, а системностью таких рефлексов и их сочетаемостью с другими системными же рефлексами в тех же языках, то есть совокупностью системных инноваций, отличающих данную группу языков от других языков этой же семьи, менее близких генетически.

Однако привлечение других, не иранских, языков ареала Восточного Гиндукуша показало, что здесь переход $*d > l$ — ареальное явление, присущее разным языкам данного региона, независимо от их генетической принадлежности, не только иранским. Кроме того, он связан с определенной типологией фонологических структур и фонетических изменений³: в тех иранских языках, где отсутствует или неустойчив фонологический статус глухой $*\theta$, неустойчивым оказывается и статус звонкого звукотипа $*\delta$ (наследующего $*d$ и $*t$ в определенных позициях). В результате дентальные смычные легко переходят в l (как в указанных языках), в r (в татском, кумзари, диалектах башкарди, талышского и других языков), в z (в языке яссов, в части персидских диалектов), в d (в персидском, таджикском и др.). Ареальный фактор в виде артикуляционных навыков способствует фонологическому закреплению того или иного фонетического рефлекса.

Особую роль в «нарушении» или трансформации спонтанного изменения иранских языков принадлежит влиянию субстратов, особенно явно прослеживаемому в осетинском языке [АБАЕВ ОЯФ; 1956] и в языках Центральноазиатского языкового союза, в разных его регионах [ЭДЕЛЬМАН 1980]; [ГРЮНБЕРГ, СТЕБЛИН-КАМЕНСКИЙ 1974].

Таким образом, в изучении истории иранских языков особую наглядность приобретает положение о том, что использование результатов, полученных при помощи приемов, относящихся к другим методам сравнительного изучения разных языков — типологическому, ареальному, — способствует уточнению и совершенствованию сравнительно-исторического метода. Это дает возможность сделать его наблюдения более реалистичными и объяснительными.

Как уже говорилось выше, сравнительно-историческое исследование языков иранской семьи, проводившееся в последние десятилетия в плане изучения истории как отдельных языков, так и основных генетических групп, создало достаточно прочную базу для создания описания в этом аспекте всей иранской семьи в целом. Грамотное и продуманное использование в этом описании результатов, полученных иными методами — типологическим и ареальным, будет способствовать более полному и реалистичному анализу непростой истории этой языковой семьи.

³ К типологии переходов дентальных смычных согласных в l , r см. также: в разных иранских языках [ЭДЕЛЬМАН 2006], в разговорном английском языке [СКАЕР, SOSEI ANIYA 2007], в русских диалектах и в других языках мира [КАСАТКИНА 2004].

Литература

- АБАЕВ ОЯФ — В. И. АБАЕВ. *Осетинский язык и фольклор*. М., 1949.
- АБАЕВ 1956 — В. И. АБАЕВ. О языковом субстрате // Доклады и сообщения Института языкознания АН СССР. IX. М.
- АБАЕВ ИЭСОЯ — В. И. АБАЕВ. *Историко-этимологический словарь осетинского языка*. Т. I: М. – Л., 1958; т. II: Л., 1973; т. III: Л., 1979; т. IV: Л., 1989; Указатель: М., 1995.
- АБАЕВ 1962 — В. И. АБАЕВ. Сравнительно-историческое иранское языкознание // *Очерки по истории изучения иранских языков*. М.
- АБАЕВ 1965 — В. И. АБАЕВ. *Скифо-европейские изоглоссы: На стыке Востока и Запада*. М., 1965
- БЕЛИКОВ 1972 — В. И. БЕЛИКОВ. К употреблению указательных местоимений в шугнанском языке // *Памирские языки и фольклор* I. Душанбе.
- ГРЮНБЕРГ, СТЕБЛИН-КАМЕНСКИЙ 1974 — А. Л. ГРЮНБЕРГ, И. М. СТЕБЛИН-КАМЕНСКИЙ. Этнолингвистическая характеристика Восточного Гиндукуша // *Проблемы картографирования в языкознании и этнографии*. Л.
- ЕЛИЗАРЕНКОВА 1974 — Т. Я. ЕЛИЗАРЕНКОВА. *Исследования по диахронической фонологии индоарийских языков*. М.
- ЕФИМОВ 1986 — В. А. ЕФИМОВ. *Язык ормури в синхронном и историческом освещении*. М., 1986.
- ЕФИМОВ 1997 — В. А. ЕФИМОВ. Парачи // *ОИЯ* 1997.
- КАСАТКИНА 2004 — Р. Ф. КАСАТКИНА. Эти странные вибранты (некоторые проявления апикоальвеолярных артикуляций в русском языке) // *Семиотика, лингвистика, поэтика: К столетию со дня рождения А. А. РЕФОРМАТСКОГО*. М.
- КЛИМОВ 1973 — Г. А. КЛИМОВ. *Очерк общей теории эргативности*. М.
- КЛИМОВ 1977 — Г. А. КЛИМОВ. *Типология языков активного строя*. М.
- КОГАН 2000 — А. И. КОГАН. *Генетическая характеристика дардских языков*. АКД. М.
- КОГАН 2003 — А. И. КОГАН. Отношения дардских и индоарийских языков: тесное родство или языковой союз // *Сравнительно-историческое исследование языков: современное состояние и перспективы*. Тез. докл. междунар. научн. конференции. МГУ. М.
- КОГАН 2005 — А. И. КОГАН. *Дардские языки. Генетическая характеристика*. М.
- ОИТИИЯ 1975 — *Опыт историко-типологического исследования иранских языков*. Т. I–II. М..
- ОИЯ 1979–2008 — *Основы иранского языкознания. Древнеиранские языки*. М., 1979; *Основы иранского языкознания. Среднеиранские языки*. М., 1981; *Основы иранского языкознания. Новоиранские языки. Западная группа*. М., 1982; *Основы иранского языкознания. Новоиранские языки. Восточная группа*. М., 1987; *Основы иранского языкознания. Новоиранские языки: Северо-западная группа. I*. М., 1991; *Основы иранского языкознания. Новоиранские языки: Северо-западная группа. II*. М., 1997; *Основы иранского языкознания. Среднеиранские и новоиранские языки*. М., 2008.
- ОРАНСКИЙ 1979 — И. М. ОРАНСКИЙ. *Иранские языки в историческом освещении*. М.
- РАСТОРГУЕВА 1964 — В. С. РАСТОРГУЕВА. *Опыт сравнительного изучения таджикских говоров*. М.
- РАСТОРГУЕВА 1990 — В. С. РАСТОРГУЕВА. *Сравнительно-историческая грамматика западноиранских языков: Фонология*. М.
- РАСТОРГУЕВА, МОЛЧАНОВА 1981 — В. С. РАСТОРГУЕВА, Е. К. МОЛЧАНОВА. Среднеперсидский язык // *ОИЯ* 1981.
- РАСТОРГУЕВА, ЭДЕЛЬМАН 1981 — В. С. РАСТОРГУЕВА, Д. И. ЭДЕЛЬМАН. Иранские языки // *Сравнительно-историческое изучение языков разных семей: Современное состояние и проблемы*. М.
- СЕМЕРЕНЬИ 1980 — О. СЕМЕРЕНЬИ. *Введение в сравнительное языкознание*. М.
- СТЕБЛИН-КАМЕНСКИЙ ЭСВЯ — И. М. СТЕБЛИН-КАМЕНСКИЙ. *Этимологический словарь ваханского языка*. СПб., 1999.
- ЭДЕЛЬМАН 1968 — Д. И. ЭДЕЛЬМАН. *Основные вопросы лингвистической географии (на материале индоиранских языков)*. М.
- ЭДЕЛЬМАН 1976 — Д. И. ЭДЕЛЬМАН. К истории язгулямских и шугнано-рушанских указательных местоимений // *Иранское языкознание. История, этимология, типология (к 75-летию проф. В. И. АБАЕВА)*. М.

- ЭДЕЛЬМАН 1980 — Д. И. ЭДЕЛЬМАН. К субстратному наследию Центральноазиатского языкового союза // *Вопросы языкознания*, № 5.
- ЭДЕЛЬМАН 1982 — Д. И. ЭДЕЛЬМАН. К перспективам реконструкции общеиранского состояния // *Вопросы языкознания*. № 1.
- ЭДЕЛЬМАН СГВЯ-Ф — Д. И. ЭДЕЛЬМАН. *Сравнительная грамматика восточноиранских языков. Фонология*. М., 1986
- ЭДЕЛЬМАН СГВЯ-М — Д. И. ЭДЕЛЬМАН. *Сравнительная грамматика восточноиранских языков. Морфология. Элементы синтаксиса*. М., 1990.
- ЭДЕЛЬМАН СГВЯ-Л — Д. И. ЭДЕЛЬМАН. *Сравнительная грамматика восточноиранских языков. Лексика* [в печати].
- ЭДЕЛЬМАН 1992 — Д. И. ЭДЕЛЬМАН. Еще раз об этапах филиации арийской языковой общности // *Вопросы языкознания*. № 3.
- ЭДЕЛЬМАН 1996 — Д. И. ЭДЕЛЬМАН. Свидетельства «малых» языков об истории языковой семьи (на примере «Этимологического словаря иранских языков») // *Гуманитарная наука в России: Соросовские лауреаты. Филология, литературоведение, культурология, лингвистика, искусствоведение*. М.
- ЭДЕЛЬМАН 1999 — Д. И. ЭДЕЛЬМАН. Индоиранские языки // *Языки мира: Дардские и нуристанские языки*. М.
- ЭДЕЛЬМАН 2002 — Д. И. ЭДЕЛЬМАН. *Иранские и славянские языки. Исторические отношения*. М.
- ЭДЕЛЬМАН 2006 — Д. И. ЭДЕЛЬМАН. Артикуляционная фонетика и типология историко-фонетических процессов (к ламбдаизму и ротацизму в иранских языках) // *Индоиранское языкознание и типология языковых ситуаций. Сборник статей к 75-летию проф. А. Л. ГРЮНБЕРГА (1930–1995)*. СПб.
- ЭСИЯ — В. С. РАСТОРГУЕВА, Д. И. ЭДЕЛЬМАН. *Этимологический словарь иранских языков*. Т. 1. М., 2000; Т. 2. М., 2003; Т. 3. М., 2007.
- ЮСУФБЕКОВ 1998 — Ш. П. ЮСУФБЕКОВ. *Дейктичность в языках шугнано-рушанской группы*. М.
- BAILEY DKS — H. W. BAILEY. *Dictionary of Khotan Saka*. Cambridge–London–New York–Melbourne, 1979.
- BARTHOLOMAE 1895–1901 — Chr. BARTHOLOMAE, *Vorgeschichte der iranischen Sprachen // Grundriss der iranischen Philologie*. Bd I, Abt. I. Strassburg.
- BUDDRUS 1977 — G. BUDDRUS. Nochmals zur Stellung der Nulristan-Sprachen des afghanischen Hindukusch // *Münchener Stud. Sprachwiss.* H. 36.
- CLI 1989 — *Compendium Linguarum Iranicarum*. Wiesbaden.
- KAZAMA 1986 — K. KAZAMA. A Phase of the History of Indic // *JIES*. Vol. 14, № 3–4.
- KNOBLOCH 1965 — J. KNOBLOCH. Concetto storico di protolingua e possibilità e limiti di applicazione ad esso dei principi strutturalistici // *Le «Protolingue»*: Atti del IV Convegno Internazionale di linguisti. Milano.
- KUIPER 1978 — F. B. J. KUIPER. Rec.: G. MORGENSTIERNE. Irano-Dardica // *IJ*. 1978, Vol. 20, № 1/2.
- MORGENSTIERNE 1945 — G. MORGENSTIERNE. Indo-European *k̂* in Kafiri // *NTS*. Bd 13.
- MORGENSTIERNE EVSh — G. MORGENSTIERNE. *Etymological Vocabulary of the Shughni Group*. Wiesbaden, 1974.
- MORGENSTIERNE ID — G. MORGENSTIERNE. *Irano-Dardica*. Wiesbaden, 1973.
- MORGENSTIERNE IFL — G. MORGENSTIERNE. *Indo-Iranian Frontier Languages*. Vol. I–IV. Oslo, 1929–1973.
- SKAER, SOSEI ANIYA 2007 — P. M. SKAER, SOSEI ANIYA. Alveolar Tap Allophones in English: a Phonology-Phonetics Interface View of Consonant Lenition // *Фонетика сегодня*. Материалы докладов и сообщений V Международной конференции 8–10 октября 2007 года. М.

Summary

Languages of the Iranian family, whose speakers over the centuries have dispersed throughout various regions of Eurasia and contacted both among themselves and with speakers of other language families, have undergone significant typological changes; this presents a serious challenge for their historical comparative treatment. The present article discusses such problems as (1) areal and temporal stratification of Proto-Iranian, regarded as a dialectal continuum; (2) stages of branching of the Aryan subgroup of Indo-European which eventually led to the appearance of Iranian as an autonomous entity; (3) genetic classification of the Iranian branch; and (4) various related methodological issues of comparative linguistics.

Proto-Afrasian Lexicon Confirming West Asian Homeland: Pastoralism

The article presents one more step towards the equation of the culture of speakers of Proto-Afrasian, reconstructed on the basis of paleolinguistic data, with the early Neolithic Post-Natufian culture of the Levant. According to the glottochronological method of S. A. STAROSTIN, Proto-Afrasian is dated back to approximately 10 000 BC — the same period as Post-Natufian (supposed to be the cradle of agriculture and livestock breeding on the planet), as far as radiocarbon dating tells us. The article offers evidence for the presence of a layer of pastoral lexicon in Proto-Afrasian, in the form of 26 reconstructed names for large and small cattle and various other pastoral terms. The lexical data are preceded with a brief summary of the current state of affairs in Afrasian historical linguistics, as well as a description of the author's methodology of linguistic analysis and his approach to combining linguistic and archaeological data in order to solve the "homeland" issue for proto-languages.

Introduction

The objective of the present paper is to present further evidence, this time referring to pastoralism presumably practised by the Proto-Afrasian (Afroasiatic, Semito-Hamitic) speaking community, for the identification of this community with the early Levantine villagers associated with the early Neolithic Post-Natufian culture. These villagers left some of the earliest known archaeological evidence for the cultivation of domesticated crops (cereals and pulse) and the raising of domestic livestock (cf., for example, [BAR-YOSEF]; [HASS.]; [PELT.]). It is for archaeologists to evaluate the correspondences between the archaeological evidence from the Levant, as well as adjacent regions, and the reconstructed terminology referring to incipient agro-pastoralism in the Proto-Afrasian language, dated by the new version of the linguistic method of glottochronology to approximately the same period (12 000–10 500 BP) and presumably the same area.

This is part of a broader project aimed at drawing a most comprehensive picture featuring practically all aspects of life of Early Neolithic people in the Near East which can be drawn from the reconstructed Proto-Afrasian lexicon, namely, terms referring to people and society; economic life and technology; intellectual culture; and natural and physical environments.

While the archaeology of the Levant is one of the most advanced fields in the domain of world prehistory, Afrasian comparative linguistics has long been lagging behind such fields as Altaic or North Caucasian, to say nothing of Indo-European comparative studies. As to comparative Afrasian lexicology and etymology, their history and present state of knowledge can be described in short as follows. Sound correspondences and etymologies proposed in the pioneer work by M. COHEN ([COH.]) postulating the Semito-Hamitic family are, as a whole, outdated. A lot of Common Afrasian lexemes were collected by J. GREENBERG in *The Languages of Africa* ([GR.]) and other works, but his method of "mass comparison", opposed to the established comparative-historical method and aiming exclusively at genealogical classification, does not care for establishing sound correspondences or reconstructing protoforms. A number of reliable phonetic and lexical Afrasian correspondences were proposed in [I-S] and other studies by V. ILLICH-SVITYCH, who included Semito-Hamitic resp. Afrasian into his Nostratic macrofamily (the present author considers Afrasian and Nostratic two "sister" entities on the same taxonomic level) and, especially, by A. DOLGOPOLSKY ([DOLG. Cush.]), who adduced Afrasian parallels to his Cushito-Omotoc protoforms and paved the way to the elaboration of Proto-Afrasian phonological system.

The first study, however, to reconstruct the latter and establish regular sound correspondences between the primary branches and languages of Afrasian (its main bulk is still valid even today), was carried out by a team of scholars headed by I. DIAKONOFF, of which the present author was a member; it also adduced a few hundred Proto-Afrasian lexemes ([HCVA]). Although an important step for comparative Afrasian back in its day, now that over 20 years have passed, its many drawbacks are obvious to everyone including its authors; some are due to many publication sources that were inaccessible in Russia back then, others — to a lack of experience in dealing with such vast and heterogeneous material, still others — to rather loose semantic criteria. Two comparative Afrasian dictionaries both published in 1995 ([HSED] and [EHR. PA]) enriched the field with more lexical data, yet again, each of the

two had its own flaws. [HSED], while replete with new and stimulating etymologies, was compiled rather hastily and carelessly; C. EHRET's method, on the other hand, involved postulating improbably sophisticated proto-phonemes in combination with far-fetched semantic comparisons, such as his attempts to relate words with meanings like 'armpit' and 'to thatch' ("the armpit is a covered area of the body"), or 'forest' and 'thirst' (with the reconstructed meaning 'waterless place, desolation'), &c.

Important contributions to the study of Afrasian lexicon have also been made by two hard-working comparative linguists, V. BLAŽEK (in many papers) and G. TAKÁCS ([EDE I, II and III] and various other papers). Invaluable and enormous Afrasian lexical data are presented by one of the world's leading macrocomparativists, A. DOLGOPOLSKY, in his massive Nostratic Dictionary (still unpublished on paper but now available online at <http://www.dspace.cam.ac.uk/handle/1810/196512> courtesy of the McDONALD Institute for Archaeological Research at Cambridge); many Afrasian reconstructions offered there are, however, in our opinion, debatable, due to DOLGOPOLSKY's adherence to the idea of Afrasian being an integral part of Nostratic, which accounts for numerous cases of "forced" cognations. Anyway, the most comprehensive collection of Afrasian etymological data (containing some 3500 entries in the main database and some 15000 in subdatabases containing lexica of all the branches and lesser groups within Afrasian), some of it representing properly verified etymologies and some serving as "raw material" open to further research, can be found today in the general Afrasian database ([AADB]), accessible electronically at <http://starling.rinet.ru>. The database has been compiled by the present author and O. STOLBOVA within the framework of the "Evolution of Human Languages" project of the Santa Fe Institute.

As to how this relates to all the work in progress elsewhere, it is worth mentioning that practically all the researchers who tried to deal with the problem of original habitat, or "homeland", of the speakers of Proto-Afrasian, put forth arguments in favor of an African homeland. These arguments, relying on general considerations like "economy of movement", as well as scarce, chaotic and carelessly compiled lexical examples, look very weak (see, for example, [EHR. EEA] and studies by R. BLENCH, such as [ALAP] and others); an exception can be made for I. DIAKONOFF's study ([DIAK. ESA]) which is methodologically impeccable, but was outdated already at the moment of publication, as far as lexical materials that underlie its conclusions are concerned. By that time, a lot of new data contradicting these conclusions had already been accumulated — partly due to ongoing work on comparative Afrasian lexicon, initiated and headed by DIAKONOFF himself. Later, he recognized the validity of our arguments in favor of a West Asian homeland ([DIAK. Sum.]).

Since the present paper is designed for a new periodical edition, devoted to issues of historical and comparative linguistics, the author thinks it appropriate to precede the data with some theoretical and methodological considerations. Namely, three different methods are used for different aspects of the author's research on Afrasian lexicon, including the one fragment that is represented by the present paper, and on the investigation of the homeland of Proto-Afrasian speakers:

(1) The main method is, as in any other standard comparative study dealing with lexical reconstruction of a protolanguage, the classic comparative-historical method elaborated for Indo-European languages by the Neogrammarian School in the late 19th century. Within this method, several principles are strictly observed, some of them slightly innovative, some being universally accepted as something that goes without saying — yet far from always followed either in etymological dictionaries for individual Afrasian languages or in studies on Common Afrasian. These principles are as follows:

(1.1.) Selection of lexical terms to be labelled Proto-Afrasian. According to the author's genetic classification of Afrasian (first branching dated to the mid-10th mill.), this macrofamily consists of the following presumed branches and universally recognized families:

1. North Afrasian (NAA) (first branching dated to the mid 9th mill. BC):

1.1. Semitic.

1.2. African North Afrasian (ANAA):

1.2.1. Egyptian.

1.2.2. Chado-Berber:

1.2.2.1. Berber-Canarian.

1.2.2.2. Chadic.

2. South Afrasian (SAA):

2.1. Cushitic.

2.2. Omotic.

In accordance with this classification, PAA terms are those attested at least in one family belonging to the NAA branch and one family belonging to the SAA branch; PNAA terms are those attested in Semitic and at least one family belonging to the ANAA subbranch, provided the compared terms are not suspected to have been borrowed (see below for criteria for borrowings). Terms attested only in ANAA or only in SAA (both branching in late 8th mill.) are not included, as they are irrelevant for the present study.

(1.2.) Ruling out borrowings.

To avoid reconstruction of “false” PAA or PNAA forms, the terms selected for inclusion should not be suspected of having been borrowed, with all controversial and debatable cases marked as such. First of all, this principle makes the inclusion of Semitic cognates highly desirable; a lack of Semitic parallels will make any form claimed to represent PAA less reliable, since cognate forms that are only attested in African Afrasian languages, even in both ANAA and SAA, may have been borrowed from a non-Afrasian African substratum. Apart from that, there can be several other situations with their specific problems requiring individual treatment. Most difficult ones involve identifying Arabisms in most spoken African Afrasian languages; Ethiopian and SAA interborrowings; Cushitic-Omotiic, Berber-Chadic, and Egyptian-Semitic interborrowings. In order to distinguish between inherited and borrowed lexemes, the following criteria are proposed (cf. [SED I and II: Introduction, 1.11]):

(1.2.1.) A term may be reasonably claimed a loanword or suspected of having been borrowed only if areal contacts between the languages in question are attested historically and linguistically (e.g. between Arabic and Berber) or, in absence of historical evidence, only linguistically (e.g. between Central Cushitic and Common Ethiopian) or are at least likely to have taken place for geographic proximity (as between Egyptian and Chadic).

(1.2.1a.) Conversely, if in languages whose ethno-linguistic contacts are unknown, there occur instances of matches unlikely to be either cognates or look-alikes, this can be only accounted for by borrowing to serve as a basis for presuming and further investigating such contacts.

(1.2.2.) Instances of borrowing are often, though far from always, characterized by irregular correspondences between consonantal phonemes of the recipient and source languages.

(1.2.3.) An identical morphological pattern in two languages that is typical of one of them, but uncommon of the other, suggests borrowing.

(1.2.3a.) Conversely, difference in morphological patterns between the two terms speaks against borrowing, save for a clear secondary change in a recipient language (e. g. pluralization).

(1.2.4.) A potential indication of borrowing is attestation of the term in question in the presumed source and recipient languages only (i.e., the word is missing in other languages of the genetic unit to which the recipient language belongs).

(1.2.4a.) On the contrary, if a term is attested in other branches of the family, it is expected to have been inherited by all the daughter languages from the family proto-language. Qualifying this term as a loan-word in the presumed recipient language implies a theoretically possible but somewhat less feasible “double” process — loss and later reappearance as a borrowing.

(1.2.4b.) Attestation in other languages within the compact genetic unit to which the presumably recipient language belongs speaks against borrowing under the following conditions:

- the languages of this compact genetic unit are presumed to have diverged prior to the period(s) of contacts between the suspected recipient language and the source language;
- the languages in question have never undergone influence from the would-be source language;
- the languages in question did not undergo influence from the suspected recipient language during and/or after the period(s) of the latter’s contact with the source language.

(1.2.5.) If the term in question belongs to certain semantic groups that are more open to borrowings, this may be an argument in favour of such a borrowing (one must, however, warn against an uncritical application of this criterion, which, in previous works, has sometimes led to an unwarranted assumption of borrowing of a great part of the cultural lexicon in such languages as Arabic).

(1.2.6.) Unmotivated difference in vocalism between the two terms is an argument against borrowing. Thus, Tigre *nib* ‘tooth’ can hardly be a borrowing from Arabic, where the attested form is *nāb-*. Not only does the Arabic vocalism leave Tigre *-i-* unexplained, but the latter form perfectly corresponds to Hebrew and Aramaic forms that also have *-i-*.

(1.2.7.) Semantic difference: if a secondary semantic development cannot be proved in a recipient language, difference in meaning between the two terms is a strong argument against borrowing.

(1.3.) Reconstruction of the meaning of the protoform.

Provided that regularity of phonetic correspondences is observed for a reconstructed protoform, identifying its most feasible meaning (we omit the easier cases when the meanings of all cognates in the daughter languages, on which the reconstructed protoform relies, are uniform) is of crucial importance for convincing extralinguistic interpretations. Although it goes without saying that a certain meaning is ascribed to each protoform based on close comparison of the complete scope of meanings in individual languages, such an operation can hardly be called proper semantic reconstruction, since, unlike the relatively strict, if not infallible, procedure of phonetic reconstruction, it relies not on a solid method, still conspicuously absent in historical semantics, but rather on the etymologist's intuition and common sense. Anyway, while a dubious choice of a meaning for a protoform may be acceptable in a regular comparative study, it is certainly unpardonable in a study that claims to draw extralinguistic information from linguistic comparison. Thus, an ungrounded, forced assignment of a "cultural" notion to a protoform makes a bias towards picturing a more advanced prehistoric society than it may have been in reality. Recognizing that more "cultural" notions usually go back to more "primitive" notions (at least on a pre-proto-language level), we accept the following guideline in our semantic reconstruction:

— faced with the choice between a "primitive" and a "cultural" meaning, e.g. between that of a wild or domesticated animal or plant species, for a given protoform, the "cultural" meaning, i.e. that of a domesticated species, is proposed only if this meaning is present in the cognate forms of all or nearly all daughter languages. This principle is based on the assumption that independent shift from a "primitive" meaning (wild species) to a "cultural" meaning (domesticated species) in each Afrasian branch and individual language, while theoretically possible, is a far less probable process than the same shift as early as in Proto-Afrasian, from which the "cultural" term was duly inherited by all the daughter languages. The ambivalent cases, i. e. those when a term in question conveys a more "cultural" notion (refers to a domesticated species) in some of the daughter languages, and a more "primitive" notion (refers to wild species) in the others, cannot be used as arguments for ascribing a "cultural" meaning to the protoform.

(2) Another method used for dating the Proto-Afrasian language on the eve of its branching into daughter languages is glottochronology, proposed by the American linguist MORRIS SWADESH in the 1950s ([Sw. 1952] and [Sw. 1955]) and radically improved, updated and tested on many languages belonging to various language families by the recently deceased Russian linguist SERGEI STAROSTIN ([STAR.]) and his successors.

According to SWADESH's method, the most essential, representative, commonly used and, hence, rarely borrowed lexemes are selected for each of the diagnostic 100 wordlist items, which convey some of the most fundamental notions presumed to be present in any human language (personal pronouns, numerals 1 and 2, certain body parts, natural objects, main color terms, several most current verbs and adjectives, &c.). These lexemes are to be compared by means of the lexicostatistical procedure to determine a percentage of etymologically identical units common to any pair of related languages. The principle implies a preliminary stage of compiling a diagnostic wordlist that requests a carefully measured selection of terms. In the Afrasian case, this involves (a) thorough philological analysis of written monuments both in extinct Semitic languages, such as Akkadian, Ugaritic, Biblical Hebrew, Syriac, Classical Arabic, Sabaic and Ge'ez, and in Egyptian, and (b) equally detailed analysis of lexical sources on modern living Afrasian sources, including, where possible, work with active language speakers. For the most part, this preliminary stage has already been completed.

At the same time, unlike SWADESH, who paid little attention to precision and reliability of individual etymologies, and avoided any detailed treatment of the complicated problem of borrowing, STAROSTIN in his method requires meticulous etymological analysis, not merely aimed at accurate and well-grounded establishment of cognate terms, but also one that is supposed to disembarass the list of potential cognates from loanwords — which violate the "natural" algorithm of substitutions in the core lexicon. Tracing loanwords and cogently distinguishing them from inherited lexemes implies high standards of etymological procedure, as well as recurring to sociolinguistic and ethnocultural data. Sometimes, this operation also leads to identifying certain "obscure" lexical items, which we cannot normally trace back to the proto-language or to a reliable source of borrowing due to a lack of data, as potential borrowings from unknown sources.

(3) The third method is that of cross-checking linguistic and archaeological data. As applied to Afrasian linguistics, it has been elaborated by the author in his previous publications, and is based on the following main criteria of identifying "homelands", or original habitats, of reconstructed proto-language communities characterized by a specific archaeological culture (or several cultures):

— one *sine qua non* condition of plausible identification is that dates estimated by both linguistic and archaeological methods should basically coincide;

— the other *sine qua non* condition is that the general outlines of the material culture (as well as elements of intellectual culture and social organization) and natural environment of the presumed homeland, one reconstructed on the basis of the evidence of the proto-language lexicon, the other through archaeological data, should be compatible;

— one strong argument for a particular homeland consists in revealing traces of linguistic contacts between the proto-language in question and its early daughter dialects, on one hand, and other reconstructable proto-languages or ancient languages, likely spoken in the area of the presumed homeland and/or along the migration routes of daughter dialects during the corresponding periods, on the other;

— another strong argument is being able to show that the proposed routes of the daughter dialects' movement towards their historically attested habitats correspond to the directions of cultural expansion or artefact spreads that have been established archaeologically, and/or to the directions of population migrations that have been established genetically.

This study has been carried out within the general framework of projects in comparative Semitic and Afrasian linguistics, supported by the Russian Foundation for Sciences (Project 03-06-80435a), the Russian Foundation for the Humanities (Project 06-04-00397a), The Santa Fe Institute (The "Evolution of Human Languages" Project), and the Ariel Group (The "Tower of Babel" Project). My gratitude goes to these institutions, as well as to my colleagues and collaborators in different projects — Prof. O. STOLBOVA and Drs. L. KOGAN and G. STAROSTIN for consultations and discussions.

The Data

The following 26 entries is an incomplete selection of data that demonstrate, in our opinion, the most reliable or promising Proto-Afrasian terms related to cattle-breeding. A lot of common Afrasian terms referring both to domesticated and wild species in daughter languages, or attested in African Afrasian branches only, are not included on purpose in accordance with the principles and considerations presented above. Undoubtedly, more terms can be adduced and the quoted ones can be complimented and strengthened by more data. I will be grateful to my colleagues for any additions, corrections and criticisms.

1. Livestock

1.1. Small cattle

1.1.1. *mar- 'lamb; ram'

Sem. **ʔimmar-* 'lamb': Akk. ('sheep; sheep and goats; ram'); Ugr.; Phoen., Hbr., Aram.; Arab. Chad. W. **mar-*: Tangale *mara* 'castrated goat', Diri *marì* 'goat', Bokkos *maray*, Tala *màar*, Buli *maro*, Polchi *mar* 'goat'.

Cush. E.: Saho, Afar *márū* 'ram'.

Omot. N.: Wolayta *mára* 'lamb', Male *màràyi* 'ram', *mármáro* 'lamb', Koyra *mará* 'ram', Bworo *meréà* 'sheep' (Blench OLT 72).

[] Cf. [HSED: №1729]; [SED II: №5]; [AADB].

1.1.2. *kar(w/y)- 'ram, goat; lamb, kid'

Sem. **ka/ir(r)-* 'ram, goat': Akk. *kirru* (or *girru*) 'a breed of sheep (?)'; Ugr. *kr* 'ram'; Hbr. *kar* '(young) ram', Aram. *kr* 'sheep'.

(?) Egyp. (NK) *kr-ty* 'horns' (dual).

Berb. **kVrr* 'ram, goat', **kVrw* 'lamb, kid'.

Chad. W. **kwar-/karw-* ~ **kir-*: Kofyar *koor* 'castrated goat', Angas *kīr* 'fattening ram', Dera *kwàrà* 'goat', Zaar *karò* 'sheep', Wangday *kérò* 'ram', &c.; E.: Tobanga (N. Gabri) *karəŋ* 'goat'.

Cush. E.: Arbore *kaaríy* 'heifer goat', *korat* 'male goat', Dobase *koren-te* 'female goat', Yaaku *kurum-* 'goat, young; lamb'.

[] Cf. [HSED: №1432]; [SED II: №118]; [AADB].

1.1.3. *ʕaʔ/ʔp- ~ *pʕ/aʕ- 'kid; goat; ram'

Sem.: Arab. *faʕfaʕ-*; ESA (Min.) *fyʕ*, Soq. *ʕéyʕif* 'kid'.

Egyp. (20th Dyn.) *ʔpwy* 'name of a holy ram'.

Chad. C.: Hwona *wuŋ-rā* ‘she-goat’, Logone (Kotoko) *húfu* ‘goat’.

Cush. E.: Dobase *piŋa-čē* ‘female goat’.

☐ See [SED II: №49]; [AADB].

1.1.4. *ʔayl- ~ *ʔal(l)Vy- ‘ram, sheep’

Sem. *ʔayl- ‘ram’: Akk. (?); Ugr.; Hbr., Aram.; ESA; Tña. (perhaps < Saho-Afar).

Berb. *ti-Hilay- ‘sheep’: Ahaggar *té-helé*, Ghat *či-hali*, Taneslemt *t-ilāy*, &c.

Cush. N.: Beja *alli*, pl. *illi* ‘long-haired sheep’; E.: Saho *ille*, Afar *illi* ‘small cattle’, Arbore *ʔellém*, Elmolo *ʔélem* ‘ram’ (both with *-m* suffixed); S.: Maʔa *iʔalé* ‘ram’, *iʔalú* ‘sheep’.

☐ Cf. [HSED: №67]; [BLA. Beja: 233–234]; [SED II: №24]; [AADB].

1.1.5. *(ya-)bVlaʷy- ‘ram, goat’

Sem. *yābil- (perhaps < *wābil-) ‘ram’: Phoen., Hbr., Aram.; Arab. (?).

Egyp. (OK, MK) *ibʷ* ‘ovis tragelaphus’.

Berb. *b/bal(l)i ‘sheep, ram’: Ghadames *ta-bali*, Audjila *te-bel*, Gurara, Tuat, Tidikelt *belli* (pl.).

Chad. W.: Geji *mbila* ‘sheep’ (cf. *baHil-Vm- ‘horn’: Montol *bulu*, Bolewa *bòlúm*, Galambu *bàlí*, Maha *belem*); C.: Boka *bwələ* ‘goat’, Matakam *báláw* ‘race de mouton sp.’ (cf. Chibak *tə-mbəl* ‘horn’); E.: Lele *bùlóbùlò* ‘he-goat’, Kabalai *bâl*, Migama *bòlíyo*, Sokoro *bàl* ‘goat’.

Cush. E.: Oromo *bulāl-ē* ‘lamb’, Hadiya *ambula* ‘ram’, Kambatta *ambula* ‘goat, ram’.

(?) Omot. S.: Dime *bal-tu*, Galila *baali* ‘horn’.

☐ Cf. [HSED: №2570]; [SED II: №245]; [AADB].

1.1.6. *čaʔw- ~ *ʔačVw- ‘small cattle; meat’

Sem. *taʔ(w)-at- ‘ewe’: Ugr.; Aram.; Arab.; MSA.

(?) Berb.: Canarian (all islands) *chivato*, *chiva* ‘kid’ (rather < Spanish *chibo*).

Chad. W.: Kariya *čiči* ‘goat’, Gera *čača* ‘she-goat’ (redupl.); E.: Kwang *čúwī* ‘he-goat’.

Cush. E.: Somali *soʔ*, Oromo *fooni*, Baiso *soʔo*, Konso *sowa*, &c. ‘meat’ (Oromo *f-* points to *č).

Omot. N. *ʔačVw- ‘meat, flesh’: Koyra *ʔaččo*, Wolayta *ʔašuwa*, Ganjule *ʔačo*, Chara *ačča*, Gimirra *ʔač*, Yamma *aša*, Dizi *ʔač-ku*.

☐ See [SED II: №236]; [AADB].

1.1.7. *(ʔa-)wʷyVn- ‘sheep, goat’

Sem.: Gur.: Cha. *onā*, Ezha *onnā* ‘young male goat or sheep’ (though isolated in Sem., no tenable source of borrowing observed).

Berb.-Can.: Tenerife *ana*, *haña*, *jana* ‘sheep’.

(?) Egyp. (NK) *wny* ‘calf (as a representation of Osiris)’.

Chad. W.: Siri *yàní* ‘she-goat’ (cf. also *nVyVw- ‘horn’: Geji *nowo*, Boghom *nyaw*, Tule *nyewò*, &c.); E.: Migama *íná*, Jegu *té-éné* (pl. *ʔéén*), Birgit *ʔàynéy* ‘she-goat’.

Cush. N.: Beja *ano* ~ *naaʔ* ‘sheep’; E.: Afar *anaŋ-to* ‘lamb (female)’ (*-ŋ* in Auslaut is unexpected as it is not confirmed by Som.), Somali *wan*, pl. *wanan* ‘ram’, Rendille *onó* ‘sheep’.

Omot. S.: Dime *iin*, (?) Ongota *hoona* ‘sheep’.

☐ Cf. [BLA. Beja: 234–235].

1.2. Large cattle

1.2.1. *lawiʔ- ‘large cattle’

Sem. *lVʔ- (m.), *lVʔ-at- (f.) ‘head of large cattle’: Akk. *littu* (*lītu*); Ebl. *lí-a-núm* ‘cow’; Arab. *lāʔa* ‘wild bull, buffalo’; Mhr. *ləháytən* ‘cows’, Jib. *léʔ*, Soq. *ʔélheh* ‘cow’.

Egyp. (Pyr.) *iwʷ* ‘bull’ (if < *lVwVʔ-).

(?) Berb. Tuareg *əlu* ‘bull’ (quoted in [EDE I: 86] as “Tamasheq”, dialect name and source not specified).

Chad. W.: Dera *lāà* ‘cow’, Pero *ló* ‘animal, meat’ (cf. also *laʔu ‘meat’); C.: Gude *la* ‘cow’.

Cush. C. *luway ‘cow’ (Bilin *luwí*, &c.); E.: Saho, Afar *lā* ‘cow, cattle’, LEC *loʔ(loʔ)- ‘cows’ (Somali *loʔ*, Konso *low-aa*, &c.), HEC *lal- ‘cows, cattle’ (Sidamo *lalo*, &c.), Dullay *loʔ-, pl. *leʔ- ‘cow’ (Tsamay *lōʔō*, pl. *lēʔē*, &c.), S.: Qwadza *leʔa-mu-ko* ‘bull’.

☐ Cf. [HSED: №1632]; [SED II: №142]; [AADB]. Cf. Austric *lVw ‘ox, cattle’ ([GIDB]). Cf. metathetic *waʔVl- ‘calf, bull’: Cush. E.: Somali *weeyl*, Hadiya *woʔl-a* ‘calves’; Berb.: Izayan *ta-wala* ‘troupeau de boeufs, sangliers’ [LOUB.: 583]. Cf. [EDE I: 86]; [HSED: №2595].

1.2.2. **(ʕ/pi-)gʷal-* ‘calf; bull, cow’

Sem. **ʕigl-* ‘calf’: Ebl. (?); Ugr.; Phoen., Hbr., Aram.; Arab.; Gez. (*ʕ/ʔəgʷəl*, Tgr. *ʔəgal*)

Egyp. *ʕny*, cow depicted (very likely < **ʕVgʷl-*).

Chad. W.: Sayanchi *gāl*, Geji *gal* ‘cow’; C.: Bura *gyel* ‘bull’.

Cush. S.: Dahalo *ngólome* ‘male buffalo’ (< **nV-gʷVl-Vm-*?).

Omot. N.: Wolayta *gallua*, Zayse *galó* ‘calf’.

[] Cf. metathetic **ʕVlag-* ~ **IVgʕ/ʔ-* ‘calf, bull; (young of small cattle?)’: Sem.: Arab. *ʕulʕūm-* ‘old bull’; Tgr. *läga* ‘male calf’, Tna *lägaʕ* ‘cow close to calving’; Chad. C. **IVgʕ/ʔ-* ‘bull’: Hidkala *álghə*, Bachama *lugüley*, &c.; Cush. N.: Beja *legha* ‘calf’; (?) S. **lagiʔ-* ‘goat’: Alagwa *lagay*, Burunge *legeʔi*. Cf. [HSED: №1100]; [SED II: №28]; [AADB].

1.2.3. **bVr-* ‘(young) bull’

Sem.: Akk. *būru* (*pūru*) ‘young calf’, *bīru* ‘bull (for breeding); young cattle (up to three years)’; Mand. *bira* ‘domestic cattle’; Arab. (Yem. dial.) *bārah* ‘cow’; Tgr. *bara* ‘ox’, Amh. *bare*, Har. *bāra* ‘ox, bull’, Gur. **bawr-* ‘ox (for farming)’. Cf. **bVʕVr-* ‘household animal; beast of burden’ ([SED II: №53]), perhaps derived with a secondary *-ʕ-*.

(?) Egyp. *bʕwy* ‘arena, battlefield for bull-fight’ (presumably a *nisba* < **bVr-* ‘bull’, cf. [EDE II: 53–54]).

Chad. W.: Gera *bara* ‘buffalo’; E.: Mokilko *búru* ‘bull’.

Cush. N.: Beja *beʔráy* ‘bull, cow’ (< Eth.?); C. **bir-* ‘ox, bull’ (Bilin *birā*, &c. < Eth.?); E.: Afar *abur* ‘bull, ox’, HEC **bōr-* ‘young bull’ (Sidamo *boor-to*, &c. < Eth.?)

Omot. N. **bariy-* ‘bull’: Wolayta *bóora* ‘ox’, Gamo *bóora* ‘bull’, Zala *bōrā* ‘ox’ (acc. to Blench OLT 68, all three < Gur.), Chara *bíira* (acc. to Blench *ibid.*, < Agaw), Kafa *bariyō* ‘calf’, Mocha *bariyo* ‘steer’, Bworo *berō*, Sheko *bariyo* ‘bull’.

[] Cf. [HSED: №183]; [BLA. Beja: 238]; [SED II: №53 notes.]; [EDE II: 54–55]; [AADB]. Cf. Nostr. **bVrV-*: Alt. **bīōru* ‘calf, lamb’, Drav. **paḍḍ-/baḍ-* ‘heifer’ ([GIDB]).

1.2.4. **ʔi/qrw-* ‘calf; bull’

Sem. **ʔi/qrw-ān-* ~ **ʔawr-* ‘calf, bull’: Syr. *ʔarwān-* ‘calf’; Arab. *ʔirān-* ‘male oryx’ (*šātu ʔirānin* ‘bull’); Amh. *awra* ‘male (animal), dominant or alpha male; bull’; Gur. **ʔaraʔ/y-* ‘cows’.

Egyp. (MK) *ir-t* ‘calf’, (Dem.) *iry-t* ‘milking cow’.

Berb. **-HirVy* ‘calf’: Ayr *ehəri*, Shilh *irey*, &c.

Chad. W.: Dera *wóré*, *ara* ‘meat’, Sha *ʔarwà* ‘ox’.

Cush. N.: Beja *oreo* ‘bull, steer’, *rēw* ‘cow’; E.: Saho, Afar *awr* ‘bull’, LEC: Somali *awr* ‘he-camel’, Rendille *or* ‘he-camel, bull’, Oromo *oor-oo* ‘burden camel’, Arbore *ʔaar*, Dasenech *ar* ‘bull’, &c., HEC: Burji *arráy*, *arʔáy* ‘bull’, *ʔre* ‘calf’; Yaaku *reheʔ* ‘calf’.

(?) Omot. N.: Malo *hāri* ‘cattle’, Oyda (*h*)*arr* ‘cow’.

[] Cf. [BLA. Beja: 236, 269]; [SED II: №16]; [AADB].

1.2.5. **maray-* ‘calf, (young) bull, steer’

Sem.: Akk. *mīru* ‘young bull’, *mīrtu* ‘cow’ (Hbr. *mārī*(?)) ‘fatted steer’ is not necessarily related being probably derived from the verbal root *mrʔ* ‘to be fat’).

Egyp. (MK) *mr(y)* ‘fighting bull’.

Chad. C.: Matakam *maray* ‘bull (for sacrifice)’, Mofu-Gudur *maray* ‘bull fattened in a stable’.

Cush. E.: LEC: Rendille *máar* ‘male calf’, *maár* ‘fem. calf’, Arbore *máar* ‘calves’ (coll.), HEC: Hadiya *moora* ‘older calf’, Dullay: Harso, Dihina *maar-akkó*, Tsamay *maare* ‘heifer’; S.: Maʔa *-moro* ‘steer’.

Omot. N.: Wolayta *mārā*, Dorze *mar*, Male *marro* ‘calf’, Yamma *omoru* ‘bull’; S.: Ongota *marte* ‘calf (she)’.

[] Cf. Chad. W.: Hausa *mārīrī* ‘white oryx’, E.: Toram *múro* ‘gazelle’. Cf. [HSED: №1728]; [SED II: №206]; [EDE III: 390–392]; [AADB]. Perhaps related to **(ʔV-)mar-* ‘lamb; ram’ on the Pre-Proto-Afrasian level. Cf. Drav. **mūr-* ‘buffalo; cow’ ([GIDB]).

1.3. Common or mixed terms for small and large cattle**1.3.1. **(ʔa-)fVr-* ‘small and large cattle’**

Sem. **parr-* (1) ‘lamb’ (Akk. *parru*; Syr. *parr-*, Mnd. *par-*; Arab. *furār-*, *farūr-*), (2) ‘cattle’ (Ugr. *pr*; Hbr. *par* ‘bull, steer’, *pārā* ‘cow’, Aram. (Sam.) *pr* ‘bull’, *prh* ‘cow’; Tgr. *fərrət* ‘pasturing herd’,

māfrār ‘herd (of cows)’, Amh. *afwarä* ‘to become a yearling ox’; (?) MSA: Mhr. *fōr* (pl. *fāhārīn*) ‘young bull’, Jib. *fǝǝr* ‘young bull, male calf’, Soq. *fāfhar* ‘young bull’ (with a secondary -f-).

Chad. C. **faray-* ‘buffalo; cattle’: Mbara *fàrày* ‘cattle; dot (bride wealth)’; there are also Bura *fir*, Kilba *fur*, Margi *fūr* ‘buffalo’, but they are considered < **fun-*, about which I have some doubts.

Cush. E.: Yaaku *apur* ‘sheep’; S.: Asa *ʔeferet*, *ʔoforok*, Qwadza *afulatu* ‘he-goat’.

☐ Cf. [HSED: №1950]; [SED II: №181]; [AADB].

1.3.2. **čayw/ɣ-* ‘(meat of) small or large cattle’

Sem. **šaw-* ‘head of small cattle’: Akk.; Ugr.; Phoen, Hbr.; Arab.; ESA (Sab.).

Egyp. (MK) *šy* ‘pig’ (cf. also *šw* ‘ass’).

Chad. W. **ča-* ‘cow’: Siri *záà-tí*, Jimi, Polchi *šaa*, Dwot, Buli, Zul, Ngizim *šaa* (cf. also **čaw-* ‘meat’); C. **ša-* ‘cow’: Tera *ža*, Bura, Margi, Gisiga *ša*, &c. (cf. also **šuw-* ‘meat’)

Cush. S. **šaʔe-* ‘cow’: Iraqw, Alagwa, Burunge *šee*, Asa *še-ok*, Qwadza *šae-ko*.

Omot. N. **ʔayš-* ~ **šaʔ-* ‘goat’: Bworo *eyššá*, Mao (Hozo) *šaa*, (Ganza) *saʔa*, Dizi *es-ku*.

☐ Cf. [HSED: №517]; [SED II: №217]; [AADB].

1.3.3. **p/ɸVl-* ‘(young of) hoofed domestic animals’

Sem. **pVlW/ɣ-* ‘foal, small of domestic animals’: Arab. *filw-*, *faluww-*, *fuluww-* ‘a yearling foal or ass already weaned’; Tgr. *fəlit* ‘calf’, *fəluɣ* ‘calf weaned’; Soq. *fólhi* ‘a yearling calf’ (in Tgr. and Soq. borrowing from Arabic is possible).

(?) Chad. W.: Guruntum *fwull* ‘cow’ (isolated term).

Cush. N.: Beja *filay* ‘she-camel just foaled’ (borrowing from Arabic or Tigre not to be ruled out); C. **fiyal-* ‘goat’ (Aungi *fəyāli*, &c.); E.: HEC **fillaʔ-* ‘goat’ (Kambatta *felle-čču*, pl. *felleʔu*, &c.).

Omot. N.: Kafa *fell* ‘goat’ (Blench OLT 71) < HEC?

☐ Cf. [BLA. Beja: 246]; [SED II: №174]; [AADB].

1.3.4. **dVbal-* ‘pig/boar, ram, goat, calf’

Sem.: Arb. *dawbal-* ‘wild boar, suckling pig’, Gez. *dābelā* ‘billy goat, bull, male of any animal’, Tgr. *dābela* ‘ram’, Tña. *dibāla*, Amh. *dabela*, *dābāl* ‘billy goat’ ([LGz.: 120–121]; in view of a tenable Arab. parallel, less likely < Cush. as LESLAU asserts, while E. Cush.: Saho, Afar *dabeéla* ‘billy goat’ are rather borrowed from Eth.).

Cush. N.: Beja *debala* ‘yearling cow’; E.: Baiso *dabaalo* ‘heifer’ (cf. in [BLA. Beja: 243]).

☐ Cf. [BLA. Beja: 269]; [AADB].

2. Pasturing, foraging and tending livestock

2.1. **g(ʷ)Vč-* ‘to tend, drive livestock’ ~ **gičac-* ‘pasture’

Sem. **gʷVšVɣ/ɣ-* ‘to tend cattle’: (?) Arb. *ǧšʔ* ‘produce vegetation (soil); emigrate (tribe)’; *ǧtašša* ‘be covered with dense grass (soil)’ (to be interpreted as ‘to migrate with the animals to grassy pastures?’); Gez. *gʷāsaya* ‘to tend cattle’ ([LGz.: 205]; < Tña.?), Tña. *gʷasäyā* id., (?) Endegeñ, Gyeto *gʷišašā* ‘field, plain, meadow, open space’ (according to [LGur.: 299], < Hadiya); Soq. *geš* ‘to pasture, drive cattle’, *gésiš* ‘pasture’.

Cush. C.: Kemant *gəšəš* ‘pasture’; E.: Hadiya *gišaša*, Burji *giiš-* ‘to graze’. Cf. E.: Somali *goš-* ‘to ply between two places’.

Omot. N.: Kafa *gaš-* ‘drive one’s cows to pasture’.

☐ [AADB].

2.2. **rVɸVɣ/ɣw-* ‘to pasture, tend livestock; chase; accompany, follow’

Sem. **rVɸVɣ-* ‘to pasture, herd (trans.); be herdsman, friend, companion’: Akk. *reʔu* ‘to pasture, guard livestock, herd, graze (trans.)’, *rūʔu* ‘friend, companion’; Ugr. *ršy* ‘herdsman’, *rš* ‘friend, companion’; Hbr. *ršy* ‘to pasture, guard livestock, herd, graze; join, associate with’, *rōšā* ‘herdsman’, *rēʔ* ‘friend, companion’; Syr. *rəšā* ‘to pasture, herd’; ESA (Sab.) *ršy* id.; Arab. *ršy* id., *rāš-* ‘herdsman’; Gez. *rəšya* ‘herdsman’, *rašawa* ‘to yoke, join’, Tigre *rāša* ‘to pasture, herd’, *rāwša* ‘have intercourse’; Mhr. *rō* ‘to pasture, herd’, *rēfi* ‘herder’, Jib. *raše* ‘to pasture, herd’, *rēfi* ‘friend, companion’, Soq. *reše* ‘to pasture, herd’.

(?) Egyp. (OK) *yš*?, verb connected with handling calves ([EG I: 27]; [EDE III: 50]; related if < **yrš*).

Chad. W.: Diri *rawā(n)*, Tsagu *rāʔa* ‘to herd, graze’ (perhaps < Arab.); E.: Bidiya *ʔaraw* ‘chase’ (< *ʕaraw-, with metathesis ?).

Cush. E.: Somali *raaʕ-* ‘to accompany, go together’ (in view of other LEC parallels, hardly < Sem.), Boni *rāà*, Rendille *raḥ* ‘to follow’, Oromo *ariʔa* ‘to pursue, chase’; S.: *ʕārō (met.?) ‘ruminated fodder, grass’ (Iraqw *ʕaaroo*, &c.)

☐ [AADB]. The original meaning may be ‘to pasture’ or ‘to chase, follow (wild ungulates)’.

2.3. *(na-)kīd- ‘shepherd of small cattle’

Sem.: Akk. *nākidu* ‘herdsman’; Ugr. *nkd* ‘shepherd’; Hbr. *nōkēd* ‘shepherd, sheep-breeder’, pB. ‘lamb’; Arab. *naḳḳād-* ‘shepherd’ ([HALOT]; not in [BK]), *naḳad-* ‘kind of ram’ (BK 2 1321; cf. *naḳd-* ‘silver, money’ *ibid.*).

Berb.: Ayr, E. Tawllemmet *ə-yadyad* ‘herd of goats’.

Omot. N.: Moča *qiddo* ‘shepherd’, Kafa *qidō* ‘guardian’.

☐ [AADB]. A promising root, though not quite reliable because of scarce data.

2.4. *cVḥ- ‘pasture, to pasture, to herd’

Sem.: Akk. *sahhu* (*sāhu*) ‘meadow, waterlogged land’ (-ḥ- < *-ḥ- is possible--cf. Kog.); Arab. *shḥ* ‘to be very fat (of rams)’; Tña. *sāwhi* ‘meadow, ever-green pasture’.

(?) Egyp. (OK) *shʔ-t* ‘herd of donkeys’.

Chad. W.: Wrj. *čiyə*, Siri *cagu*, Sha *čó* ‘to herd, graze’, (?) Dera *ǰíwà* ‘herd’.

Cush. C.: Kemant *sēḥā* ‘prairie’.

☐ Cf. [HSED: №385]; [AADB]. Not quite reliable because of phonetic problems and semantic diversity.

2.5. *k^wal- ‘forage, fodder; pasture; mowing, collecting, gleaning’

Sem. *k^walāʔ- ‘forage’, *k^wVIVw- ‘to mow’: Akk. *ukullū* ‘cattle fodder, forage’; Arab. *klʔ* ‘abound in forage (area)’, *kalaʔ-* ‘forage (dry or fresh)’; Gez. *k^wālawa* ‘to reap, mow’ ([LGz.: 284]), *makala*, *makkola* (acc. to LESLAU, for *makk^wala*; secondary derivation with *m-* prefixed) ‘to cut with a sickle, mow’ ([*ibid.*: 339]), Tgr. *māklay* ‘halm of durra, halm of corn’, Tña. *māḳälä* ‘to mow, cut’, Amh. *kalkəl* ‘pasture’; cf. Jib. *kélet* ‘bush with edible twigs’.

Berb. *kVlkVl- ‘pick up, collect’: Ahaggar *kelukelu*, Ayr *kələnkilet* (with a secondary -n-).

Chad. W.: Hausa *kālā* ‘gleaning’; E.: Lele *kél* ‘pick up, collect’.

Cush. C. (?) Khamir *kilkil* ‘pasture’ (likely a loan of Amh. *kalkəl*): E.: Oromo *kalō* ‘pasture land’, Kambatta *kalu*, Sidamo, Burji *kalo* ‘pasture’ (probably borrowed from Oromo).

☐ [AADB]. To separate from *kal(aʔ)- ‘earth, land’ (see [MIL.: №23]).

2.6. *ʔVry- ‘cattle-shed’

Sem. *ʔu/arVry- ‘stall’: Akk. *urū* ‘stall’; Hbr. *ʔurwā (HALOT: “from Akk.-Sum. *urū* > Aram > Arab.”) ‘stall’; Jud. *ʔūryā*, Syr. *ʔōryā*; Arb. *ʔariy-*, *ʔiry-* id.

Cush. E. *ʔari-t- ‘gate of animal pen’: Somali *irid-i*, Rendille *ar’it* (< *arit Hei. 74).

☐ [AADB]. Not A promising root, though not quite reliable because of scarce data.

Products of stock-raising

*sim-an- ‘fat milk; to milk, suck; butter, oil, fat’

Sem. *šamVn-* ‘fat, oil, butter’: Akk. *šamnu* ‘oil, fat’; Ugr. *šmn*; Hbr. *šämän* ‘oil, fat’, Syr. *šumn*, Mnd.šamin ‘fat’, Maʕlula *šomna* ‘butter’; Arab. *samn-* ‘fat; melted) butter’; Jib. *šēn* ‘fat’.

Egyp. (Med.) *smy* ‘fat milk, cream’.

Berb.-Can.: Ahaggar *ésim* ‘melted fat’, Ghat *isim* ‘animal fat’, Qabyle *ta-ssəm-t* id., *summ* ‘to suck’; Can. (Ferro) *achemen* ‘milk’.

Chad. W.: Somrai *šī-šēm*, Sokoro *ʔā-sīmà* ‘to suck (sugar)’.

Cush. N: Beja *simuum* ‘suet, fatty covering of kidneys’; E.: LEC *siḥim- with a secondary -h- (?): Somali *siḥin-* ‘curds’, Rendille *siḥime* ‘butter’, Dullay *šīnam-(met.?): Dalpena *šiinan-ko*, pl. *šiinan-aane* ‘butter’, Gollango *šiinan-ko*, pl. *šinam-aane* ‘fat’; S.: Iraqw *ismoo* ‘nipple’, Qwadza *sum-* ‘to milk’, Maʔa *semu* ‘breast’.

☐ Cf. [HSED: №2247]; [SED I: №248]; [AADB]; [BLA. Review: 505].

***sVP/fVy- ‘churned milk, curds’**

Sem.: Hbr. *šəpōt* ‘cheese or curds (made of cow’s milk)’.

Cush. N.: Beja *šəfi* ‘drink milk’; E.: HEC **šaff-* ‘to churn’ (Kambatta *šaffo*, &c.).

□ [AADB].

Livestock as a socio-economic category***g(ʷ)iʒʒ- ‘domestic animals as possessions, property’**

Sem.: Arb. *šwz* ‘go, march, drive beasts of burden and riding animals, take them to watering place’, IV ‘give so. a certain sum of money’; Gez. *gāz*, *gāzā*, *gizān* ‘treasury, wealth, money’ (acc. to [LGz.: 210]: ‘< Greek γᾶζᾶ; also occurs in Aram.-Syr. *gazzā* going back to Median *ganza*’, which is questionable in view of Semitic and Afrasian parallels), Gafat *gəzzā* ‘cattle, money’, *gəzzā* ‘to master’, Wol., Zw. *gəzat*, Sel. *gəzāt* ‘cows, domestic animals, cattle’, Muh., Gog., Sod. *gəzz* ‘cows, domestic animals, cattle’, Sod. *gəzzoday* ‘shepherd’ (Acc. to LESLAU: ‘probably passed into Cushitic... The root could also be common to Semitic-Ethiopic and Cushitic’; the latter suggestion is more likely than the former as the meaning ‘cattle’ is not attested in Amharic, a plausible source of borrowing into HEC and N. Omot., which can hardly borrow terms meaning ‘cattle’ and ‘money’ from Gafat or Gurage dialects).

Berb.: Zenaga *a-guzzīh* ‘herd of sheep’.

Cush. E.: Kambatta *gizza* ‘cattle’, Tembaro *gəzza* ‘money, cattle’, Hadiya *godda* ‘cattle, property, wealth’.

Omot. N.: Kafa *giʒʒō* ‘livestock, money, welfare, possessions’, Chara *gizā*, *giʒʒā* id., She *giz* ‘welfare, possessions’, Mao (Sezo) *gizzi*, (Hozo) *gitza* ‘cattle’.

□ [AADB]. Perhaps related to Eth. **giz(z)-* ‘cattle, money’ is Eth. **gzʔ* ‘to master, possess’: Gez. *ʔəgzʔ* ‘master’, Tgr. *gəzʔa* ‘to possess, dominate’, Tña. *gəzʔa* ‘to possess, buy’, Amh. *gəzza* ‘to possess, buy, govern’ ([LGz. 210]), Gur. **gāza* ‘to own, possess, govern, &c.’ ([LGur. 304]).

***mal- ‘livestock (as a source of milk or meat, or as a capital); tend livestock’:**

Sem.: Arab. *mwl* ‘be rich, esp. in livestock’, *māl-* ‘herd of camels; richness’; ESA (Sab.) *mly* ‘to get, win, obtain as booty’, *mlt* ‘loot, booty, prize of war’; Tgr. *mal* ‘money, fortune, property’, Tña. *mal* ‘herd (of livestock); goods, property, wealth’; Mhr. *mōl* ‘livestock, capital’, *mālēt* ‘she-camel’, Jib. *mol* ‘livestock, capital’, Soq. *māl* ‘richness’. The N. Eth. and MSA forms meaning ‘livestock, capital’ are most probably Arabisms while Mhr. *mālēt* ‘she-camel’ is not necessarily so. Cf. also Hrs. *melēt* ‘bride-price’ and Muh. *muli* (however, Chaha, Eža, Endegeñ, Gyeto *muri*) ‘boy to whom a girl is given by her parents without being asked for by his parents’.

(?) Egypt. (ME) *mmmn-t* ‘herd’ (if < **mVlmVl-*); *mny* (MK or NK) ‘herdsman’ (if < **mVly*; cf., however, Coptic Fayumic *mani* id., with *-n-* instead of the expected *-l-*). Cf. *mny* ‘marry; endow with’ (in [FAUL.: 104] combined with ‘to moor’, ‘attach’, ‘save’ and ‘die’, semantic connections not quite clear).

Berb. **-malVy-* ‘camel, stallion, not castrated animal’: Ghadames *amāli* ‘stallion camel’, Ghat *a-mali* ‘stallion’, Ahaggar *āmāli*, Ayr *əmaləy*, E. Tawllemmet *əmeləy* (cf. Ahaggar *əmhəl* ‘to push ahead, drive (animals, livestock)’, Ayr *əmhəl* ‘to advance, push ahead’ < **mVʔVl* ?).

Chad. W.: Jimi *màalo*, Geji *máal* ‘goat’ (acc. to [EDE III: 42], < **mar-*, see 1.1.1); C.: Masa *mòl-* ‘to assemble (a herd of animals)’; (?) E.: Mokilko *máàlà* ‘wealth, dowry, property’ (*máàlàdò*, *màldò* ‘my property, my herd’), W. Dangla *málá*, E. Dangla *màllē* ‘herds, animals, cattle; riches’ (would be undoubtedly labelled Arabisms if not for W. and C. Chad. forms that are evidently not, which leaves room for some doubts).

Cush. C.: Kemant, Qwara *mält*, Kailiña *mil-d-* ‘to look after (cattle), tend flocks’ (Bilin *mal* ‘herd, livestock, richness’ is, acc. to [APPL.: 83], from Tgr. or Tña.); E.: LEC: Somali *màal* ‘livestock that provide milk’ (*máal*, Rendille *m’aal* ‘to milk’), (?) Oromo *mil-* ‘to guard’; HEC: Burji *malāl-* ‘to herd’, *maalát-e* ‘lending out of cattle’ (cf. also Sidamo, Darasa, Qabenna *maal-a*, Burji *máal-a* ‘meat’); (?) Dullay: Gollango *mila* ‘fresh, cool milk’.

Omot. N.: Haruro *mālināy* ‘herdsman’, Koyra *mālē* ‘cow’.

□ Cf. [LAMB. 474]: Koyra (comp. to unrelated Omot. forms); HEC; Som. (Saho-Afar ‘money’ is evidently < Arab.). Cf. [EDE III: 254, 294 and 72] (acc. to TAKÁCS, Berb. Tuareg forms meaning ‘stallion, camel’ ‘‘look rather like **m-* prefix participial derivations from Berb. **l-y* ‘to mount’ = **ā-hlū/īy*, which looks rather far-fetched for several reasons, one being lack of *h-* in these forms, another, strange semantic development).

Abbreviations of languages and language periods

Akkadian; **Amharic**; **Arabic**; **Aramaic**; **Berber**; **Central**; **Canarian**; **Chadic**; **Cushitic**; **Demotic**; **Dynasty**; **East**; **Egyptian**; **ESA** — Epigraphic South Arabian; **Ethiopian**; **Gafat**; **Gez.** — Ge'ez; **Gurage**; **Harari**; **HEC** — Highland East Cushitic; **Hbr.** — Hebrew; **Hrs.** — Harsusi; **Jibbali**; **Judaic** Aramaic; **LEC** — Lowland East Cushitic; **Medical Texts**; **Mhr.** — Mehri; **MK** — Middle Kingdom; **Mnd.** — Mandaic Aramaic; **MSA** — Modern South Arabian; **North**; **NK** — New Kingdom; **OK** — Old Kingdom; **Omoti**; **Phoenician**; **Pyramid Texts**; **South**; **Sabaic**; **Selti**; **Semitic**; **Soddo**; **Soqotri**; **Syrian Aramaic**; **Tña.** — Tigriña (= Tigray); **Tgr.** — Tigre; **Ugr.** — Ugaritic; **West**; **Wolane**.

Bibliographic References and Abbreviations

- AADB — *Afrasian Database*, sites: <http://ehl.santafe.edu> and <http://starling.rinet.ru>.
- APPL. — D. A. APPLEYARD. *Comparative Dictionary of the Agaw Languages // Kuschitische Sprachstudien ~ Cushitic Language Studies*, Band 24. Köln. 2006.
- BAR-YOSSEF — O. BAR-YOSSEF. *The Natufian Culture and the Early Neolithic: Social and Economic Trends in South-Western Asia // Examining the Farming/Language Dispersal Hypothesis / eds. P. BELLWOOD & C. RENFREW. McDonald Institute Monographs, Cambridge. 2002; pp. 113–126.*
- BK — A. DE BIBERSTEIN-KAZIMIRSKI. *Dictionnaire arabe-français*. Paris. 1860.
- BLA. Beja — V. BLAŽEK. *Fauna in Beja Lexicon: A Fragment of a Comparative-Etymological Dictionary of Beja // Studia Semitica: FS for A. MILITAREV (= Orientalia: Труды Института восточных культур, III [Papers of the Oriental Institute, III])*. Moscow. 2003; pp. 230–294.
- BLA. Review — V. BLAŽEK. *Semitic Etymological Dictionary I. Archiv orientální*, Vol. 69. 2001; pp. 495–510.
- BLENCH ALAP — R. BLENCH. *Archaeology, Language and the African Past*. Lanham. 2006.
- BLENCH OLT — R. BLENCH. *Omoti Livestock Terminology and Its Implication for the History of Afroasiatic. Semito-Hamitic Festschrift for A. B. DOLGOPOLSKY and H. JUNGRAITHMAYR / Ed. G. TAKÁCS. 2008; pp. 63–78.*
- COH. — M. COHEN. *Essai comparatif sur le vocabulaire et la phonétique du chamito-sémitique*. Paris. 1947.
- DIAK. ESA — I. DIAKONOFF. *Earliest Semites in Asia. Agriculture and Animal Husbandry According to Linguistic Data (VIIIth–IVth Millennia B.C.) // Altorientalische Forschungen*, 8. 1981; pp. 23–74.
- DIAK. Sum. — И. М. ДЯКОНОВ. *Шумеры и афразийцы глазами историка // Вестник древней истории*, № 4. 1996; стр. 81–86 [I. M. DIAKONOFF. *Sumerians and Afrasians through the Eyes of a Historian // Journal of Ancient History*, № 4. 1996]; pp. 81–86.
- DOLG. Cush. — А. Б. ДОЛГОПОЛЬСКИЙ. *Сравнительно-историческая фонетика кушитских языков*. М., 1973 [A. B. DOLGOPOLSKY. *Comparative-Historical Phonetics of Cushitic*. Moscow. 1973].
- EDE I — G. TAKÁCS. *Etymological Dictionary of Egyptian, Volume One: A Phonological Introduction*. Leiden-Boston (Ma) & Cologne. 1999
- EDE II — G. TAKÁCS. *Etymological Dictionary of Egyptian, Volume Two*. Leiden-Boston (MA) & Cologne. 2001.
- EDE III — G. TAKÁCS. *Etymological Dictionary of Egyptian, Volume Three*. Leiden-Boston. 2008.
- EHR. EEA — Ch. EHRET. *Ethiopians and East Africans. The Problems of Contacts*. Nairobi. 1974.
- EHR. LFE — Ch. EHRET. *Language Family Expansions: Broadening our Understanding of Cause from an African Perspective // Examining the Farming/Language Dispersal Hypothesis / eds. P. BELLWOOD & C. RENFREW. McDonald Institute Monographs, Cambridge. 2002; pp. 163–176.*
- EHR. PA — Ch. EHRET. *Reconstructing Proto-Afroasiatic (Proto-Afrasian). Vowels, Tone, Consonants, and Vocabulary*. Berkeley, Los Angeles. 1995.
- EG — *Wörterbuch der ägyptischen Sprache / Im Auftrage der deutschen Akademien hrsg. von Adolf ERMAN & Hermann GRAPOW. Erster Band: 1961 [1971]; Zweiter Band: 1955 [1971]; Dritter Band: 1954 [1971]; Vierter Band: 1957 [1971]; Fünfter Band: 1954 [1971]; Sechster Band (Deutsch-Aegyptisches Wörterverzeichnis. In alphabetischer und sachlicher Ordnung. Nebst Verzeichnissen der koptischen, semitischen und griechischen Wörter): 1950 [1957]; Siebenter Band (Rückläufiges Wörterverzeichnis / bearb. von W. F. REINEKE): 1963 [1971]. Berlin: Akademie Verlag.*
- FAUL. — R. O. FAULKNER. *A Concise Dictionary of Middle Egyptian*. Oxford. 1962.
- GIDB — *Global Database*, sites: <http://ehl.santafe.edu> and <http://starling.rinet.ru>.
- GR. — J. GREENBERG. *The Languages of Africa*. The Hague. 1963.

- HALOT — L. KOEHLER and W. BAUMGARTNER. *The Hebrew and Aramaic Lexicon of the Old Testament* I–III. Leiden, New York & Köln. IV–V. Leiden, Boston & Köln. 1994–1996, 1999–2000.
- HASS. — F. HASSAN. Archaeology and Linguistic Diversity in North Africa // *Examining the Farming/Language Dispersal Hypothesis* / Eds. P. BELLWOOD & C. RENFREW. McDonald Institute Monographs, Cambridge. 2002; pp. 127–133.
- HCVA — DIAKONOFF et al. 1994–1997 — Historical-Comparative Vocabulary of Afrasian // *St. Petersburg Journal of African Studies*, 2 (1994): pp. 5–28; 3 (1994): pp. 5–26; 4 (1995): pp. 7–38; 5 (1995): pp. 4–32; 6 (1997): pp. 12–35.
- HEI. — B. HEINE. The Sam Languages: A history of Rendille, Boni and Somali // *Afroasiatic Linguistics*, 6/2. 1978. 23–116.
- HSED — V. OREL and O. STOLBOVA. *Hamito-Semitic Etymological Dictionary. Materials for a Reconstruction*. Leiden–New York–Köln. 1995.
- I-S — В. М. ИЛЛИЧ-СВИТЫЧ. *Опыт сравнения ностратических языков (семитохамитский, картвельский, индоевропейский, уральский, дравидийский, алтайский): Введение. Сравнительный словарь*. М.: «Наука» [*An Attempt at Comparative Dictionary of the Nostratic Languages (Semitic-Hamitic, Kartvelian, Indo-European, Uralic, Dravidian, Altaic)*]. Moscow: “Nauka” publishers]. V. 1 (b–k): 1971; V. 2 (l–j): 1976; V. 3 (p–q): 1984.
- KOG. — Л. Е. КОГАН. О нерегулярных рефлексах семитских ларингалов в аккадском языке [On Irregular Reflexes of Proto-Semitic Laryngeals in Akkadian] // *Вестник древней истории* [*Journal of Ancient History*] 2. 1995; pp. 156–162.
- LAMB. — M. LAMBERTI and R. SOTTILE. *The Wolaytta Language*. Köln. 1997.
- LGur. — W. LESLAU. *Etymological Dictionary of Gurage (Ethiopic)*. Vol. III. Wiesbaden. 1979.
- LGz. — W. LESLAU. *Comparative Dictionary of Ge'ez (Classical Ethiopic)*. Wiesbaden. 1987.
- MIL. — A. MILITAREV. The Prehistory of a Dispersal: The Proto-Afrasian (Afroasiatic) Farming Lexicon // *Examining the Farming/Language Dispersal Hypothesis* / eds. P. BELLWOOD & C. RENFREW. McDonald Institute Monographs, Cambridge. 2002; pp. 135–150.
- PELT. — E. PELTENBURG & al. Agro-Pastoralist Colonization of Cyprus in the 10th Millennium BP: Initial Assessments // *Antiquity*, 74, 2000; pp. 844–853.
- SED — Alexander MILITAREV & Leonid KOGAN. *Semitic Etymological Dictionary (= Alter Orient und Altes Testament: Veröffentlichungen zur Kultur und Geschichte des Alten Orients und des Alten Testaments)*. Vol. I (*Anatomy of Man and Animals*): 2000 (= AOAT 278/1). Vol. II (*Animal Names*): 2005 (= AOAT 278/2). Münster: Ugarit-Verlag.
- STAR. — S. STAROSTIN. Comparative-Historical Linguistics and Lexicostatistics // *Time Depth in Historical Linguistics*, Vol. 1 / eds. C. RENFREW, A. McMAHON & L. TRASK. Papers in the Prehistory of Languages, Cambridge. 2000; pp. 223–265.
- SW. 1952 — M. SWADESH. Lexico-Statistical Dating of Prehistoric Ethnic Contacts: With Special Reference to North American Indians and Eskimos // *Proceedings of the American Philological Society*, 96; pp. 452–463.
- SW. 1955 — M. SWADESH. 1955. Toward Greater Accuracy in Lexicostatistical Dating // *International Journal of American Linguistics*, 21; pp. 121–137.

Р е з ю м е

Статья — очередной шаг в направлении отождествления автором картины жизни носителей праафразийского (ПАА), или прасемито-хамитского, языка, реконструируемой по общафразийской лексике, с ранне-неолитической постнатуфийской археологической культурой Восточного Средиземноморья. ПАА язык накануне распада датируется автором по глоттохронологическому методу С. А. Старостина 10 тыс. до н. э. — тем же временем, что и постнатуф (предполагаемая родина земледелия и скотоводства на планете) по радиоуглеродным датировкам. Статья посвящена доказательствам наличия в ПАА языке скотоводческой лексики. Приводятся 26 реконструированных названий мелкого и крупного рогатого скота и хозяйственных терминов. Лексическим данным предшествует краткое описание ситуации в современном сравнительно-историческом афразийском языкознании и изложение авторских принципов и приемов этимологического анализа и реконструкции праязыковой лексики, а также разработанных им методов сопоставления лингвистических и археологических данных для установления прародины языковых семей.

On the Problem of the Proto-Mande Homeland¹

When dealing with the internal classification of the Mande language family (Niger-Congo phylum), short-range groups can be singled out easily; the main problem concerns mid-range stages of the internal classification. The present article suggests a new version of such a classification. For the first time, it is based on the standard SWADESH 100-word list. Lexicostatistical data are then confronted with the evidence of paleoclimatology and archaeology. I conclude that speakers of Proto-Mande most probably lived around the second half of the 4th millennium BC in Southern Sahara, somewhere between 3° and 12° Western longitude, to the North of the 16° or even 18° latitude. This hypothesis is supported through analysis of the cultural vocabulary that can presumably be reconstructed for the Proto-Mande language.

The Mande family² belongs to the hypothetical Niger-Congo phylum whose genetic depth, according to tentative estimations, lies in the range of 12 to 14 millennia. The inner classification of Mande has a long history, expounded in [KASTENHOLZ 1997]. As is often the case with language classifications, there is a relative consensus here concerning short-range grouping. At this level, the following groups can be established:

1. **Southern Mande** (Dan, Guro, Yaure, Tura, Mano, Mwan, Wan, Beng, Gban, Gbin).
2. **Eastern Mande** (San, Sane, Bisa, Boko, Bokobaru, Busa, Kyenga, Shanga).
3. **Samogo** (Jo, Seenku, Banka, Duun, Dzuun, Kpan, Kpeen).
4. **Bobo**.
5. **Soninke-Bozo**.
6. **South-Western Mande** (Mende, Loko, Bandi, Looma, Kpelle).
7. **Soso-Jalonke**.
8. **Jogo** (Jogo, Jeri, Numu, Ble).
9. **Vai-Kono**.
10. **Mokole** (Mogofin, Kakabe, Koranko, Lele).
11. **Manding** (Bamana, Maninka, numerous varieties of Jula, Mandinka, Xasonka, Marka-Dafin, Mau, &c.).

However, when it comes to combining smaller groups into middle-range ones, the story is long and controversial. I am not going to delve into the comparison of all the preceding variants of the inner classification of Mande, as this job has already been done by KASTENHOLZ; only two most recent classifications, one by Claire GRÉGOIRE and Bernard DE HALLEUX [1994] and the other by Raimund KASTENHOLZ [1997], will be considered here.

The former represents the most recent and, until now, the most successful application of Morris SWADESH's glottochronology. Claire GRÉGOIRE and Bernard DE HALLEUX modified the wordlist, reducing it to 92 items by eliminating "bad" words ("bark", "swim", "feather", "round", "to fly") and replacing them with more suitable equivalents. At the same time, the majority of the 43 languages handled in their study were represented by even shorter lists, due to poverty of the available sources. The results of the calculation are represented in the form of a chart with the percentage of cognates for each pair of languages and a genealogical tree.

Such reticence might be a reaction to the too straightforward evaluations of the time distances during the "romantic period" of glottochronology and the subsequent scathing criticism of the method, which compelled many linguists to regard statistical data as relevant exclusively for relative, not for absolute dating of language divisions.

Unfortunately, the wordlist used in [GRÉGOIRE & DE HALLEUX 1994] for their lexicostatistics has not been published, which makes it impossible for other scholars to verify or update their results.

Claire GRÉGOIRE & Bernard DE HALLEUX's classification of Mande languages (for abbreviations see *Appendix 1*; the scale of this chart reflects the percent of cognates in the comparative list, and not time depth):

¹ The current study is a part of the project "Integral description of the Southern Mande languages: Dictionaries, grammars, glossed corpora of texts" supported by a grant of the Russian Foundation for Human Studies (№ 08-04-00144A).

² A detailed language map of the Mande family can be found at <http://www.sil.org/SILESR/2000/2000-003/silesr2000-003.htm>.

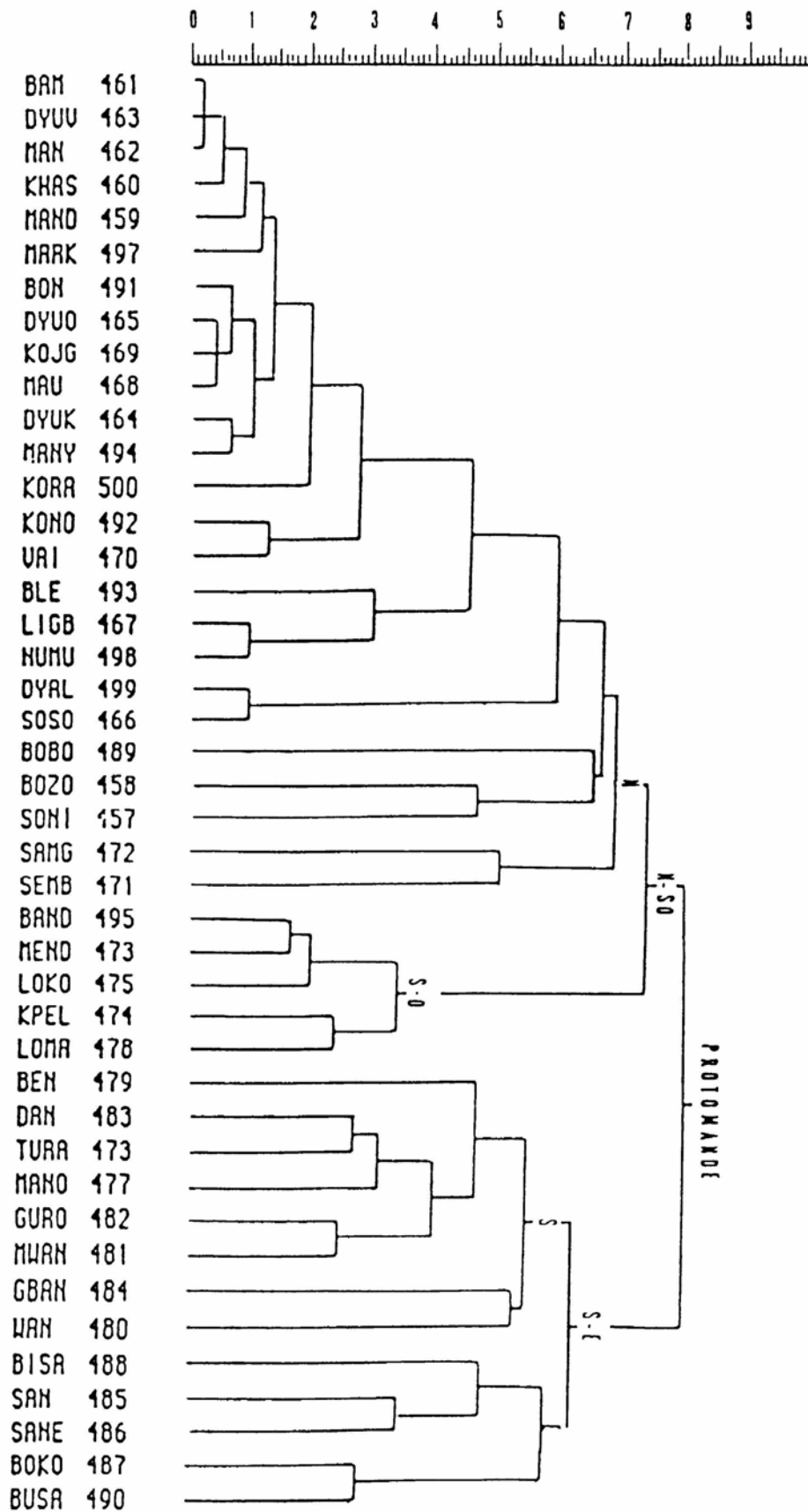


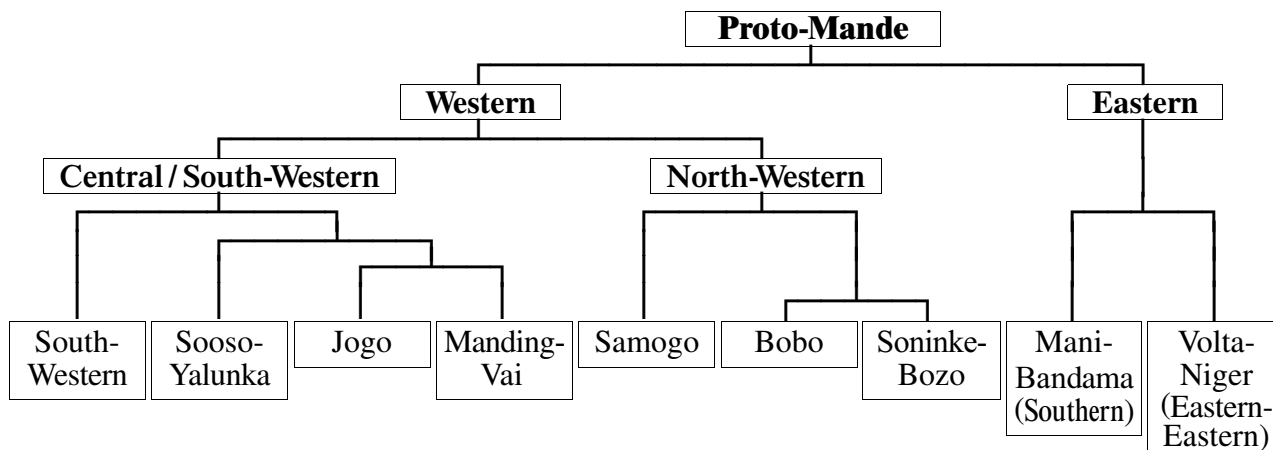
Tableau 1.- Classification des langues

Claire GRÉGOIRE & Bernard DE HALLEUX's lexicostatistical data (see the next page):

The work by Raimund KASTENHOLZ [1997] represents the next stage of the “anti-glottochronological reaction” in Mande comparative linguistics: he rejects glottochronology altogether, relying exclusively upon the method of shared lexical innovations. In his classification, KASTENHOLZ leaves aside the “Eastern” group (his only innovation in this sphere is a label change: “Eastern” for what was previously “South-Eastern”, “Mani-Bandama” for “Southern Mande” and “Volta-Niger” for the ex-“Eastern Mande”) and restricts his scope to Western Mande. The classification of the latter generally follows those of the earlier authors, but differs in a few minor points. His major innovations consist in grouping together:

- 1) Samogo, Bobo and Soninke + Bozo languages into a “North-Western group”, and
- 2) South-Western Mande and “Central Mande”, the latter including Soso, Jogo and “Great Manding” (Manding + Vai-Kono + Mokole groups).

Chart 1. Mande classification according to R. KASTENHOLZ



It is Raimund KASTENHOLZ’s classification that has been accepted by the compilers of one of the most popular sources on language classification, the *Ethnologue*, which makes it something of a current “standard”; its major outlines were also reproduced in MY OWN *Manding Dictionary* and in some other publications.

However, KASTENHOLZ recognizes that “...the described above method of segmentation, finally, is not based on an objective criterion. Therefore, a classification that follows cannot be considered as genetic in the narrow sense”. Being highly skeptical about lexicostatistics, KASTENHOLZ believes that verification should be performed by comparing the results of the method of shared lexical innovations with evidence provided by historical sound change [KASTENHOLZ 1997: 41]. In fact, KASTENHOLZ recognizes the shortcomings of his method: a lack of quantification combined with difficulties in distinguishing between true innovations and retention, which make it unverifiable. In my opinion, despite all the imperfections of lexicostatistics, it is still preferable to the method of shared lexical innovations.

Today, after Sergei STAROSTIN’s improvement of glottochronology³ has addressed numerous problems of the method, it is time to apply it again to the data of Mande languages. It should be mentioned that Mande languages are much better documented now than they were fifteen years ago, when GRÉGOIRE & DE HALLEUX’s paper was written; this is particularly relevant in relation to Southern Mande, Eastern Mande, Samogo and Mokole languages. Apart from this, certain progress has been made in the phonological and lexical reconstruction of some branches of Mande, especially South-Western and Southern, which is a very important prerequisite for the precision of glottochronological calculations. In my study, I use the data of 54 languages and dialects (for a complete list with sources, see *Appendix I*).

Before discussing the results of my calculation, some preliminary remarks are in order:

1) My calculation, unlike GRÉGOIRE & DE HALLEUX, is based on the standard 100-wordlist of SWAD-ESH. It is true that a “correction” of the wordlist by eliminating “bad” words may produce an impression of improvement when dealing with the data of one single language family, but it restricts the potential use of the wordlist in higher level comparative studies. It is well known that every language family has certain “bad words” of its own, and if we clear out all of them, we risk having nothing at all left to compare.

³ [S. STAROSTIN 1989/2000]; an electronic version of the paper is available at [http://www.nostratic.ru/books/\(140\)Starostin_Glottol.pdf](http://www.nostratic.ru/books/(140)Starostin_Glottol.pdf). For further discussion of STAROSTIN’s “improved glottochronology”, see [VASILYEV & MILITAREV 2008].

On the other hand, it turns out that even “very bad words”, such as “green” and “yellow”⁴, are sometimes not as bad as one could imagine. Thus, the word for “yellow”, although a compound (most often = “like the powder of the fruits of the locust tree”), is unexpectedly stable in Central Mande languages, in Bozo and Samogo. As for the argument that words for “bark” and “feather” are very often the same as for “skin” and “hair” respectively (for this reason, “bark” and “feather” were excluded from the list of GRÉGOIRE & DE HALLEUX), it is somewhat irrelevant, since lexicostatistics (and glottochronology) does not deal merely with roots, but takes into account the meanings of the words as well. So, if in language *A* two notions (for example, “bark” and “skin”) are expressed by one single word, while in language *B* these are two different words (one of which may be the same as the word in the language *A*), this fact is quite meaningful and does not hurt the principle of lexicostatistics.

2) I accepted an approach to compounds different from that of GRÉGOIRE & DE HALLEUX: they considered each component of the compound word separately, so that a pair of words received half a point when only one component was the same. In my calculation, I award such cognates a full point.

3) When compiling a standard comparative wordlist, one becomes aware of the fact that the degree of proximity of two languages (according to the method of glottochronology) may depend, to some extent, on the degree of elaboration of the sources⁵. A very simple example will suffice: let us imagine that we are comparing two dialects belonging to the Manding branch. In both cases, we have only 100-wordlists collected specially for lexicostatistical purposes, with no extra data available. In the list for the variety *A*, ‘small’ is rendered as *dógo*, and in the list for the variety *B*, in the same position we have *fitinin*. We have to treat those forms as unrelated, which adds about 200 years to the time depth. However, we know that in better documented Manding variants, both of these stems often coexist, and it is highly probable that when we get more data on the varieties *A* and *B*, we’ll have to mark this position with a “+”. This rule can be formulated as follows: the better the languages in question are documented, the closer they become⁶.

There is a countertendency which comes into the picture at a more advanced stage: a highly elaborated language description may “decrease” the degree of closeness of a certain language to others. Let us imagine that we have small (or even mid-size) dictionaries for Maninka and Bamana. For “egg”, we will find Maninka *kíli*, while in Bamana there are two synonyms, *fán* and *kíli*. According to the rules of STAROSTIN’s lexicostatistics, we should mark this position with a “+”. However, if a large-size Bamana dictionary is available, with a well elaborated system of stylistic and other markings, we then find out that *fán* is the main word, while *kíli* is only rarely used in this meaning, its principal meaning in standard Bamana being ‘testicle’. Therefore, we will have to mark “egg” as a “-” for the pair “Maninka — Bamana”.

4) STAROSTIN’s method implies serious preliminary etymological handling of the languages in question — unfortunately, there is no other way one can increase the reliability of the calculations. In the case of the Mande family, we are now in a situation where only certain branches (South-Western, Southern, Manding) are provided with phonological and lexical reconstructions based on reliable data⁷. But even there, reconstruction has primarily been carried out for the system of initial consonants, to a lesser extent — inlaut consonants, while vowels and tones remain almost unexplored.

As far as lexicostatistics is concerned, the above means that fairly often we have to work with resemblances rather than established cognates. In some cases, making the choice between including forms into a cognate series or splitting them into two (or more) series becomes a true puzzle. A good example can be made by quoting the forms for “bone”: are South-Western forms of the **kálénj* series (Mende *kálé*, Looma *kai(g)*, Liberian Kpelle *káo*, &c.) related to the series represented by Soso *xóri*, Jogo *yéli*, Mokole **kólo*, Manding **xólo*, Soninke *xótti*, Dzuun *xó*, Guro *wólé* &c.? In the present

⁴ Unfortunately, the SWADESH 100-word list had been created long before BERLIN and KAY’s book *Basic Color Terms, their Universality and Evolution* appeared (1969); otherwise, these two color names would hardly be there.

⁵ I am not speaking here of situations when certain words of the wordlist are missing altogether in the used sources; in such cases, we just shorten the comparative wordlist. Theoretically, shortening of the list should not influence the results of lexicostatistical calculation.

⁶ For a real example, see the discussion concerning the degree of divergence between Bozo languages/dialects: [SMELZER & SMELZER 1995] and [BLECKE & BLECKE 1997].

⁷ See, for South-Western Mande: [Dwyer 1974]; [Vydrine 1989]; [KASTENHOLZ 1997]; [Vydrin 2006a]; for Southern Mande: [Vydrin 2007]; [Vydrin 2006b]; for Manding: [CREISSELS 1979]; [GALTIER 1980]; [POZDNIakov & Vydrin 1986–1988]; [KASTENHOLZ 1997]; [CREISSELS 2004]. There is also a dissertation by Konstantin POZDNIakov [1978] attempting a reconstruction of the system of initial consonants for the entire Mande family, but this reconstruction is based mainly on outdated sources.

work I have given them as reflexes of two different roots, but I am by no means sure about it. Otherwise the genetic tree and the percentage chart given below are based on a “maximalist” list, where all doubtful cognates (resemblances) are put on the same footing as reliable cognates.

Another requirement of STAROSTIN’s glottochronology is that loanwords should not be counted, but tracked down and eliminated from the wordlist before we proceed to the calculations. In some cases, the borrowed nature of a word is quite evident. For example, Soso *kobole*, *kobele* and Jallonke *kóbólé* ‘bark’ are undoubtedly borrowed from Mokole languages (or the ancestor language of Mokole and Vai-Kono), where similar forms exist and can be easily etymologized, cf. Lele *k̀-gbolo* (with the referential article: *k̀-gbole*) ‘bark’, lit. “tree-skin”. There are also no doubts concerning the borrowed character of such forms as *wíre* ‘to fly’ or *yáage-nden* ‘big’ in Kakabe (cf. Pular *wii-ru-gol* ‘to fly’ and *yaaja* ‘be broad’). However, the situation is much less clear when we come to such Kakabe forms as *f̂* ‘all’ and *dúule* ‘cloud’: on one hand, we have Pular *fow* ‘all’⁸ and *duule-re* ‘cloud’ (since all Kakabe speakers are bilingual in Pular as the second language, such borrowings are obviously possible), but, on the other hand, we have Gban *f̂f̂* ‘all’, Jo *f̂* ‘all’, &c., and for ‘cloud’ we find such forms as Lele *dúye*, Koranko *dúru* ‘dew, fog’, &c., which imply a native Mande origin for the Kakabe forms. There are sometimes “hidden” influences which may distort the picture, such as retention of forms close to Maninka in Kakabe⁹.

One of KASTENHOLZ’s criticisms of glottochronology is that it is rather difficult to evaluate and compare the results achieved by different authors [KASTENHOLZ 1997: 38]. I believe that this flaw can be fixed by sticking as close as possible to a standard and unified methodology (the more we deviate from it, the less comparable our results become), and by publishing the complete wordlists used for calculations, so that everyone can check both the data and the procedure. That is why I provide here not only the genetic tree of the Mande family (see *Insert 1*) and the chart of the percentage of cognates for each pair of languages (see *Insert 2*), both generated by the STARLING software, but also the full 100-word list as well (see *Appendice 1*).

Looking at the resulting genealogical tree, we find some interesting differences from both KASTENHOLZ’s and GRÉGOIRE & DE HALLEUX’s classifications. Contrary to the former one, Bobo does not constitute one group with Soninke and Bozo; Soso and Jalonke constitute one large group with South-Western Mande, not with Manding-Jogo; and Mokole turns out to be closer to Vai-Kono than to Manding. Contrary to the latter classification, South-Western Mande languages do not constitute the very first split of the Western (“Northern/South-Western”) branch, since their separation follows that of numerous other groups; the internal arrangement of the Southern group is different, &c.¹⁰

On the other hand, when we confront the tree and the chart, some discrepancies between them also catch the eye. Below I list only a few of them:

a) Bobo is placed together with the Samogo group. This means that the Bobo — Southern Mande split took place sometime around 3100 BC (i. e., the period during which Proto-Mande split into Eastern and Western branches). However, in the chart, the separation dates of Bobo and individual languages of the Southern group are all in the range between 1900 and 2400 BC (i. e., more or less the same as the divergence between Bobo and South-Western Mande);

b) the same is true for the relations of the Southern group with South-Western Mande and Central Mande (Manding + Mokole + Jogo): here again, the chart gives us about the same range of dates for Southern — South-Western, Southern — Manding, Southern — Mokole, Southern — Jogo divisions (roughly, between 2000 and 2400 BC), while, according to the tree, the split should be dated back to 3100 BC.

⁸ Cf. also the form *fow* ‘all’ attested in the Jalonke dialect of Futa-Jallon described by Friederike LÜPKE [2005], most probably a Pular loan; in the Jalonke of Faleya we find *bírín*, cf. Soso *bírín*.

⁹ A “Maninka attraction” most probably explains an above “expectations degree” of Kakabe-Maninka cognate percentage in the 100-word list, while the respective degrees for Kakabe-Mogofin and Kakabe-Lele are lower. According to historical evidence, the Kakabe-Mogofin split must have taken place 250 to 300 years ago, whereas glottochronology gives us 780 years.

¹⁰ A special issue is the considerable discrepancy in the figures resulting from the calculation by GRÉGOIRE & DE HALLEUX and by myself. Here are some of them:

Pair of languages	GRÉGOIRE & DE HALLEUX	VYDRIN
Bamana — Kpelle	43	36
Bamana — Dan	43	35 to 38
Gban — Beng	43	57.

Unfortunately, the wordlist used by those authors has not been published, which makes impossible its verification.

I think that the explanation lies in the simplistic nature of the genealogical tree model, based on the “closest neighbor” method which reflects correctly only one type of language divergence, i. e. a division of a language community followed by an immediate cessation of contacts between its offsprings. It is true that such cases happen from time to time (the migration of the ancestors of Mogofin to the West in the 18th century probably constitutes one example), but much more usual is a situation of gradual divergence, where contacts between descendants continue during a long period. Such contacts further the retention of common vocabulary in one of the languages or in both of them. There are also horizontal contacts among related languages, and, quite often, loans resulting from such contacts cannot be easily identified because of similar phonological and syllabic structures of the languages.

If we take the above factors into account, the increased cognate percentages between Southern Mande and the majority of the Western groups can be explained through the fact that the “Proto-Southern” dialects of the Proto-South-Eastern language might have remained in close contact with the Proto-Western language long after the Western—South-Eastern split of about 3100 BC (some facts show that those contacts could continue even after the Eastern—Southern split at about 1560 BC, see below), while the “Proto-Eastern” dialects severed their direct contacts with the Proto-Western language immediately after the split.

Within this context, an explanation can be found for certain particularly anomalous cases. E. g., Gban, Wan and Beng are not only the most eccentric languages of the Southern group (56 to 67% of cognates with the other languages of the group¹¹, while other pairs of languages oscillate mainly between 66 and 78%); they often reveal lower percentages of cognates with individual languages of other groups as well (Samogo, Bobo, Great Manding, Eastern and, to a lesser degree, Soso-Jalonke), if compared with the data of other Southern languages. However, their cognate percentages with the languages of the South-Western group and Soninke-Bozo are more or less the same as for the other Southern Mande languages. Within the framework of the “abrupt split” model, this situation remains unaccountable, but it can be rationally interpreted within the model of a “soft drift”: presumably, after their separation from Proto-Mande, the Proto-Southern language remained in contact with the proto-languages of Samogo, Bobo, Great Manding and Eastern groups; there were probably less intense contacts with the Soso-Jalonke ancestor language as well. Somewhere between 700 and 300 BC Gban, Wan and Beng split apart, and this division was rather abrupt¹², resulting in the end of all contacts with the abovementioned ancestor languages, while all the other Southern languages maintained those contacts even after their subsequent divergence from the Proto-Southern Mande language. As for the South-Western and Soninke-Bozo groups, their contacts with Southern Mande might have stopped before Gban, Wan and Beng split apart.

The increase in cognate percentage for the pair Seenku — Bobo (in comparison with other Samogo languages) can undoubtedly be explained through horizontal contacts: these are neighboring language communities. In fact, the relative closeness of Bobo with the entire Samogo group may result from horizontal links rather than genetic proximity (all Samogo languages are spoken in the vicinity of the Bobo-speaking area). The same explanation might be valid for the pair Bozo-Kelinga — Bamana.

A two-dimensional model advanced by Konstantin POZDNIakov for the Nostratic phylum [1993: 270–272] would represent the process of divergence in a more appropriate way. However, even if we retain the genetic tree model, horizontal links and “soft drift” should still be taken into account.

Another important objection to my automatically generated genetic tree can be formulated as follows. It turns out that the genetic tree reproduces the eternal problem of classification at the intermediate level: seemingly, the reliability of groups whose age exceeds the threshold of about 50% of cognates (i. e., the time depth of more than 3200 years, according to Starostin’s glottochronology) is inferior compared to others.¹³ It is true that Starling produces a tree which seems to solve the problem, but a question remains: how seriously should we take decisions taken by the computer program (e. g., splitting apart or lumping together Soninke + Bozo and Bobo + Samogo) if we know that imparting a status of cognates to extra two or three doubtful pairs of words (more or less arbitrary, at

¹¹ An increased percentage of cognates in the pairs Wan-Guro and Wan-Mwan can be easily explained by horizontal influences: today the majority of Wan are bilingual in Guro and/or in Mwan.

¹² It could be an early southward migration. In today’s Côte-d’Ivoire, Gban has a reputation of “the very first ethnic group of the country”; all other people are believed to come there after Gban.

¹³ One could assume that the explanation of this fact lies in our lack of a true phonological and lexical reconstruction of the Mande language family: the margins of error at that level are comparable with the time distances between different branchings. However, it seems that in much better explored language families (such as Indo-European) we have the same problem of intermediate grouping, which makes me think that this problem may be of a more general order.

the current stage of our knowledge) or adding another dialect of Dan or Looma to the database may result in a cardinal change of the entire configuration of the intermediate level of the classification?

An alternative solution, at the current stage, would be to put all the major groups numbering less than 50% cognates among them at the same level; in this case, all the instances in which the percentage of cognates is rising will be explained by horizontal contacts. I am not sure we should take such a radical path, but we should keep this option in mind.

* * *

After this survey of the questions related to the internal classification of the Mande family, let us now turn to the problem of the homeland. Several hypotheses were formulated so far by different authors:

1) William WELMERS [1971]: the homeland of Mande is in Northern Benin (where the languages Kyenga, Shanga, Bokobaru, Boko and Busa are spoken today). First, the ancestors of Bobo moved to the West, followed by the ancestors of the Northern + South-Western groups; finally, one by one, different groups of South-Eastern Mande left that area as well.

2) According to BIMSON [1976], the location of Proto-Mande was “between the Black-Volta and Upper Niger rivers, in the general vicinity of Bobo-Dioulasso more than 4.000 years ago”.

3) David DALBY [1971] cautiously expresses the opinion that the Mande homeland was probably located in present day Sierra Leone and Liberia.

4) Claire GRÉGOIRE [1994: 63], even more cautiously, subscribes to DALBY’s opinion, in a modified form: for this author, it should rather be the area of Northern Côte-d’Ivoire, Southern Guinea and adjacent regions of Sierra Leone and Liberia.

My opinion is that the answer to this problem lies in our being able to combine linguistic evidence with evidence from other branches of knowledge, such as archaeology and paleoclimatology.

Thus, according to glottochronology, the forking of Proto-Mande began some time during the second half of the 4th millennium BC, and its further splitting into the main groups took place before 1500 BC. What do other sources tell us about this period?

According to Robert VERNET [2004: 37–39], during the wet period between 5000 and 3000 BC Sahara was densely populated by hunters and cattle-breeders; there were also some agricultural areas in the region as well. At the end of the 4th millennium BC, an “arid crisis” took place in southern Sahara, resulting in a considerable shift of the population to the South. Roderick MCINTOSH [2004: 19] provides more precise dates of “dry incursions” in West Africa: 6400, 5500, 4800 and 4200 years ago.

Archaeological evidence tells us that during the second millennium BC, numerous human groups of Southern Sahara were engaged in hunting, agriculture (millet, sorghum, rice) and cattle breeding; in some areas also in fishing. Ferrous and non-ferrous metallurgy came to existence (most probably, independently of the Mediterranean world) in the area to the North-West of the lake Chad about 1500 BC¹⁴ [VERNET 2004: 38]; [PARIS & al. 1992] and gradually spread to the other areas. During the same period (second millennium BC), Sahel was still too wet to have a dense population, but there were already some agriculturalists and cattle-breeders along the rivers.

As for the forested areas to the South of Sahel, they were scarcely populated, their inhabitants did not know cattle breeding, and their Neolithic culture was different from the one in Southern Sahara [VERNET 2004: 39].

After an “arid crisis” of ca. 1000 BC, a drastic shift of population to the South took place. After that date, 18° of the Northern latitude became the northern limit of the settled population in West Sudan. Thus, the dense settlement of the Sahelian valleys dates back to the first millennium BC. Before the 3rd century BC, Southern Sahara was still relatively wet, so that cattle breeding by transhumance remained possible.

Following a harsh drop in precipitation between the 3rd century BC and 5th century AD, Southern Sahara was mainly abandoned by the Negroid population (except for some oases) to Berbers. The desert reaches the 18° latitude, and to the South of that line, we see flourishing ferrous metallurgy; in Southern Mauritania, copper is also produced.

¹⁴ That is, more or less simultaneously with the Hittites. Until recently, the beginning of ferrous metallurgy in Africa was reported to the middle of the first millennium BC. Archaeological discoveries in Sahel and Southern Sahara pushed its dating by 1000 back.

Between the 6th and 11th centuries AD, an increase in rainfall allowed the Sahelian population to go back to 20° or even 21°, establishing contacts with Berbers and Arabs [VERNET 2004: 39]. It is the time of Ancient Ghana and Tekrur; further on in time, we are in the epoch of the written (or rather “semi-written”) history of that region.

Let us try to put the climatological, archaeological, and linguistic evidence together in a chart.

Time	Climate	Archaeology	Linguistics
5000 to 3000 BC	Wet period; an arid episode about 3500 BC	Sahara is densely populated: hunting, cattle-breeding, some agriculture. Sahel is scarcely populated	3500–3000 BC: division of Proto-Mande into Western and South-Eastern, contacts continue.
3000 to 2000 BC	“Arid crisis” about 2800 and especially 2200 BC; otherwise, the rainfall is high.	Sahara is densely populated: hunting, cattle-breeding, agriculture. Sahel is scarcely populated (agriculture along river valleys).	About 2100 and afterwards: Division of Western Mande into Proto-Soninke-Bozo, Proto-Bobo, Proto-Samogo and Proto-Central.
2000 to 1000 BC	Rainfall is at the same level as before.	Sahara is densely populated: hunting, cattle-breeding, agriculture. Beginning of metallurgy ca. 1500 BC to the NW of the lake Chad. Sahel is scarcely populated (agriculture along river valleys).	1900–1800: Division of Proto-Western into Proto-Central (Gr. Manding + Jogo) and Proto-SW + Soso. 1600–1500 BC: split of SE into Southern and Eastern; Proto-Southern remains in contact with Proto-Western. 1500–1400: splitting of Jo from Proto-Samogo. 1400–1300: separation of Proto-Soso + Jalonke and Proto-SW. 1150–1100 BC: separation of Soninke and Bozo. 1100–1000 BC: split of Eastern Mande into 2 branches.
1000 BC to 0	An arid crisis ca. 1000 BC. A wet period till 300 BC. A sharp drop in the rainfall after 300 BC.	About 1000 BC: a drastic shift of the population to the South. Cattle breeding to the North of 18°, dense agriculture population in the Sahel valleys, spread of metallurgy. After 300 BC: Sahara is abandoned by the Negroid population.	900–800 BC: separation of Gr. Manding and Proto-Jogo-Jeri. Separation of Seenku from Proto-Samogo. 700–600 BC: Separation of Gban and Beng from Proto-SM, migration of Gban to the South. Division of Proto-Kyenga + Shanga from Proto-Boko-Busa. 600–200 BC: Further division of Eastern and Southern. 100 BC–0: Splitting of the Great Manding.
0 to 1000 AD	Arid period till ca. 500 AD. Humid period after 500 AD.	After 500 AD, settled population of Sahel occupies the territory up to 20°–21°. Contacts with Berbers and Arabs.	Splitting of the major groups of the Mande family into individual languages.

The two main questions concerning the homeland of the Mande (i. e., the localization of the speakers of the Proto-Mande language in the second half of the 4th millennium BC and before) can be formulated as follows:

1) The North — South axis: was it Southern Sahara, with a subsequent southward movement, or the river valleys of Sahel (or even the humid area further to the South)?

2) The East — West axis: was it East (NW Nigeria — Northern Benin — Western Niger), Center (Central-Eastern Mali — Western Burkina Faso — NW Ghana) or West (SE Mauritania? Eastern Guinea? Liberia — Sierra-Leone?).

Let us try to answer the first question¹⁵.

The Southern Hypothesis: If we assume that the Proto-Mande speakers lived in the South, we should identify them with the scarce population of the Sahel river valleys whose main source of living

¹⁵ Cf. George BOOKS’ ([1998] and earlier papers) analysis of the influence of climate on the ancient and medieval history of the Mande-speaking groups. However, his study uses the linguistic data to a lesser extent, and it is focused mainly on the two last millennia.

was hunting, as they were (at that time) hardly acquainted with agriculture. We do not know what climatic changes could have triggered the division of Proto-Mande into Eastern and Western Mande. Their further divisions (towards the end of the 3rd millennium BC) might have been provoked by the arrival of migrants from the North during the “arid incursions” (although there is no clear evidence about any considerable migrations from the North to areas South of 16°, where Mande people live today). When the culturally advanced agriculturalists and cattle breeders, probably acquainted with metal, fled the drought in the first millennium BC, they might have met numerous ethnic groups speaking different Mande languages in the Sahel.

According to this hypothesis, we should suppose that migrants from the North, despite their incontestable cultural, technical and numerical superiority, were somehow assimilated by the scarce indigenous Mande-speaking population of Sahel, who took over their technological achievements.

The Northern Hypothesis: Proto-Mande speakers lived in Southern Sahara. The Western/South-Eastern split was probably triggered by the arid episode of ca. 3500 BC, and the arid crisis of ca. 2200 BC brought forth the further divisions of Proto-Western Mande. The subsequent divergence of the proto-languages may be explained by the population growth (which is a natural corollary of the neolithic revolution). After 1000 BC, climatic changes pushed all the Mande-speaking groups to the South of 18°, where they easily assimilated and/or chased away the scarce, culturally and technologically backward population of the Sahel river valleys.

In any case, localization of the Proto-Mande speakers as far to the South as Sierra Leone and Liberia rainforest (suggested by David DALBY) or the adjacent area further to the East (Claire GRÉGOIRE) seems highly improbable: that would mean that during the entire period in question, Mande migrations would have to take a northward direction, completely opposite to those which are actually postulated by archaeologists and justified by the history of climatic changes.

Taking into account the fact that before the first millennium BC Sahel was a swampy land, difficult to traverse, infected with tsetse fly, we can hardly postulate long-range eastward or westward migrations prior to ca. 1000 or even 500 BC, unless they passed through the North. Therefore, the great expansion of Mande-speakers along the longitudinal axis is by itself a strong argument in favor of the Northern Hypothesis.

This means that the second question does not make much sense if we assume the Southern Hypothesis, and while discussing the Western, Central and Eastern Hypotheses of the Proto-Mande homeland, in all of these cases I assume that it was localized in Southern Sahara (16–18° or further to the North).

Considering the axis East — West, so far, I do not see any archaeological or climatological facts which, confronted with linguistic data, could serve as an argument for either of the two localizations¹⁶. One might suppose that the emergence of metallurgy in the vicinity of Lake Chad could provide such kind of facts. However, 15th century BC is a relatively belated date; around that moment, all the major subgroups of Mande had already diverged and, most probably, moved along the longitudinal axis toward their current areas.

However, there is an elementary argument which can be taken into account in this respect. If we sort out historically documented migrations of Mande-speaking peoples (i. e., dating back to the 11th century AD at the most), it turns out that languages belonging to all the 11 groups of the Mande family are located between 3° and 12° of Western longitude. This sector can be considered as the geographical “center of gravity” of the Mande family. It seems logical to suppose that the Proto-Mande homeland was somewhere within this sector, rather than somewhere else, unless we have very strong arguments for the contrary¹⁷.

My conclusion is: the Mande homeland at the second half of the 4th millennium BC was located in Southern Sahara, somewhere to the North of 16° or even 18° of Northern latitude and between 3° and 12° of Western longitude.

* * *

If the Proto-Mande homeland was located in southern Sahara, one can expect that terms related to agriculture and/or cattle breeding will be reconstructed for this protolanguage; on the other hand, it would hardly be possible to reconstruct terms for metallurgy. Let us try to explore the common Mande lexicon in this respect. My study is based on a comparative database which counts today a little less than 2200 roots that can presumably be reconstructed for Proto-Mande and/or its daughter protolanguages. So far, this database is not a true comparative etymological dictionary, as there is no underlying

¹⁶ I am not an archaeologist, and my competence in this area is limited; thus, I will be thankful if scholars that are more advanced in this field of knowledge inform me of the existence of such evidence.

¹⁷ Such an argument could, for example, constitute information that the eastern part of Sahara underwent desertification, which could have pushed people to the West. However, to my knowledge, no factors of this kind are attested for the second half of the 4th millennium BC.

phonological and lexical reconstruction of all the stages, and large quantities of available data have not yet been entered into it. Therefore, an absence of a word in my data may either mean its non-existence in the language(s), or merely reflect the fact that it has not been added yet. We should also keep in mind that available data for many Mande languages are still insufficient (although the situation has improved considerably compared to the state of 10 or 20 years ago). Therefore, this part of our study should be regarded as just a pilot survey, without any pretension to definitive conclusions.

One objective difficulty is that it is often hard to discern between true reflexes of the proto-forms and results of horizontal influences. The great majority of Mande languages lies in the area of Manding cultural influence¹⁸, and it is often not easy at all to decide whether the item in question is borrowed from Bamana or Jula or dates back to the protolanguage of the Mande family. Let us consider the following set of forms:

***well (water hole)**

Soso *kələn*

Jogo *kəlɔ́(n)*

P.-Manding **xələn*, Mandinka *kòloŋ*, Jaxanka *koloŋ*, Xasonka *xòloŋ*, Kagoro *kòlóló*, Maninka *kəlɔ́*, Bamana *kəlɔ́*, Marka-Dafin *kərɔ*

Samogo: Dzuun *kərən*

P.-Southern-Mande: Dan (Gwèèta) *klɔ́ŋ*, Dan (Kla) *klɔ́ŋ*, Dan (Blo) *klɔ́ŋ*, Tura *kəlɔ́ŋ*, Guro *kəlɔ́*, Mwan *klɔ́ŋ*, Wan *klɔ́ŋ*.

As we can see, this root is represented in 5 (out of 9) languages of the Southern Mande group — a good argument for its reconstruction on the Proto-SM level. It is also well attested in the Central group (Jogo, Manding), in Soso and in the Samogo group, which makes it a good candidate for reconstruction at the Proto-Mande level.

The problem is that ethnography of the Tura and Dan people tells us that traditionally they drew water from brooks and rivers, and where there were none, from shallow “water holes”. Well sinking began quite recently, under the influence of the Jula people (i. e., Manding), and this makes us think that the word *kəlɔ́ŋ* in Tura and *klɔ́ŋ* &c. in Dan might be a Manding loan. The same may be true for the other Southern Mande languages.

Here is another example:

***milk**

South-Western **nónɔ́*: Mende *nónɔ́* ‘milk (of cow, goat)’, Looma *nɔ́nɔ*

Vai-Mokole: Kono *nónɔ́*, Koranko *nónɔ́*, *nónɔ́* ‘milk (of animals)’

P.-Manding **nónɔ́*: Mandinka *nóno* ‘breast milk; sour milk’, Jaxanka *nono*, Xasonka *nóno*, Kagoro *nóno*, Maninka Bamana, Jula *nónɔ́* ‘milk (of an animal)’, Marka-Dafin *nɔ́nɔ*

Samogo: Dzuun *nónɔ́n*

Southern Mande: Dan (Blo, Kla) *nónɔ́*, Dan (Gwèètaa) *nónɔ́*, *nónɔ́*, Tura *nónɔ́*, *nónɔ́ŋ*, Guro *nónɔ́*, Mwan *nónɔ́*, *nrɔ́*, Wan *nónɔ́*, Beng *nónɔ́*.

Here again, the word is very well represented in both of the main branches of Mande, yet we know that speakers of Southern Mande languages apply it to the tinned milk bought in grocery stores. In fact, primarily they hardly ever consume milk “naturally”; the main proprietors of livestock in the area are the Fulbe.

How shall we interpret this situation? Option one is that the word was borrowed into SM from Manding. Option two is that, when the ancestors of the speakers of SM languages moved from the northern area to the forest, they abandoned cattle-breeding, but kept the word for milk. When Manding merchants brought milk as an item of their commerce, the old word was revived.

There are many similar cases for which we do not have sufficient cultural information, so our decisions will be inevitably flawed. Nevertheless, we can still try to survey the stock of roots for cultural items whose reflexes are represented in the main branches of the Mande family¹⁹. Below I will provide only a tentative reconstruction of the meaning of the item and a “pre-reconstruction” of the form (no phonological reconstruction of Proto-Mande has been performed so far!). Whenever possible, comments about the reliability of the reconstruction are given.

¹⁸ With the exception of some Eastern Mande languages, Bisa (Burkina Faso), Kyenga–Shanga and Busa-Boko-Bokobaru (northern Benin and NW Nigeria).

¹⁹ Unfortunately, I have not yet been able to analyze the data for Eastern Mande languages — another serious shortcoming of the current study.

Social life.

Central, Southern and South-Western: ‘age-mate’ **N-bó*, ‘ancestor’ **bénbá*, ‘(paternal?) aunt’ **téné*, ‘damn, curse’ **dángá*, ‘friend’ **N-bórà*, ‘grandmother’ **mámá*, ‘to greet’ **fǒ*, ‘initiation’ **kéné*, ‘law, interdiction’ **tǔ*, ‘to marry’ **fúdíú* (borrowed into SM?), ‘educate / obey’ **kóló*.

Manding — Southern Mande: ‘albino’ **gbàñ-bèlè*, ‘bachelor’ **gbálan*, ‘to bury’ **bìrì*, ‘dance’ **tan?* **dǎn?*, ‘deceive’ **nége*, ‘evil’ **jáxu*, ‘griot’ **jèlí* (Manding — Dan, Mano), ‘heritage’ **xéjé*, ‘laugh’ **jélé*, ‘means’ **fèeré*, ‘namesake’ **tóxo-ma* (SM: only Dan), ‘play’ **tóló*, ‘steal’ **kúmá* (?) (Jalonke, Bozo, SM), ‘truth’ **tùjá* (Manding — Dan), ‘war’ **Kèlè*.

South-Western Mande — Southern Mande: ‘respect’ **bèlè*, ‘slave’ **dūwá*.

Spiritual life.

Central, Southern and South-Western: ‘life / soul’ **ní*.

Manding — Southern Mande: ‘fetish’ **jò*, ‘totem’ **tèná* (Manding — Dan).

Trade, money relations.

Central, Southern and South-Western: ‘to borrow’ **dóndó*, ‘hundred’ **kémé*, ‘market’ **dóxo*, ‘price’ **sòngó*, ‘to sell / trade’ **fèeré*, ‘thousand’ **wúlú*, ‘wealth’ **nàafóló*.

Manding — Southern Mande: ‘money’ **?ódi* (Southern Mande < Manding?), ‘profit’ **tónó*, ‘thousand’ **wágá* / **wálá*, ‘trade’ **jàagó*, ‘twenty’ **mùgán*.

South-Western Mande — Southern Mande: ‘to count’ **dódoñ*.

Weaving.

Central, Southern and South-Western: ‘harness of loom’ **níirí*, ‘loom’ **xòlé*, ‘loincloth’ **N-pèrèñ*, ‘spindle’ **N-kèndán*, ‘cotton / thread’ **gèsá*.

Manding — Southern Mande: ‘indigo’ **gálá*, ‘sew’ **xálá* (Manding — Gban), ‘thread’ **gáari*, ‘wax’ **kàná*, ‘weave, plait’ **dán* / **tán*, ‘ribbon’ **jálá*.

South-Western Mande — Southern Mande: ‘sew’ **sólòñ* (?) .

Metal work.

Manding — Southern Mande: ‘chain’ **jòlòkó* (Southern Mande < Manding?), ‘copper’ **N-sìrà* (cf. San: *zan*), ‘gold’ **sánV* (both Southern Mande terms are from Manding?).

Fishing, hunting.

Central, Southern and South-Western: ‘hook’ **N-dóolèñ*, ‘hunter’ **dònsó*, fishing hook **kóñ* (Soso — Mogofin — Southern Mande).

Agriculture.

Central, Southern and South-Western: ‘millet’ **nǒ* (in South-Western Mande and Vai it means ‘maize, corn’, which indicates its borrowed nature from Maninka; also represented in some Kru languages: Bete, Godie in the meaning of ‘millet’); ‘rice’ **màló* (a very reliable root; cf. in Eastern Mande: San *mele*, Busa *múrē*, Bokobaru *mórēé*, *múrēé*, Boko *móé*); ‘cooked rice’ **báxáa*; ‘fonio’ **pòjdíñ* (a reliable root; also well attested in Atlantic languages); ‘granary’ **bòndó*, ‘onion’ **N-jàbá*.

Manding — Southern Mande: ‘field work / field’ (**xí*; Southern Mande — Dan and Tura only); ‘banana’ **bàràndá* (?), ‘cut, mow’ **ká*, ‘platform (to chase birds; for a hunter)’ **N-gbàlá*, ‘ripen’ **mó*, ‘shed (drier)’ **gbá*, ‘sow (to)’ **dá?* **tá?*, ‘sow (broadcast)’ **púyì*, ‘winnow’ **fíyè*, ‘cola nut’ **gÒrÓ*, ‘hibiscus’ **dǎ*, ‘pumpkin’ **gè*, ‘sweet potato’ **wísé*, ‘yam’ **kiù* (Soso, Manding, Dan).

South-Western Mande — Southern Mande: ‘hoe’ **kálì* (cf. in Eastern Mande: Lebir *kàá*), ‘oil palm nut’ **tówú*.

Cooking.

Central, Southern and South-Western: ‘paste’ **dèkE*, ‘spoon’ **mìntá*.

Manding — SM: ‘burnt part of food in a pot’ **sàná*, ‘grind’ **wúgù* (Bamana — Dan), ‘mortar’ **wúli*, ‘mush’ **tòó*, ‘pancake’ **ñòmí*, ‘powder’ **bún*.

Artefacts.

Central, Southern and South-Western: ‘bag’ **bòtót*, ‘balaphone’ **bálan* (probably borrowed from Manding into South-Western Mande and Southern Mande), ‘drum’: **bàdá*, ‘hat’ **N-kpàbélé*, ‘pot’ **dVGV*, ‘spoon’ **mìntá*.

Manding — Southern Mande: ‘basket’ **butu-ru* (? Manding + Dan only); ‘boat’ **gúlú*, ‘dug-out canoe’ **kóñ*, ‘bow (weapon)’ **xálá* (SM: only Wan), ‘broom’ **SELAN*, ‘basket’ (?) **gbòngbòñ*, ‘calabash’ **fě* (SM: only Gban), ‘cowrie’ **kòolón*, ‘whistle’ **fílé* (Bamana + Gban), ‘lute’ **N-kòní*

(Southern Mande < Manding?), ‘leatherbag’ **fóroko* (Manding — Tura), ‘mirror’ **dùṅalé* (SM < Manding?), ‘net’ **jò*, ‘jar’ **fágà* (Manding — Guro).

South-Western Mande — Southern Mande: ‘axe’ **jòofá*, ‘bottle’ **sàṅdè*, ‘mat’ **sábá*.

House, village.

Central, Southern and South-Western: door **kóṅ*, ‘shelter / hut’ **bùgú*, ‘fence’ **jásá* (Southern Mande < Manding?), ‘henhouse / sheep barn’ **kúlù*.

Manding — Southern Mande: ‘gatehouse’ **bùlón* (Southern Mande — only Beng), ‘house’ **kzè*, ‘village’ **só*, ‘well’ **kòlón*.

South-Western Mande — Southern Mande: ‘house / village’ **péré*.

Wild animals.

Central, Southern and South-Western: ‘baboon’ **n-gbòṅ*, ‘hippopotamus’ **màlé*, ‘leopard’ **kóli*, ‘lion’ **jàlá*, ‘squirrels’: **N-kèrén*, **N-gànsá*, ‘vulture’ **dùbá*.

Manding — Southern Mande: ‘eagle’ **bíbi* (< Manding?), ‘guinea fowl’ **Kàn* (not in Manding), ‘hare’ **káj* (SM — only Dan), Agama ‘lizard’ **báka* (Manding — Soninke — Eastern Mande), ‘monitor lizard’ **N-koòDo*, ‘porcupine’ **bàlá*, ‘sparrow / hawk / eagle’ **ségè*, ‘flying termite’ **N-bíLí*, ‘turtle’ **kùtá*, ‘wasp’ **N-dòndòlí*.

South-Western Mande — Southern Mande: ‘crocodile’ **fàalé*, **zúlù*.

Domestic animals.

Central, Southern and South-Western: ‘horse’ **sòó* (certainly a borrowing in Southern Mande and Eastern Mande), ‘camel’ **nógóma* (a borrowing, not later than the 1st millennium AD), ‘pig’ **bóyì*.

Manding — Southern Mande: ‘cat’ **N-jàkùamá*, ‘chicken’ **siisée* (Southern Mande: only Wan), **tàxó* (Soso, Jogo — Dan, Tura, Mano)

South-Western Mande — Southern Mande: ‘dog’ **gbírà*, ‘sheep’ **bála* (well attested in Kru and Atlantic languages; also in Eastern Mande).

We should also mention that Southern and Eastern Mande groups have the same root for “cow”: Eastern Mande — San *dii*, Boko *zū*, Busa and Bokobaru *zù*; Southern Mande: Dan *dū*, *dū*, Tura *dilì*, Mano *dii*, Guro *dilì*, Yaure *tirì*, Gban *dòdò*, *dòò*, Mwan *drī*, Wan *dī*, Beng *zǒ*.

Even if we sort out items spread out through borrowing (which is evident, for example, for ‘horse’ **sòó*, ‘camel’ **nógóma*, ‘sweet potato’ **wisé* and some others), we still find a rather advanced society, one that is well acquainted with agriculture (rice, fonio, probably millet and yams), living in villages with dogs, cats, chickens, small and possibly large cattle.

However hasty and imperfect this survey of cultural vocabulary, hypothetically reconstructed for Proto-Mande, may appear, it still seems to provide more evidence for a Southern Saharan rather than Sahelian homeland.

References

- AMS [Atlas mandé-sud] = HALAOUI & al. 1983.
 BAILLEUL 2000cCh. BAILLEUL. *Dictionnaire Bambara–Français*. Bamako : Donniya, 2000, 494 p.
 BIMSON 1978 — Kent David BIMSON. *Comparative Reconstruction of Proto-Northern-Western Mande*. Ph. D. Thesis, University of California, Los Angeles. 321 p.
 BLECKE & BLECKE 1997 — Thomas BLECKE, Bärbel BLECKE. Les parlers bozo : révision d’une enquête dialectologique // *Mandenkan*, 32; pp. 23–34 [<http://l1acan.vjf.cnrs.fr/PDF/Mandenkan32/32Blecke.pdf>].
 Boo 2003 — (Anne-Marie GIMÉNEZ). *Boo nen séwe san-fransi, fransi-san* [*Lexique san-français, français-san*]. Ouagadougou: SIL. 120 p.
 BRACONNIER *ms.* — Cassian BRACONNIER. *Lexique Jógó*. [Ms.]
 BROOKS 1998 — George E. BROOKS. Climate and History in West Africa // Graham CONNAH (ed.) / *Transformations in Africa: Essays on Africa’s later past*. London & Washington: Leicester Univ. Press, 1998; pp. 139–159.
 CARLSON 1993 — Robert CARLSON. A sketch of Jo: A Mande Language with a Feminine Pronoun // *Mandenkan*, 25. 109 p. [<http://l1acan.vjf.cnrs.fr/PDF/Mandenkan25/25carlson.pdf>].
 CREISSELS 1979 — Denis CREISSELS. Étude comparative de consonantisme de deux parlers Manding (Mandinka — Bambara) // *Bulletin de phonétique de Grenoble*, vol. VIII, pp. 99–155.
 CREISSELS 1982 — Denis CREISSELS. Document lexical Maukakan (Parler Manding du Maou) // *Publications du Centre de dialectologie africaine*, n° 1. Université de Grenoble, 1982.

- KASTENHOLZ 1987 — Raimund KASTENHOLZ. *Materialen zum Koranko. Afrikanistische Arbeitspapiere*, Sondernummer, 1987. 181 S.
- KASTENHOLZ 1997 — Raimund KASTENHOLZ. *Sprachgeschichte im West-Mande. Methoden und Rekonstruktionen*. Köln: Rüdiger Köppe Verlag, 281 S.
- KASTENHOLZ 2001 — Raimund KASTENHOLZ. Le jèri.kùò (langue mandé des Jéri) : répartition géographique et matériel lexical // *Mandenkan*, 37; pp. 49–88 [<http://llacan.vjf.cnrs.fr/PDF/Mandenkan37/37kastenholz.pdf>].
- KOELLE — Sigismund Wilhelm KOELLE. *Polyglotta Africana, or a Comparative Vocabulary of Nearly Three Hundred Words and Phrases, in More than One Hundred Distinct African Languages*. London, Church Missionary House. 188 p. 1854
- KONOSHENKO ms. — Maria KONOSHENKO. *Dictionnaire kpelle–français–anglais–russe*. [Ms.]
- KOITÉ-HERSCHEL 1981 — Ute KOITÉ-HERSCHEL. *Le xàsonga (Mali) : phonologie, morphosyntaxe, lexique xàsonga–français et textes*. Thèse de doctorat de 3^e cycle. Université de Grenoble, 1981, 265 + 106 p.
- KUZNETSOVA & KUZNETSOVA ms. — Natalia KUZNETSOVA, Olga KUZNETSOVA. *Dictionnaire gouro–français–anglais–russe*. [Ms.]
- LE BRIS & PROST 1981 — Pierre LE BRIS, André PROST. *Dictionnaire bobo–français, précédé d'une introduction grammaticale et suivi d'un Lexique français–bobo*. Paris, SELAF, 415 p.
- LE SAOUT 1973 — Joseph LE SAOUT. *Etude descriptive du Gban : phonétique et phonologie*. Paris : LACITO, 1973.
- LEGER 1975 — Jean LEGER. *Dictionnaire guéréz (kpele.wo)*. (Nzerékoré), 1975. 289 p.
- LEIDENFROST & MCKAY 2005 — Theodore E. LEIDENFROST & John S. MCKAY. *Kpelle–English Dictionary with a Grammar Sketch and English–Kpelle Finder List*. Liberia, 2005.
- LESSAU & KASTENHOLZ 1989 — Andreas LESSAU & Raimund KASTENHOLZ. An English–Kɔ̀nɔ̀ Comparative Word List // *Afrikanistische Arbeitspapiere* (Köln), 20, SS. 49–93.
- Lexique Susu-Français*. Rep. de Guinée : Service Alphabétisation et Education des Adultes, 88 p.
- LÜPKE, FRIEDRIKE. *A Grammar of Jalonke Argument Structure*. PhD Diss., Radboud Universiteit Nijmegen, 2005, 482 p.
- MAKEEVA ms. — Nadezda MAKEEVA. *Dictionnaire dan (kla)–français–anglais–russe*. [Ms.]
- MCINTOSH 2004 — Roderick MCINTOSH. Chasing Denkéjugu over the Mande Landscape: Making Sense of Prehistoric and Historic Climate Change // *Mande Studies*, 6; pp. 11–28.
- MOROZOVA ms. — Maria MOROZOVA. *Loko–English–Russian Dictionary*. [Ms.]
- NEAL et al. 1946 — Vern C. NEAL, William C. SINCLAIR, Bartholomew J. FINN, Marilyn COMPTON. *Beginning Mano. A Course for Speakers of English*. San Francisco. [Ms.]
- NIKITINA ms. — Tatiana NIKITINA. *Dictionnaire wan–français–anglais–russe*. [Ms.]
- PAPERNO ms. — Denis PAPERNO. *Dictionnaire beng–français–anglais–russe*. [Ms.]
- PARIS et al. 1992 — A. S. PARIS, G. QUÉCHON, J. F. SALIEGE. Les débuts de la métallurgie au Niger septentrional : Aïr, Azawagh, Ighazer, Termit // *Journal des Africanistes*, 62; pp. 55–68.
- PEREKHVALSKAYA ms. — Elena PEREKHVALSKAYA. *Dictionnaire mwan–français–anglais–russe*. [Ms.]
- PERSSON & PERSSON 1976 — Andrew PERSSON & Janet PERSSON. *Word List of Brawhani Ligbi (Numu)*, 1976, 26 p. [Ms.]
- POZDNIakov 1978 — К. И. Поздняков. *Языки манде: Сравнительно-исторический анализ*. Кандидатская диссертация. М.: Ин-т языкознания АН СССР, 1978, 149 + 210 с. [*Mande Languages: A historical comparative analysis*. PhD Thesis. Moscow: Institute of Linguistics, 1978, 149 + 210 p.]
- POZDNIakov 1980 — К. И. Поздняков. Языки манде (Результаты сравнительно-исторического анализа) // *Африканский этнографический сборник XII* [*Mande Languages (Results of a Historical Comparative Analysis)* // *Africana XII*]. Труды Института этнографии им. Н. Н. Миклухо-Маклая, Новая серия, т. 109. Ленинград: Наука, с. 173–180.
- POZDNIakov, 1993 — К. Поздняков. *Сравнительная грамматика атлантических языков*. М., «Наука», 375 с. [Konstantin POZDNIakov. *Comparative Grammar of the Atlantic Languages*. Moscow, “Nauka”, 1993, 375 p.]
- POZDNIakov & VYDRIN, 1986–1988 — Konstantin POZDNIakov & Valentin VYDRIN. Die Entwicklung des phonetischen Systems des Prämanden: Nasalität und Entwicklung des phonologischen Systems der Mandensprachen // *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung*, Bd. 39, H. 5, SS. 549–561; Rekonstruktion des phonologischen Systems des Prämanden // *ZPSK*, Bd. 41, H. 3, SS. 353–371.
- PROST 1971 — André PROST. *Eléments de sembla: phonologie, grammaire, lexique (Haute Volta: groupe mandé)*. (*Afrique et Langage*, Documents, 5). Lyon, 1971.

- RAIMBAULT 1923 — R. P. RAIMBAULT. *Dictionnaire français-soso et soso-français* (2^e ed.). Rome : Sodalité de St. Pierre CLAVER, 1923, 165 p.
- SADLER 1949/2006 — Wesley L. SADLER. A Complete Analysis of the Looma Language, Interior Liberia, West Africa // *Mandenkan*, 42, 2006, 109 p. [<http://llacan.vjf.cnrs.fr/PDF/Mandenkan42/sadler.pdf>].
- SANGARÉ 1984 — Aby SANGARÉ. *Dioula de Kong (Côte-d'Ivoire) : Phonologie, Grammaire, Lexique et Textes*. Vol. 1 – 3. Université de Grenoble. Thèse du doctorat de III^e cycle, 1984.
- SMELTZER & SMELTZER 1995 — Brad SMELTZER & Susan SMELTZER. Compte-rendu des enquêtes cartographique, lexicostatistique et sociolinguistique des parlers bozo // *Mandenkan*, 30, pp. 55 – 101 [<http://llacan.vjf.cnrs.fr/PDF/Mandenkan30/bozoRapport.pdf>].
- SMELTZER & SMELTZER 2001 — Brad SMELTZER & Susan SMELTZER. *Lexiques Soninké-Français-Anglais / Soninke-French-English Dictionary. Index Français-Soninké, English-Soninke*. Bamako. 234 p.
- STAROSTIN 1989/2000 — С. А. СТАРОСТИН. Сравнительно-историческое языкознание и лексико-статистика // С. А. СТАРОСТИН. *Труды по языкознанию*. М.: «Языки Славянской Культуры» [*Works on Linguistics*. Moscow: “Languages of Slavic Culture”], 1989; pp. 407–447. [English version: Comparative-historical linguistics and lexicostatistics // *Time Depth in Historical Linguistics* / Ed. by Colin RENFREW, April McMAHON & Larry TRASK. McDonald Institute for Archaeological Research, Cambridge, 2000; pp. 223–265.]
- TOURÉ 1994 — Aboubacar TOURÉ. *Éléments de description de la langue soso*. Thèse de Doctorat. Université STENDHAL — Grenoble III, 1994, 311 p.
- TOURÉ 2004 — Aboubacar TOURÉ. *Parlons soso. Langue et culture du peuple de la Guinée Maritime*. Paris : l'Harmattan, 2004, 208 p.
- TRAORE 1978 — Karim TRAORE. *Dafing-Deutsch Wörterbuch*. Maîtrise-Arbeit. Saarbrücken, 293 SS.
- TRAORÉ & Traoré 1998 — Ali TRAORÉ & Fabé TRAORÉ. *Lexique orthographique dzùùngoo-français*. Rédacteur: Paul SOLOMIAC. Ougadougou: SIL, 1998, 120 p.
- TRÖBS 1998 — Holger TRÖBS. *Funktionale Sprachbeschreibung des Jeli (West-Mande)*. (*Mande Languages and Linguistics*; Vol. 3). Köln: KÖPPE, 241 p.
- TSYURUPA ms. — Maria TSYURUPA. *Dictionnaire yaouré-français-anglais-russe*. [Ms.]
- TVEIT & DANSOKO 1993 — Harald TVEIT & Gagny DANSOKO. *Petit dictionnaire khashonké-français*. Ousoubidiagna, Mission Protestante Norvégienne, 1993, 105 p.
- VANHOUDT 1999 — Bettie VANHOUDT. Lexique bisa-français // *Mandenkan*, 34 [http://llacan.vjf.cnrs.fr/PDF/Mandenkan34/Vanhoudt_1999-bisa.pdf].
- VASILYEV & MILITAREV 2008 — М. Е. ВАСИЛЬЕВ, А. Ю. МИЛИТАРЁВ. Глоттохронология в сравнительно-историческом языкознании. Модели дивергенции языков // *Аспекты компаративистики*. [т.] 3 (= *Orientalia et Classica: Труды Института восточных культур и античности*. Вып. XIX. М.: Российский государственный гуманитарный университет); стр. 509–536 [Glottochronology in Comparative-Historical Linguistics and the Models of Linguistic Divergence // *Aspects of Comparative Linguistics*. [v.] 3 (= *Orientalia et Classica: Papers of the Institute of Oriental and Classical Studies*. Issue XIX. Moscow: RSUH Publishers); pp. 509–536].
- VERNET 2004 — Robert VERNET. Évolution du peuplement et glissement des isohyètes à la fin de la pré-histoire et au début de l'histoire en Afrique de l'ouest sahélienne // *Mande Studies* 6, 2004, pp. 29–48.
- VYDRINE 1987 — В. Ф. ВЫДРИН. *Язык лоома*. Москва: «Наука», 121 с. [Valentin VYDRIN. *The Looma language*. Moscow, “Nauka” Publishers, 121 p.].
- VYDRIN 1989 — Valentin VYDRINE. Reflection of the Nominal Classification in Manden and South-Western Mande Languages: the Connection Category // *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschungen*, Bd. 42, H. 1, 1989, SS. 91–101.
- VYDRIN 2006a — В. Ф. ВЫДРИН. К реконструкции фонологического типа и именной морфологии пра-манде // *Труды Института лингвистических исследований*. Т. 2, Ч. 2. СПб: «Наука»; с. 3–246 [Valentin VYDRIN. Toward the Reconstruction of the Phonological Type and the Noun Morphology of the Proto-Mande // *Acta Linguistica Petropolitana*. Vol. 2, Part 2. St. Petersburg: “Nauka”; pp. 3–246].
- VYDRIN 2006b — В. Ф. ВЫДРИН. Личные местоимения в южных языках манде // *Труды Института лингвистических исследований*. Т. 2, Ч. 2. СПб: «Наука»; стр. 327–413. [Valentin VYDRIN. Personal Pronouns in the Southern Mande Languages // *Acta Linguistica Petropolitana*. Vol. 2, Part 2. St. Petersburg: “Nauka”, 2006; pp. 3–246].
- VYDRIN 2007 — Valentin VYDRIN. South(ern) Mande Reconstruction: Initial Consonants // *Аспекты компаративистики*. Вып. 2 (= *Orientalia et classica XI: Труды Института восточных культур и античности*). М.: Издательство РГГУ; стр. 409–498.

- VYDRINA *ms.* — Alexandra VYDRINA. *Dictionnaire kakabé-français-anglais-russe*. [*Ms.*]
 WELMERS 1971 — William E. WELMERS. Niger-Congo — Mande // *Current Trends in Linguistics*, 7.
 The Hague: MOUTON, 1971, pp. 113–140.
 WELMERS, KANDAKAI 1974 — William E. WELMERS, C. K. KANDAKAI. *A Vai-English Dictionary*
 (*Preliminary Draft*). Monrovia: The Institute for Liberian Languages, 1974, 118 p.
 ZWERNEMANN 1996 — Jürgen ZWERNEMANN. Documents kpeego // *Cahiers Voltaïques / Gur Papers*
 1, 1996, pp. 147–164.

Appendix 1

List of the languages/dialects and sources²⁰

- Bamana, Bambara** [BAM] (Manding), [BAILLEUL 2000].
Bandi [BAND] (South-Western Mande), [GROSSMANN & al. 1991]; [HEYDORN 1940–1941], my personal data. *Loopholes*: bark, cloud, green, horn, liver, (not), swim.
Banka (Samogo [SAMG]), [HOCHSTETLER 1996]. *Loopholes*: all, ashes, big, die, fat, many.
Beng [BEN] (Southern Mande), [PAPERNO *ms.*]. *Loopholes*: that.
Bisa-Lebir [BISA] (Eastern Mande), [VANHOUDT 1999]. *Loopholes*: fly.
Bobo (Madaré) [BOBO] (Bobo), [LE BRIS & PROST 1981]. *Loopholes*: round, yellow.
Boko [] (Eastern Mande), [JONES 2004a]. *Loopholes*: (not).
Bokobaru (Eastern Mande), [JONES 2004b]. *Loopholes*: (not).
Bozo-Jenaama, -Soroga(a)ma, -Sorko [BOZO] (Soninke-Bozo): [CREISSELS 1992]; data received from Christiane LAUSCHITZKY. *Loopholes*: not, this.
Bozo-Kelinga, Xainyaxo [BOZO] (Soninke-Bozo): [CREISSELS 1992]. *Loopholes*: bark, green, not, seed, stand, that, this, yellow.
Bozo-Tigemaxo, -Tiéyaho [BOZO] (Soninke-Bozo): [CREISSELS 1992]; [Guide bozo 1982]; all the loopholes were filled by Thomas BLECKE.
Busa [] (Eastern Mande), [JONES 2004c]. *Loopholes*: (not).
Dafing, Marka (Dafing) [MARK] (Manding), [TRAORE 1978], occasionally [DIALLO MOHAMADOU 1988]. *Loopholes*: horn, many, round, swim.
Dan-Blo, Western Yacuba [DAN] (Southern Mande), [ERMAN *ms.*]. *Loopholes*: (not), swim.
Dan-Gwèetaa, Eastern Yacuba [DAN] (Southern Mande), personal data. *Loopholes*: (not), yellow.
Dan-Kla [DAN] (Southern Mande), [MAKEEVA *ms.*]. *Loopholes*: bite, (not), swim, who?
Duun (Samogo [SAMG]), [HOCHSTETLER 1996]. *Loopholes*: fat, fly, stand.
Dzuun (Samogo), [TRAORÉ & TRAORÉ 1998]. *Loopholes*: ashes, fat, stand.
Gban, Gagu [GBAN] (Southern Mande), [LE SAOUT 1973], AMS, some data collected by Alexandre ZHELTOV and Ekatherina KOTCHMAR. *Loopholes*: bark, full, good, green, (not), swim, that, this.
Guro, Kweni [GURO] (Southern Mande), [KUZNETSOVA & KUZNETSOVA *ms.*]. *Loopholes*: swim.
Jalonke, Yalunka [DYAL] (Soso-Jalonke), [CREISSELS 1988]. *Loopholes*: big, many, round, yellow.
Jeri (Jogo), [KASTENHOLZ 2001]; [TRÖBS 1998]. *Loopholes*: cloud, fat, green, swim, yellow.
Jo (Samogo [SAMG]), [CARLSON 1993]; [DJILLA, EENKHOORN & EENKHOORN-PILON 2004]. *Loopholes*: stand.
Jogo, Ligbi [LIGB] (Jogo), [BRACONNIER *ms.*]. *Loopholes*: all, bite, come, feather, full, swim, that, this
Jula-Kong [DYUK] (Manding), [SANGARÉ 1984].
Kakabe (Mokole), [VYDRINA *ms.*]. *Loopholes*: swim, yellow.
Kono [] (Vai-Kono), [LESSAU & KASTENHOLZ 1989] and some personal data.
Koranko, Kuranko [KORA] (Mokole), [KASTENHOLZ 1987].
Kpan (Samogo [SAMG]), [HOCHSTETLER 1996]. *Loopholes*: ashes, stand.
Kpeen (Samogo [SAMG]), [ZWERNEMANN 1996]. *Loopholes*: ashes, black, burn, die, dry, fat, fly, full, good, green, know, many, name, neck, new, not, round, seed, sleep, smoke, sand, swim, tail, that, this, what, who, yellow.
Kpelle-Guinea [KPEL] (South-Western Mande), [LEGER 1975, Konoshenko *ms.*]. *Loopholes*: (not).
Kpelle-Liberia [KPEL] (South-Western Mande), [LEIDENFROST & MCKAY 2005]. *Loopholes*: (not).
Kyenga (Eastern Mande), [JONES 1998]. *Loopholes*: green, lie, liver, (not).
Lele (Mokole), personal data.

²⁰ Information for each language includes: the group which it belongs to within the Mande family; the source of lexical data (with comments, if necessary); the list of loopholes, if any. For many languages, “not” is put in brackets, which means that this word is missing in the language (being incorporated into personal pronouns). The most current alternative names are also given (those abbreviated by [GRÉGOIRE & DE HALLEUX 1994] in brackets: []).

- Loko** [] (South-Western Mande), [MOROZOVA *ms.*]. *Loopholes*: bark, cloud, green, long, louse, many, not, round, yellow.
- Looma, Toma** [LOMA] (South-Western Mande), [Dwyer, Bodegie & Baque 1981]; [Sadler 1949/2006]; [Vydrin 1987]²¹. *Loopholes*: (not).
- Mandinka** [MAND] (Manding), [Creissels, Jatta & Jobarteh 1983].
- Maninka** [MAN – *Maninka of Kankan*] (Manding), [Kánte 1962/1992].
- Mano** [] (Southern Mande), [DeZeeuw & Kruah 1981]; [Neal & al. 1946]²². *Loopholes*: all, bite, claw, feather, fly, green, neck, swim.
- Mau, Mahou, Mauka** [MAU] (Manding), [Creissels 1982]. *Loopholes*: die, round, swim.
- Mende** [MEND] (South-Western Mande), [Innes 1969]. *Loopholes*: feather, liver, many, not, swim, yellow.
- Mogofin, Mixifore** (Mokole), personal data. *Loopholes*: cloud.
- Mwan, Mona** [MWAN] (Southern Mande), [Perkhvalskaya *ms.*]. *Loopholes*: yellow.
- Numu, Brawhani Ligbi** [NUMU] (Jogo), [Persson & Persson 1976]. *Loopholes*: not, see.
- San** [] (Eastern Mande), [Boo 2003]. *Loopholes*: fly, (not).
- Seenku, Seeku, Sembla** [SEMB] (Samogo [SAMG]), [Prost 1971]. *Loopholes*: cloud, fat, fly, green, I, louse, not, road, root, round, smoke, that, warm, yellow.
- Shanga** (Eastern Mande), [Jones 1998]. *Loopholes*: green, liver, (not).
- Soninke, Sarokole** [SONI] (Soninke-Bozo), [Smeltzer & Smeltzer 2001].
- Soso, Susu** [SOSO] (Soso-Jalonke), [*Lexique Susu-Français*]; [Rimbault 1923]; [Touré 1994]; [Touré 2004]. *Loopholes*: yellow.
- Tura, Wen** [TURA] (Southern Mande), [Idiatov *ms.*]. *Loopholes*: (not), swim.
- Vai** [] (Vai-Kono): [Welmers & Kandakai 1974]. *Loopholes*: round.
- Wan** [] (Southern Mande), [Nikitina *ms.*]. *Loopholes*: that, yellow.
- Xasonka, Khas(s)onké** [KHAS] (Manding), [Tveit & Dansoko 1993; Koité-Herschel 1981].
- Yaure** (Southern Mande), [Tsyurupa *ms.*], AMS. *Loopholes*: ashes, bark, green, heart, moon, not, small, swim, that, this, what, yellow.

Mande languages not in the list:

- Bisa-Barka** (Eastern Mande),
- Blé, Jalkuna(n)** [BLE] (Jogo),
- Bolon, Bon** [BON] (Manding),
- Bozo-Tièma-Cièwe** (Bozo-Soninke),
- Gbin** (Southern),

interethnic **Jula** of Côte-d'Ivoire [DYUV] (Manding),

Jula-Odienné [DYUO] (Manding),

Koyaga, Kojaga (Jula) [KOJG] (Manding),

Manya (Kan), Mandingo [MANY](Manding),

Sane languages [SANE] (“northern San”, probably 4 different languages — Eastern Mande).

For large language/dialect clusters, such as **Manding** or **Dan**, only sample variants were taken.

Appendix 2

100-wordlist for 54 Mande languages and dialects²³

all: Dan-G *pěpě* [1], *gbà* [15]; Dan-B *kpékpé* [11]; Dan-K *gbà* [15]; Tura *páí, páípáí* [1]; Guro *fáá* [1]; Yaure *pèrǒ* [11]; Gban *fǒfǒ* [1]; Mwan *kpè* [11]; Wan *tēñ* [22], *tú* [23]; Beng *kpà* [11]; Kpeen *sá* [18]; Kpan *sqq-pope* [1], *sqq-pope* [18]; Dzuun *sá-pé* [1], *sá-pé* [18]; Duun *bie* [17], *sa* [18]; Seenku *kpòkpò* [11]; Banka *nogo* [21]; Jo *fǒ* [19], *bàtàà* [20]; Bobo *pěpé* [1]; Bozo-J *sqq* [18]; Bozo-K *súnú* [18]; Bozo-T *syen* [18]; Soninke *sú* [18]; Mau *bé* [17]; Jula-K *byé* [17]; Dafing *bié* [17]; Bamana *bée* [17]; Maninka *bée* [17]; Xasonka *bée* [17]; Mandinka *bée* [17]; Koranko *bée* [17]; Lele *bée* [17]; Mogofin *bée* [17]; Kakabe *fǒ* (< Pular?) [–2]; Kono *gbée* [11]; Vai *gbí* [14]; Soso *bírín* [16]; Jalonke *bírín* [16]; Jeri *kpó, gbón* [15]; Numu *kpó* [15]; Mende *pépé* [1], *kpélé* [11], *gbí* [14]; Loko *kperé* [11]; Bandi *kpelee* [11]; Looma *pe-g* [1]; Kpelle-L *kpélee* [11], *kélee* [12]; Kpelle-G

²¹ I am thankful to Daria MISCHENKO for her assistance in filling in the loopholes in the Looma list.

²² I am thankful to Maria KHACHATURYAN for her assistance in filling in the loopholes in the Mano list.

²³ Numbers in brackets following each form are meant to indicate cognates: all cognates have one and the same number. As a rule, cognates appearing in different entries (e. g., Dan forms for “feather” and “hair”) have the same number; however, the latter rule may be not respected in some instances (which does not influence the results of lexicostatistical calculation). Borrowings are provided with negative numbers and are not taken into account in the calculation.

kélee [12]; Bokobaru *pingi* [1]; Boko *píi* [1]; Busa *pínki* [1]; Lebir *fáay* [2], *kár* (*kári*) [3], *wóo* [4]; San *lenlen* [5], *tumaa*, *tumu* [6], *yii* [7], *kuturu* [8]; Kyenga *suka* [9]; Shanga *wanni* [10].

ashes: Dan-G *yóó* [34]; Dan-B *yóó* [34]; Dan-K *yóó* [34]; Tura *yúè* [34]; Mano *yúbé* [34]; Guro *yíi*, *yíbi* [34]; Gban *fiè* [34]; Mwan *yiè* [34]; Wan *ūú* [34]; Beng *yépé* [34]; Seenku *búkàn* [29], *ɲógó* (?) [34]; Bobo *yībē* [34]; Bozo-J *sibo* [33]; Bozo-K *húlúndú* [32]; Bozo-T *xoo* [31]; Soninke *xama/e*, *xema/e* [30]; Mau *bùùlù* [29]; Jula-K *bùguri* [29]; Dafing *búká* [29]; Bamana *bùguri* [29]; Maninka *bùudi gbé*, *bùuri gbé* [29]; Xasonka *bùguti xoyi* [29]; Mandinka *bùutu*, *bùuti* [29]; Koranko *bùura* [29]; Lele *bùura*; *bùuya* [29]; Mogofin *bògotá* [29]; Kakabe *bùgutá* [29]; Kono *búa*, *buwa* [29]; Vai *bùá* [29]; Soso *te-xube* (“dust of fire”) [30]; Jalonke *qúbé* [30]; Jogo *bògó* [29]; Jeri *bùgu* [29]; Numu *bògo* [29]; Mende *ndùvú* [24]; Loko *ndùhú* [24]; Bandi *nduhu* [24]; Looma *zuvu* [24]; Kpelle-L *lūū* [24]; Kpelle-G *luyū* [24]; Bokobaru *túbu* [24]; Boko *túfu* [24]; Busa *túbu* [24]; Lebir *súu* [25]; San *ton* [26]; Kyenga *céci* [27], *kpaso* (SP) [28]; Shanga *kwasy* [28].

bark: Dan-G *kěě* [40]; Dan-B *kiā* [40]; Dan-K *k̄* [40]; Tura *kělē* [40]; Mano *kēlē* [40]; Guro *pālā* [37], *kālē* [40]; Mwan *klōō* [40]; Wan *klónj* [40]; Beng *kóh* [40]; Kpeen *gɔ̀-kaa* [41]; Kpan *gɔ̀-kaa* [41]; Dzuun *gɔ̀-kaa* [41]; Duun *ka'a* [41]; Seenku *kà* [41]; Banka *ka'a* [41]; Jo *kúlìj* [40]; Bobo *kiēbē*, *sing. kōbā* (?) [43]; Bozo-J *gopo* [43]; Bozo-T *gɔ̀lɔ* [40]; Soninke *fatama* [37], *korom/e*, *kurum/e* [40]; Mau *wòmā* [43]; Jula-K *fārā* [37]; Dafing *páa*, *párá* [37]; Bamana *fārā* [37], *ɲómó* [43]; Maninka *fādá* [37], *wònbo* [43]; Xasonka *fāta* [37], *ɲòmo* [43]; Mandinka *fāta* [37], *wònbo*, *hònbo* [43]; Koranko *gbólo* [40], *kɔ̀-ɔ̀nbe* [43]; Lele *kón-kpóle* [40]; Mogofin *kɔ̀-bole* [40], *ɔ̀nbo* [43]; Kakabe *ɔ̀nbe* [43]; Kono *wónbo* [43]; Vai *kóngó* [44]; Soso *kobebe*, *kobole* (< Mokole *kon-bole*) [-3]; Jalonke *kóbólé* (< Mokole *kon-bole*) [-3]; Jogo *sógó* [42]; Jeri *fòroj* [37]; Numu *sògò* [42]; Mende *kóló* [40]; Looma *kɔ̀lɔ-g* [40]; Kpelle-L *kālɔ* [40]; Kpelle-G *kɔ̀lɔ* [40], *kala* [41]; Bokobaru *tèe* [35]; Boko *tèe* [35]; Busa *tèkè* [35]; Lebir *kyí* [36]; San *ki* [36], *fɔ̀rɔ* [37]; Kyenga *gbálá* [38]; Shanga *báí* [39].

belly: Dan-G *gwí* [45], *gú* [47]; Dan-B *gú* [47]; Dan-K *gú* [47]; Tura *gwí* [45]; Mano *gí* [45]; Guro *bɔ̀* (?) [48]; Yaure *cì*, *cì* [45]; Gban *gɔ̀* [47]; Mwan *vī* [45]; Wan *blū* [48]; Beng *nɔ̀n* [46]; Kpeen *dó* [46]; Kpan *dɔ̀* [46]; Dzuun *dón*, *dúùn* [46]; Duun *do* [46]; Seenku *sɔ̀gɔ̀* [51]; Banka *donɔ̀* [46]; Jo *cìi* [45]; Bobo *wūrù* [48]; Bozo-J *nun* [46]; Bozo-K *nùjú* [46]; Bozo-T *nun* [46], *kusu* [47]; Soninke *noxo/θ*, *nuxu-* [46]; Mau *kɔ̀ɔ̀* [50]; Jula-K *kónó* [50]; Dafing *kòò* [50]; Bamana *kónɔ* [50]; Maninka *kónɔ* [50]; Xasonka *xóno* [50]; Mandinka *kóno* [50]; Koranko *búu* [49]; Lele *bú* [49]; Mogofin *bó* [49]; Kakabe *bú* [49]; Kono *bú* [49]; Vai *bú* [49]; Soso *fùrí* [48]; Jalonke *kwíi* [45], *fùrí* [48]; Jogo *kó(n)* [47], *fùrú* [48]; Jeri *kóh* [47]; Numu *kòn* [47], *fùrú* [48]; Mende *kó* [47]; Loko *ko* [47]; Bandi *ko* [47]; Looma *ko-g* [47]; Kpelle-L *kōo* [47]; Kpelle-G *ko* [47]; Bokobaru *gbéé* [45], *nɔ̀sɛ* [46]; Boko *gbéé* [45], *nɔ̀sɛ* [46]; Busa *gbere* [45], *nɛsɛ* [46]; Lebir *nó* [46]; San *nò* [46]; Kyenga *gbèlè* [45]; Shanga *gwèè* [45].

big: Dan-G *gbé* [69], *kpíi* [71]; Dan-B *và* (?) [60]; Dan-K *gbàñ* [64], *kpíi* [71]; Tura *gbéé* [69]; Mano *gbùò* [64]; Guro *bèñè* [69]; Yaure *dã* [70]; Gban *blé* [60]; Mwan *gbèñè*, *gbètè* [69]; Wan *bálè* [60], *gbètè* [69]; Beng *béé* (?) [69]; Kpeen *brá* [60]; Kpan *bla* [60]; Dzuun *blá* [60]; Duun *bio-ma* [60]; Seenku *bārè* [60]; Jo *gba-na* [68], *gbè* [69]; Bobo *gbágá* [64]; Bozo-J *banu*, *burɔ* ? [67]; Bozo-K *hùlò* [66]; Bozo-T *xolo* [66]; Soninke *xòor-é* [66]; Mau *bèlèbélé* [65], *bwónj* [67]; Jula-K *bèlebéle* [65], *bòn* [67]; Dafing *bó* [67]; Bamana *bèlebéle* [65], *bɔ̀* [67]; Maninka *bèlebéle* [65], *bɔ̀* [67]; Xasonka *bèlebele-ba* [65], *bùn* [67]; Mandinka *búnj* [67]; Koranko *bèlebele* [65], *kólo* [66], *bɔ̀* [67]; Lele *bèlebele* [65], *kólo* [66]; Mogofin *kólo* [66]; Kakabe *bèlebele* [65]; Kono *kóo* [66]; Vai *kòò* [66]; Soso *gbo* [64], *bèlebéle* [65]; Jogo *gbùu* [64]; Jeri *bála(n)* [60]; Numu *gbóón* [64]; Mende *mbàwà* [60]; Loko *mba* [60]; Bandi *wála-wála* [60]; Looma *bala* [60]; Kpelle-L *há-wāā* [60], *kéte* [61], *kpàyakpaya* [62], *wúlénj* [63]; Kpelle-G *kene* [61], *kpeakpea* [62]; Bokobaru *bitá* [53], *isi* [54]; Boko *zɔ̀* [52]; Busa *zɔ̀kɔ̀* [52]; Lebir *búrgú* [55]; San *dadara* [56], *gole* [57], *kaara* [58]; Kyenga *yàsì* [54]; Shanga *gwà* [59].

bird: Dan-G *mää* [72]; Dan-B *má* [72]; Dan-K *mää* [72]; Tura *màà* [72]; Mano *mòò* [72]; Guro *mīā-né* [72]; Yaure *lò-bà* [72]; Gban *sã* [74]; Mwan *māā* [72]; Wan *bòlè* [72]; Beng *sièh* [74]; Kpeen *kàn* [73]; Kpan *kaqzi* [73]; Dzuun *kèyn* [73]; Duun *ka* [73]; Seenku *kállè* [73]; Banka *kana* [73]; Jo *jànání* [73]; Bobo *yàlò* [73]; Bozo-J *konu* [73]; Bozo-K *hònòh-ó* [73]; Bozo-T *xóni* [73]; Soninke *yeliN* / *yèlìnhé* [73]; Mau *kɔ̀ɔ̀* [73]; Jula-K *kónó* [73]; Dafing *kwóni* [73]; Bamana *kónó* [73]; Maninka *kónó* [73]; Xasonka *xòno* [73]; Mandinka *kùnu* [73]; Koranko *kòndɛ* [73]; Lele *kònde* [73]; Mogofin *kòndén* [73]; Kakabe *kònde* [73]; Kono *kònde* [73]; Vai *kòndé* [73]; Soso *xóní* [73]; Jalonke *qón*, *qòn-díí* [73]; Jogo *kòndere* [73]; Jeri *kóndi* [73]; Numu *kòndèré* [73]; Mende *ɲóní* [73]; Loko *ɲòì*, *wèì* [73]; Bandi *wònií* [73]; Looma *wóni*, *ɲóni* [73]; Kpelle-L *ɲòní* [73]; Kpelle-G *weni*, *wóni* [73]; Bokobaru *baá* [72]; Boko *ba* [72]; Busa *ba* [72]; Lebir *bèem* [*bèèni*] [72]; San *bàan* [72]; Kyenga *ba* [72]; Shanga *baqani* [72].

bite: Dan-G *kú sɔ̀ɔ̀* ‘catch with teeth’ [77]; Dan-B *kū s̄ s̄ ká* ‘catch with teeth’ [77]; Tura *sɔ̀ dɔ̀* ‘put teeth’ [77]; Guro *súé dɔ̀* [77]; Yaure *šé-kà* [77]; Gban *kióú* [79]; Mwan *sɔ̀ kpá* [77]; Wan *wèjù* [78]; Beng *só-bí* [77]; Kpeen *kí* [75]; Kpan *xí-ma* [75]; Dzuun *cín* [75]; Duun *cyi* [75]; Seenku *kōn* ~ *kwōn* [75]; Banka *kyi* [75]; Jo *ɲíí* [75]; Bobo *ɲwí* [75]; Bozo-J *sí* (?) [75]; Bozo-K *xwi* [75]; Bozo-T *xoy* [75]; Soninke

xíjní [75]; Mau *kín* [75]; Jula-K *kín* [75]; Dafing *cì* [75]; Bamana *cí, kí* [75]; Maninka *kí* [75]; Xasonka *kín* [75]; Mandinka *kín* [75]; Koranko *kí* [75]; Lele *kin, cin* [75]; Mogofin *kí* [75]; Kakabe *kin* [75]; Kono *cí* [75]; Vai *kín* [75]; Soso *xín* [75]; Jalonke *qín* [75]; Jeri *jín, jén* [75]; Mende *ní* [75]; Loko *ní(η)* [75]; Bandi *ní* [75]; Looma *gí* [75]; Kpelle-L *jín* [75]; Kpelle-G *jín* [75]; Bokobaru *té* [75]; Boko *kí* [75]; Busa *só* [76]; Lebir *sò dá* “put teeth” [77]; San *sɔn dɔ* “stand teeth” [77]; Kyenga *sɔ ká* [77]; Shanga *yí* [75].

black: Dan-G *tí* [80]; Dan-B *tí* [80]; Dan-K *tí* [80]; Tura *tí* [80]; Mano *tí* [80]; Guro *tí* [80]; Yaure *tíí-dí* [80]; Gban *tíé* [80]; Mwan *tí* [80]; Wan *tí* [80]; Beng *tí* [80]; Kpan *φw̄eey* [80]; Dzuun *tsiù* [80]; Duun *siu-ma* [80]; Seenku *tíé* [80]; Banka *sii* [80]; Jo *mìn* (?) [84]; Bobo *dūngū* [85]; Bozo-J *pí* [83]; Bozo-K *finá* [83]; Bozo-T *pin* [83]; Soninke *finna, fina* (‘become black’) [83], *binné* [84]; Mau *fín* [83]; Jula-K *fin* [83]; Dafing *pí* [83]; Bamana *fí* [83]; Maninka *fí* [83]; Xasonka *fín* [83]; Mandinka *fín* [83]; Koranko *fí* [83]; Lele *fin-dè* [83]; Mogofin *fí* [83]; Kakabe *fin* [83]; Kono *fí* [83]; Vai *fín* [83]; Soso *fóró* [82]; Jalonke *fóórè* [82]; Jogo *gbùú* [81]; Jeri *gbɔgɔ* [81]; Numu *gbùu* [81]; Mende *téli* [80]; Loko *tei* [80]; Bandi *tei(η)* [80]; Looma *tei(g)* [80]; Kpelle-L *tēi* [80]; Kpelle-G *teyi* [80]; Bokobaru *si-a* [80]; Boko *sí-a* [80]; Busa *sí-ra* [80]; Lebir *tindà* [80]; San *tí* [80]; Kyenga *ci* [80]; Shanga *sii* [80].

blood: Dan-G *yɔ̀* [88]; Dan-B *yɔ̀* [88]; Dan-K *yɔ̀* [88]; Tura *yùé* [88]; Mano *bai* [90]; Guro *yēé* [88]; Yaure *yé* [88]; Gban *wɔ̀* [88]; Mwan *yēé* [88]; Wan *yémí* [88]; Beng *wàh* [88]; Kpeen *túmó / tùmù* [94]; Kpan *thij* [94]; Dzuun *tsiin* [94]; Duun *timi* [94]; Banka *mila* [95]; Jo *cùh* [94]; Bobo *tò* [94]; Bozo-J *jolo* [91]; Bozo-K *jèrì* [91]; Bozo-T *jɔlɔ* [91], *deu* [93]; Soninke *fàré* [92]; Mau *yèè* [91]; Jula-K *jòrì* [91]; Dafing *yúru* [91]; Bamana *jòlí* [91]; Maninka *jèlí* [91]; Xasonka *jélu* [91]; Mandinka *yèle* [91]; Koranko *yèle* [91]; Lele *jìlí* [91]; Mogofin *báse* [90]; Kakabe *jèli* [91]; Kono *wii* [91]; Vai *wòí* [91]; Soso *wùlí* [89]; Jalonke *wùlí* [89]; Jogo *jìní* [88]; Jeri *jìní* [88]; Numu *jìní* [88]; Mende *hámá* [88]; Loko *hàwà* [88]; Bandi *hàawɔ* [88]; Looma *hámá* [88]; Kpelle-L *hāā, hāmā* [88]; Kpelle-G *nama* [88]; Bokobaru *yàrú* [87]; Boko *yàú* [87]; Busa *yàrú* [87]; Lebir *mà* [86]; San *mà* [86]; Kyenga *ɣama* [88]; Shanga *yàma* [88].

bone: Dan-G *gā* [96]; Dan-B *gā* [96]; Dan-K *gà* [96]; Tura *wéé* [98]; Mano *wélé* [98]; Guro *wólé* [98]; Yaure *wélé* [98]; Gban *gbū* (?) [96], *wlí, wlí* [98]; Mwan *bélé* [99]; Wan *z̄h* [98]; Beng *wlé* [98]; Kpeen *kùwɔ* [98]; Kpan *xɔ* [98]; Dzuun *xó* [98]; Duun *kwɔ* [98]; Seenku *kɔ* [98]; Banka *kwɔq* [98]; Jo *wó* [98]; Bobo *wōnón* [98]; Bozo-J *kwɔ* [98]; Bozo-K *hwɔ* [98]; Bozo-T *xɔn* [98]; Soninke *xóttí/e* [98]; Mau *kóó* [98]; Jula-K *kóró* [98]; Dafing *kùrù* [98]; Bamana *kólo* [98]; Maninka *kólo* [98]; Xasonka *xúlu* [98]; Mandinka *kúlu* [98]; Koranko *kólo* [98]; Lele *kólo* [98]; Mogofin *kólo* [98]; Kakabe *kólo* [98]; Kono *kúu* [98]; Vai *kúú* [98]; Soso *xórí* [98]; Jalonke *qórí* [98]; Jogo *yélé* [98]; Jeri *yèlí* [98]; Numu *lígé, lígén* [98]; Mende *kálé* [97]; Loko *káé* [97]; Bandi *kàalé* [97]; Looma *kai(g)* [97]; Kpelle-L *káo* [97]; Kpelle-G *kɔw* [97]; Bokobaru *wá* [96]; Boko *wá* [96]; Busa *wá* [96]; Lebir *yàá* [96]; San *we* [96]; Kyenga *yèle* [96]; Shanga *yáwé* [96].

breast: Dan-G *yɔ̀* [100]; Dan-B *yɔ̀* [100]; Dan-K *yɔ̀* [100]; Tura *yɔ̀* [100]; Mano *yɔ̀* [100]; Guro *yɔ̀* [100]; Yaure *yɔ̀* [100]; Gban *yí* [100]; Mwan *yɔ̀* [100]; Wan *yɔ́h* [100]; Beng *jón* [100]; Kpeen *cá* [100]; Kpan *ce* [100]; Dzuun *cén* [100]; Duun *ce* [100]; Seenku *kyè* [100]; Banka *kɛna* [100]; Jo *néí* [100]; Bobo *nīngí* [100]; Bozo-J *kanu* (?) [100]; Bozo-K *húnɔ* [100]; Bozo-T *xen* [100]; Soninke *kòmbé* [103]; Mau *síh* [100]; Jula-K *sín* [100]; Dafing *sí* [100]; Bamana *sí* [100]; Maninka *sí* [100]; Xasonka *súnjii* [100]; Mandinka *súnju* [100]; Koranko *kí, sí* [100]; Lele *sùuse* [102]; Mogofin *síngè* [100]; Kakabe *síngi* [100]; Kono *súsu* [102]; Vai *sùsù* [102]; Soso *xíné* [100]; Jalonke *qínjé* [100]; Jogo *cén* [100]; Jeri *cíé, kíé* [100]; Numu *kyè* [100]; Mende *jíní* [100]; Loko *ní* [100]; Bandi *jìní* [100]; Looma *jìni, jini* [100]; Kpelle-L *jìni* [100]; Kpelle-G *jini* [100]; Bokobaru *nó* [100]; Boko *nó* [100]; Busa *nó* [100]; Lebir *nó* [100]; San *nɔ* [100]; Kyenga *yɔ̀* [100]; Shanga *diyófàà* [101].

burn: Dan-G *gǔ* [107]; Dan-B *gúx* [107]; Dan-K *gǔ* [107]; Tura *pàí-da* [104]; Mano *gélé* [107]; Guro *wāq* [105]; Yaure *té-bìlí* [104]; Gban *kè* [107]; Mwan *kpá-cíé* [104]; Wan *lígi, té-lò* [104]; Beng *hūān* [105]; Kpan *zuu-ma* [108]; Dzuun *dú* [108]; Duun *zu* [108]; Seenku *dyòn* [108]; Banka *ju* [108]; Jo *bilitúru* [117]; Bobo *h̄wènè, h̄wùnà* [107]; Bozo-J *boi* [116]; Bozo-K *bu* [116]; Bozo-T *kele* [107], *buo* [116]; Soninke *buyi / -ni, bi / biini, buya* [116]; Mau *máq̄h* [113], *swèèh* [114], *kín* [115]; Jula-K *jèní, jèní* [114]; Dafing *zéí* [114]; Bamana *méne, mána* [113], *jèní* [114]; Maninka *méne, ménε* [113], *jèní, jàní* [114]; Xasonka *jàni* [114]; Mandinka *mála* [113], *jàni* [114]; Koranko *minda* [113]; Lele *minda* [113]; Mogofin *bintá* [113]; Kakabe *bintan* [113]; Kono *minda* [113]; Vai *binda* [113]; Soso *gan* [112]; Jalonke *gàn* [112]; Jogo *dō(n)* [108]; Jeri *siran* [111]; Numu *kónóh* [110]; Mende *mó* [109]; Loko *mó(η)* [109]; Bandi *mò* [109]; Looma *gála* [107]; Kpelle-L *kéren* [107]; Kpelle-G *kéren* [107]; Bokobaru *té ky* [104]; Boko *té ky* [104], *kpásá* [106]; Busa *té ky* [104], *kpátá* [106]; Lebir *sé sa* [104]; San *te koa* [104]; Kyenga *cé kú* [104]; Shanga *le ky* [104].

claw, nail: Dan-G *sūll* [123]; Dan-B *sòò* [123]; Dan-K *sūř* [123]; Tura *sùè* [123]; Guro *sònè* [123]; Yaure *sɔ̀rè* (AMS) [123]; Gban *kēs̄j* (AMS) [126]; Mwan *srò* [123]; Wan *sròh* [123]; Beng *klōh* [119]; Kpeen *nàh-kàà* [124]; Kpan *w̄egj-kaa* [124]; Dzuun *jéyn-káá* [124]; Duun *sii-ma’a* [124]; Seenku *kpèn-lèn-ká* [124]; Banka *kyè’edika’a* [130]; Jo *gbèná-kul̄h* [124]; Bobo *kiébé, pl. kóbá* [129]; Bozo-J *kobε* [129]; Bozo-K *kòfò*

[129]; Bozo-T *kewen* [129]; Soninke *ségéné* [128]; Mau *swègèŋ* [123]; Jula-K *sòní* [123]; Dafing *swèj* [123]; Bamana *sòní*, *sòní* [123]; Maninka *sòrí* [123]; Xasonka *sòniŋ* [123]; Mandinka *ŋònsiŋ*, *ŋòriŋ*, *ŋòoriŋ*, *ŋòoliŋ* [126]; Koranko *suri* [123]; Lele *jànsí* [126]; Mogofin *jànsín* [126]; Kakabe *sòrin* [123]; Kono *yáansen* [126]; Vai *kènji* [126]; Soso *xálé* [124]; Jalonke *qàli* [124]; Jogo *bólodé-ŋàì* (“finger’s thorn”) [124]; Jeri *kàà-niŋ* [125]; Numu *ŋáín* [124]; Mende *ngé-ngáú* [124]; Loko *nge-ngau* [124]; Bandi *sai(ŋ)*, *sà-wóó(ŋ)* [123]; Looma *saj(g)* [123]; Kpelle-L *ŋàlèŋ* [124]; Kpelle-G *paniŋ* [124]; Bokobaru *ɣɔŋkəkəkə* [121]; Boko *ɣɔ-kəɔ* [121]; Busa *yárúkəkəkə* [121]; Lebir *ŋikóntim* [118]; San *gòon tón* [120]; Kyenga *ɣəkəkəkə* [121]; Shanga *hji* [122].

cloud: Dan-G *dúú* [141], *dā-kpòŋ* [143]; Dan-B *dā-kp̄* [143]; Dan-K *d̄* [141]; Tura *d̄d̄* [141]; Mano *lúélúé* [132]; Guro *lólú* [141]; Yaure *lólú-kwèí* [141]; Gban *l̄* [141]; Mwan *lā-flé-pú* [143]; Wan *nàŋ-gbàni-pú* [143]; Beng *écí* [142]; Kpeen *niàngà* [149]; Kpan *xuɔ* [145]; Dzuun *ŋáá-múnún* [148]; Duun *kabakaba* [145]; Banka *kava* [145]; Jo *lòŋ* [141]; Bobo *dàbà* (?) [145]; Bozo-J *kaba* [145]; Bozo-K *habaa-foluu* [145]; Bozo-T *tile* (?) [141]; Soninke *xaba*, *kaba* [145], *tabulle*, *tebulle* [146]; Mau *kàwà* [145]; Jula-K *kàwá* [145]; Dafing *kábá* [145], *sà* [147]; Bamana *bírinka*, *máranka* [140], *kàbá* [145], *múnta*, *múnta* [148]; Maninka *bídika*, *bírika*, *midika*, *midigba* [140], *kàbá* [145]; Xasonka *xàba* [145]; Mandinka *mìnaayí* (?) [140], (*sán*)-*kàba* [145]; Koranko *bánda-fj̄* (“black sky”) [144]; Lele *dúye* [141], *bánda-nɔɔ* (“sky dirt”) [144]; Kakabe *dúule* [141]; Kono *bánda gbé* (“sky-white”) [144]; Vai *bánda b̄ià'é* [144]; Soso *luxu* (?) [141]; Jalonke *tùnègwií* (“sky excrement”) [565]; Jogo *bírígá(n)* [140]; Numu *kyí* (“rain, cloud”) [139]; Mende *ngèlé-gata* “compound of the sky” [137]; Looma *tona-bi(g)* (“rain-heavy”?) [565]; Kpelle-L *yēlē-kəl̄ŋ kōlē* “the white of the sky” [137]; Kpelle-G *tulo* [565]; Bokobaru *luguí* (?) [141], *lòu* [132]; Boko *lèu* [132], *luabe-puana* [133]; Busa *luku* (?) [141], *lègy* [132]; Lebir *wàrkv* [134]; San *la-bara* [135]; Kyenga *búlú* [136]; Shanga *le-kísínà* [135].

cold: Dan-G *néně* [151], *sə̄ə̄* [154]; Dan-B *néné* [151], *sā̄* [154]; Dan-K *něně* [151], *š̄iàà* [154]; Tura *n̄n̄í* [151]; Mano *néné* [151], *dóóló* [153]; Guro *dùnù* [153]; Yaure *t̄r̄ɔ̄-d̄r̄ɔ̄* [153]; Gban *β̄j̄* [152]; Mwan *n̄n̄í* [151]; Wan *lèlè*, *lrè* [151], *d̄l̄* [153]; Beng *ngéně* [151]; Kpeen *m̄i* [151], *kiimá* [156]; Kpan *makai* [151]; Dzuun *mùnkééy* [151]; Duun *mij̄kɛ* [151]; Seenku *m̄in̄í* [151]; Banka *mixya* [151]; Jo *š̄ií* [156]; Bobo *lū* [159]; Bozo-J *muun* [158]; Bozo-K *mùnún* [158]; Bozo-T *munan* [158]; Soninke *mullé* [158]; Mau *ngéně* [151], *swǎǎ* [156]; Jula-K *néne* [151]; Dafing *nèè* [151], *swòɔ* [156]; Bamana *néne* [151], *súma* [156]; Maninka *néne* [151], *súma* [156]; Xasonka *néne* [151], *súma* [156]; Mandinka *néne* [151], *súmaa* [156]; Koranko *kíma* [156]; Lele *kyímaa* [156]; Mogofin *kíma-rɛ* [156]; Kakabe *cúma*, *kíma* [156]; Kono *címa* [156]; Vai *kímá* [156]; Soso *xinbeli* [157]; Jalonke *qínbèlí* [157]; Jogo *kómá* [156]; Jeri *kíma*, *kúma* [156]; Numu *kpòǎ* [156]; Mende *kólé* [150]; Loko *kɔɔle* [150]; Bandi *kólé* [150]; Looma *kole* [150]; Kpelle-L *kólé* [150]; Kpelle-G *kwεele* [150]; Bokobaru *ȳú* [151]; Boko *niē* [151]; Busa *ȳi-da* [151]; Lebir *ɲèèni* [151]; San *ney*, *nej*, *naŋ* [151]; Kyenga *nini* [151]; Shanga *yáyà* [155].

come: Dan-G *nū* [164]; Dan-B *nū* [164]; Dan-K *nù* [164]; Tura *nú* [164]; Mano *nū* [164]; Guro *lú* [164]; Yaure *lú* [164]; Gban *nú* [164]; Mwan *nū / nú* [164]; Wan *zò* [165]; Beng *nū-lè* [164]; Kpeen *dánà* (?) [168]; Kpan *naa-ma* [168]; Dzuun *nà* [168]; Duun *na-m̄i* [168]; Seenku *nà* [168]; Banka *na* [168]; Jo *nā* [168]; Bobo *nā* [168]; Bozo-J *be* [169]; Bozo-K *bye* [169]; Bozo-T *be* [169]; Soninke *r̄ii* [170]; Mau *n̄* [168]; Jula-K *nà* [168]; Dafing *ná* [168]; Bamana *nǎ* [168]; Maninka *nǎ* [168]; Xasonka *nà* [168]; Mandinka *nà* [168]; Koranko *nà* [168]; Lele *nà* [168]; Mogofin *nà* [168]; Kakabe *nà* [168]; Kono *nà* [168]; Vai *ná* [168]; Soso *fáá* [166]; Jalonke *fáá* [166]; Jeri *sá* [163]; Numu *yá* [167]; Mende *pá* [166]; Loko *pá* [166]; Bandi *pa* [166]; Looma *pa* [166]; Kpelle-L *pá* [166]; Kpelle-G *pa* [166]; Bokobaru *mó* [160]; Boko *mó* [160]; Busa *mó* [160], *su* [161]; Lebir *b̄r̄* [162], *zá* [163]; San *daa / die* [163]; Kyenga *nú* [164]; Shanga *nu* [164].

die: Dan-G *gā* [171]; Dan-B *gā* [171]; Dan-K *gà* [171]; Tura *gá* [171]; Mano *gà* [171]; Guro *gā* [171]; Yaure *kà* [171]; Gban *gà* [171]; Mwan *gā* [171]; Wan *ká* [171]; Beng *gā* [171]; Kpan *sa-ma* [171]; Dzuun *sán* [171]; Duun *voo* [176]; Seenku *kyi / kyie / kyiinè* [174]; Jo *bèení*, *bèenú* (?) [176]; Bobo *sir̄* [175]; Bozo-J *kale* [174]; Bozo-K *sá* [171]; Bozo-T *kale* [174]; Soninke *kàrà* [174]; Jula-K *sà* [171], *fàgá* [173]; Dafing *sá* [171], *páá* [173]; Bamana *sǎ* [171]; Maninka *sǎ* [171]; Xasonka *sà* [171]; Mandinka *sàa* [171], *fàa* [173]; Koranko *fàga*, *fàa* [173]; Lele *fàa* [173]; Mogofin *fàgá* [173]; Kakabe *fàgá* [173]; Kono *fàa*, *fà* [173]; Vai *fàá* [173]; Soso *faxa* [173]; Jalonke *báá* [172]; Jogo *kpá* [171]; Jeri *kpà* [171]; Numu *kpà* [171]; Mende *hǎ* [171]; Loko *hǎ* [171]; Bandi *sà* [171]; Looma *saa* [171]; Kpelle-L *sāā* [171]; Kpelle-G *ha* [171]; Bokobaru *ga* [171]; Boko *gá* [171]; Busa *ga* [171]; Lebir *gà* [171]; San *ga / giɛ* [171]; Kyenga *ga* [171]; Shanga *ga* [171].

dog: Dan-G *gb̄* [177]; Dan-B *gb̄* [177]; Dan-K *gb̄* [177]; Tura *gb̄* [177]; Mano *gb̄* [177]; Guro *b̄ēn̄* [177]; Yaure *pl̄* [177]; Gban *kp̄kp̄* [178]; Mwan *gb̄* [177]; Wan *gb̄* [177]; Beng *j̄ēŋ* [177]; Kpeen *bȳw* [177]; Kpan *bī* [177]; Dzuun *v̄* [177]; Duun *vie* [177]; Seenku *dyigè*, *dyigè* [177]; Banka *via* [177]; Jo *kf̄ū* (?) [181]; Bobo *gb̄ḡ* [177]; Bozo-J *koŋgo-ro* [182]; Bozo-K *mùnù* [181]; Bozo-T *guun* [181]; Soninke *wùllé* [181]; Mau *wù* [181]; Jula-K *wùrú* [181]; Dafing *wùrú* [181]; Bamana *wùlú* [181]; Maninka *wùlú* [181]; Xasonka *wùlu* [181]; Mandinka *wùlu* [181]; Koranko *wùlu* [181]; Lele *wùlu* [181]; Mogofin *wùlú* [181]; Kakabe *wùlu* [181]; Kono *wù* [181]; Vai *wù* [181]; Soso *bàré* [180]; Jalonke *bàré* [180]; Jogo

dùmu [226]; Kpan *dúj* [226]; Dzuun *dzuùn* [226]; Duun *dyy* [226]; Seenku *kyē* [221]; Banka *juna* [226]; Jo *nò* [221]; Bobo *wéré* [221]; Bozo-J *bye* [225]; Bozo-K *fàlá* [223]; Bozo-T *pala* [223]; Soninke *yellⁱ/e* [221], *xabaN* / *xábaané* [224]; Mau *kíí* [221]; Jula-K *kírí* [221]; Dafing *cù* [221]; Bamana *fán* [222]; Maninka *kíli* [221]; Xasonka *kíli* [221]; Mandinka *kíli* [221]; Koranko *kíli* [221]; Lele *káli* [221]; Mogofin *kéle* [221]; Kakabe *kíli* [221]; Kono *cíi* [221]; Vai *kéí* [221]; Soso *xele* [221]; Jalonke *qèlé* [221]; Jogo *yélé* [221]; Jeri *yèli* [221]; Numu *ligé, ligén* [221]; Mende *ngálú* [221]; Loko *ngáú* [221]; Bandi *ngàlù* [221]; Looma *káí, kái, yai, ηε* [221]; Kpelle-L *yáloη* [221]; Kpelle-G *janinη* [221]; Bokobaru *né-gbèrè, né-gbèè* [221]; Boko *gbē* [221]; Busa *né-gbìnì* [221]; Lebir *gyir* [gìrì] [221]; San *gìi* [221]; Kyenga *gbi* [221]; Shanga *gye* [221].

eye: Dan-G *yǎ* [229]; Dan-B *yá* [229]; Dan-K *yá* [229]; Tura *yǎ* [229]; Mano *yǎ* [229]; Guro *yié, ywē* [229]; Yaure *yùè* [229]; Gban *wǎ* [229]; Mwan *yrè* [229]; Wan *lé* [230]; Beng *yō-wlé* [229]; Kpeen *jààn* [229]; Kpan *jaa-zí* [229]; Dzuun *jàá* [229]; Duun *ja'a-di* [229]; Seenku *dyúni* [229]; Banka *ja-dina* [229]; Jo *nà-néí* [229]; Bobo *nónō* [229]; Bozo-J *nō* [229]; Bozo-K *náná* [229]; Bozo-T *na* [229]; Soninke *yáa-x^a/e* [229]; Mau *né* [229]; Jula-K *ná, né* [229]; Dafing *nà* [229]; Bamana *né* [229]; Maninka *ná* [229]; Xasonka *náa* [229]; Mandinka *náa* [229]; Koranko *yá* [229]; Lele *yáa / jáa* [229]; Mogofin *náa* [229]; Kakabe *náa* [229]; Kono *yá / já* [229]; Vai *já* [229]; Soso *yáá* [229]; Jalonke *jáá* [229]; Jogo *náa* [229]; Jeri *ná* [229]; Numu *háa-dé* [229]; Mende *ngá* [229]; Loko *ngà-ú* [229]; Bandi *ngàa-hù* [229]; Looma *gaa-zu* [229]; Kpelle-L *hé-í* [229]; Kpelle-G *ηε* [229]; Bokobaru *wéé* [227]; Boko *wé* [227]; Busa *wé* [227]; Lebir *mim* [mìni], *míyāa* [228]; San *yíi* [229]; Kyenga *wélé* [227]; Shanga *wíiyé* [227].

fat: Dan-G *yǎ* [232]; Dan-B *yǎ* [232]; Dan-K *yǎ* [232]; Tura *yǎ* [232]; Mano *yǎnǎ* [232]; Guro *yǎnǎ* [232]; Yaure *yrǎ* [232]; Gban *wǎ* [232]; Mwan *yrǎ* [232]; Wan *yèη* [232]; Beng *yrǎ* [232]; Kpan *nii* [232]; Jo *kéη* [234]; Bobo *ní* [232]; Bozo-J *cεε* [234]; Bozo-K *cé* [234]; Bozo-T *tye* (?) [234]; Soninke *tée* (?) [234]; Mau *céη* [234]; Jula-K *cén* [234]; Dafing *tùrù, tùu* [235]; Bamana *ké, ké* [234]; Maninka *ké, cé* [234], *tulu* [235]; Xasonka *kéη* [234]; Mandinka *kéη, céη* [234]; Koranko *ké* [234], *tulu* [235]; Lele *ké* [234]; Mogofin *ké* [234]; Kakabe *kèη-ε* [234]; Kono *túu* [235]; Vai *túú* [235]; Soso *ture* [235]; Jalonke *turé* [235]; Jogo *náa* [233], *cén* [234]; Numu *ná* [233], *kilé* [235]; Mende *ngúló* [232]; Loko *ngúó* [232]; Bandi *ngúló* [232]; Looma *kpóló* [232]; Kpelle-L *wúló* [232]; Kpelle-G *ulo* [232]; Bokobaru *nósi* [231]; Boko *nísi* [231]; Busa *nísi* [231]; Lebir *nóó* [232]; San *nó* [232]; Kyenga *yǎnǎ* [232]; Shanga *nǎ* [232].

feather: Dan-G *kāā* [328]; Dan-B *kāā* [328]; Dan-K *kāā* [328]; Tura *kúā* [328]; Guro *cēē* [328]; Yaure *jé, cē* [328]; Gban *mā* [237]; Mwan *cīē* [328]; Wan *kāη* [328]; Beng *céη* [328]; Kpeen *súlù* [243]; Kpan *sui* [243]; Dzuun *siú* [243]; Duun *suu* [243]; Seenku *tān* [242]; Banka *šulq* [243]; Jo *káq* [328]; Bobo *dēgē, dēgē* [238]; Bozo-J *tíin* [239]; Bozo-K *túnǎ* [239]; Bozo-T *tiin* [239]; Soninke *yinⁱ/e* [239]; Mau *yóó* [241]; Jula-K *šé, syé* [239]; Dafing *tí* [239]; Bamana *sí* [239], *jólo* [241]; Maninka *sí* [239], *jólo* [241]; Xasonka *tí* [239]; Mandinka *tii* [239]; Koranko *kòndε-tigi* [239]; Lele *yònbe* [240]; Mogofin *kòndε-téi* [239]; Kakabe *kònde tí* [239]; Kono *yònbo* [240]; Vai *kùn-dii* [239]; Soso *xabe* [328]; Jalonke *qábé* (?) [328]; Jeri *tí, cí* [239]; Numu *tín* [239]; Loko *deya* [238]; Bandi *ndeya* (KOELLE) [238]; Looma *déyá, dege* [238]; Kpelle-L *léya* [238]; Kpelle-G *leya* [238]; Bokobaru *ká* [328]; Boko *ká* [328]; Busa *ká* [328]; Lebir *ká* [328]; San *ka* [328]; Kyenga *káká* [328]; Shanga *kaka* [328].

fire: Dan-G *síř* [245], *pēη* [246]; Dan-B *sír* [245]; Dan-K *sé* [245]; Tura *síē* [245], *pàì* [246]; Mano *tíé* [244]; Guro *cé* [244]; Yaure *té* [244]; Gban *sé* [245]; Mwan *té* [244]; Wan *té* [244]; Beng *síé* [245]; Kpeen *sàà, sàà* [245]; Kpan *sa* [245]; Dzuun *sà* [245]; Duun *sa* [245]; Seenku *tā* [244]; Banka *sa* [245]; Jo *túu* (?) [248]; Bobo *tógó* [248]; Bozo-J *tau* [248]; Bozo-K *túbá* [248]; Bozo-T *tuu* [248]; Soninke *yinⁱ/e* [247]; Mau *tá* [244]; Jula-K *tá* [244]; Dafing *tà* [244]; Bamana *tá, tá-suma* [244]; Maninka *tá* [244]; Xasonka *tá* [244]; Mandinka *dinbaa, kimaá* [247]; Koranko *tá* [244]; Lele *táa* [244]; Mogofin *tá* [244]; Kakabe *tá* [244]; Kono *tá* [244]; Vai *tá* [244]; Soso *téé* [244]; Jalonke *téé* [244]; Jogo *tá* [244]; Jeri *tá* [244]; Numu *tá* [244]; Mende *ngònbú* [247]; Loko *ngòmbù* [247]; Bandi *ngònbù* [247]; Looma *hábu* [247]; Kpelle-L *ηōη* [247]; Kpelle-G *wonη* [247]; Bokobaru *té* [244]; Boko *té* [244]; Busa *té* [244]; Lebir *sé* [245]; San *te* [244]; Kyenga *cé* [244]; Shanga *cé* [244].

fish: Dan-G *yúřř* [250]; Dan-B *yúř, yúř* [250]; Dan-K *pǎ* [249]; Tura *vé* [251]; Mano *kpáq* [249]; Guro *pō* [249]; Yaure *pǎ* [249]; Gban *zǎ* [250]; Mwan *kāā* [249]; Wan *kpó* [249]; Beng *zinǎ* [250]; Kpeen *gí* [250]; Kpan *žee* [250]; Dzuun *jí* [250]; Duun *ži* [250]; Seenku *kwēē* (?) [249]; Banka *jidina* [250]; Jo *yérii* [250]; Bobo *zē* [250]; Bozo-J *yuō* [250]; Bozo-K *nóηó* [250]; Bozo-T *nuōn* [250]; Soninke *néxiⁱ/e* [250]; Mau *yéē, yǎ* [250]; Jula-K *yígén* [250]; Dafing *yié* [250]; Bamana *jége* [250]; Maninka *jée* [250]; Xasonka *nége* [250]; Mandinka *née, yée* [250]; Koranko *yége* [250]; Lele *née* [250]; Mogofin *yége* [250]; Kakabe *yége* [250]; Kono *née* [250]; Vai *nié* [250]; Soso *yéxé* [250]; Jalonke *jéqé* [250]; Jogo *yéyé(n)* [250]; Jeri *yége, yéga(n)* [250]; Numu *yígén* [250]; Mende *né* [250]; Loko *nié* [250]; Bandi *nè* [250]; Looma *feze(g), feza(g)* [253]; Kpelle-L *ηéε* [250]; Kpelle-G *hòηòlèè* [252]; Bokobaru *kpǎ* [249]; Boko *kpō* [249]; Busa *kpǎ* [249]; Lebir *zǎ* [250]; San *zò* [250]; Kyenga *zǎ* [250]; Shanga *zǎ* [250].

fly: Dan-G *wl̥* [267]; Dan-B *l̥* [267]; Dan-K *wl̥* [267]; Tura *walá* “fly up” [267]; Guro *sí* (AMS) [269]; Yaure *sú* [269]; Gban *tl̥* [268]; Mwan *sí* [269]; Wan *gl̥n̥* [267]; Beng *p̥el̥ōñ* [260]; Kpan *tsén* [266]; *fiu-ma* [272]; Dzuun *teruye*, *šɛ* [266]; Banka *fú* [272]; Jo *m̥ín* [271]; Bozo-J *piri-ti* [260]; Bozo-K *piri-ndi* [260]; Bozo-T *piri-ndi* [260]; Soninke *kanpi* [270]; Mau *t̥ín* [266]; Jula-K *wíri*, *wúri* [267]; Dafing *tú* [266]; Bamana *pá* [264]; Maninka *gbá* [264]; Xasonka *bá* (?) [264]; Mandinka *tii* [266]; Koranko *gbá* [264]; Lele *la-sò* [265]; Mogofin *tú* [266]; Kono *t̥ín* (? “jump”) [266]; Vai *p̥íi* [260]; Soso *tuban*, *tugan* [262]; Jalonke *tùgán* [262]; Jogo *fínín* [260]; Jeri *tàn* [262]; Numu *yèlò kánká* (“stand in the sky”) [261]; Mende *wùwù* [257], *kòwò* [259]; Loko *p̥iré* [260]; Bandi *kòvó* [259]; Looma *puwu(g)*, *powo(g)* [257], *gove* (?) [259]; Kpelle-L *kòōñ* [259]; Kpelle-G *koñ* [259]; Bokobaru *v̥já* [254]; Boko *vùá* [254]; Busa *v̥inà*, *v̥urá* [254]; Kyenga *kp̥ete* [256]; Shanga *kwàt̥e* [256].

foot: Dan-G *g̥è* [273]; Dan-B *g̥è* [273]; Dan-K *j̥è* [273]; Tura *g̥èè* [273]; Mano *gà* [273]; Guro *gàné* [273]; Yaure *c̥èi*, *k̥èi* [273]; Gban *gè* [273]; Mwan *g̥ā* [273]; Wan *c̥é* [273]; Beng *gā* [273]; Kpeen *sànddà* [273]; Kpan *sq̥a* [273]; Dzuun *sàn* [273]; Duun *sa-n̥á* [273]; Seenku *kyèn* [273]; Banka *sa-na'a* [273]; Jo *k̥àñ-dà* [273]; Bobo *kàn* [273]; Bozo-J *taba* [275]; Bozo-K *tá* [275]; Bozo-T *tó* [275]; Soninke *tàá* [275]; Mau *sèñ* [273]; Jula-K *sèn* [273]; Dafing *sé* [273]; Bamana *sě* [273]; Maninka *sě* [273]; Xasonka *s̥iñ* [273]; Mandinka *s̥iñ* [273]; Koranko *k̥i* [273]; Lele *c̥è* [273]; Mogofin *k̥è* [273]; Kakabe *kèn* [273]; Kono *c̥è* [273]; Vai *k̥èñ* [273]; Soso *san* [273]; Jalonke *sàn* [273]; Jogo *kpó* [274]; Jeri *kpó* [274]; Numu *kpó* [274]; Mende *kówó* [274]; Loko *k̥og̥* [274]; Bandi *kòwò* [274]; Looma *k̥og̥* [274]; Kpelle-L *kó* [274]; Kpelle-G *k̥og̥* [274]; Bokobaru *gbá* [273]; Boko *gbá* [273]; Busa *gbá* [273]; Lebir *gám* [gàni] [273]; San *gò* [273]; Kyenga *gbé* [273]; Shanga *gi* [273].

full (fill): Dan-G *pā* [276]; Dan-B *pā* [276]; Dan-K *pā* [276]; Tura *pá* [276]; Mano *pā* [276]; Guro *f̥ā* [276]; Yaure *fā* [276]; Mwan *pá* [276]; Wan *pá* [276]; Beng *pā* [276]; Kpan *t̥e-ma* [277]; Dzuun *tsyén* [277]; Duun *teo-ma* [277]; Seenku *f̥āg̥* [276]; Banka *t̥e-ma* [277]; Jo *f̥áá-m̥à* [276]; Bobo *pāgā* [276]; Bozo-J *pa* [276]; Bozo-K *fagi* [276]; Bozo-T *pay* [276]; Soninke *f̥ágá* [276]; Mau *fá* [276]; Jula-K *fá* [276]; Dafing *pà* [276]; Bamana *fā* [276]; Maninka *fá* [276]; Xasonka *fá* [276]; Mandinka *fāa* [276]; Koranko *fá* [276]; Lele *fá* [276]; Mogofin *fāa* [276]; Kakabe *fāa* [276]; Kono *fá* [276]; Vai *fá* [276]; Soso *ra-fe* [276]; Jalonke *féé* [276]; Jeri *f̥i* [276]; Numu *f̥éñ / f̥éñè* [276]; Mende *ndá-vé* [276]; Loko *f̥é(ñ)* [276]; Bandi *nda-he* [276]; Looma *f̥é(g)*, *da-f̥é(g)* [276]; Kpelle-L *f̥é* [276]; Kpelle-G *hve* [276]; Bokobaru *pá* [276]; Boko *pá* [276]; Busa *pá* [276]; Lebir *pà* [276]; San *p̥á / pi* [276]; Kyenga *fà* [276]; Shanga *dàk̥i* [278].

give: Dan-G *gbā* [279], *nū* [280]; Dan-B *gbā* [279], *nū* [280]; Dan-K *gbà* [279], *nù* [280]; Tura *gbá* [279], *nú* [280]; Mano *gbā* [279], *nō* [280]; Guro *nó* [280]; Yaure *nò* [280]; Gban *nò* [280]; Mwan *nū / ná* [280], *p̥ē* [282]; Wan *ké* [281]; Beng *gbā* [279]; Kpeen *tèná* [283]; Kpan *s̥ii-ma* [283]; Dzuun *kò* [281], *s̥in* [283]; Duun *t̥im̥i* [283]; Seenku *kyē* [281]; Banka *š̥i* [283]; Jo *wòr̥i* [286]; Bobo *p̥ēr̥ē* [282]; Bozo-J *do*, *dw̥ana* [283]; Bozo-K *ky* [281], *lò* [283]; Bozo-T *ko* [281], *l̥oña* [283]; Soninke *k̥iú* [281], *k̥in̥i* [284]; Mau *só* [281], *dí* [283]; Jula-K *són* [281], *dí* [283]; Dafing *sò* [281], *dì* [283]; Bamana *só* [281], *dí* [283]; Maninka *só* [281], *dí* [283]; Xasonka *só*, *s̥oñ* [281], *dí* [283]; Mandinka *só* [281], *dii* [283]; Koranko *kó* [281], *dí* [283]; Lele *kó* [281], *dí* [283]; Mogofin *dí* [283]; Kakabe *kó* [281], *dí* [283]; Kono *kó* [281], *bè* [285]; Vai *kó* [281], *bee* [285]; Soso *f̥ii* [282]; Jalonke *f̥ii* [282]; Jogo *kó* [281]; Jeri *ku* [281]; Numu *kó* [281]; Mende *kó* [281], *f̥è* [282]; Loko *f̥è* [282]; Bandi *k̥o(ñ)* [281], *fe* [282]; Looma *kó* [281], *fe* [282]; Kpelle-L *kó* [281], *f̥è* [282]; Kpelle-G *k̥oñ*, *kò* [281], *hvə* [282]; Bokobaru *kpá* [279]; Boko *kpá* [279]; Busa *kpá* [279]; Lebir *kà* [279]; San *kó / koe*, *g̥o / goe* [281]; Kyenga *gba* [279]; Shanga *ky* [281].

good: Dan-G *dī* [296], *t̥i̥āādē* [301], *sā* [306]; Dan-B *dī* [296], *sā* [306]; Dan-K *t̥i̥āā* [301]; Tura *s̥ēgi* [306]; Mano *yīē* [295], *sè* [306]; Guro *zímá* [300]; Yaure *yí* [295]; Mwan *lr̥è* [292]; Wan *zr̥ā* [298], *kpl̥ā-kpl̥ā* [299]; Beng *tè* [296], *g̥ēñ* [297]; Kpan *zii* [296], *wo* [305]; Dzuun *dz̥i* [296], *wù* [305]; Duun *wo-ma* [305]; Seenku *dz̥* [296], *s̥ùg̥* [306]; Banka *wo* [305]; Jo *d̥i* [296], *š̥òr̥iñ* (?) [303]; Bobo *d̥ià* [296], *f̥r̥ō* [304]; Bozo-J *m̥ē* [287], *d̥e* [296]; Bozo-K *mai / moño* [287], *le* [296]; Bozo-T *main* [287]; Soninke *l̥iññé* [296], *s̥ir̥i/e* [303]; Mau *n̥iñ* [295], *dí* [296], *béé*, *béébéé* [302]; Jula-K *n̥i* [295], *dí* [296], *bébébé* [302]; Dafing *n̥i* [295]; Bamana *n̥i*, *n̥i* [295], *dí* [296], *béere*, *béere* [302]; Maninka *n̥i* [295], *dí* [296], *bédé*, *bèré* [302]; Xasonka *n̥iñ* [295], *dí* [296], *béte* [302]; Mandinka *n̥iñ* [295], *dí-yaa* [296], *béte* [302]; Koranko *n̥i* [295], *dí* [296]; Lele *n̥in*, *n̥uma* [295], *dí* [296]; Mogofin *n̥i*, *k̥áñi* [295]; Kakabe *n̥uma* [295]; Kono *n̥i* [295], *dí* [296]; Vai *n̥i* [295], *bé'è* [302]; Soso *f̥an* [294]; Jalonke *f̥án* [294]; Jogo *n̥èa(n)* [295]; Jeri *n̥an*, *n̥èan* [295]; Numu *n̥èanè* [295]; Mende *kp̥ékp̥é* [293], *p̥àndà* [294], *n̥àndé* [295]; Loko *kp̥ékp̥é(ñ)* [293], *n̥àndé(ñ)* [295]; Bandi *panda* [294]; Looma *pa(g)* [294]; Kpelle-L *l̥éle* [292]; Kpelle-G *lele* [292]; Bokobaru *maa* [287], *na* [288]; Boko *maa* [287], *na* [288]; Busa *mana* [287], *na* [288]; Lebir *m̥iññv* [289], *w̥s̥i* [290]; San *k̥oñ* [291], *s̥iñ* [306]; Kyenga *n̥o* [288]; Shanga *naa* [288].

green: Dan-G *gbl̥èy̥idē* [315]; Dan-B *gbl̥èb̥iz̥l̥*, *bh̥ud̥éy̥iz̥l̥*, *bèd̥éy̥iz̥l̥* [315]; Dan-K *fl̥è* [314]; Tura *kp̥òlò* [318]; Mano *béil̥éy̥ize* [315]; Guro *t̥i̥ūl̥ū* [313], *d̥òñò* (?) [324]; Mwan *yr̥èy̥r̥è* [312]; Wan *g̥in̥ōñ*

[311], *yri-yri* [312], *kpōnūŋ* [318]; Beng *gbé* [310]; Kpan *vucirama* [323]; Dzuun *vutsirama* [323]; Duun *bɔfwema* [323]; Banka *vuplanama* [323]; Jo *dàràràŋ* [324]; Bobo *sàrà-zū* (“like juice of the leaves of beans”) [326]; Bozo-J *sɔgu kaa* (“grass-fresh”) [327]; Bozo-T *sáúxáámú* [327]; Soninke *xallí/e* [320]; Mau *gbèèŋ* (“green fruit”) [318]; Jula-K *bín-gédédámán* [323]; Dafing *gwlé* (“unripe”) [318], *bùcèŋmà* [323]; Bamana *gèrén* (“unripe”) [318], *bínkenelama* [323]; Maninka *gbèdén* (“unripe”) [318], *bínkendelama* [323]; Xasonka *kéren* [318], *bínkerekɛŋ* [323]; Mandinka *jànbakere* [319]; Koranko *yànbaku-ye* (“like leaves”) [319]; Lele *yànbalama* [319]; Mogofin *yànbakutannama*, *bùlulama* [319]; Kakabe *kùtaŋ-ε* [321]; Kono *yanbakua* [319]; Vai *jàmbà-kùà* [319], *jìŋ* [322]; Soso *xinde* (“unripe”) [320]; Jalonke *qíndè* (“unripe”) [320]; Jogo *jànkómà* (“fresh leaf”) [319]; Numu *gyáŋkpòà* (“fresh leaf”) [319]; Mende *kpólé* [318]; Looma *kpoe* [318]; Kpelle-L *kpólé* (“unripe”) [318]; Kpelle-G *kpele* (“unripe”) [318]; Bokobaru *lákí* [307], *yísí* [308]; Boko *yísí* [308]; Busa *lákí* [307], *yísí* [308]; Lebir *vannu mi* [309], *zár mi* [316]; San *burun* [317].

hair: Dan-G *kāā* [328], *wū* [329]; Dan-B *kāā* [328], *wū* [329]; Dan-K *kāā* [328], *yǎǎ* [330]; Tura *kúá* (AMS) [328], *wə-wə* [329]; Mano *wīī*, {*wīī*-} *wū* [329]; Guro *cēē* [328], *wī*, *wūī* [329]; Yaure *jě*, *cě* [328]; Gban *mlě*; *mǎ* [329]; Mwan *cíē* [328], *wī* / *wí* [329]; Wan *kāŋ* [328]; Beng *céŋ* [328]; Kpeen *kùŋwànzúlù* [333]; Kpan *kɔɔ* [333]; Dzuun *kùŋgwèyn* [333]; Duun *kùŋkwei* [333]; Seenku *kònsùgí* [332]; Banka *kwá'q-sulu* [333]; Jo *ŋúžŋ* [332]; Bobo *ŋwòn-sógó* [332]; Bozo-J *tíin* [332]; Bozo-K *túnò* (“body hair”) [332]; Bozo-T *tiin* [332]; Soninke *yìn-tí* [332]; Mau *kùŋžyé* [332]; Jula-K *syé*, *šé* [332]; Dafing *tí* [332]; Bamana *sí* [332]; Maninka *sí* [332]; Xasonka *tí* [332]; Mandinka *tíi* [332]; Koranko *síi* [332]; Lele *kùn-tii* [332]; Mogofin *kùn-sigi* [332]; Kakabe *kùn-sigi* [332]; Kono *kùndii* [332]; Vai *kùn-dii* [332]; Soso *xun-sexé* [332]; Jalonke *qùn-séqè* [332]; Jogo *tín* [332]; Jeri *tí*, *cí* [332]; Numu *tigi* [332]; Mende *ndéwá* (“body hair”) [331]; Loko *déya* [331]; Bandi *ndéya* [331]; Looma *déyá* [331]; Kpelle-L *léya* [331]; Kpelle-G *leya* [331]; Bokobaru *ká* [328]; Boko *ká* [328]; Busa *ká* [328]; Lebir *ká* [328]; San *ka* [328]; Kyenga *káká* [328]; Shanga *siékaka* [328].

hand: Dan-G *kò* [334], *gbè* [339]; Dan-B *kò* [334], *gbíř* [339]; Dan-K *kò* [334], *gbè* [339]; Tura *kòò* [334]; Mano *kò* [334]; Guro *bē* [339]; Yaure *pè* [339]; Gban *kɔ* [334]; Mwan *kóó* [334], *gbē* [339]; Wan *ɔ* [334]; Beng *wɔ* [334]; Kpeen *sí-táà* [340]; Kpan *si-taqrɔ* [340]; Dzuun *sí-táà* [340]; Duun *si-ra'a* [340]; Seenku *kyè*, *kyie* [340]; Banka *ši-rá'q* [340]; Jo *gbé* [339], *je* [340]; Bobo *sōrō* [337]; Bozo-J *sugu* [337]; Bozo-K *kiri* [338]; Bozo-T *suu* [337]; Soninke *kitti/e* [338]; Mau *boó* [336]; Jula-K *bóro* [336]; Dafing *bòo* [336]; Bamana *bólo* [336]; Maninka *bólo* [336]; Xasonka *búlu* [336]; Mandinka *búlu* [336]; Koranko *bóló* [336]; Lele *bóló* [336]; Mogofin *bóló* [336]; Kakabe *bóló* [336]; Kono *bóó* [336]; Vai *boó* [336]; Soso *yíi* [335]; Jalonke *yíi* [335]; Jogo *bóló* [336]; Jeri *bulu* [336]; Numu *bóló* [336]; Mende *ngé-yà* [335]; Loko *nge-ya* [335]; Bandi *ngea* [335]; Looma *ze* [335]; Kpelle-L *yée* [335]; Kpelle-G *ye* [335]; Bokobaru *ɔ* [334]; Boko *ó* [334]; Busa *ɔ* [334]; Lebir *wó* [334]; San *gɔn* [334]; Kyenga *ɔ* [334]; Shanga *yòò* [334].

head: Dan-G *gò* [342]; Dan-B *gò* [342]; Dan-K *gò* [342]; Tura *wú* [342]; Mano *wú-kēlē* [342]; Guro *wūō* [342]; Yaure *wùlò* [342]; Gban *bò* [343]; Mwan *ŋgblò*, *ŋgblò* [343]; Wan *bó-ŋglò* [343]; Beng *wlū* [342]; Kpeen *kùŋgɔ* [342]; Kpan *kɔɔ* [342]; Dzuun *kùn-gɔ* [342]; Duun *kùŋkɔ* [342]; Seenku *kòn* [342]; Banka *kwá'q* [342]; Jo *ɲǎnǎ* [342]; Bobo *ŋwòn* [342]; Bozo-J *ŋe* (?) [342]; Bozo-K *šyāhwɔ* [345]; Bozo-T *tɔmɔ* [344]; Soninke *yin* / *yímmé* [342]; Mau *kùŋ* [342]; Jula-K *kùn* [342]; Dafing *kú* [342]; Bamana *kú* [342]; Maninka *kú* [342]; Xasonka *kùŋ* [342]; Mandinka *kùŋ* [342]; Koranko *kù* [342]; Lele *kù* [342]; Mogofin *kù* [342]; Kakabe *kùn* [342]; Kono *kù* [342]; Vai *kùŋ* [342]; Soso *xùn* [342]; Jalonke *qùn* [342]; Jogo *wú* [342]; Jeri *wù* [342]; Numu *wú* [342]; Mende *ngü* [342]; Loko *ngù-kókó* [342]; Bandi *ngü(ŋ)* [342]; Looma *ŋu(g)*, *wu(g)* [342]; Kpelle-L *ŋūŋ* [342]; Kpelle-G *wuŋ* [342]; Bokobaru *mì* [341]; Boko *mī* [341]; Busa *mì* [341]; Lebir *minì* [341]; San *mii* [341]; Kyenga *mì* [341]; Shanga *míi* [341].

hear: Dan-G *mā* [346]; Dan-B *mā* [346]; Dan-K *mà* [346]; Tura *má* [346]; Mano *mā* [346]; Guro *má* [346]; Yaure *mà* (ASM) [346]; Gban *mà* (ASM) [346]; Mwan *mā* / *má* [346]; Wan *mə* [346]; Beng *mā* [346]; Kpan *ŋmɛɛ-ma* [346]; Dzuun *myèn* [346]; Duun *míɛnɔ* [346]; Seenku *mèè* / *míé* [346]; Banka *mie* [346]; Jo *dá-mé* [346]; Bobo *mǎ* [346]; Bozo-J *myee* [346]; Bozo-K *bogi* (?) [346]; Bozo-T *buge* (?) [346]; Soninke *mùgú* [346]; Mau *myéŋ* [346]; Jula-K *mén*, *myén* [346]; Dafing *mwíi* [346]; Bamana *mé*, *mé* [346]; Maninka *mé* [346]; Xasonka *mé* [346]; Mandinka *mòyi* [346]; Koranko *lá mé*, *lá mé* [346]; Lele *lá-mín* [346]; Mogofin *mèè* [346]; Kakabe *mén* [346]; Kono *mì* [346]; Vai *lán* [347]; Soso *mèè* [346]; Jalonke *mèè* [346]; Jogo *mén* [346]; Jeri *mén*, *mé* [346]; Numu *mén* [346]; Mende *mèni* [346]; Loko *mě(ŋ)* [346]; Bandi *mèni* [346]; Looma *mèni(g)* [346]; Kpelle-L *mēni* [346]; Kpelle-G *man* [346]; Bokobaru *ma* [346]; Boko *ma* [346]; Busa *ma* [346]; Lebir *mà* [346]; San *má*, *méɛ* [346]; Kyenga *ma* [346]; Shanga *màà* [346].

heart: Dan-G *zūā*, *zò* [349]; Dan-B *zùř* [349]; Dan-K *zūā* [349]; Tura *zùv* [349]; Mano *zò* [349]; Guro *zùlù* [349]; Gban *gbí* [351], *klò* [352]; Mwan *sōŋ* [348], *zrū* [349]; Wan *zòŋ* [349], *tólì* [350]; Beng *blòniníŋ* (?) [359]; Kpeen *sòŋ* [348]; Kpan *sɔ* [348]; Dzuun *sòn* [348]; Duun *sono* [348]; Seenku *sò* [348]; Banka *sona* [348]; Jo *sòŋ* [348]; Bobo *wōrō* [357], *dógó* [358]; Bozo-J *buguliɔ* [359]; Bozo-K *bori* [359]; Bozo-T *boojɔ* [359]; So-

ninke *sondoN / sòndomme* [348]; Mau *sòŋ* [348]; Jula-K *sòn* [348]; Dafing *sɔ́* [348]; Bamana *sɔ́* [348]; Maninka *sɔ́, sòndòmɛ* [348]; Xasonka *sò, sòndomme* [348]; Mandinka *sòndome* [348]; Koranko *sɔ́, sò* [348], *yìsi* [354]; Lele *nii* [356]; Mogofin *dùsɛ* [354]; Kakabe *jùsu* [354]; Kono *sònsòn* [348], *fàa* [355]; Vai *fàá* [355]; Soso *sondon* [348]; Jalonke *sòndòn* [348]; Jogo *sò(n)* [348]; Jeri *gbógɔn* (?) [359]; Numu *són* [348]; Mende *ndíí* [353]; Loko *ndíí* [353]; Bandi *ndíí* [353]; Looma *zii* [353]; Kpelle-L *líi* [353]; Kpelle-G *li* [353]; Bokobaru *swɛ* [348]; Boko *sɔ́* [348]; Busa *swɛ* [348]; Lebir *sètɔ, sétò* [348]; San *koron* [352]; Kyenga *zù* [349]; Shanga *zunú* [349].

horn: Dan-G *sɔ́* [361]; Dan-B *sɔ́* [361]; Dan-K *sɔ́* [361]; Tura *sèlè* [361]; Mano *bé* [360]; Guro *béí ~ bélí* [360]; Yaure *béí* (AMS) [360]; Gban *bí* [360]; Mwan *mlè* [360]; Wan *mlɛŋ* [360]; Beng *béj* [360]; Kpeen *kjúú* [368]; Kpan *bíj* [360]; Dzuun *bín* [360]; Duun *bi* [360]; Seenku *bì(n)* [360]; Banka *viq* [360]; Jo *péŋ* [360]; Bobo *bɛ́* [360]; Bozo-J *buuon* [360]; Bozo-K *hómó* [368]; Bozo-T *bolon* (?) [360]; Soninke *beN / bénné* [360]; Mau *gɛ́* [367]; Jula-K *gbàn* [365]; Bamana *bíjɛ, byé* [360]; *gwére* [367]; Maninka *bíjɛ* [360]; *kére* [367]; Xasonka *bíjɛ* [360]; Mandinka *bína* [360]; Koranko *kíri, kére* [367]; Lele *bíjɛ* [360]; Mogofin *bína* [360]; Kakabe *bíjɛ* [360]; Kono *béna* [360]; Vai *bíná* [360]; Soso *feri* [366]; Jalonke *féerí* [366]; Jogo *gɔ́(n)* [365]; Jeri *gbòŋ, gbòŋ* [365]; Numu *gòn* [365]; Mende *ndówó* [363]; Loko *ndógó* [363]; Looma *mena* [360]; Kpelle-L *mílá* [360]; Kpelle-G *minɛ* [360]; Bokobaru *bɛ́ɛ* [360], *kóbá* [362]; Boko *béná* [360], *kóbá* [362]; Busa *béné* [360]; Lebir *míní* [360]; San *ben* [360]; Kyenga *kaho* (< Hausa *kàhó*, < *kàfó*) [-4]; Shanga *bèè* [360].

I: Dan-G *ma* [369], *a* [370], *n* [372]; Dan-B *ma* [369], *a* [370], *n* [372]; Dan-K *ma* [369], *n* [372]; Tura *ma* [369], *a* [370], *n* [372]; Mano *ma* [369], *n* [372]; Guro *ma* [369], *a* [370]; Yaure *ma* [369], *a* [370]; Gban *mi* [369], *i* [372]; Mwan *mi* [369], *ŋ* [372]; Wan *mí-nà* [369], *na* [371], *n* [372]; Beng *ma* [369], *ŋ* [372]; Kpeen *ma* [369], *n* [372]; Kpan *my* [369]; Dzuun *mun* [369]; Duun *ma* [369]; Banka *mu* [369]; Jo *má* [369], *n* [372]; Bobo *ma, me* [369], *ne* [372]; Bozo-J *n* [372]; Bozo-K *n* [372]; Bozo-T *n* [372]; Soninke *n* [372]; Mau *né* [372]; Jula-K *n* [372]; Dafing *n* [372]; Bamana *n* [372]; Maninka *n* [372]; Xasonka *n* [372]; Mandinka *n* [372]; Koranko *n* [372]; Lele *n* [372]; Mogofin *n* [372]; Kakabe *n* [372]; Kono *n* [372]; Vai *n* [372]; Soso *n* [372]; Jalonke *n* [372]; Jogo *n* [372]; Jeri *nà* [372]; Numu *n* [372]; Mende *ngá, ná* [371]; Loko *ngi* [371]; Bandi *ngí* [371], *ŋ* [372]; Looma *gè* [371], (*g*) [372]; Kpelle-L *nán* [371], *n* [372]; Kpelle-G *nán* [371], *n* [372]; Bokobaru *ma* [369]; Boko *ma* [369]; Busa *ma* [369]; Lebir *ma* [369]; San *ma* [369]; Kyenga *ma* [369]; Shanga *ma* [369].

kill: Dan-G *zā* [373]; Dan-B *zā* [373]; Dan-K *zè* [373]; Tura *zé* [373]; Mano *zē* [373]; Guro *jē* [373]; Yaure *tē* [373]; Gban *zè* [373]; Mwan *dē* [373]; Wan *té* [373]; Beng *dē* [373]; Kpeen *bùó* [377]; Kpan *boo-ma* [377]; Dzuun *vú* [377]; Duun *vo* [377]; Seenku *bwò / bwò / bwönè* [377]; Banka *bo* [377]; Jo *kílɛ́, klà* [376]; Bobo *yè* [373]; Bozo-J *waa* [375]; Bozo-K *waya* [375]; Bozo-T *gwan* [375]; Soninke *kàrí* [376]; Mau *fàà* [374]; Jula-K *fàgá* [374]; Dafing *páá* [374]; Bamana *fàgá, fàá* [374]; Maninka *fàá* [374]; Xasonka *fàxa* [374]; Mandinka *fàa* [374]; Koranko *fàga, fàa* [374]; Lele *fàa* [374]; Mogofin *fàgá* [374]; Kakabe *fàgá* [374]; Kono *fàa, fà* [374]; Vai *fàá* [374]; Soso *fàxá* [374]; Jalonke *fɔ́qɔ́* [374]; Jogo *kpá* [375]; Jeri *kpá* [375]; Numu *kpá* [375]; Mende *pàá* [374]; Loko *pàá* [374]; Bandi *pàa* [374]; Looma *paa* [374]; Kpelle-L *pāā* [374]; Kpelle-G *pa* [374]; Bokobaru *dɛ* [373]; Boko *dé* [373]; Busa *dɛ* [373]; Lebir *zé* [373]; San *dɛ / dii* [373]; Kyenga *dɛ* [373]; Shanga *dɛ* [373].

knee: Dan-G *kpīřř* [378]; Dan-B *kpɔ́gɔ́* [378]; Dan-K *kpīřřgɔ́* [378]; Tura *kpóŋ* [378]; Mano *kpūó* [378]; Guro *pó* [378]; Yaure *pó* [378]; Gban *kpó* [378]; Mwan *kpó* [378]; Wan *kpóŋ-gbóli* [378]; Beng *kpó* [378]; Kpeen *kùmú* [378]; Kpan *kuj* [381]; Dzuun *kúnún* [381]; Duun *kumu* [378]; Seenku *tyìmi* [378]; Banka *kuna* [381]; Jo *kànkwòvò* [378]; Bobo *fūngūnù* [382]; Bozo-J *kɔ́bɔ́* [378]; Bozo-K *húní* [378]; Bozo-T *xumun* [378]; Soninke *xúrúngí/o* [381]; Mau *kùmɛ́ɛ́ŋ* [378]; Jula-K *kúnbiri* [378]; Dafing *kùŋwii* [378]; Bamana *kúnbere, -kú, -kuru* [378], *kúnbere, -kú, -kuru* [381]; Maninka *kúnberen, -kudu* [378], *kúnberen, -kudu* [381]; Xasonka *xúnbalij-kuy* [378]; Mandinka *kúmbalij* [378]; Koranko *kúnbelɛ* [378]; Lele *kúnbili* [378]; Mogofin *kúnbelɛ* [378]; Kakabe *kùngbelén* [378]; Kono *kúnbai* [378]; Vai *kúmbéé* [378]; Soso *xinbi* [378]; Jalonke *qínbí* [378]; Jogo *gbóó* [378]; Jeri *gbó-wù* [378]; Numu *gbó* [378]; Mende *ngónbí* [378]; Loko *ngónbí* [378]; Bandi *ngùnbù* [378]; Looma *nibi* [378]; Kpelle-L *ŋúme* [378]; Kpelle-G *wungbe* [378]; Bokobaru *kosoó* [379]; Boko *koso* [379]; Busa *koso* [379]; Lebir *kùŋkóo, kùŋkoo, kùŋkoom* [378]; San *musi* [380]; Kyenga *kusú* [379]; Shanga *kùsù* [379].

know: Dan-G *dɔ́* [383]; Dan-B *dɔ́* [383]; Dan-K *dɔ́* [383]; Tura *dɔ́* [383]; Mano *dɔ́* [383]; Guro *dòò* [383]; Yaure *tɔ́* [383]; Gban *dɔ́* [383]; Mwan *dɔ́* [383]; Wan *tɔ́* [383]; Beng *dɔ́* [383]; Kpan *tɔɔ-ma* [383]; Dzuun *tò* [383]; Duun *tɔ* [383]; Seenku *tò / tyò* [383]; Banka *tɔ* [383]; Jo *tɔ́* [383]; Bobo *tò* [383]; Bozo-J *two* [383]; Bozo-K *to* [383]; Bozo-T *toy* [383]; Soninke *tùú* [383]; Mau *lón* [383]; Jula-K *lɔ́* [383]; Dafing *lɔ́* [383]; Bamana *dɔ́* [383]; Maninka *lɔ́* [383]; Xasonka *lón* [383]; Mandinka *lón* [383]; Koranko *lɔ́* [383]; Lele *són, lón* [383]; Mogofin *lɔ́* [383]; Kakabe *lón* [383]; Kono *sɔ́* [383]; Vai *só* [383]; Soso *kolon* [384]; Jalonke *kólón* [384]; Jogo *só* [383]; Jeri *só* [383]; Numu *sò* [383]; Mende *kóó* [384]; Loko *kóó(ŋ)* [384]; Bandi *kòlò* [384]; Looma *kòj(g), kwɛ(g)* [384]; Kpelle-L *kólɔŋ* [384]; Kpelle-G *kólɔŋ* [384]; Bokobaru *dɔ́* [383]; Boko *dɔ́* [383]; Busa *dɔ́* [383]; Lebir *dó* [383]; San *dɔ́ / doɛ* [383]; Kyenga *dɔ́* [383]; Shanga *dɔ́* [383].

leaf: Dan-G *dě* [385]; Dan-B *dě* [385]; Dan-K *lě* [385]; Tura *dě-kwě* [385]; Mano *léé* [385]; Guro *lää* [385]; Yaure *lá* [385]; Gban *lã* [385]; Mwan *làà* [385]; Wan *lāh* [385]; Beng *láná* [385]; Kpeen *gǒǒ-tírà*, *pl. gǒǒ-tírà-rì* [390]; Kpan *gɔɔ-ciraa* [390]; Dzuun *tsírá* [390]; Duun *cira* [390]; Seenku *tiri* [390]; Banka *fla* [390]; Jo *dá* [385]; Bobo *dā* [385]; Bozo-J *duo* [386]; Bozo-K *lálá* [385]; Bozo-T *lele* [385]; Soninke *dari / dáré* [385]; Mau *búú* [389], *fyáá* [390]; Jula-K *filá-búru, fúlá-búru* [390]; Dafing *plà* [390]; Bamana *búlu* [389], *fúra* [390]; Maninka *búlu* [389], *fída, fíra* [390]; Xasonka *fíta* [390]; Mandinka *jànba* [388], *fíta* [390]; Koranko *yànba* [388]; Lele *yànba, yènba* [388]; Mogofin *yànba* [388]; Kakabe *fíta* [390]; Kono *yànba* [388]; Vai *jànbá* [388]; Soso *búra-xe, búre-xe* [389]; Jalonke *búru-qé* [389]; Jogo *jà(n)* [388]; Jeri *jàn* [388]; Numu *gyán* [388]; Mende *ndáwá* [385]; Loko *ndága(η)* [385]; Bandi *ndáyá* [385]; Looma *daa(g)* [385]; Kpelle-L *lāa* [385]; Kpelle-G *la* [385]; Bokobaru *lá* [385]; Boko *lá* [385]; Busa *lá* [385]; Lebir *lèé* [385]; San *luu* [386]; Kyenga *lálá* [385]; Shanga *daàkò* [387].

lie: Dan-G *wǒ* [393]; Dan-B *wō* [393]; Dan-K *wò* [393]; Tura *kpáá* [396]; Mano *wō* [393]; Guro *yíá* [397]; Yaure *yìrǎ* [397]; Gban *wǎ-tǔ* (AMS) [392]; Mwan *yílà* [397]; Wan *gòdì* [395]; Beng *bèdǎ* [394]; Kpeen *né* [397]; Kpan *ɲεε-ma* [397]; Dzuun *nén* [397]; Duun *na* [397]; Seenku *sínǎ* [399]; Banka *na* [397]; Jo *bílí* [398]; Bobo *dē* (?) [391]; Bozo-J *saa* [391]; Bozo-K *saya* [391]; Bozo-T *saa* [391]; Soninke *sáxú* [391]; Mau *lá* [391]; Jula-K *lá* [391]; Dafing *là* [391]; Bamana *dá* [391]; Maninka *lá* [391]; Xasonka *lá* [391]; Mandinka *lāa* [391]; Koranko *sá* [391]; Lele *sá* [391]; Mogofin *sá* [391]; Kakabe *lāa* [391]; Kono *sá* [391]; Vai *sá* [391]; Soso *sá* [391]; Jalonke *sāa* [391]; Jogo *sá* [391]; Jeri *sá* [391]; Numu *sà* [391]; Mende *ndá* [391]; Loko *ndá* [391]; Bandi *ndà* [391]; Looma *da* [391]; Kpelle-L *lāa* [391]; Kpelle-G *laa* [391]; Bokobaru *da-é* [391]; Boko *da-é* [391], *wílé* [392]; Busa *wíté* [392]; Lebir *wòtá* [392]; San *waa* (?) [392]; Shanga *intee* [393].

liver: Dan-G *blúú* [410]; Dan-B *blú* [410]; Dan-K *búú* [410]; Tura *búllú* [410]; Mano *búlú* [410]; Guro *bílí* [410]; Yaure *fǔvǔ* (?) [410]; Gban *gbǐ* [411]; Mwan *zrú* [412]; Wan *tólì* [413]; Beng *blǒh* [410]; Kpeen *bé* [405]; Kpan *buyi* [410]; Dzuun *vínún* [410]; Duun *bu* [410]; Seenku *bòròn* [410]; Banka *vuna* [410]; Jo *tǔh* [413]; Bobo *fègèrè* [416]; Bozo-J *tuo* [413], *tasa* (< Songai) [-1]; Bozo-K *ndǒlǒ* [413]; Bozo-T *tolǒ* [413], *sere* [414]; Soninke *buttu / bútté* [410]; Mau *byé* [405]; Jula-K *bíyán, bíyen* [405]; Dafing *bíéé* [405]; Bamana *bíne, byé* [405]; Maninka *bíne* [405]; Xasonka *búna* [405]; Mandinka *júsu* [409]; Koranko *bíya, béya* [405]; Lele *fúyiwa* [407]; Mogofin *sòndǒ* [408]; Kakabe *búna* [405]; Kono *fàa* [406]; Vai *fǎá* [406]; Soso *bǒné* [405]; Jalonke *bǒné* [405]; Jogo *bǒgǒ(n)* [404]; Jeri *bǒgǒh, bǒgǒ* [404], *bíyè* [405]; Numu *bògǒn* [404]; Loko *né* [402]; Looma *zi* [403]; Kpelle-L *líi* [403]; Kpelle-G *yele* [401]; Bokobaru *pǒ, pǒǒ* [400]; Boko *pǒ, pǒá* [400]; Busa *pǒ* [400]; Lebir *fǒǒ* [400]; San *foo* [400].

long: Dan-G *gblèè* [417]; Dan-B *gbèè, gbèèh, gbǐ, gbèih* [417], *lǒh* [418]; Dan-K *dró* [419]; Tura *dǒǒ* [419]; Mano *gbēī* [417]; Guro *tōrō-ī* [421]; Yaure *pǒlǐ* [417]; Gban *gǎgǎ* (?) [426]; Mwan *gbǒǒ* [417]; Wan *blǒ* [420]; Beng *gblēh* [417], *jǎ* [426]; Kpeen *gbǎ* [417]; Kpan *gbǎ* [417]; Dzuun *gbààn* [417]; Duun *gban-ma* [417]; Seenku *dyù* [426]; Banka *gbǎ* [417]; Jo *ɲmà* [417]; Bobo *kú* [424]; Bozo-J *kosǎ* [424]; Bozo-K *mwǒnǒ* (?) [417]; Bozo-T *xuyon* [424]; Soninke *gille* (?) [417]; Mau *jǎh, yǎh* [426]; Jula-K *jàn* [426]; Dafing *záqzǎ* [426]; Bamana *jǎ* [426]; Maninka *jǎ* [426]; Xasonka *jàn* [426]; Mandinka *jàn* [426]; Koranko *yǎ* [426]; Lele *yàn* [426]; Mogofin *ká-yǎ* [426]; Kakabe *ján* [426]; Kono *yàn-sǎ* [426]; Vai *jǎh* [426]; Soso *kuya* [424]; Jalonke *kújá* [424]; Jogo *sǒǒ-mán* [425]; Jeri *sùǒ, sǐǒ* [425]; Numu *sòǒ-* [425]; Mende *kúhàn* [424]; Bandi *kòhó* (“tall”) [424]; Looma *kooza(g)* [424]; Kpelle-L *kóyá* [424]; Kpelle-G *kwea* [424]; Bokobaru *gbǎá* [417]; Boko *gbǎá* [417]; Busa *gbǎná* [417]; Lebir *kyǎ* [422]; San *sasa* [423]; Kyenga *sà* [423]; Shanga *sàà* [423].

louse: Dan-G *wèè* [427], *kèfá* [430]; Dan-B *lǐr* [427]; Dan-K *wèè* [427], *kààngbā* [430]; Tura *kǎábǎ* [430]; Mano *wèè* [427]; Guro *wēi ~ wōi* [427]; Yaure *wè* [427]; Gban *glǎgbǎ, glǎǎ, gbǎ* [430]; Mwan *klǎgbǎ* [430]; Wan *wèh* [427]; Beng *yídǎ* [431]; Kpeen *kpàngbǎ* [430]; Kpan *fífú* [434]; Dzuun *kpèyngbe* [430], *fínfín* [434]; Duun *fífi* [434]; Banka *finá* [434]; Jo *kpàrǐngǎ* [430]; Bobo *kpàngbǎ* [430]; Bozo-J *sǒmǒ* [432]; Bozo-K *hámú* [432]; Bozo-T *sǒmǒ* [432]; Soninke *duuyi / diúyné* [433]; Mau *kààngbǎ* [430], *nyúyú* [432]; Jula-K *kààngbǎ* [430], *nyúmúgú* [432]; Dafing *nyámì* [432]; Bamana *nkàrangá* [430], *nyími* [432]; Maninka *kàrangbá* [430], *nyímí* [432]; Xasonka *xàranxa* [430], *nyámi* [432]; Mandinka *kàranka* [430], *dúyna* [433]; Koranko *kàrangbǎ* [430], *nyenke* [432]; Lele *kàingbǎ* [430], *nyánka* [432]; Mogofin *kàranbá* [430], *nyàngá* [432]; Kakabe *kàran* [430], *nyànga* [432]; Kono *nyànga* [432]; Vai *nyàá* [432]; Soso *karangbe* [430]; Jalonke *gàràngǎ* [430]; Jogo *wùlú* [427]; Jeri *kàrangbǎ* [430]; Numu *wùlú* [427]; Mende *kàrangbǎ* [430]; Bandi *kàlabá* [430], *nèni* [432]; Looma *kalabá* [430]; Kpelle-L *yǎū* [427]; Kpelle-G *yǒw* [427], *kelaa, kalaa* [430]; Bokobaru *gǎí* [427]; Boko *gǎí* [427]; Busa *sàa* [428]; Lebir *sìhǐr / sìngìrì* [429]; San *korǎnkien* [430]; Kyenga *sili* [428]; Shanga *sé* [428].

man: Dan-G *gǒǒ* [435]; Dan-B *gǒǒ* [435]; Dan-K *gǒ* [435]; Tura *gǒǒ* [435]; Mano *gǒ* [435]; Guro *gǒnē* [435]; Yaure *kwǐǐ-mǒ* [435]; Gban *kò* [435]; Mwan *gwǐǐ* [435]; Wan *kǒ-lé* [435]; Beng *gǒh* [435]; Kpeen *pǐh* [440]; Kpan *kpíi* [440]; Dzuun *kpíri* [440]; Duun *kwíi* [440]; Seenku *šùrù* [435]; Banka *kpila, kpíri* [440]; Jo *šùrù* [435]; Bobo *sìn* [435]; Bozo-J *kaygu* [438]; Bozo-K *hálú* (?) [438]; Bozo-T *xalu* (?) [438];

Soninke *yúgó* [439]; Mau *cè* [438]; Jula-K *cè* [438]; Dafing *cié* [438]; Bamana *cě* [438]; Maninka *kě, cě* [438]; Xasonka *xè* [438]; Mandinka *kèe* [438]; Koranko *kè* [438]; Lele *cè-mɔɔ* [438]; Mogofin *kàyi* [438]; Kakabe *kàyi* [438]; Kono *kàj* [438]; Vai *kàí* [438]; Soso *xáméé, xeme* [438]; Jalonke *qémé* [438]; Jogo *cíni* [435]; Jeri *kíni* [435]; Numu *kyíni* [435]; Mende *híná* [435]; Loko *híè* [435]; Bandi *sina* [435]; Looma *zina* [435]; Kpelle-L *sūrôŋ, pl. sînâa* [435]; Kpelle-G *hîlè-nû* [435]; Bokobaru *gɔ́* [435]; Boko *gɔ* [435]; Busa *gɔ* [435]; Lebir *sà* [436]; San *gīī* [435]; Kyenga *nigzni* [435]; Shanga *ezaa* [437].

many: Dan-G *súmǎ* [453], *gbé* [464]; Dan-B *plāzà* [452]; Dan-K *plàá* [452]; Tura *béé* [451]; Mano *kogbuòze* [1007]; Guro *dì* (?) [450]; Yaure *kágá* [448]; Gban *nǔú* [446], *gbǔú* [447]; Mwan *sréé* [444], *būū* [445], *gbògbò* [447], *bèbè* [449]; Wan *bèbè* [449], *tí* [450]; Beng *blèh* [443], *tè* [450]; Kpan *gi* [473]; Dzuun *gírí* [473]; Duun *ŋeŋe* [472], *gíru-ma* [473]; Seenku *mùmà, mummà* [471], *gúrù* [473], *bùgù* [474]; Jo *mwòðnnà* [471]; Bobo *fīngīnī* [463], *mū* [471], *ján* [472], *bùgū* [474]; Bozo-J *kusa* [469]; Bozo-K *kóná* [470]; Bozo-T *kuse* [469], *kɔ* [470]; Soninke *gàbé* [468]; Mau *šyáá* [465]; Jula-K *šyá* [465]; Bamana *cá, cǎ* [465]; Maninka *síya* [465]; Xasonka *sá, síya* [465], *kiika* (?) [466]; Mandinka *síyaa* [465]; Koranko *síya* [465]; Lele *síya* [465]; Mogofin *síyaa* [465]; Kakabe *síya-man* [465]; Kono *kókó, kókó* [466]; Vai *kúúŋ* [467]; Soso *gbegbe* [464]; Jogo *féénámá* [463]; Jeri *fié-nè* [463]; Numu *féénè* [463]; Bandi *kpihí* [462]; Looma *támaa(g)* [461]; Kpelle-L *támaa* [461]; Kpelle-G *támq* [461]; Bokobaru *pàri* [441]; Boko *dàsí* [442], *bílá* [454]; Busa *dàsí* [442], *bítá* [454]; Lebir *góta* [455], *nwií ~ wií* [456]; San *damatan* [457], *didi* [458], *gigiq* [459]; Kyenga *gbà* [460]; Shanga *dààki* [442].

meat: Dan-G *wū* [477]; Dan-B *wū* [477]; Dan-K *wù* [477]; Tura *wií* [477]; Mano *wií* [477]; Guro *wī* [477]; Yaure *wì* [477]; Gban *wè* [477]; Mwan *wī* [477]; Wan *wì* [477]; Beng *sōŋ* [478]; Kpeen *só* [478]; Kpan *šo* [478]; Dzuun *sú* [478]; Duun *šwo* [478]; Seenku *kèkà* [480]; Banka *šoq* [478]; Jo *šǔǔ* [478]; Bobo *kikā, kiègā, kiègè* [480]; Bozo-J *teu* [479]; Bozo-K *túgú* [479]; Bozo-T *teu* [479]; Soninke *tíyè, ciye* [479]; Mau *sòð* [478]; Jula-K *sògó* [478]; Dafing *swó* [478]; Bamana *sògó* [478]; Maninka *sòbó* [478]; Xasonka *sùbu* [478]; Mandinka *sùbu* [478]; Koranko *sògo, sòo* [478]; Lele *sòo* [478]; Mogofin *sòwo* [478]; Kakabe *sòbo* [478]; Kono *sùì* [478]; Vai *sùé* [478]; Soso *sube* [478]; Jalonke *sùbé* [478]; Jogo *siè* [478]; Jeri *sibe* [478]; Numu *sìè* [478]; Mende *hùán, hùá* [478]; Loko *hòà* [478]; Bandi *sùwá* [478]; Looma *suwa* [478]; Kpelle-L *sūā* [478]; Kpelle-G *hyuo* [478]; Bokobaru *nòbò* [475]; Boko *nòbō ~ nòḅ* [475]; Busa *nòbò* [475]; Lebir *sím |siní* [476]; San *sij* [476]; Kyenga *wemɔ* [477]; Shanga *sɔ̀* [478].

moon: Dan-G *sú* [482]; Dan-B *sú* [482]; Dan-K *sú* [482]; Tura *méè* [481]; Mano *mēnē, mlé* [481]; Guro *mōnē* [481]; Gban *mé* [481]; Mwan *mlēē* [481]; Wan *bōlēŋ* [481]; Beng *m̄* [481]; Kpeen *kébé* [488]; Kpan *xuɔ* [487]; Dzuun *kúú* [487]; Duun *ko'o* [487]; Seenku *kè ?* [483]; Banka *kogwq* (?), *koko* (?) [487]; Jo *né* (?) [483]; Bobo *yū* [486]; Bozo-J *keu* [483]; Bozo-K *háyo* [483]; Bozo-T *xay* [483]; Soninke *xasu / xásó* [483]; Mau *káá* [483]; Jula-K *kárí* [483]; Dafing *káru* [483]; Bamana *kálo* [483]; Maninka *káro* [483]; Xasonka *xáru* [483]; Mandinka *kári* [483]; Koranko *kári* [483]; Lele *kári, káyi* [483]; Mogofin *káre* [483]; Kakabe *káru* [483]; Kono *káo* [483]; Vai *káó* [483]; Soso *kike, kige* [485]; Jalonke *kiiké* [485]; Jogo *kái* [483]; Jeri *káyi* [483]; Numu *kànkànsí* (“sky charcoal”) [484]; Mende *ngálú* [483]; Loko *ngáú* [483]; Bandi *ngálú(ŋ)* [483]; Looma *yálo(g)* [483]; Kpelle-L *yáloŋ* [483]; Kpelle-G *paniŋ* [483]; Bokobaru *mɔ́* [481]; Boko *mɔ* [481]; Busa *mɔ* [481]; Lebir *mòní* [481]; San *mui* [481]; Kyenga *mɔ* [481]; Shanga *mɔɔ* [481].

mountain: Dan-G *t̄* [497]; Dan-B *t̄* [497]; Dan-K *t̄* [497]; Tura *t̄̀* [497]; Mano *t̄̀* [497]; Guro *gò* [498]; Yaure *p̄* [499]; Gban *kpi* [489]; Mwan *ḡɔ̄* [498]; Wan *gòŋ* [498]; Beng *dèŋ* (?) [496]; Kpeen *fálá* [502]; Kpan *je* [492]; Dzuun *jèn* [492]; Duun *zeni* [492]; Seenku *dyūun* [492], *túnnó* [497]; Banka *gena* [492]; Jo *t̄ŋ* [497]; Bobo *tòlò* [497]; Bozo-J *sire-kutu* [501]; Bozo-K *koro-sie* (?) [493], *koro-sie* (?) [501]; Bozo-T *serè* [501]; Soninke *gídí/e* [492], *segenda* [500]; Mau *kùù* [493], *t̄n̄ŋ* [496]; Jula-K *kùrú* [493]; Dafing *kùrú* [493], *kùŋwù* [495]; Bamana *kùlú* [493]; Maninka *kùrú* [493], *kònké* [494]; Xasonka *kùru* [493]; Mandinka *kònkò-baa* [494]; Koranko *kònkɔ* [494]; Lele *kòngɔ* [494]; Mogofin *kònkó* [494]; Kakabe *kòнке* [494]; Kono *kòngɔ* [494]; Vai *kòŋɔ* [494]; Soso *gèyáá* [492]; Jalonke *gèjáá* [492], *kólí* [493]; Jogo *kólé* [493]; Jeri *t̄iri, t̄uru* [496]; Numu *kólé* [493]; Mende *ngíyè* [492]; Loko *ngíhé* [492]; Bandi *ngéhé, ngéhé(ŋ)* [492]; Looma *gizi(g), gizi* [492]; Kpelle-L *yéè* [492]; Kpelle-G *yé* [492]; Bokobaru *kpiii* [489], *gbè-s̄s̄i* [490]; Boko *gbè-s̄s̄i* [490]; Busa *kpi* [489]; Lebir *kyi* [489]; San *kele* [489]; Kyenga *kàsa* [491]; Shanga *kàsà* [491].

mouth: Dan-G *dí* [503]; Dan-B *dí* [503]; Dan-K *lí* [503]; Tura *dě* [503]; Mano *lé* [503]; Guro *lēé* [503]; Yaure *lē-bò* [503]; Gban *lū* [503]; Mwan *lū* [503]; Wan *lā-gá* [503]; Beng *yé* [503]; Kpeen *dígí* [503]; Kpan *dii* [503]; Dzuun *dzií, dzú* [503]; Duun *ŋa'a* [504]; Seenku *dzò* [503]; Banka *jigq* [503]; Jo *t̄ŋ* [503]; Bobo *dō* [503]; Bozo-J *la* [503]; Bozo-K *lógó* [503]; Bozo-T *la* [503]; Soninke *raxa / ráqqé* [503]; Mau *lā* [503]; Jula-K *dá* [503]; Dafing *dà* [503]; Bamana *dá* [503]; Maninka *dá* [503]; Xasonka *dá* [503]; Mandinka *dáa* [503]; Koranko *dá* [503]; Lele *dá* [503]; Mogofin *dá* [503]; Kakabe *dàa* [503]; Kono *dá* [503]; Vai *dá* [503]; Soso *déé* [503]; Jalonke *déé* [503]; Jogo *dáa* [503]; Jeri *dá* [503]; Numu *dá* [503]; Mende *ndá* [503];

Loko *ndá* [503]; Bandi *ndà* [503]; Looma *da, da(g)* [503]; Kpelle-L *lá* [503]; Kpelle-G *la* [503]; Bokobaru *lé* [503]; Boko *lé* [503]; Busa *lé* [503]; Lebir *lé (lè)* [503]; San *lé* [503]; Kyenga *lé* [503]; Shanga *lé* [503].

name: Dan-G *tó* [505]; Dan-B *tó* [505]; Dan-K *tó* [505]; Tura *tó* [505]; Mano *tó* [505]; Guro *tó* [505]; Yaure *tó* [505]; Gban *tó* [505]; Mwan *tó* [505]; Wan *tó* [505]; Beng *tó* [505]; Kpan *thuɔ* [505]; Dzuun *tsóo* [505]; Duun *šwɔ́ɔɔ* [505]; Seenku *tōgò* [505]; Banka *šwá'g* [505]; Jo *tó* [505]; Bobo *tōgō* [505]; Bozo-J *tube* (?) [505]; Bozo-K *týó* [505]; Bozo-T *tɔ* [505]; Soninke *taxi / tóxó* [505]; Mau *tóó* [505]; Jula-K *tógó* [505]; Dafing *twò* [505]; Bamana *tógó* [505]; Maninka *tóɔ* [505]; Xasonka *tóxo* [505]; Mandinka *tóo* [505]; Koranko *tógɔ, tóo, tóó* [505]; Lele *tóɔ* [505]; Mogofin *tógè* [505]; Kakabe *tógɔ* [505]; Kono *tó* [505]; Vai *tóó* [505]; Soso *xili* [507]; Jalonke *qíllí* [507]; Jogo *tógó* [505]; Jeri *tógɔ* [505]; Numu *tògò* [505]; Mende *ndá* [506]; Loko *ndàà, ndá* (?) [506]; Bandi *ndàa-héi(η)* [506]; Looma *daa-zei(g), da-se(g)* [506]; Kpelle-L *lóa* [506]; Kpelle-G *la* [506]; Bokobaru *tó* [505]; Boko *tó* [505]; Busa *tó* [505]; Lebir *tó* [505]; San *tɔ* [505]; Kyenga *tó* [505]; Shanga *toó* [505].

neck: Dan-G *b̄* [511]; Dan-B *b̄* [511]; Dan-K *b̄* [511]; Tura *vèlè* [513]; Guro *b̄l̄* [511]; Yaure *bl̄* [511]; Gban *w̄* [510], *β̄* [513]; Mwan *bl̄* [511]; Wan *ōl̄* [512]; Beng *lò* [514]; Kpan *vo* [511]; Dzuun *vú* [511]; Duun *vo* [511]; Seenku *gù* [516]; Banka *vwa* [511]; Jo *ȳ* [517]; Bobo *mògò* [519]; Bozo-J *bondo* [511]; Bozo-K *hánj* [517]; Bozo-T *xan* [517]; Soninke *xaN / xánné* [517]; Mau *kán* [517]; Jula-K *kán* [517]; Dafing *kà* [517]; Bamana *ká* [517]; Maninka *ká* [517]; Xasonka *xán* [517]; Mandinka *kán* [517]; Koranko *ká* [517]; Lele *ká* [517]; Mogofin *ká* [517]; Kakabe *kán* [517]; Kono *ká* [517]; Vai *kán* [517]; Soso *kǎn* [516]; Jalonke *kǎn* [516]; Jogo *wùrà-k̄(n)* [516], *fólin* [518]; Jeri *ku* [516]; Numu *fóli* [518]; Mende *kàngá* [516]; Loko *kǎnga* (KOELLE) [516]; Bandi *kàngà(η)* [516]; Looma *kɔ(g)* [516]; Kpelle-L *kɔ̄* [516]; Kpelle-G *kɔ̄* [516]; Bokobaru *wáà* [508]; Boko *wáà* [508], *n̄é* [509]; Busa *wákà* [508]; Lebir *w̄r* [510]; San *bóló* [511]; Kyenga *lɔkɔ* [514]; Shanga *kotoó'* [515].

new: Dan-G *d̄è* [521]; Dan-B *d̄ú, d̄ú* [521]; Dan-K *d̄è* [521]; Tura *d̄é* [521]; Mano *d̄é* [521]; Guro *d̄ē* [521]; Yaure *tèrè* [521]; Gban *dòà* [521]; Mwan *dr̄* [521]; Wan *tólé* [521]; Beng *dà-dr̄* [521]; Kpan *jei* [521]; Dzuun *yé, yéú* [521], *kyèèrón* [523]; Duun *žeo, žere* [521]; Seenku *sige* [1011]; Banka *yie* [521]; Jo *ɲùnɲɲ* (?) [522]; Bobo *ɲwún* (?) [522]; Bozo-J *toi* [521]; Bozo-K *tòri* [521]; Bozo-T *tori* [521]; Soninke *kúrumbá* [524]; Mau *kwáá* [523]; Jula-K *kúra* [523]; Dafing *kùrà* [523]; Bamana *kúra* [523]; Maninka *kúda, kúra* [523]; Xasonka *kúta* [523]; Mandinka *kúta* [523]; Koranko *kúra* [523]; Lele *kúya, kúra* [523]; Mogofin *nàma* [522]; Kakabe *nàma* [522]; Kono *nàma* [522]; Vai *námá* [522]; Soso *neene* [522]; Jalonke *néenè* [522]; Jogo *nán* [522]; Jeri *nàn* [522]; Numu *nàn* [522]; Mende *ninã* [522]; Loko *nyie* [522]; Bandi *nìina* [522]; Looma *niine* [522]; Kpelle-L *nīnā* [522]; Kpelle-G *ning* [522]; Bokobaru *dufú* [520]; Boko *dafú, defú* [520]; Busa *dufú* [520]; Lebir *dàà* [521]; San *díí* [521]; Kyenga *dele* [521]; Shanga *dè* [521].

night: Dan-G *b̄j* [527], *gb̄j* [532]; Dan-B *b̄j* [527], *gb̄j* [532]; Dan-K *gb̄j* [532]; Tura *gb̄è* [532]; Mano *b̄jé* [527]; Guro *b̄l̄, b̄ȳ* [532]; Yaure *pèi-màn* [532]; Gban *túé* [530]; Mwan *b̄* [527]; Wan *cīj* [531]; Beng *yrú* [529]; Kpeen *féngò* [536]; Kpan *fè-ne* [536]; Dzuun *fèn* [536]; Duun *fè* [536]; Seenku *gw̄* [526]; Banka *fèna* [536]; Jo *cī* [531]; Bobo *wùrū* [526]; Bozo-J *guu* [526]; Bozo-K *gúlú* [526]; Bozo-T *gulu* [526]; Soninke *wuru / wùró* [526]; Mau *sú* [533]; Jula-K *sú* [533]; Dafing *sù* [533]; Bamana *sú* [533]; Maninka *sú* [533]; Xasonka *sú* [533]; Mandinka *súu* [533]; Koranko *súu* [533]; Lele *swí* [533], *p̄npi* [534]; Mogofin *sú* [533]; Kakabe *súu* [533]; Kono *súu* [533], *p̄mbi* [534]; Vai *lifí* [535]; Soso *kǎð / kwèè* [528]; Jalonke *kwèè* [528]; Jogo *kúró* [528]; Jeri *kúron, kúrɔn* [528]; Numu *kúró* [528]; Mende *kp̄indí* [532]; Loko *kp̄indí* [532]; Bandi *kp̄indí* [532]; Looma *kp̄idi, kp̄idī* [532]; Kpelle-L *kp̄ini* [532]; Kpelle-G *kwi* [528], *kp̄ini* [532]; Bokobaru *gwá* [526]; Boko *gwá* [526]; Busa *gwá* [526]; Lebir *gùmú, gvinú* [526]; San *t̄lo* [525]; Kyenga *gáoní* [526]; Shanga *gu sí* [526].

nose: Dan-G *ȳ* [537]; Dan-B *ȳ* [537]; Dan-K *ȳ* [537]; Tura *ȳú* [537]; Mano *ȳ* [537]; Guro *ȳ* [537]; Yaure *yè-bò* [537]; Gban *ȳ* [537]; Mwan *ȳ* [537]; Wan *ȳ* [537]; Beng *ȳjé* [537]; Kpeen *kùwó* [538]; Kpan *xɔ* [538]; Dzuun *xón* [538]; Duun *xwɔ-dī* [538]; Seenku *bùnù* (?) [537]; Banka *xwɔn* [538]; Jo *ɲén-tà* [538]; Bobo *b̄n* [537]; Bozo-J *m̄ȳ* [537]; Bozo-K *m̄nú* [537]; Bozo-T *muun* [537]; Soninke *nuxuN / n̄x̄unné* [537]; Mau *nȳj* [537]; Jula-K *nún* [537]; Dafing *nùù* [537]; Bamana *n̄ȳ* [537]; Maninka *n̄ȳ* [537]; Xasonka *n̄ȳ* [537]; Mandinka *n̄ȳ* [537]; Koranko *s̄ȳ* [537]; Lele *s̄ȳ* [537]; Mogofin *s̄ȳ* [537]; Kakabe *s̄ún* [537]; Kono *s̄ȳ* [537]; Vai *s̄ȳ* [537]; Soso *ɲóó, ɲwéé* [537]; Jalonke *ɲwéé* [537]; Jogo *s̄ȳ* [537]; Jeri *s̄ȳ-gbàlà* [537]; Numu *s̄ȳn* [537]; Mende *hókpá* [537]; Loko *hókpá* [537]; Bandi *sòkpá* [537]; Looma *sokpa* [537]; Kpelle-L *s̄ȳa* [537]; Kpelle-G *hygɔ* [537]; Bokobaru *ȳ* [537]; Boko *ní* [537]; Busa *ȳ* [537]; Lebir *m̄* [537]; San *nin* [537]; Kyenga *ȳ* [537]; Shanga *í* [537].

not: Guro *lo* [539]; Mwan *la* [539]; Wan *ó* [539]; Kpan *na* [547]; Dzuun *ma, maa* [543], *fá* [546]; Duun *na* [547]; Banka *ma* [543]; Jo *kí* [1000]; Bobo *má* [543]; Bozo-T *te* [542]; Soninke *ntá* (imperfectif) [542], *má* (perfective) [543], *fè* (negative copula) [546]; Mau *te* [542], *ma, m̄* [543]; Jula-K *tí* [542], *má* [543]; Dafing *tí* [542], *mà* [543]; Bamana *té, té* [542], *má* [543]; Maninka *té* [542], *má* [543]; Xasonka *nté* [542], *mée* [543]; Mandinka *té* [542], *máj* [543]; Koranko *sá* [541], *té* [542], *má* [543]; Lele *sá* [541], *té* [542], *má* [543]; Mogofin *má* [543], *bélè* [544]; Kakabe *má* [543], *bélè* [544]; Kono *má* [543], *ní* [545]; Vai *má* [543], *bèè* [544]; Soso *mú* [541]; Jalonke *mún* [541]; Jogo *ró* [539]; Jeri *te, nde* [542]; Lebir *bí* [540].

one: Dan-G *dō* [548]; Dan-B *dō* [548]; Dan-K *dò* [548]; Tura *dó* [548]; Mano *dōo* [548]; Guro *dō* [548]; Yaure *tō ~ tū ~ tũ* [548]; Gban *dò* [548]; Mwan *dō* [548]; Wan *dō* [548]; Beng *dō* [548]; Kpeen *séélè* [554]; Kpan *sɔɔi* [554]; Dzuun *sóó, són* [554]; Duun *soe'* [554]; Seenku *sòèn* [554]; Banka *fiè* [555]; Jo *tàn* [553]; Bobo *tèlè, tàlà* [553]; Bozo-J *sanna* [554]; Bozo-K *kene* [550]; Bozo-T *kuɔn* [550], *sanna* [554]; Soninke *báané* [552]; Mau *kééŋ* [550]; Jula-K *kélen* [550]; Dafing *cè* [550]; Bamana *kéle* [550]; Maninka *kéle* [550]; Xasonka *kílin* [550]; Mandinka *kílin* [550]; Koranko *kéle* [550]; Lele *kélen, célen* [550]; Mogofin *dóndò* [548]; Kakabe *kélen* [550]; Kono *dóndó* [548], *ncélén* [550]; Vai *lòndó* [548]; Soso *kérén* [550]; Jalonke *kédén* [550]; Jogo *dié(n)* [551]; Jeri *dié(n)* [551]; Numu *dién* [551]; Mende *ngilá* [550]; Loko *ngèrà(ŋ)* [550]; Bandi *ngilàà(ŋ)* [550]; Looma *gila(g)* [550]; Kpelle-L *táan* [553]; Kpelle-G *tan* [553]; Bokobaru *do* [548]; Boko *do* [548]; Busa *do* [548]; Lebir *díim* [548]; San *goon* [549]; Kyenga *du* [548]; Shanga *dòò* [548].

man, person: Dan-G *mē* [557]; Dan-B *mē* [557]; Dan-K *mè* [557]; Tura *mèè* [557]; Mano *mī* [557]; Guro *mī* [557]; Yaure *mì* [557]; Gban *mú* [557]; Mwan *mēē* [557]; Wan *mī* [557]; Beng *sònn̄* [559]; Kpeen *mwón* [557]; Kpan *moj* [557]; Dzuun *mò-dziin* [557]; Duun *mɔ'edi* [557]; Seenku *mò* [557]; Banka *mo'o-wiri* [557]; Jo *kònn̄* [561]; Bobo *són* [559]; Bozo-J *nimi* [560]; Bozo-K *nini* [560]; Bozo-T *muɔ* [557]; Soninke *sere* [559]; Mau *mɔ̀* [557]; Jula-K *mɔ̀gɔ̀* [557]; Dafing *mó* [557]; Bamana *mɔ̀gɔ̀* [557]; Maninka *mɔ̀* [557]; Xasonka *mòxo* [557]; Mandinka *mò* [557]; Koranko *mɔ̀gɔ̀* [557]; Lele *mɔ̀* [557]; Mogofin *mɔ̀gɛ* [557]; Kakabe *mɔ̀gɔ̀* [557]; Kono *mò* [557]; Vai *mɔ̀* [557]; Soso *mìxí* [557]; Jalonke *mèqé* [557]; Jogo *mògò* [557]; Jeri *mùgú* [557]; Numu *mɔ̀gò* [557]; Mende *nú* [558]; Loko *nú* [558]; Bandi *nù* [558]; Looma *nu* [558]; Kpelle-L *núú* [558]; Kpelle-G *nuu* [558]; Bokobaru *gbé* [556]; Boko *gbé* [556]; Busa *gbé* [556]; Lebir *gwáá* [556]; San *min* [557]; Kyenga *guce* [556]; Shanga *gwéé* [556].

rain: Dan-G *dā* [562]; Dan-B *dā* [562]; Dan-K *là* [562]; Tura *dā* [562]; Mano *lē-yí* [564]; Guro *leé* [564]; Yaure *làá* [562]; Gban *gwlé* [563]; Mwan *lā* [562]; Wan *cí, cíŋ* [567]; Beng *lā* [562]; Kpeen *ɲàrà jí* ('God's water') [571]; Kpan *ɲaa* [571]; Dzuun *ɲáá* [571]; Duun *man̄kara-dj* [571]; Seenku *són* [568]; Banka *ɲga* [571]; Jo *kir̄ŋ* [567]; Bobo *wūrò* ('sky, rain') [570]; Bozo-J *kwà* [568]; Bozo-K *hàbá* [569]; Bozo-T *xɔin-ji* [568]; Soninke *kaN / kammé, kanji* [568]; Mau *sāŋ* [568]; Jula-K *sán, sájyé* [568]; Dafing *sà* [568]; Bamana *sán* [568]; Maninka *sán* [568]; Xasonka *sán-ji* [568]; Mandinka *sán-ji* [568]; Koranko *séki* [568]; Lele *sányi* [568]; Mogofin *sángei* [568]; Kakabe *sángi* [568]; Kono *sóna* [565]; Vai *sóná* [565]; Soso *tùnè* [565]; Jalonke *tùnè* [565]; Jogo *cíí* [567]; Jeri *kí, kii, cí, cíí* [567]; Numu *kyí* [567]; Mende *njá* ('water; rain') [831]; Loko *njà* ('water; rain') [831]; Bandi *njeí* ('water; rain') [831]; Looma *tíne, tune* [565]; Kpelle-L *tīnā, tūnā* [565]; Kpelle-G *tulo* [565]; Bokobaru *lòu* [564]; Boko *lèu* [564]; Busa *lègy* [564]; Lebir *lá* [562]; San *lá* [562]; Kyenga *lá* [562]; Shanga *là* [562].

red: Dan-G *z̄p̄d̄ē* [573]; Dan-B *n̄ȳāz̄á* [574]; Dan-K *tá* [572]; Tura *téé* [572]; Mano *zóló* [573]; Guro *s̄ɔ̀* [573]; Yaure *té-dé* [572]; Gban *gbá-táá* [572]; Mwan *té* [572]; Wan *té* [572]; Beng *téé* [572]; Kpeen *kée* [584]; Kpan *xeey* [584]; Dzuun *xèò* [584]; Duun *xio-ma* [584]; Seenku *sèn* [572]; Banka *xie* [584]; Jo *kilè* [584]; Bobo *pēnēn, pēnē* [583]; Bozo-J *tɔmɔ* [582]; Bozo-K *tɔŋɔ* [582]; Bozo-T *tɔmɔ* [582]; Soninke *dumbé* [581]; Mau *yw̄èŋ* [579]; Jula-K *wùlén* [579]; Dafing *wlé, lwé* [579]; Bamana *bilé* [579]; Maninka *wùlé* [579]; Xasonka *wùle* [579]; Mandinka *wùle(ŋ)* [579]; Koranko *wùle* [579]; Lele *wùlen* [579]; Mogofin *wùlé* [579]; Kakabe *wùlen* [579]; Kono *yàwa, yàa* [580]; Vai *jàá* [580]; Soso *gbeli* [577]; Jalonke *gwééli* [577]; Jogo *tàná* [572]; Jeri *tàna* [572]; Numu *térq̄màn* [572]; Mende *kpólú* [577]; Loko *kpòù* [577]; Bandi *kpèá(ŋ)* [577]; Looma *kpóei, kpoye(g)* [577]; Kpelle-L *kpōlū* [577]; Kpelle-G *kpɔlu* [577]; Bokobaru *té-aá* [572]; Boko *tééé* [572]; Busa *téé* [572]; Lebir *búmbweenda* [575], *ɲwáada* [576]; San *tq* [572]; Kyenga *té* [572]; Shanga *cé* [572].

road: Dan-G *z̄īā* [585]; Dan-B *z̄īā* [585]; Dan-K *z̄ìŋ* [585]; Tura *záà* [585]; Mano *zī* [585]; Guro *zī* [585]; Yaure *sì* [585]; Gban *dí* [585]; Mwan *zī* [585]; Wan *zī* [585]; Beng *zrē* [585]; Kpeen *sé* [585]; Kpan *še* [585]; Dzuun *shí* [585]; Duun *sira-na* [585]; Banka *šiq* [585]; Jo *sáá* [585]; Bobo *sògò* [587]; Bozo-J *sɛ* [587]; Bozo-K *kinó* [585]; Bozo-T *syen* [587]; Soninke *killi / killé* [585]; Mau *šyáá* [585]; Jula-K *sira* [585]; Dafing *sirà* [585]; Bamana *sira* [585]; Maninka *sila* [585]; Xasonka *sila* [585]; Mandinka *sila* [585]; Koranko *kéla, kila* [585]; Lele *kila* [585]; Mogofin *kila* [585]; Kakabe *kila* [585]; Kono *cia* [585]; Vai *kiá* [585]; Soso *kiráá* [585]; Jalonke *kiráá* [585]; Jogo *cilí* [585]; Jeri *kilí/ó, kulu/ó* [585]; Numu *kilí* [585]; Mende *pèlé* [586]; Loko *pèlè* [586]; Bandi *pèlé* [586]; Looma *pele* [586]; Kpelle-L *pērē* [586]; Kpelle-G *pēlē* [586]; Bokobaru *zé* [585]; Boko *zé* [585]; Busa *zé* [585]; Lebir *záá, zaa* [585]; San *zii* [585]; Kyenga *zéle* [585]; Shanga *zēē* [585].

root: Dan-G *yū* [592]; Dan-B *yū* [592]; Dan-K *mù* [592], *wú* (?) [593]; Tura *sù* [596]; Mano *gáná* [589]; Guro *sū-nī* [596]; Yaure *séí* [596]; Gban *sá* [592]; Mwan *sé-né ~ sé-né* [596]; Wan *sēnī* [596]; Beng *nīŋ* (?) [596]; Kpeen *gɔ̀ɔ̀-gú* [596]; Kpan *yü* [596]; Dzuun *jíí* [596]; Duun *gii* [596]; Banka *ji-ra* [596]; Jo *kàn, kànà* (?) [594]; Bobo *sùlù* [596]; Bozo-J *sɔndo* (?) [592]; Bozo-K *dirì* [596]; Bozo-T *sij* [596]; Soninke *n̄ime* [597]; Mau *d̄iŋ* [596]; Jula-K *lilin* [596]; Dafing *lwiitwii* [596]; Bamana *dili* [596]; Maninka *lilí* [596]; Xasonka *lilli* [596]; Mandinka *sulu* [596]; Koranko *sulu* [596]; Lele *sulu* [596]; Mogofin *lilí*

[596]; Kakabe *liili* [596]; Kono *súu* [596]; Vai *súú* [596]; Soso *sanke* (?) [592]; Jalonke *fànsá* [595]; Jogo *kún* [594]; Jeri *nòŋ* [594]; Numu *kún* [594]; Mende *hápé* [592]; Loko *hápè* [592]; Bandi *sàpè* [592]; Looma *sape* [592], *nata* [593]; Kpelle-L *sāmē* [592]; Kpelle-G *hamu* [592]; Bokobaru *kasqa* [588]; Boko *zíná* [596]; Busa *zíní* [596]; Lebir *gòmasi* [590]; San *gòon* [590]; Kyenga *misa* [591]; Shanga *kaàsi* [588].

round: Dan-G *dřřdřř* [599]; Dan-B *dřřúđřřú*, *dřřđř* [599], *klóókló* [602]; Dan-K *glúúúglúú* [602]; Tura *giniyè* [602]; Mano *kélékèlè* [602]; Guro *lilí* [599], *kèlèè* [602]; Yaure *kulú-gulú* [602]; Gban *lilè* [599], *giùù*, *gbiùù* [602]; Mwan *trúú* [599]; Wan *glíú* [602]; Beng *klilí* [602]; Kpan *xɔɔri-ma* [608]; Dzuun *kòòró* (‘round off’) [608]; Duun *kúšíromq* [610]; Banka *xworo-ma* [608]; Jo *kiùlímàq* [608]; Bozo-J *gumę* [609]; Bozo-K *kori* [608]; Bozo-T *koorina* [608]; Soninke *buudi* [601]; Bamana *kulukutú* [602]; Maninka *tindan* [605]; Xasonka *godongoda* [604], *durunderendij* [607]; Mandinka *múruu*, *múruj* [606]; Koranko *tintan-ni* [605]; Lele *latindeen* [605]; Mogofin *rádigirindén* [599]; Kono *tindan-den* [605]; Soso *digilin* [599], *gundu* [604]; Jogo *kulúkulú* [602]; Jeri *kórondó*, *kóróró* [603]; Numu *kùrùkùrúá* [602]; Mende *kélékélé* [602]; Bandi *kèléyèlè* [602]; Looma *keleyele(g)* [602]; Kpelle-L *kérékéré* [602]; Kpelle-G *kiliyilij* [602]; Bokobaru *botogo* [598]; Boko *boolo* [598]; Busa *bunduru* [598], *búkutú* [601]; San *luurii* [599]; Kyenga *lokoto* [600]; Shanga *lókoci* [600].

sand: Dan-G *yěě* [613]; Dan-B *yáh*, *yáh* [613]; Dan-K *yěě* [613]; Tura *kèè* [613]; Mano *yěě-sélé* [613]; Guro *yēnè* [613]; Yaure *yèrè* [613]; Gban *yè* [613]; Mwan *yrè* [613]; Wan *wéj* [613]; Beng *juùn-móntré* (?) [614]; Kpeen *sísijé* [613]; Kpan *šęšę* [613]; Dzuun *céncén* [613]; Duun *šig-di* [613]; Seenku *kyijē* [613]; Banka *cècè* [613]; Jo *nèññ* [613]; Bobo *yègé*, *bàn-yēgē*, *bàmà-yēgē* [613]; Bozo-J *syè* [616]; Bozo-K *tó* [618]; Bozo-T *xamu* [617]; Soninke *sèyínè*, *saye* [616]; Mau *cè* [613]; Jula-K *kèngén* [613]; Dafing *cié* [613]; Bamana *cě*, *cěcě* [613]; Maninka *kijé*, *tijé*, *cijé* [613]; Xasonka *kèpe* [613]; Mandinka *kèpe*, *kèpekepe* [613]; Koranko *cèncèn*, *kènkèn* [613]; Lele *cèeka*, *kèeka* (?) [613]; Mogofin *mènépe* [613]; Kakabe *kèpeképe* [613]; Kono *tèku* [615]; Vai *kèné* [613]; Soso *meyepi*, *meeji* [613]; Jalonke *qéne* [613]; Jogo *nàgàn* [613]; Jeri *kpè*, *kpèè*, *kpèñ* [613]; Numu *juingán* [613]; Mende *jàná* [613]; Loko *nèà* [613]; Bandi *jiyè* [613]; Looma *yáqza(g)*, *náyayè(g)* [613]; Kpelle-L *ñēyā* [613]; Kpelle-G *ñeia* [613]; Bokobaru *bùsù-yqaté* [611]; Boko *yufqa* [612]; Busa *bùsù-yáté* [611]; Lebir *juintáa* [613]; San *ɲɔɔ-sɔ* [614]; Kyenga *yaaté-fu* [611]; Shanga *iyqa* [613].

say: Dan-G *př* [619]; Dan-B *př* [619]; Dan-K *př* [619]; Tura *pé* [619]; Mano *gèè* [620]; Guro *fī* [619]; Yaure *ví* (?) [619], *tj* [630]; Gban *kè* [620]; Mwan *pē / pé* [619]; Wan *pé* [619]; Beng *pē* [619]; Kpeen *dè* [623]; Kpan *žee-ma* [623]; Dzuun *dè* [623]; Duun *že* [623]; Seenku *dyù* [623]; Banka *je* [623]; Jo *bì* (?) [619], *tú* (?) [628]; Bobo *dē* [623], *yīrā* [632]; Bozo-J *se* [631]; Bozo-K *ši* [631]; Bozo-T *se* [631]; Soninke *tī / tiini* [630]; Mau *fš* [619]; Jula-K *fš*, *fwó* [619]; Dafing *fš* [619]; Bamana *fš* [619]; Maninka *fš* [619]; Xasonka *fó* [619]; Mandinka *fó* [619]; Koranko *fš* [619]; Lele *fš* [619]; Mogofin *fš* [619]; Kakabe *fš* [619]; Kono *fš* [619]; Vai *fš* [619]; Soso *fala* [629], *tií* [630]; Jalonke *fàlá* [629]; Jogo *re* [623]; Jeri *tógo* [628], *tè* [630]; Numu *ɬkɔ* [628]; Mende *ndé* [623], *yè* [624]; Loko *ké* [625]; Bandi *ndè* [623], *ke* [625]; Looma *ké* [625], *bo(g)* [626]; Kpelle-L *yēē* [624], *ké* [625], *bó* [626], *yepè* [627]; Kpelle-G *ke* [625], *bo* [626], *hvaa* [629]; Bokobaru *be* [621]; Boko *me* [622]; Busa *pi* [619]; Lebir *pi ... m* [619]; San *pé* [619]; Shanga *hi* [619].

see: Dan-G *yř* [633]; Dan-B *yř* [633]; Dan-K *yè* [633]; Tura *yé* [633]; Mano *gè* [634]; Guro *yī* [633]; Yaure *yì* [633]; Gban *yè* [633]; Mwan *yē / yé* [633]; Wan *é* [633]; Beng *yè* [633]; Kpeen *je* [633]; Kpan *jaa-ma* [633]; Dzuun *jáá* [633]; Duun *si* [639]; Seenku *dyò* [633]; Banka *ja* [633]; Jo *jà* [633]; Bobo *zá* [633]; Bozo-J *kai* [637]; Bozo-K *he* [637]; Bozo-T *xay* [637]; Soninke *yingi / -ni* (?) [633], *wàrí* [635]; Mau *yé* [633]; Jula-K *yé* [633]; Dafing *yè* [633]; Bamana *yé* [633]; Maninka *yé*, *yé* [633]; Xasonka *je* [633]; Mandinka *je* [633]; Koranko *yé* [633]; Lele *yén* [633]; Mogofin *yé* [633]; Kakabe *yén* [633]; Kono *yé* [633]; Vai *fě'é* [638]; Soso *tó* [636]; Jalonke *tóó* [636]; Jogo *ji* [633]; Jeri *ji* [633]; Mende *tó* [636]; Loko *tó* [636]; Bandi *tó* [636]; Looma *ka* [637]; Kpelle-L *káa* [637]; Kpelle-G *ka* [637]; Bokobaru *ye* [633]; Boko *ye* [633]; Busa *ye* [633]; Lebir *yí* [633]; San *yē* [633]; Kyenga *wa* [635]; Shanga *wa* [635].

seed: Dan-G *gā* [649]; Dan-B *gā* [649]; Dan-K *gā* [649]; Tura *wéé* [645]; Mano *wélé* [645], *sīī* [648]; Guro *wólé* [645]; Yaure *wélé* [645]; Gban *wlí*, *wlí* [645]; Mwan *bē* [643]; Wan *ḡŋ* [645]; Beng *wlé* [645]; Kpan *gee-zij* [647]; Dzuun *gè-dziin* [647]; Duun *dī* (‘child’) [647]; Seenku *dè* [647]; Banka *di-na* [647]; Jo *dán* [642]; Bobo *tògò* [651]; Bozo-J *suuro* [644], *sí* [648]; Bozo-T *ɔɔri* [644]; Soninke *sòxòodé* [644]; Mau *sí* [648]; Jula-K *sí* [648], *kisè* [650]; Dafing *sí* [648], *césé* [650]; Bamana *sí* [648], *kisè* [650]; Maninka *sí* [648], *kisè* [650]; Xasonka *sí* [648], *kèse* [650]; Mandinka *kèse* [650]; Koranko *kólo* [645]; Lele *sú* [644], *kólo* [645]; Mogofin *kóló* [645]; Kakabe *súgu* [644]; Kono *kiu* [645]; Vai *siú* [644]; Soso *san-si* [648]; Jalonke *sàn-si* [648]; Jogo *súgu* [644], *dé* (‘child’) [647]; Jeri *dí* (‘child’) [647], *sí* [648]; Mende *húwú* [644]; Loko *káé*, *káái* [646]; Bandi *súuwú* [644]; Looma *suwu* [644]; Kpelle-L *kāyā* (?) [646]; Kpelle-G *huyu* [644]; Bokobaru *pó* [640]; Boko *pó* [640]; Busa *pó* [640]; Lebir *yáá* [641]; San *nàn* [642]; Kyenga *bèè* [643]; Shanga *bùì* [643].

sit: Dan-G *yà* [655]; Dan-B *yá* [655]; Dan-K *yà* [655]; Tura *yàà* [655]; Mano *yà* [655]; Guro *yā-nq* [655]; Yaure *yā-nq* [655]; Gban *yè* [655]; Mwan *yā-là* [655]; Wan *yá* [655]; Beng *yrä* [655]; Kpeen *tògò* [656]; Kpan

tuu-ma [656]; Dzuun *tsùù* [656]; Duun *to'o* [656]; Seenku *tùgù, tùgò / tūgū* [656]; Banka *tugu* [656]; Jo *sóò* [656]; Bobo *tàngà* [656]; Bozo-J *pɔ* [657]; Bozo-K *fɔɔ* [657]; Bozo-T *pay* [657]; Soninke *taaxu* [656]; Mau *sii* [654]; Jula-K *sigí* [654]; Dafing *sí* [654]; Bamana *sigí* [654]; Maninka *sii* [654]; Xasonka *sigi* [654]; Mandinka *sii* [654]; Koranko *sigi, sii* [654]; Lele *sii* [654]; Mogofin *sigí* [654]; Kakabe *sigi* [654]; Kono *sii* [654]; Vai *sii* [654]; Soso *dàxó* [656]; Jalonke *dàqó* [656]; Jogo *yàgá* [655]; Jeri *sàga* (?) [654]; Numu *yákò* [655]; Mende *hèí* [654]; Loko *hèi(η)* [654]; Bandi *sèi* [654]; Looma *sei* [654]; Kpelle-L *sēē* [654]; Kpelle-G *he* [654]; Bokobaru *vèé* [653]; Boko *zɔlé* [652]; Busa *vuté* [653]; Lebir *nintá* [652]; San *giq-tán* [652]; Kyenga *ɣaté* [652]; Shanga *zutee* [652].

skin: Dan-G *kwī* [658]; Dan-B *kwī* [658]; Dan-K *kpí* [658]; Tura *kwii* [658]; Mano *kii* [658]; Guro *kólé* [658]; Yaure *kóle* [658]; Gban *lǒě* [659]; Mwan *flé* [660]; Wan *klé* [658]; Beng *cī* [658]; Kpeen *fērúñ* [660]; Kpan *sy* [664]; Dzuun *tsùn* [664]; Duun *tunɔ* [664]; Seenku *tsùn* [664]; Banka *tuna* [664]; Jo *pólí* [660]; Bobo *kòn* [658]; Bozo-J *korɔ* [658]; Bozo-K *kòròñɔ* [658]; Bozo-T *koron* [658]; Soninke *fàté* [663]; Mau *gbòò* [658]; Jula-K *gbòlò* [658]; Dafing *gòo* [658]; Bamana *gòlò, wòlò* [658]; Maninka *gbòlò* [658]; Xasonka *gùlu* [658]; Mandinka *kùlu* [658]; Koranko *gbòlo* [658]; Lele *kpòlé* [658]; Mogofin *bole* [658]; Kakabe *gbòlo* [658]; Kono *gbòo* [658]; Vai *kpò'ò* [658]; Soso *kiri* [658]; Jalonke *kiri* [658]; Jogo *kpòrò* [658]; Jeri *kpùru* [658]; Numu *kpòrò* [658]; Mende *kòlɔ* [658]; Loko *kò* [658]; Bandi *kòlɔ(η)* [658]; Looma *kòlɔ* [658]; Kpelle-L *kòlɔ* [658]; Kpelle-G *kòlɔ* [658]; Bokobaru *báá* [662]; Boko *báá* [662]; Busa *bárá* [662]; Lebir *kyí* [658]; San *ki* [658]; Kyenga *se* [661]; Shanga *sēé* [661].

sleep: Dan-G *yī* [665]; Dan-B *yī* [665]; Dan-K *yī* [665]; Tura *yí* [665]; Mano *yī, yí* [665]; Guro *yī* [665]; Yaure *yí* [665]; Gban *yí* [665]; Mwan *yī* [665]; Wan *yí* [665]; Beng *yī* [665]; Kpan *sii-ma* [665]; Dzuun *cii* [665]; Duun *ca'ara* [665]; Seenku *kyirè* [665]; Banka *xie* [665]; Jo *nini* [665]; Bobo *nì* [665]; Bozo-J *kumu* [666]; Bozo-K *kòmú* [666]; Bozo-T *kyemu* [666]; Soninke *xen-qo* [665]; Mau *sùnɔɔ* [665]; Jula-K *sindɔgɔ* [665]; Dafing *sɔ* [665]; Bamana *sù-nɔgɔ* [665]; Maninka *sù-nɔɔ, sù-nɔɔ* [665]; Xasonka *sinnoxo* [665]; Mandinka *sii-noo* [665]; Koranko *kinoo, kindɔge* [665]; Lele *kinɔɔ* [665]; Mogofin *kundɔgɔ, kindɔgɔ* [665]; Kakabe *kinɔgɔ* [665]; Kono *cinɔ, ci* [665]; Vai *kí* [665]; Soso *xi* [665]; Jalonke *qii* [665]; Jogo *nì-sá* [665]; Jeri *nin-bá* [665]; Numu *nì-sá* [665]; Mende *njií* [665]; Loko *njií* [665]; Bandi *nii* [665]; Looma *nii* [665]; Kpelle-L *nii* [665]; Kpelle-G *yi* [665]; Bokobaru *ii* [665]; Boko *yi* [665]; Busa *yi* [665]; Lebir *nin-tiim* [665]; San *niiu* [665]; Kyenga *yi* [665]; Shanga *i* [665].

small: Dan-G *sēē-ná* [673]; Dan-B *séé, séé* [673]; Dan-K *sēē-ná* [673]; Tura *pīi-né* [667]; Mano *pééti* [667]; Guro *fīné* [667], *cèné* (?) [688]; Gban *zɔɔ* [669], *kpà* [670]; Mwan *fēétrɔ* [667]; Wan *wlɛñ* [671], *kpɛlɛñ* [672]; Beng *kló* [681]; Kpeen *fúwé* [668]; Kpan *thyagani* [688]; Dzuun *myéyn* [675], *ɔyɛn* [687], *kinàànin* [688]; Duun *mij* (?) [675], *deuma* [687], *kuumei* [689]; Seenku *mùēen* [675]; Banka *niiri* [685], *kpekpɛ'ɛni* [686]; Jo *pyàà* [667]; Bobo *lóró* [683], *sálaló* [684]; Bozo-J *duɔ* [682]; Bozo-K *lɔɔ* [682]; Bozo-T *muyɔn* [675], *lɔn* [682]; Soninke *ròxé* [682]; Mau *mèsɛ* [675], *lɔɔ* [682]; Jula-K *fítinín* [667], *mèsén* [675], *dɔgɔ* [682], *nciní* [688]; Dafing *méséí* [675]; Bamana *fiti-ni* [667], *mìsɛ* [675], *dɔɔ* [682]; Maninka *mìsɛ* [675], *dɔɔ* [682]; Xasonka *mèse* [675], *dóxo* [682]; Mandinka *mésɛñ* [675], *dóoyariñ* [682]; Koranko *mènsɛ* [675], *dóo, dɔgɔ* [682]; Lele *mèse* [675], *dɔɔ* [682]; Mogofin *mèsɛ* [675]; Kakabe *dɔgɔ* [682]; Kono *dɔɔ* [682]; Vai *mésémésé* [675], *dɔɔ* [682]; Soso *xurun* [681]; Jalonke *quíurún* [681]; Jogo *fyáán* [667], *dɔgɔ* [682]; Jeri *dɔgá* [682], *diri* [683]; Numu *fìàà* [667], *dèré* [683]; Mende *kulò* [681]; Loko *kòlò(η)* [681]; Bandi *kulò* [681]; Looma *péné, pélevèle(g), penɛ(g)* [667], *kolo* [681]; Kpelle-L *pèlè* [667], *kòlò* [681]; Kpelle-G *pele* [667], *kolo* [681]; Bokobaru *fété* [667], *néngó* [674]; Boko *fénéná* [667], *néná* [674]; Busa *fítí* [667]; Lebir *pòorè, pwèéré* [667], *mí* [675]; San *bisin* [676], *boeene* [677], *boro* [678]; Kyenga *ziu* [679]; Shanga *mògòní* [680].

smoke: Dan-G *pēñ-tēē* [692]; Dan-B *sís-táá* [692], *síj* (?) [693]; Dan-K *sē-gbē* [691]; Tura *gbēē* [691]; Mano *gbēi* [691]; Guro *cé-gwēi, gōi* [691]; Yaure *té-kwēi* [691]; Gban *gīmí* [691]; Mwan *té-gblēē* [691]; Wan *gbàní* [691]; Beng *gbíá* [691]; Kpan *sa-tui* [694]; Dzuun *sà-tsúru* [694]; Duun *sa-tuu* [694]; Banka *sa-sula* [694]; Jo *cikilí* [696]; Bobo *tiòrò* [694]; Bozo-J *titi* [694]; Bozo-K *sísí* (< Bamana?) [-2]; Bozo-T *piran* [695]; Soninke *tiidi / tiidé* [694]; Mau *sisi* [693]; Jula-K *sisi* [693]; Dafing *sisi* [693]; Bamana *sisi* [693]; Maninka *sisi* [693]; Xasonka *sisi* [693]; Mandinka *sisi* [693]; Koranko *sisi* [693]; Lele *sisi* [693]; Mogofin *sisei* [693]; Kakabe *sisi* [693]; Kono *sisi* [693]; Vai *sisi* [693]; Soso *turi* [694]; Jalonke *tùqùti* [694]; Jogo *sisi* [693]; Jeri *sigí, sigu, sugu* [693]; Numu *sisi* [693]; Mende *ndulú* [694]; Loko *nduri* [694]; Bandi *ndulí* [694]; Looma *duulu, duwulu* [694]; Kpelle-L *lūri* [694]; Kpelle-G *lulu* [694]; Bokobaru *té-súkpé* [690]; Boko *súélé* [690]; Busa *té-súkpé* [690]; Lebir *singa* [693]; San *te-maa* [690]; Kyenga *tékpésé* [690]; Shanga *tesongi* [690].

stand: Dan-G *dɔ* [702]; Dan-B *dɔ* [702]; Dan-K *dɔ* [702]; Tura *dɔ* [702]; Mano *dɔ* [702]; Guro *dɔ* [702]; Yaure *tò* [702]; Gban *dò* [702]; Mwan *dɔ* [702]; Wan *dì-nà* [703]; Beng *dɔ* [702]; Seenku *tigi* [1012]; Bobo *tà* [702]; Bozo-J *táá* [702]; Bozo-T *táá* [702], *sii, kii* [704]; Soninke *sigi / sikki* [704]; Mau *lɔ* [702]; Jula-K *lò, lwò* [702]; Dafing *lɔ* [702]; Bamana *jɔ* [702]; Maninka *lɔ* [702]; Xasonka *lò* [702]; Mandinka *lòo* [702]; Koranko *sò, sò* [702]; Lele *sò* [702]; Mogofin *sɔ* [702]; Kakabe *lɔ* [702]; Kono *sò* [702]; Vai *só* [702]; Soso *ti* [1012]; Jalonke *tii*

[1012]; Jogo *yǎ(n)* [699], *jǎ(n)* [702]; Jeri *sò* [702]; Numu *yónj / yónè* [699], *dó / dórè* [702]; Mende *tó* [702]; Loko *tò(ŋ)* [702]; Bandi *to, to(ŋ)* [702]; Looma *to(g)* [702]; Kpelle-L *tɔ̄* [702]; Kpelle-G *tɔ* [702]; Bokobaru *ze* [697]; Boko *ze* [697]; Busa *ze* [697]; Lebir *gyim |gini|* [698]; San *yoo* [699]; Kyenga *kpeté* [700]; Shanga *gɔɔ* [701].

star: Dan-G *súsɔ̄j* [705]; Dan-B *súsɔ̄j* [705]; Dan-K *fij̄ ~ fīj* [708]; Tura *mééé* [709]; Mano *siéŋle* [705]; Guro *mānē gwēnén* [709]; Yaure *mènè cènè* [709]; Gban *nú-yǎná* [707]; Mwan *mlē-klēnē* [709]; Wan *pèŋpèŋ* [706]; Beng *mān-léj* [709]; Kpeen *sòwá* [705]; Kpan *sɔɔy* [705]; Dzuun *swéyn* [705]; Duun *kwo'o-dini* (“child of moon”) [718]; Seenku *káalle* [718]; Banka *sɛsɛ-dina* [705]; Jo *sòrnij̄* [705]; Bobo *mùmúlù* [717]; Bozo-J *tonon* [716]; Bozo-K *twé* [716]; Bozo-T *tɔɔ* [716]; Soninke *saana / sàané* [705]; Mau *lòlò, lòd-dyèŋ* [715]; Jula-K *lòlò* [715]; Dafing *lòlò* [715]; Bamana *dòlò, lòlò* [715]; Maninka *lòlò* [715]; Xasonka *lòolo* [715]; Mandinka *lòoloo* [715]; Koranko *lòlo* [715]; Lele *tànbuyá, tánbura* [712]; Mogofin *lòolo* [715]; Kakabe *lòolo* [715]; Kono *tónboáká* [712]; Vai *tóomà* (?) [712]; Soso *tunbu* [712]; Jalonke *sàré* (?) [705]; Jogo *tétènkòlà* [714], *lòlón* [715]; Jeri *lòlò* [715]; Numu *tétènkólá* [714]; Mende *túnbékà* [712]; Loko *téléga* [712]; Bandi *tèmuleyá* [712]; Looma *sɔmɔ deka* (“coal of a torch”) [1010]; Kpelle-L *pémélèŋ-kāō* [706]; Kpelle-G *peleŋ, pemeneŋ-gɔw* [706]; Bokobaru *sòsòné* [705]; Boko *saana* [705]; Busa *sùsùnè* [705], *sura* [710]; San *leere* [711]; Kyenga *sólò* [710]; Shanga *sɔsɔni* [705].

stone: Dan-G *gwà* [719]; Dan-B *gúŋ* [719]; Dan-K *gùŋ* [719]; Tura *gwèè* [719]; Mano *gèlè* [719]; Guro *gɔ̄lè* [719]; Yaure *kɔ̄lè* [719]; Gban *gwà* [719]; Mwan *vlé* [719], *kpōtī* [721]; Wan *klé* [719], *kpótī* [721]; Beng *glé* [719]; Kpeen *kpèè* [719]; Kpan *kpai* [719]; Dzuun *kpáá* [719], *kóóná* [721]; Duun *kwaj̄-ni* [719]; Seenku *kùnùn* [721]; Banka *kpá'a-ra* [719]; Jo *kfɔ* [721]; Bobo *dībī* [729], *dɔgò* [730]; Bozo-J *sire* [728]; Bozo-K *fára-húrúŋ* (< Bamana) [727]; Bozo-T *sere* [728]; Soninke *gidí/e* [719]; Mau *bèè* [725], *káwá* [726], *fáá* [727]; Jula-K *bèré* [725], *fára* [727]; Dafing *béé* [725], *kàbà* [726]; Bamana *bélé* [725], *kába* [726], *fára* [727]; Maninka *bèré* [725], *kába* [726]; Xasonka *bère* [725], *fáta* [727]; Mandinka *bère* [725]; Koranko *kùru* [720]; Lele *gbéré* [719], *kùru, kuyi* [720]; Mogofin *kùrèi* (def.) [720]; Kakabe *kùru* [720]; Kono *sè* [724]; Vai *sèŋ* [724]; Soso *gème* [723]; Jalonke *gèmé* [723]; Jogo *kpíin* [721]; Jeri *kùngu* [722]; Numu *kpíin* [721]; Mende *kòtù* [721]; Loko *kòtù* [721]; Bandi *kòtù, kòti* [721]; Looma *kwótí, kwotu* [721]; Kpelle-L *kōnī* [721]; Kpelle-G *kweni* [721]; Bokobaru *gbè* [719]; Boko *gbè* [719]; Busa *gbè* [719]; Lebir *gyàa, dyàa* [719]; San *kòlò* [720]; Kyenga *gbèlè* [719]; Shanga *gwèè* [719].

sun: Dan-G *yǎj̄* [732]; Dan-B *lāj̄* [732]; Dan-K *yéj̄* [732]; Tura *yáá* [732]; Mano *yénē, ñlé* [732]; Guro *yīlī* [732]; Yaure *yí-dé* [732]; Gban *yè-vè, yè-è* [732]; Mwan *yērè-té* [732]; Wan *yè-té* [732]; Beng *yī-gbié* [732]; Kpeen *béè* [740]; Kpan *boi* [740]; Dzuun *bwèy* [740]; Duun *boj̄* [740]; Seenku *són-tá* (“sky fire”) [737]; Banka *bi-ra* [740]; Jo *súnj̄, sún-ta* [737]; Bobo *sín* ? [736]; Bozo-J *kubu* [738]; Bozo-K *kwò* [738]; Bozo-T *kuu-jɔ* [738]; Soninke *kíyè* [739]; Mau *tèè* [735]; Jula-K *tèrè* [735]; Dafing *tée* [735]; Bamana *tìlè* [735]; Maninka *tèlè* [735]; Xasonka *tìli* [735]; Mandinka *tìli* [735]; Koranko *tèlè* [735]; Lele *tèlè* [735]; Mogofin *tèlè* [735]; Kakabe *tèli* [735]; Kono *tèè* [735]; Vai *tèé* [735]; Soso *soge* [736]; Jalonke *sógé* [736]; Jogo *télé* [735]; Jeri *tèlè, tèli* [735]; Numu *télé, talé* [735]; Mende *fólò* [734]; Loko *fɔ* [734]; Bandi *fólò* [734]; Looma *folo* [734]; Kpelle-L *fólo* [734]; Kpelle-G *hvolo* [734]; Bokobaru *yɔfáté* [731]; Boko *yíaté* [731]; Busa *yifá* [731]; Lebir *wòsò, wùsò* [733]; San *wasa* [733]; Kyenga *yásɔ* [733]; Shanga *ísɔɔ* [733].

swim: Dan-G *yí ká* (“do water”) [741]; Mwan *yí-ká* (“cut water”) [741]; Wan *yí kɔ* (“cut water”) [741], *tij̄ lɔ* [743]; Beng *dòŋ* [742]; Kpan *nɔmu* [753]; Duun *mij̄-na* [753]; Seenku *nònnà* [753]; Banka *nene* [753]; Jo *teri* [757]; Bobo *nīnā* [756]; Bozo-J *nuȳ, nij̄* [756]; Bozo-K *nuȳ* [756]; Bozo-T *nij̄* [756]; Soninke *banbe* [755], *xini* [756]; Jula-K *nómí* [753]; Bamana *néwuy, nɔ, nó* [753]; Maninka *néwuy, námuy, náwuy* [753]; Xasonka *nèwuy* [753]; Mandinka *nóŋ, nóo, néwuy* [753]; Koranko *bá rá matège* [752]; Lele *bá matee* (?) [752]; Mogofin *jei masa* (< Soso?) [750], *fúulè* [751]; Kono *nòè cé* [753]; Vai *jí bì'á* (“catch water”) [754]; Soso *ye masa* [750]; Jalonke *sáfáá* [749]; Numu *bùán kyó* (“swimming-do”) [748]; Loko *ndembu* [745]; Looma *dase(g)* (?) [744]; Kpelle-L *yá ná kélè* (“roam through water”) [741]; Kpelle-G *ya kele* [741]; Bokobaru *í kpá* [741], *dàdaá o* [744]; Boko *kpáa o* [741]; Busa *í kpá* [741]; Lebir *dá* [744], *pí zé* [746]; San *mu baa* [747]; Kyenga *yí kpá* [741]; Shanga *dè-i* [744].

tail: Dan-G *wèj̄* [760]; Dan-B *wèj̄, wèj̄* [760]; Dan-K *wèj̄* [760]; Tura *wíi* [760]; Mano *wɔ̄ȳ* [760]; Guro *wòlì* [760]; Yaure *wèi* [760]; Gban *wí* [760]; Mwan *wèè* [760]; Wan *wèŋ* [760]; Beng *pínáŋ* [762]; Kpan *kpla* [760]; Dzuun *kplá* [760]; Duun *kura* [760]; Seenku *kúrà, kúrwà* [760]; Banka *kpla-na* [760]; Jo *jij̄* [761]; Bobo *pègè* [762]; Bozo-J *pié, piq* [762]; Bozo-K *hóri* [760]; Bozo-T *piŋe* (?) [762]; Soninke *xoqqi / xóqqé* (?) [760]; Mau *kwó* [760]; Jula-K *kó, kwó* [760]; Dafing *kwò* [760]; Bamana *kú* [760]; Maninka *kó* [760]; Xasonka *xú* [760]; Mandinka *fèje* [762]; Koranko *kú* [760], *fj̄é* [762]; Lele *kɔɔ* [760]; Mogofin *fj̄nè* [762]; Kakabe *fj̄nɔ* [762]; Kono *fèe* [762]; Vai *fèjè* [762]; Soso *xúlí* [760]; Jalonke *qúlí* [760]; Jogo *já(n)* [761]; Jeri *ján* [761]; Numu *gyàn* [761]; Mende *ngólí* [760]; Loko *ngó* [760]; Bandi *ngò(ŋ)* [760]; Looma *ŋj̄(g)* [760]; Kpelle-L *wónj̄* [760]; Kpelle-G *wonj̄* [760]; Bokobaru *vláá* [758]; Boko *vlá* [758]; Busa *vlá* [758]; Lebir *mòm |mòni|* [759]; San *mui* [759]; Kyenga *mɔnɔ* [759]; Shanga *buyu* [759].

that: Dan-G *bā* [764]; Dan-B *bā* [764]; Tura *lāà* [765]; Mano *e yía* [1009]; Guro *bē* [764]; Mwan *bē* [764], *lāālē* [765], *nū* [773]; Kpan *nie* [773]; Dzuun *nii* [773]; Duun *nydulu* [773]; Banka *ny-kere* [773]; Jo *kàà* [771]; Bobo *yóó* [772]; Bozo-J *kwɔ* [771]; Bozo-T *ɲɔ* [770], *ku* [771]; Soninke *kén* [771]; Mau *ò* [770]; Jula-K *wò*, *ò* [770]; Dafing *wó* [770]; Bamana *wǒ*, *ǒ* [770]; Maninka *wǒ*, *ǒ* [770]; Xasonka *wò* [770]; Mandinka *wò* [770]; Koranko *wò* [770]; Lele *wo* [770]; Mogofin *wǒ* [770]; Kakabe *wó* [770]; Kono *ce* [771]; Vai *mè* [769], *kè* [771]; Soso *ná* [766]; Jalonke *ná* [766]; Jeri *wo* [770]; Numu *má* [769]; Mende *ná* [766]; Loko *nà* [766]; Bandi *nà* [766]; Looma *na* [766]; Kpelle-L *tí, ní-tí* [767]; Kpelle-G *ti* [767]; Bokobaru *kèkqà* [763]; Boko *bee* [764]; Busa *dire* [768]; Lebir *bí* [764]; San *bé* [764].

this: Dan-G *yā*, *yř* [777]; Dan-B *nè* [778]; Tura *ě* [778]; Mano *e kēa* [780]; Guro *kú* [780], *bē* [764]; Mwan *lèè* [778], *gè* [780]; Wan *kē* [780]; Beng *bì* [764]; Kpan *nie* [778]; Dzuun *nii* [778]; Duun *ny* [778]; Seenku *lèè* [778]; Banka *fó* [781]; Jo *kàà* [780]; Bobo *béè* [764]; Bozo-T *ku* [780]; Soninke *kée* [780]; Mau *mīññ* [779]; Jula-K *mìn* [779]; Dafing *mí* [779]; Bamana *nǚ* [778]; Maninka *nǚ, nǚ* [778], *mǝ* [779]; Xasonka *nìñ, nǚñ, nǚñ* [778]; Mandinka *nǚñ* [778]; Koranko *kè, cè* [780]; Lele *ke, ce* [780]; Mogofin *ke* [780]; Kakabe *ké, cé* [780]; Kono *ce* [780]; Vai *kē* [780]; Soso *yí* [777]; Jalonke *jí* [777]; Jeri *mē* [779]; Numu *má* [779]; Mende *jí* [777]; Loko *ndí* [777]; Bandi *sí* [777]; Looma *ni* [777]; Kpelle-L *ní* [777]; Kpelle-G *yi* [777]; Bokobaru *bee* [764]; Boko *bee* [764]; Busa *di* [775]; Lebir *bí* [764], *nàndv* [776]; San *bé* [764]; Kyenga *di* [775]; Shanga *idii* [775].

thou (you sg.): Dan-G *u* [783], *ba, br* [784]; Dan-B *u* [783], *ba, br* [784]; Dan-K *r* [783], *ba* [784]; Tura *i, ya* [783], *ba* [784]; Mano *i* [783], *ba, bi, bi* [784]; Guro *i* [783], *bi* [784]; Yaure *i, yia* [783]; Gban *εε* [783], *be* [784], *ni* [786]; Mwan *é* [783], *bí* [784]; Wan *é* [783], *bí-lāā* [784], *lā* [785]; Beng *mi* [784]; Kpeen *wò* [788]; Kpan *wo* [788]; Dzuun *wo* [788]; Duun *va* [788]; Seenku *à* [787]; Banka *wɔ* [788]; Jo *n* [782], *mí* [784]; Bobo *be* [784]; Bozo-J *q* [787]; Bozo-K *á* [787]; Bozo-T *an* [787]; Soninke *án* [787]; Mau *í* [783]; Jula-K *í* [783]; Dafing *èe* [783]; Bamana *í* [783]; Maninka *í* [783]; Xasonka *í* [783]; Mandinka *í* [783]; Koranko *í* [783]; Lele *í* [783]; Mogofin *í* [783]; Kakabe *í* [783]; Kono *í* [783]; Vai *í* [783]; Soso *í* [783]; Jalonke *í* [783]; Jogo *é* [783]; Jeri *í* [783]; Numu *é* [783]; Mende *bi* [784]; Loko *bi* [784]; Bandi *í* [783]; Looma *è, ì* [783]; Kpelle-L *é, í* [783]; Kpelle-G *é, í* [783]; Bokobaru *n* [782]; Boko *n* [782]; Busa *n* [782]; Lebir *ì, ibí* [783]; San *n* [782].

tongue: Dan-G *néè* [789]; Dan-B *nē* [789]; Dan-K *néè* [789]; Tura *né-ěé* [789]; Mano *nānà* [789]; Guro *nēnē* [789]; Yaure *nēnē* [789]; Gban *nē* [789]; Mwan *néné ~ nré* [789]; Wan *lēnī* (?) [789]; Beng *nàñáñ* [789]; Kpeen *nén* [789]; Kpan *ngɛw* [789]; Dzuun *nèyn* [789]; Duun *ne* [789]; Seenku *nèen* [789]; Banka *neena* [789]; Jo *náná-í, néné-í* [789]; Bobo *nìq* [789]; Bozo-J *neere* [789]; Bozo-K *nēñq* [789]; Bozo-T *myemu* [789]; Soninke *neeni / néené* [789]; Mau *nyéñ* [789]; Jula-K *nèn* [789]; Dafing *nīē* [789]; Bamana *nǝ, nē* [789]; Maninka *nǝ* [789]; Xasonka *nèñ* [789]; Mandinka *néñ* [789]; Koranko *nème* [789]; Lele *nèen* [789]; Mogofin *nén* [789]; Kakabe *nèn* [789]; Kono *nénén* [789]; Vai *nèé* [789]; Soso *nén* [789]; Jalonke *nèñ* [789]; Jogo *nèn-dé, lélè(n)* [789]; Jeri *nèn-dí* [789]; Numu *nén-dé* [789]; Mende *néé* [789]; Loko *nè* [789]; Bandi *nèè(ñ)* [789]; Looma *ne(g)* [789]; Kpelle-L *néñ* [789]; Kpelle-G *nəñ* [789]; Bokobaru *néné* [789]; Boko *lé-ná* [789]; Busa *néné* [789]; Lebir *nèém |nèèni* [789]; San *le-wɔn* [789]; Kyenga *néné* [789]; Shanga *léèni* [789].

tooth: Dan-G *sǚ* [791]; Dan-B *sǚ* [791]; Dan-K *só* [791]; Tura *sǚ* [791]; Mano *sǚ* [791]; Guro *só* [791]; Yaure *sóó* [791]; Gban *sǚ* [791]; Mwan *súé* [791]; Wan *své* [791]; Beng *sé* [791]; Kpeen *sǚ* [791], *nín-dé* [793]; Kpan *sǚñ* [791], *nǚ* [793]; Dzuun *só* [791], *nín* [793]; Duun *ni* [793]; Seenku *çèn* [793]; Banka *nǚna* [793]; Jo *nǚñáñ* [793]; Bobo *níní* [793]; Bozo-J *ni* [793]; Bozo-K *nínó* [793]; Bozo-T *nimin* [793]; Soninke *kambé* [795]; Mau *nǚñ* [793]; Jula-K *nín* [793]; Dafing *nì* [793]; Bamana *nǚ, ní* [793]; Maninka *nǚ* [793]; Xasonka *nǚñ* [793]; Mandinka *nǚñ* [793]; Koranko *nǚñ* [793]; Lele *nín* [793]; Mogofin *nǚ* [793]; Kakabe *nín* [793]; Kono *nǚ* [793]; Vai *nǚñ* [793]; Soso *nín* [793]; Jalonke *nín* [793]; Jogo *nín* [793]; Jeri *nín-di* [793]; Numu *nín* [793]; Mende *ngóngólù, ngóngóú* [794]; Loko *ngóngóú* [794]; Bandi *ndàa-ngólui* [794]; Looma *ni(g), yi(g)* [793]; Kpelle-L *nǚñ* [793]; Kpelle-G *ni* [793]; Bokobaru *swáá* [792]; Boko *swáá* [792]; Busa *sáká* [792]; Lebir *só* [791]; San *só* [791]; Kyenga *sɔ* [791]; Shanga *sǚó* [791].

tree: Dan-G *diú* [800]; Dan-B *hú* [800]; Dan-K *hú* [800]; Tura *yíli* [800]; Mano *yíli* [800]; Guro *yíli* [800]; Yaure *yí-bá* [800]; Gban *yǚ-kwì* (?) [800]; Mwan *yli / ylí* [800]; Wan *yléé* [800]; Beng *yli* [800]; Kpeen *gǚó-kú* [801]; Kpan *gɔɔ-xu* [801]; Dzuun *jinin* [800], *gǚó-kú* [801]; Duun *gɔ-hu* [801]; Seenku *gǚó* [801]; Banka *gwo-kura* [801]; Jo *jírín* [800]; Bobo *sǚ-nó* [802]; Bozo-J *jugu* [800]; Bozo-K *jírí* [800]; Bozo-T *jiri* [800]; Soninke *yitti / yítte* [800]; Mau *yíí, yírí, yíli, jíí, jírí, jíli* [800]; Jula-K *yírí* [800]; Dafing *yí* [800]; Bamana *jiri, yírí* [800]; Maninka *yírí* [800]; Xasonka *yírí* [800]; Mandinka *yírí* [800]; Koranko *kǚ* [797], *kǚloma* [798]; Lele *kǚ* [797], *kǚloman* [798]; Mogofin *lógɔ* [802]; Kakabe *lógɔ* [802]; Kono *kǚ* [797]; Vai *kǚñ* [797]; Soso *wúrí* [800]; Jalonke *wúrí* [800]; Jogo *gwá* [801]; Jeri *gbà* [801]; Numu *gòá* [801]; Mende *ngúlú* [800]; Loko *ngúru* [800]; Bandi *ngúlú* [800]; Looma *gúlú* [800]; Kpelle-L *wúru* [800]; Kpelle-G *ulu* [800]; Bokobaru *lí* [796]; Boko *lí* [796]; Busa *lí, nè* [796]; Lebir *gǚ, gǚ* [801]; San *da* [799]; Kyenga *yakɔ* [797]; Shanga *agy* [797].

two: Dan-G *plè, pèèdā* [803]; Dan-B *plè, pèedā* [803]; Dan-K *plè, pèèdā* [803]; Tura *pīllé* [803]; Mano *pèèlè* [803]; Guro *fié* [803]; Yaure *fli, filí* [803]; Gban *fǿí* [803]; Mwan *plē* [803]; Wan *pīlōŋ* [803]; Beng *plāŋ* [803]; Kpeen *figi* [803]; Kpan *fii* [803]; Dzuun *fii* [803]; Duun *fii* [803]; Seenku *fi* [803]; Banka *figi* [803]; Jo *fiúúli* [803]; Bobo *pālā* [803]; Bozo-J *pende* [803]; Bozo-K *fyenu* [803]; Bozo-T *pende* [803]; Soninke *filló* [803]; Mau *fyàà* [803]; Jula-K *filá, fúlá* [803]; Dafing *flá* [803]; Bamana *filá* [803]; Maninka *filá* [803]; Xasonka *fùla* [803]; Mandinka *fùla* [803]; Koranko *fèla* [803]; Lele *fèla* [803]; Mogofin *filá* [803]; Kakabe *fila* [803]; Kono *fèa* [803]; Vai *fè'á* [803]; Soso *firin* [803]; Jalonke *fidin* [803]; Jogo *fàlá* [803]; Jeri *fàla* [803]; Numu *filà, fàlá* [803]; Mende *fèlé* [803]; Loko *fele* [803]; Bandi *féle* [803]; Looma *fele* [803]; Kpelle-L *fēerē* [803]; Kpelle-G *hvile* [803]; Bokobaru *pláa* [803]; Boko *pla* [803]; Busa *pla* [803]; Lebir *pīiya* [803]; San *pāā* [803]; Kyenga *fia* [803]; Shanga *hwaa* [803].

walk: Dan-G *tā* [804]; Dan-B *tá* [804]; Dan-K *tā* [804]; Tura *tāā* [804]; Mano *tàá* [804]; Guro *táá* [804]; Yaure *tá* [804]; Gban *tā* [804]; Mwan *táá* [804]; Wan *tá* [804]; Beng *tá* [804]; Kpeen *jigi* [809]; Kpan *jii* [809]; Dzuun *jii* [809]; Duun *jii-ra* [809]; Seenku *ká* [1013]; Banka *jεε* [809]; Jo *tiri* [808]; Bobo *yā* [810]; Bozo-J *ni* [809]; Bozo-K *nini* [809]; Bozo-T *nini* [809]; Soninke *tere* [808]; Mau *táá, táámǎ* [804]; Jula-K *tágá, tágámá* [804]; Dafing *tà* [804], *wà* [807]; Bamana *tága, táa, tágama, táama* [804]; Maninka *táa, táama* [804], *wā* [807]; Xasonka *táxa, táxama* [804]; Mandinka *táa* [804]; Koranko *tá, táa* [804]; Lele *táa* [804]; Mogofin *tága* [804]; Kakabe *táa, tága* [804]; Kono *tá* [804]; Vai *táá* [804]; Soso *sìgá* [805]; Jalonke *sìgá* [805]; Jogo *tágámá* [804]; Jeri *tága* [804]; Numu *tàgámá* [804]; Mende *njiá* [805]; Loko *sía* [805], *mbee(ŋ)* [806]; Bandi *sia(ŋ)* [805]; Looma *siè, siè(g)* [805]; Kpelle-L *sīa* [805]; Kpelle-G *hie* [805]; Bokobaru *tá* [804]; Boko *tá* [804]; Busa *tá* [804]; Lebir *tá* [804]; San *tō* [804]; Kyenga *tá* [804]; Shanga *tá* [804].

warm: Dan-G *wōsū* [814]; Dan-B *wōzā* [814], *sī* [815]; Dan-K *wōsū* [814]; Tura *zóló* [816]; Mano *gágá, gáká, gáà* [817]; Guro *cé* [818]; Yaure *kálágálá* [824]; Gban *fóá* [813]; Mwan *té-lē* [818]; Wan *té* [818]; Beng *bátú* [819], *pápá* [820], *wlùwlù* [821]; Kpeen *táá-má* [826]; Kpan *wōonta* [829]; Dzuun *wōdzùuntá* [829]; Duun *wata* [829]; Banka *fīte-na* [813]; Jo *māfwū* [830]; Bobo *tóbá* [826]; Bozo-J *pyεε* [828]; Bozo-K *gulena* [827]; Bozo-T *gula* [827]; Soninke *táwú, tayi, taye* [826]; Mau *gbèèŋ* [825]; Jula-K *gbàn* [825]; Dafing *gwá* [825]; Bamana *kálá, gǎ* [825]; Maninka *kálamán* [824], *gbǎ* [825]; Xasonka *xàndi* [825]; Mandinka *kàndi* [825]; Koranko *kála-ma* [824]; Lele *kálama* [824]; Mogofin *bàndi-ré* [825]; Kakabe *gbàndin-den* [825]; Kono *gbandi* [825]; Vai *kpàndi* [825]; Soso *fúrá* [813]; Jalonke *fúrá* [813]; Jogo *dě* [822]; Jeri *dēε, diε* [822]; Numu *dó / dèré* [822]; Mende *fúlá* [813]; Loko *fura* [813]; Bandi *fúlá* [813]; Looma *fila, file* [813]; Kpelle-L *wúleŋ* [823]; Kpelle-G *hvúlo* [813]; Bokobaru *lógólógó* [811]; Boko *lòlòlò* [811]; Busa *lógólógó* [811]; Lebir *yéé* [812]; San *fu* [813]; Kyenga *fula* [813].

water: Dan-G *yí* [831]; Dan-B *yí* [831]; Dan-K *yí* [831]; Tura *yí* [831]; Mano *yíi* [831]; Guro *yí* [831]; Yaure *yí* [831]; Gban *yí* [831]; Mwan *yí* [831]; Wan *yí* [831]; Beng *yí* [831]; Kpeen *jó* [831]; Kpan *zo* [831]; Dzuun *yú* [831]; Duun *zo* [831]; Seenku *jó* [831]; Banka *zo* [831]; Jo *jú* [831]; Bobo *ziō / zō* [831]; Bozo-J *ji* [831]; Bozo-K *jú* [831]; Bozo-T *jii* [831]; Soninke *ji* [831]; Mau *jí, yí* [831]; Jula-K *ji, jyé* [831]; Dafing *zié* [831]; Bamana *ji* [831]; Maninka *ji* [831]; Xasonka *ji* [831]; Mandinka *jii* [831]; Koranko *yí* [831]; Lele *yíi* [831]; Mogofin *jié, géi* [831]; Kakabe *jii* [831]; Kono *yíi* [831]; Vai *ji* [831]; Soso *yéé* [831]; Jalonke *jié* [831]; Jogo *yí* [831]; Jeri *yí* [831]; Numu *yí* [831]; Mende *njá* [831]; Loko *njá* [831]; Bandi *njá* [831]; Looma *zié, zié* [831]; Kpelle-L *yá* [831]; Kpelle-G *ya* [831]; Bokobaru *i* [831]; Boko *i* [831]; Busa *i* [831]; Lebir *pi* [832]; San *mu* [833]; Kyenga *i* [831]; Shanga *ii* [831].

we: Dan-G *kwa* (incl.) [837], *yi* (excl.) [844]; Dan-B *kwa* (incl.) [837], *yi* (excl.) [844]; Dan-K *kwa* (incl.) [837], *yi* (excl.) [844]; Tura *ko* [836]; Mano *kō* (excl.) [836], *kōa* (incl.) [837]; Guro *kv* (excl.) [836], *kaa* (incl.) [837]; Yaure *kv* (excl.) [836], *kaa* (incl.) [837]; Gban *u* [834]; Mwan *ó* (excl.) [834], *kòó* (incl.) [836]; Wan *ká* (excl.) [836], *ka* (incl.) [837]; Beng *qŋ* [843]; Kpeen *mò* [835]; Kpan *ne* [844]; Dzuun *mùn* [835], *ne* [844]; Duun *nīe* [844]; Seenku *wí* [835]; Banka *ne* [844]; Jo *yí* [844]; Bobo *mε* (excl.) [835], *ke* (incl.) [838]; Bozo-J *ke* [838]; Bozo-K *ngí* [838]; Bozo-T *ki* [838]; Soninke *ó* [834]; Mau *qŋ* [843]; Jula-K *á* [843]; Dafing *á* [843]; Bamana *á* [843]; Maninka *ñ* [842], *á* [843]; Xasonka *ñ* [842]; Mandinka *ñ* [842]; Koranko *mà* [835]; Lele *mà* [835], *ñ* [842]; Mogofin *móó* [835]; Kakabe *má* [835]; Kono *mó* (incl.) [835]; Vai *mú* [835]; Soso *mùxú* (excl.) [840], *wōn* (incl.) [841]; Jalonke *nxo* [840], *on* [841]; Jogo *á* [839]; Jeri *á* [839]; Numu *á* [839]; Mende *mú* [835]; Loko *mú* (incl.) [835]; Bandi *mú* (incl.) [835]; Looma *gé* (excl.) [838]; Kpelle-L *kú* (excl.) [836]; Kpelle-G *kú* (excl.) [836]; Bokobaru *wa* [834]; Boko *wa* [834]; Busa *o* [834]; Lebir *v, wó* [834]; San *wō* [834]; Kyenga *wa* [834]; Shanga *wà* [834].

what: Dan-G *mēē* [849]; Dan-B *mā* [849]; Dan-K *mā, mlā* [849]; Tura *mèè* [849]; Mano *kō* [850]; Guro *mēè* [849]; Gban *bé* (?) [849]; Mwan *mè* [849]; Wan *mē* [849]; Beng *pó* [845]; Kpan *fee* [851]; Dzuun *fée* [851]; Duun *fōε* [851]; Seenku *kūrè* [854]; Banka *fó* [851]; Jo *vε* [851]; Bobo *ŋwōnōn* [853]; Bozo-J *mwq* [849]; Bozo-K *màná* [849]; Bozo-T *man* [849]; Soninke *màní* [849]; Mau *mīŋ* [849]; Jula-K *mùn, mìní* [849]; Dafing *mú* [849]; Bamana *mù* [849]; Maninka *mù* [849]; Xasonka *mūŋ* [849]; Mandinka *mūŋ* [849]; Koranko *nfé*

[851]; Lele *fě-se* [851]; Mogofin *fěn-dè* [851]; Kakabe *jùmande* [852]; Kono *fě* [851]; Vai *mé* [851]; Soso *mun* [849]; Jalonke *nèñ* [848]; Jogo *mú* [849]; Jeri *mi* [849]; Numu *mèn* [849]; Mende *gbé* [847]; Loko *gbé* [847]; Bandi *bèe* [847], *ndè* [848]; Looma *à bè* [847]; Kpelle-L *gbéè* [847], *lé* [848]; Kpelle-G *le* [848]; Bokobaru *bó* [845]; Boko *bó* [845]; Busa *bó* [845]; Lebir *wàa* [846]; San *bó* [845]; Kyenga *bó* [845]; Shanga *bu* [845].

white: Dan-G *púú* [855]; Dan-B *púú* [855]; Dan-K *púú* [855]; Tura *púú* [855]; Mano *púlú* [855]; Guro *fiúú* [855]; Yaure *fǔ-vǔ* [855]; Gban *βútù* [855]; Mwan *pú* 1 [855]; Wan *pú* 1 [855]; Beng *púú* [855]; Kpeen *kpé* [856]; Kpan *φwēey* [855]; Dzuun *fyé* [855]; Duun *fiɔ-ma* [855]; Seenku *kàn* (?) [856]; Banka *xue* [856]; Jo *fǔ* [855]; Bobo *furu* [855]; Bozo-J *kuo* [856]; Bozo-K *hoona* [856]; Bozo-T *xon* [856]; Soninke *xulli / xullé* [856]; Mau *gbé* [856]; Jula-K *gbé* [856]; Dafing *kwì* [856]; Bamana *jé* [856]; Maninka *gbé* [856]; Xasonka *xóyi* [856]; Mandinka *kóyi* [856]; Koranko *gbé* [856]; Lele *gbé* [856]; Mogofin *bée* [856]; Kakabe *gbé* [856]; Kono *gbé* [856]; Vai *kpé* [856]; Soso *fikke* (?) [855]; Jalonke *fíiqè* (?) [855]; Jogo *kpé* [856]; Jeri *kpé* [856]; Numu *kpé* [856]; Mende *kòlè* [856]; Loko *kùè* [856]; Bandi *kòlè* [856]; Looma *kwéle(g)* [856]; Kpelle-L *kōlē* [856]; Kpelle-G *kwele* [856]; Bokobaru *pú-ra* [855]; Boko *pú-a* [855]; Busa *pú-rā, pú-lā* [855]; Lebir *fú* [fù] [855]; San *fu* [855]; Kyenga *fu* [855]; Shanga *húú* [855].

who: Dan-G *dē* [857]; Dan-B *dǔ* [857]; Dan-K *dè* [857]; Tura *wáá* [858]; Mano *dēī* [857]; Guro *dī* [857]; Yaure *tǐ* [857]; Gban *dě* [857]; Mwan *dè-lè, dè-là* [857]; Wan *dè, dèè* [857]; Beng *dé* [857]; Kpan *sii* [857]; Dzuun *dzi* [857]; Duun *si* [857]; Seenku *tē* [857]; Banka *ši* [857]; Jo *nkēη* [868]; Bobo *kǎ* [864]; Bozo-J *wura* [865], *gira* [866]; Bozo-K *jǎ* [861]; Bozo-T *gwa* [867]; Soninke *kóo* [864]; Mau *yéη, jéη* [861]; Jula-K *jóón* [861]; Dafing *zòn* [861]; Bamana *jó'* [861]; Maninka *jó'* [861]; Xasonka *jón* [861]; Mandinka *jón* [861]; Koranko *yó, yǔ* [861]; Lele *yón* [861]; Mogofin *jò-mɔ* [861]; Kakabe *yón* [861]; Kono *jó* [861]; Vai *jó, jǔ* [861]; Soso *ndé* [857]; Jalonke *ndé* [857]; Jogo *mǎ* [863]; Jeri *ma* [863]; Numu *máà* [863]; Mende *yóó* [861]; Loko *e* [862]; Bandi *e-le* [862]; Looma *bēe* [860]; Kpelle-L *gbéè* [860]; Kpelle-G *gbe* [860]; Bokobaru *dé* [857]; Boko *dé* [857]; Busa *dí* [857]; San *die* [857]; Kyenga *dí* [857]; Shanga *yí* [859].

woman: Dan-G *dē-bǎ* [869]; Dan-B *dē-bǎ* [869]; Dan-K *dē* [869]; Tura *dě* [869]; Mano *lēē* [869]; Guro *lē* [869]; Yaure *li-mò* [869]; Gban *lè* [869]; Mwan *lē* [869]; Wan *lē* [869]; Beng *lēη* [869]; Kpeen *táà* [873]; Kpan *taq* [873]; Dzuun *tàán* [873]; Duun *taq* [873]; Seenku *míni* (?) [869]; Banka *taq* [873]; Jo *màrè* (?) [869]; Bobo *yà* [870]; Bozo-J *jugó* [870]; Bozo-K *jànànú* [870]; Bozo-T *yá-lò* [870]; Soninke *yà-xàri* [870]; Mau *mòsò* [872]; Jula-K *mùsò* [872]; Dafing *mósó* [872]; Bamana *mùsò* [872]; Maninka *mùsò, mòsò* [872]; Xasonka *mùsu* [872]; Mandinka *mùsu* [872]; Koranko *mùsu* [872]; Lele *mùsu, mùso* [872]; Mogofin *mùsú* [872]; Kakabe *mùsu* [872]; Kono *mùsu* [872]; Vai *mùsú* [872]; Soso *gìné* [871]; Jalonke *gìné* [871]; Jogo *ná* [870]; Jeri *ná* [870]; Numu *ná* [870]; Mende *nà-há* [870]; Loko *nà-há* [870]; Bandi *nà-há* [870]; Looma *na-za* [870]; Kpelle-L *nēní, néí* [869]; Kpelle-G *nenu* [869]; Bokobaru *nɔ-gbé* [869]; Boko *nɔ* [869]; Busa *nɔ* [869]; Lebir *lú* [lǔ] [869]; San *lɔ* [869]; Kyenga *ne-gbí* [869]; Shanga *nénèni* [869].

yellow: Dan-B *sóηgòzǎ, víávíǎ* [875]; Dan-K *gǔēyǐǐdǎyǎ* [876]; Tura *zàinlé* [877]; Mano *zóló* [573]; Guro *kán* [878]; Gban *βúǎǎ* [879]; Beng *dálé* [881]; Kpan *nefaama* [886]; Dzuun *nééfámá* [886]; Duun *nefama* [886]; Banka *nerefarama* [886]; Jo *náfánámáη* [886]; Bozo-J *nàlòndúú* (“powder of nere”) [886]; Bozo-T *nááhúúmǎǎ* [886]; Soninke *boobo* [890], *makkajinma* (“juice of corn”) [891], *puuru* [892]; Mau *nèēmùùlámáη* [886]; Jula-K *nèrè̀mùgùrà̀má̀n, nèrè̀mùgùtígí* [886]; Dafing *néémúúmá* [886]; Bamana *nèremuguláma* [886]; Maninka *nèremuuláma* [886]; Xasonka *nètemunguma* [886]; Mandinka *nètemunku* [886]; Koranko *neremunkulame* [886]; Lele *nèemuulamá* [886]; Mogofin *nètekutalama* [886]; Kono *mɔin dene* [889]; Vai *suηa* (KOELLE) [887], *waηη* [888]; Jogo *yè̀fògò* (“nere flour”) [886]; Numu *yíré fúkò* (“nere flour”) [886]; Bandi *báhi* [884]; Looma *kpazi* (“corn”) [1009]; Kpelle-L *gbēlēηjǐ* [885]; Kpelle-G *bahi* [884]; Bokobaru *kpà̀íkpàí* [874]; Boko *kpà̀í* [874]; Busa *tǎra* [881]; Lebir *ká̀r̀nǐsi* [882]; San *kùsi* [883]; Shanga *kwǎ̀í* [874].

Р е з ю м е

В семье манде (макросемья нигер-конго) достаточно однозначно выделяются группы небольшой глубины родства; основные споры ведутся по поводу промежуточных ступеней внутренней классификации. В статье предлагается новый вариант такой классификации, впервые выполненный на основании обчёта полного 100-словного списка М. Сводеша. Данные глоттохронологии сопоставляются с данными палеоклиматологии и археологии, делается предположение, что территория обитания людей, говоривших на праязыке семьи манде во второй половине IV тысячелетия до н. э., локализуется в Южной Сахаре, между 3° и 12° западной долготы и к северу от 16° или даже 18° северной широты. Это предположение подтверждает и анализ культурной лексики, предположительно реконструируемой для праманде.

*Происхождение индоевропейских показателей лица:
Исторический анализ и данные внешнего сравнения*
К. В. БАБАЕВ
М. – Калуга: «Эйдос». 2008. 295 стр.

А. В. ДЫБО
(Российский государственный гуманитарный университет)

Книга Кирилла Владимировича БАБАЕВА представляет собой очень сильно переработанный и расширенный вариант защищенной им весной 2008 г. кандидатской диссертации. Работа была начата под руководством С. А. Старостина, а продолжена официально под моим руководством, но хотелось бы подчеркнуть, что и окончательный выбор темы, и ряд концептуальных моментов работы еще в ее диссертационном виде автор, с самого начала показавший себя сложившимся исследователем, произвел абсолютно самостоятельно. Тема, рассмотренная в книге, безусловно, представляет значительный интерес для проблематики дальнего родства. Действительно, личные местоимения по праву входят в стословник Сводеша и зачастую при подсчетах в рамках отдаленного родства остаются единственными достоверными «плюсами», позволяющими говорить, например, о родстве айнского языка и австрических (или, при другом анализе, ностратических). Они общепринято считаются самой устойчивой и самой непроницаемой из лексических групп, и факт заимствования личных местоимений зачастую ставит в тупик специалистов по языковой конвергенции, готовых в этом случае утверждать невозможность определения генеалогической принадлежности, допустим, медновского пиджина (с точки зрения компаративиста, совершенно определенно аллоторского по происхождению). Необходимость внести определенность в работу с личными местоимениями в историческом плане, расставить методологические точки над «i» побуждает с вниманием отнестись к рецензируемой работе. Ниже излагается ряд моих соображений и замечаний, из которых, я надеюсь, станет понятно, почему я, хотя и признаю, что книга **а)** написана в «правильной», т. е. классической компаративистской идеологии, **б)** является весьма полезным сводом материала и **в)** представляет ряд очень интересных и важных обсуждений по проблеме, — но не считаю, что в ней дано «полное и окончательное» решение проблемы происхождения индоевропейских личных местоимений и показателей лица, и даже что в ней получен значительный прорыв в решении тех вопросов, которые встают перед каждым «ностратическим» этимологом, когда он смотрит на морфемы со значением лица в ностратических языках.

Автор с первых страниц проявляет себя как хорошо образованный индоевропеист и ностратист, что позволяет ему прежде всего дать прекрасный подробный обзор по истории изучения личных местоимений и морфологических показателей лица в индоевропеистике и исследованиях по дальнему родству. Практически со всеми оценками работ предшественников в этом обзоре можно согласиться. Надо сказать, что выводы из этого обзора для непредвзятого читателя выглядят скорее мрачно. Действительно, в основном лингвисты, пытавшиеся заниматься глубокой историей личных местоимений и показателей лица, упираются в то, что система местоимений и местоименных показателей в конкретных языковых группах представляет чересчур скудный и трансформированный морфологическими процессами материал для строгой фонетической и лексической реконструкции. Вследствие этого они зачастую сворачивают на путь глоттогонических построений, который не подходит для сравнительно-исторического языкознания из-за отсутствия доказательности в рамках компаративистской процедуры.

Но автор не склонен к унынию. Он полагает, что улучшить ситуацию (ввести какие-то добавочные доказательные аргументы в построение) можно, обратившись к типологии языков мира. Этот способ расширения доказательной базы уже неоднократно предлагался в индоевропеистике (прежде всего, в книге [ГАМКРЕЛИДЗЕ, ИВАНОВ 1984]); автор рецензируемой работы считает, что, оперируя строго типологией языковых изменений (а не просто синхронной типологией) и придерживаясь при этом общих принципов диахронического анализа, он сможет сильно уточнить результаты исследования.

Действительно, ему удастся вполне разумным образом привлечь обширную имеющуюся к настоящему времени литературу по типологии грамматики, на основании которой он прежде всего пытается построить исчисление различных возможностей происхождения для различ-

ных морфем с личным значением. Сразу оговорю, что качество этого исчисления меня не удовлетворяет, я думаю, из-за многочисленных недочетов, свидетельствующих о плохом знакомстве с формально-грамматической теорией. Не совсем ясно, только ли это свойство автора, хорошо образованного как диахронист и типолог, но недостаточно образованного как грамматист, или даже уже свойство и той типологической литературы, на которую он опирается.

Недочеты этого типа появляются уже во введении, в подразделе о применении типологии в диахроническом исследовании. Здесь они не влияют на идеологию работы, в целом, повторю, производящую весьма положительное впечатление. Тем не менее, неточность выражения грамматических мыслей в работе, посвященной дальнему родству, я бы считала опасной в том плане, что она может запутать и даже отвратить от предмета обсуждения не вполне компетентного читателя. На с. 31 видим рассуждение об общепринятом среди лингвистов-теоретиков мнении, что существует фундаментальная разница между первым и вторым лицом, с одной стороны, и третьим, с другой. Цитата: «Эта разница обычно объясняется тем, что первое и второе лицо обозначаются личными показателями, в то время как третье лицо в языке может быть выражено любой именной лексической единицей». Взгляд на вещи очень наивный: наоборот, постулируемое глубокое семантическое различие между лицами, задействованными в дискурсе, и «третьим», незадействованным лицом теоретики полагают, в частности, причиной того, что для выражения 1-го и 2-го лиц языка развивают специфические лексические единицы, в то время, как для третьего лица зачастую специфическое личное местоимение, отличное от указательных, отсутствует.

В 1-й главе производится упомянутая попытка построить исчисление возможностей происхождения морфем с личным значением. Сразу отмечу, что автор, кажется, напрасно называет объект своего исследования личными показателями. Для слова «показатель» в лингвистике зарезервировано значение «грамматический показатель», то есть не «лексема», а «грамматическая морфема» (в аналитической конструкции может и не входить в состав словоформы). Личные местоимения, которые автор включает в класс «независимых личных показателей»¹, как и другие местоимения, являются лексическим классом, принадлежность к которому определяется их семантикой (шифтеры различного рода), а не грамматическим (морфологическим или синтаксическим) поведением: грамматически в тех языках, где есть такой грамматический класс (это, видимо, все известные на сегодняшний день языки, кроме разве что древнекитайского²), они принадлежат к классу имен либо субстантивов. То есть, они имеют тот же набор грамматических категорий, что и другие субстантивы. При этом в ностратических языках они часто образуют специфические парадигматические классы (например, особые типы склонения), но это все же не влияет на их грамматическую характеристику. Далее К. В. БАБАЕВ пишет (с. 34–35): «Утверждения о том, что выражение 1-го и 2-го лица является языковой универсалией, по-видимому, верны. Вопрос лишь в том, являются ли их выразители в языке лексическими или обособлены в отдельный класс в системе морфологии. По этому критерию языки мира можно смело разделить на два чётких типа по способу выражения — лексические и грамматические. В первом случае морфология языка не содержит отдельного класса показателей лица (то есть личных местоимений), однако понятия «я», «ты» как в единственном, так и в других числах выражаются с помощью однозначных лексических единиц. Таким образом, широко распространённая точка зрения о существовании языков, не содержащих личных показателей, неверна, так как подменяет понятия: да, безусловно, существуют языки без личных местоимений, но языков без средств выражения категории лица — не существует». Так вот, языков без выражения категории лица, действительно, не существует, но неправ будет тот читатель, который подумает, что с категорией лица по языкам все обстоит не так, как с другими потенциально грамматическими категориями, которые в части языков действительно являются грамматическими, а уж в тех языках, где соответствующей грамматике для них не нашлось, выражаются лексически. И категория лица также кое-где выражается грамматически, а в остальных языках выражается лексически, то есть с помощью специфического класса лексем со значением лица, то есть личными местоимениями (в японском, например). Для синхронно-типологической характеристики языка не имеет никакого значения этимологический аспект, то есть, из слов с каким зна-

¹ Называя их синтаксически независимыми, автор допускает небрежность выражения. Они, конечно, синтаксически зависимы от предиката, выражая падежные роли. Имелось в виду, по-видимому, «грамматически независимые», то есть то, что они присутствуют в предложении не в составе словоформы, а как самостоятельные словоформы. Вообще-то для этого существует традиционное выражение «лексические» (другие способы выражения категории, соответственно, — грамматические).

² См. [СТАРОСТИН 2007: 499; 504–513]; [СТАРОСТИН 1993: 93; 98–107].

чением образованы эти личные местоимения. Несомненно, приводимые автором тайские фразеологизованные сочетания с внутренней формой «волос головы» и «низ ноги» (или скорее «под ногой»), служащие для выражения 1-го лица в разных регистрах респективности, являются тайскими личными местоимениями в традиционном (то есть семантическом) понимании этого слова³, вопреки замечанию автора на с. 36; они не выглядят личными местоимениями только для «ограниченного» индоевропейца, который ожидал бы, что это будут слова с давно затертой внутренней формой и предпочтительно с отдельным типом склонения (в тайском, впрочем, и вообще нет типов склонения) — но ему точно так же не понравится в качестве местоимения и японское ワタ *watákúsi*.

Соответственно, процесс, который привел к образованию этих тайских личных местоимений, а также яп. ワタ *watá(kú)si*, вопреки мнению автора на с. 41–46, не процесс грамматикализации, а процесс лексикализации (иногда универбации). Отсюда классификация типов происхождения морфем с личным значением должна в первую очередь выделять две группы: происхождение лексем, выражающих лицо, и происхождение морфологических показателей лица. Лексемы, выражающие лицо, в свою очередь должны делиться на личные местоимения, притяжательные и возвратные. Этимологии слов, относящихся ко всем этим группам, должны рассматриваться по отдельности. Вот опыт такой классификации «навскидку» по материалам, приведенным в книге.

Известные автору книги случаи прозрачного происхождения личных местоимений, соответственно, разбиваются на следующие:

1. из отдельных лексем-субстантивов с не-шифтерной семантикой:

1-е л.: *тело, лицо, существо, жизнь, основной (основа), раб, слуга*

2-е л.: *господин, хозяин, остальной;*

3-е л.: *вещь, человек, мужчина, тело;*

Примеры приведены К. В. БАБАЕВЫМ на с. 42, со ссылкой на [СУНИК 1978]⁴ и [HEINE – KUTEVA 2007]; следует заметить, что этимология нилосахарского языка нгити явно не настолько хорошо развита, чтобы уверенно говорить о том, что личное местоимение 1-го л. *le* действительно этимологически тождественно существительному *ale* ‘человек’, как утверждается в работе [HEINE – KUTEVA 2007], некритично цитируемой автором, а не имеет совершенно другое происхождение⁵. К сожалению, убедительных примеров на такое происхождение местоимений автор не приводит, но, например, список субститутов личных местоимений в корейском, приводимый в [РАМСТЕДТ ГКЯ: § 101], не противоречит приведенному. Так что можно считать, что значение личного местоимения действительно может возникать из таких значений, т. е. с помощью тропов. Вопросом для меня остается, нельзя ли все эти случаи считать эллиптическими (с эллипсисом притяжательного местоимения или показателя, уместным внутри дискурса) по отношению к случаю ниже в п-те 4. Отдельно отмечу, что к этому пункту не имеет отношения ошибочное рассуждение автора там же о том, что в качестве маркеров лица широко используются собственные имена. Конечно, слово *Ахмед* в выражении *Ахмед сделает*, даже если его произносит Ахмед, не маркирует лица. Лицо — шифтерное значение, а соотношение между словом *Ахмед* и его денотатом не меняется в зависимости от того, какой из участников дискурса его произносит. Именно поэтому и происхождение личного местоимения из имени собст-

³ И, по-видимому, экстремистским следует признать принимаемое автором утверждение Дж. Кука о том, что в тайском языке не существует «обычных» (т. е. с затертой внутренней формой) местоимений, ср. в кн. «Тайский язык» ([Морев и др. 1961]) на с. 89–90, где приведен ряд вполне деэтимологизованных личных местоимений, хотя и говорится, что употребляются они редко. Кстати, фразеологизованное сочетание เกล้ากระผม *klàw-krâ-p'hóm*, которое переводится Куком как «волос головы» = «я», в действительности содержит в себе композит с личным местоимением ผม *p'hóm* «я», букв. «волос головы меня», т. е. этот случай «происхождения» личного местоимения относится к другому классу, чем думает автор.

⁴ Почему-то при группе слов 1-го л.; в статье О. П. Суника разбирается происхождение возвратных местоимений, которые сами могут принимать личные показатели и тогда значат «я сам», «ты сам» и «он сам», — в частности, из лексем типа «тело»; мимоходом на с. 265 справедливо критикуется попытка А. Ф. Бойцовой возвести ТМ местоимение *мэн* «сам» к *бэжэ* «тело» — и это единственное место, которое могло бы быть принято автором книги за разговор о происхождении местоимения 1-го л. от полнозначного существительного, если он принял ТМ *мэн* за омонимичное тюркское *мэн* «я»; больше нигде в статье о происхождении местоимений 1-го л. речи нет, есть только сопоставления их грамматического поведения с поведением указательных и возвратных местоимений.

⁵ Это как раз весьма распространенная ошибка у типологов, занимающихся грамматикализацией, — некритическое отношение к материалу, который берется из вторых-третьих рук. Но у автора рецензируемой книги — по образованию и области интересов прежде всего компаративиста — хотелось бы видеть большую осторожность в таких случаях.

венного кажется теоретически невозможным — если только это не местоимение третьего лица (по всему, что говорилось раньше, не вполне личное местоимение), восходящее к неопределенно-личному, для которого можно себе представить источник в виде личного имени вроде санскр. *Devadamma* — абстрактный герой буддийских притч, стандартный контрагент Будды.

2. Из возвратных и под. местоимений:

сам, — в действительности, конечно, с личными показателями: «я сам», «ты сам» и под. (а вовсе не как специфический источник местоимения 1-го лица, как полагает автор). Такие сочетания могут употребляться для выражения лица в предложении эмфатически; при стирании эмфазы они должны становиться просто личными местоимениями. Примеров автор не приводит, но такие случаи известны, ср. т. наз. «усилительно-личные» местоимения в мордовском, которые успешно вытесняют простые личные. Подобным же образом получились исп. *nosotros, vosotros* — первоначально эмфатические сочетания личного местоимения с «неопределенным» *otros* ‘другие’.

3. Из указательных местоимений:

Убедительных примеров автор не приводит. Баскаковский пример с образованием тюрк. *ben* ‘я’ из *bu-nuñ* ‘этот-Gen.’, конечно, неверен, как и предполагаемая Дельбрюком связь лат. *tibi* с скр. *tam* и пр. (с. 43–35); случаи яп. 此の方 *kònó’ hō* ‘эта сторона’ = ‘я’ и арм. *utñu tēr-³s* ‘этот господин’ = ‘я’ относятся к следующему классу. С некоторой натяжкой сюда можно отнести происхождение местоимения 2-го лица вежливой формы из местоимения 3-го лица — как в нем. *Sie* ‘Вы’ из ‘они’ и у Лессинга *Er* ‘Вы’ (по отношению к слуге, вообще к низшему) из ‘он’; поскольку местоимения 3-го лица обычно восходят к демонстративам, такое умозаключение возможно.

4. Из сочетаний лексем-субстантивов с не-шифтерной семантикой с а) указательными местоимениями или местоименными показателями (яп. 此の方 *kònó’ hō* ‘эта сторона’ = ‘я’ и арм. *utñu tēr-³s* ‘этот господин’ = ‘я’), б) притяжательными местоимениями или местоименными показателями (исп. *Usted < Vuestra merced* ‘Ваша милость’ и под.). Сюда же, наверное, стоит относить и «косвенные падежные формы» личных местоимений в венгерском, представляющие собой по внутренней форме притяжательные формы от послелогов, восходящих, в свою очередь, к существительным с пространственным значением.

По-видимому, отдельной разновидностью этого типа следует считать случаи типа тайского ใต้เท้า *tàj-t’ăw*, букв. «(тот, что) под ногой», местоимение 1-го л. в почтительном регистре: это сочетание, скорее всего, содержит эллипсис личного местоимения, означающего принадлежность 2-му лицу, уместный внутри дискурса (например, ใต้เท้าพານ *tàj-t’ăw p’hôn-t’àn*, где พาน *p’hôn-t’àn* — местоимение 2-го лица, употребляемое при обращении к высшим чиновникам, по происхождению — композит существительного с местоимением 2-го лица พาน *t’àn*).

5. Из отдельной предикации, содержащей словоформу глагола с бытийным/связочным/каким-то еще энциклопедически относительно пустым значением в форме соответствующего лица; т. е. с исходной формой типа «будучи мной» или «что касается меня». Примеры приведены Бабаевым на с. 46 (семитские языки — со ссылкой на [Орел 1990], кушитские, омотские, айнский — со ссылкой на [SIEWIERSKA 2004: 255–260]); лишним является здесь указание на употребление вспомогательного глагола «быть» в рамках личной глагольной словоформы — как в бедауйе, так и в (старо)польском *padleśm*; это классификационно совсем другой случай, происхождение грамматического показателя, который, очевидно, перед тем не успел побывать личным местоимением — см. ниже.

Известные нам случаи происхождения грамматических личных показателей, со своей стороны, разбиваются на следующие (ср. подробную и не вызывающую возражений историю вопроса, которую К. В. Бабаев дает во 2-й главе, с. 66–80):

1. Самый обширный случай — из личных местоимений (можно предполагать при этом, что такое может случиться и с личными местоимениями, возникшими в одном из указанных выше случаев). Здесь необходима более подробная классификация, так как разные типы показателей могут (и должны) происходить из разных форм личных местоимений, поскольку словоформа с показателем, очевидно, должна формироваться из сочетания основной словоформы со словоформой личного местоимения, а они могли находиться в разных синтаксических отношениях, по крайней мере, в случаях притяжательности и личного согласования. Если мы будем говорить только о личных показателях предикации, то просматривается четыре основных потенциальных типа возникновения.

1) Из личной формы глагола-связки, прицепленной к именной форме смыслового глагола [старопольск. *padleśm* (1-е л. ед. ч. перфекта) «я упал» = *padł* (причастие прош. вр. от «пасть»)]

+ *eśm* (Псл. **esť*, 1-е л. ед. ч. глагола «esse»)⁶]; происхождение этой личной формы глагола-связки, соответственно, может быть возведено к одному из следующих случаев.

2) Из прямой формы личного местоимения (т. е. формы подлежащего) — при стандартной диатезе (подобным образом возникают личные показатели большинства глагольных форм и именных сказуемых в тюркских языках: *-*ben*, *-*sen*, *-*ol/-θ*, *-*biz*, *-*siz*, *-*θ* в общетюркском, аффиксы первого лица *-*bi*, *-*bir* в чувашском (остальные лица замещены аффиксами принадлежности вследствие аналогических процессов); в части монгольских языков: бур. -*б* (**bi* ‘я’), -*ш* (**či* ‘ты’), -*бди* (**bide* ‘мы’), -*м* (**ta* ‘вы’) и т. д.).

3) Из косвенных форм личного местоимения (т. е. различных дополнений) — при трансформированной диатезе (примеры теоретически мыслимы в языках с вторичной эргативностью; ср. в мунджанском личные показатели в прошедшем времени от переходных глаголов (т. е. с эргативным типом построения предложения), восходящие к клитическим формам дательного падежа личных местоимений: ед. ч. 1-е л. -*əm* < **mai*, 2-е л. -*ət* < **tai*, мн. ч. 2-е л. -*āf* < **vah*, поскольку в целом конструкция восходит к древней пассивной *ima mai kartam* ‘то у меня/мною сделано’, см. [ЭДЕЛЬМАН 1990: 109]).

4) Из притяжательного местоимения (которое может формально быть отыменным прилагательным от личного местоимения) или атрибутивной формы (родительного падежа) личного местоимения. По-видимому, при этом смысловой глагол должен выступать в форме имени действия (тогда его первый актанта естественным образом трансформируется в атрибут) — или из показателя притяжательности, который теоретически может развиваться из этих источников. Ср. употребление масдара на -*dlǵ* в качестве сказуемого в турецком предложении: *En kötü şey bizim uzun zaman yaz-ış-ma-dıǵ-ımız* «самое плохое — то, что мы долгое время не переписывались», букв. «самая плохая вещь наше долговременное (взаимное) не-писание» (не-писание с афф. принадлежности 1-го л. мн. ч., который отличается и от афф. сказуемости — -(y)Iz, и от собственно глагольного афф. 1-го л. мн. ч. -Ik). При определенном типе синтаксического развития такая конструкция может стать основной в языке, и тогда можно ожидать, что бывшее посессивное окончание станет стандартным показателем лица у сказуемого.

Кроме того, в системе личных окончаний могут происходить аналогические унифицирующие перестройки, которые смешивают в одной подсистеме окончания различного происхождения и иногда гибридируют их. Очень хорошо видны такие (еще относительно недавние) процессы, например, в системах личных показателей современных огузских диалектов (см. [ДЫБО 2007]). О подобных же развитиях в новоиранских языках пишет и Дж. И. ЭДЕЛЬМАН ([ЭДЕЛЬМАН 1990]).

2. Что касается возникновения личных показателей из других источников, возможность рассмотрена у П. Хайду ([Хайду 1985: 329]). Правдоподобным из рассмотренных случаев могло бы показаться разве только происхождение финских окончаний 3-го л. из суффикса причастий, встроившихся в глагольную парадигму, распространяющегося впоследствии на не-причастные по происхождению формы. Остальные случаи (венг. -*k* 1-го л. ед. ч., для которого предлагали происхождение из показателя мн. ч. и из именного суффикса; -*sz*, -*l* 2-го л. ед. ч., которые, скорее, являются временными показателями) сам Хайду считает сомнительными.

Конечно, эта наша классификация не претендует на полноту; жаль, что автор книги, специально задававшийся такой классификационной задачей, раз уж он решил ввести типологические обоснования в свою реконструкцию, не эксплицировал ее в первой главе примерно в таком виде, что значительно облегчило бы дальнейшую верификацию его гипотез⁶.

⁶ Автор дважды в этом разделе говорит о польском *chodzileśm*, но это старопольская форма, современная — *chodźilem*. В польском показателе в описываемой парадигме —

-(e/a/o по родам: м./ж./ср.)- <i>m</i>	-(i/y по категории «мужского лица»): лично-м./не-лично-м.)- <i>śmy</i>
-(e/a/o)- <i>ś</i>	-(i/y)- <i>ście</i>
-(θ/a/o)- <i>θ</i>	-(i/y)- <i>θ</i>

Видно, что это — не рефлексы Псл форм глагола ‘esse’ (было бы (*eśm*, (*eś*, (*e*)*st/śc* по диалектам, как это было в XV веке), а попросту те же окончания, что и в современном спряжении этого глагола в презенсе, *jestem*, *jestes*, *jest*, *jestesmy*, *jestescie*, *sq*. То есть, это окончания, перенесенные из форм глагола-связки, а вот в глаголе-связке это окончания, сложным путем полученные из скрещивания атематической системы окончаний (как в *wiem*, *dam*) и старых форм глагола-связки (*eśm*, (*eś*, (*e*)*st*). Последовательность переходов прослеживается исторически: сначала в памятниках появляются формы типа *jestem* (унификация в рамках атематического спряжения, которая, впрочем, далеко не заходит), а потом формы типа *padleśm* заменяются на *padlem*. Приходится отметить, что недоразобранность подобных примеров может создать впечатление недостоверности всего изложения, хотя понятно, что по большому счету автор правильно излагает основные положения теории грамматикализации.

Во второй главе К. В. Бабаев рассматривает историю реконструкции и различные гипотезы происхождения индоевропейских личных показателей. Изложение практически безупречно; отметим только пару словесных неточностей.

С. 52. «Антуан Мейе, один из наиболее ярких представителей младограмматической школы индоевропейского сравнительного языкознания...». Все-таки А. Мейе — представитель французской социологической школы, восходящей к Ф. де Соссюру, хотя и овладевший приемами младограмматиков.

С. 61. «...В индоевропейском праязыке произошел окончательный переход к флективному синтаксическому строю» — следует читать «к флективному морфологическому строю».

Главы 3 и 4 посвящены собственно основной теме работы. Это реконструкция показателей соответственно первого и второго лица для праиндоевропейского состояния. Следует помнить, что под «показателями», в соответствии со словоупотреблением 1-й главы, здесь автор подразумевает местоименные корни, а не конкретные грамматические показатели. Известно, что для индоевропейского языка периода распада восстанавливается разветвленная система личных глагольных показателей, подсистемы которой распределены по грамматическим видо-временным подсистемам, а также — в презенсе — лексически. Кроме того, для того же периода явно можно восстанавливать личные местоимения обоих (или трех) чисел, с супплетивными парадигмами склонения (все эти реконструкции подвергнуты обзору во 2-й главе книги). К. В. Бабаев пытается выделить единый местоименный корень из всех парадигм каждого лица и кроме того установить, какой набор этих местоименных корней функционировал в праиндоевропейском. Соответственно, его задача относится не к праиндоевропейскому языку периода распада, а к более раннему языковому состоянию и может решаться в основном с помощью метода внутренней реконструкции.

Предложенное в книге решение весьма правдоподобно, но остаются недоисследованными несколько моментов. Автор очень (чрезмерно, на мой взгляд) решительно предполагает для общеиндоевропейского состояния малопадежную систему, оставляя образование большей части падежей для периода распада на отдельные диалектные группы. Понятно, что эта гипотеза не особенно доказательна. Например, аблатив на **-d* трудно считать явлением, возникающим в отдельных диалектных группах (или имеется в виду распад на анатолийскую и собственно индоевропейскую группы? но и в анатолийской группе явный след такого аблатива *-aza*, как, впрочем, и большинства других падежей). Конечно, можно предположить параллельное развитие аблатива во всех диалектных группах из локативной частицы, но вообще-то сравнительное языкознание стремится избегать таких предположений. Кроме того, такое предположение приводит автора к некоторым построениям, противоречащим типологическим постулатам. Например, известно (считается типологией), что обычно личные местоимения в отдельно взятом языке имеют больше разных падежных форм, чем другие субстантивы (ср. современный английский или французский языки). Автор же восстанавливает в праиндоевропейском систему с двумя падежными формами, объектной (общекосвенной) **me* и притяжательной (генитивной) **me-ne* (плюс еще форма множ. ч. **mes*). Такая система очень хорошо «опускается» в праностратическую реконструкцию (для согласования с которой еще В. М. Иллич-Свитыч восстанавливал в праиндоевропейском **me* 'я' и **me-ne* косвенную форму, т. е. формы, практически изоморфные реконструируемым К. В. Бабаевым, см. [Иллич-Свитыч 1971: 6]), что является ее несомненным достоинством. Но, чтобы объяснить другие падежные формы в языках-потомках, автору приходится предполагать, что парадигма местоимения развивалась по аналогии с субстантивной парадигмой и, например, получила из нее окончания винительного **-t* и аблатива, что значит, что в какой-то момент парадигма обычных субстантивов содержала больше падежных форм, чем местоименная. Между тем, по традиционным представлениям, ближе к исходной субстантивной парадигме набор окончаний консонантного типа склонения, а тематический тип, вслед за прилагательными, приобрел отличающие его оконча-

⁷ В дальнейшем верификация происхождения некоторых личных местоимений с помощью типологии семантических переходов производится в работе в следующих случаях: др.-яп. *mi* 'я' сравнивается с омонимичным др.-яп. *mi* 'тело' (с. 97; здесь уместна была бы ссылка на [ITABASHI 1998: 129], где как раз обсуждается происхождение др.-яп. 自 *mi* 'я' — как специализации значения 自 *mi* 'сам' — из др.-яп. 身 *mi* [совр. *mi*] 'тело'); для ностр. **mV* 'я' осторожно предполагается происхождение от ностр. **menV* 'сам, тело' (с. 113–114); для ностр. **qV* 'я' — вслед за А. Б. Долгопольским — происхождение от слова со значением 'сам' (с. 136–137); для ИЕ **ius* 'вы' — сочетание демонстратива **i* с местоимением 2-го л. мн. ч. **us* (с. 208); ностр. местоименная основа **nV* возводится к показателю косвенной основы **nV*, или, вместе с ним, к **nV* 'сторона' (с. 215). Несомненно, все эти этимологические гипотезы вполне логичны и имеют право на существование.

ния из местоименной парадигмы, то есть предполагается обратная направленность процессов. В общем, я готова согласиться с реконструкцией для какого-то раннеиндоевропейского состояния именно такой парадигмы местоимения 1-го л., как предполагает К. В. БАБАЕВ, но для ее доказательства кажется необходимой значительно более подробная проработка вопроса о путях появления засвидетельствованных и реконструируемых позднеиндоевропейских систем.

К. В. БАБАЕВ вполне правомерно выделяет в индоевропейской парадигме склонения личного местоимения 1-го л. косвенную основу **me-*, но, например, насколько правильно считать **te* основной объектной (т. е. винительного падежа) формой, неясно; в свете общей структуры индоевропейской местоименной парадигмы вполне возможно, что это усеченная клитическая форма, а старым винительным была, например, форма с окончанием **me-(e)m* (как полноударное ПСлав. **me*, впоследствии встроившееся в подпарадигму клитик, а в подпарадигме полных местоимений замененное на вторичное **menę* — см. [ДЫБЬ 1971]). Клитические формы, наряду с полными, имеются и в «общеиндоевропейской», и в анатолийской местоименной парадигме, и в книге не проведено обоснование, например, такого положения, что клитические формы парадигм в зафиксированных языках представляют собой архаичные словоформы, а полные — инновационные. Напротив, остается не опровергнутым возможное предположение, что клитические формы — фонетически усеченные полные (при возможности того, что впоследствии клитическая форма в отдельных парадигмах могла вытеснить полную, от нее могли образоваться инновационные падежные формы, от которых могли произойти новые клитические формы). Автору удается обосновать выделение *-m-* как общей неизменяемой части глагольных окончаний 1-го л., но уже с обоснованием того, что первичная форма личного окончания 1-го л. мн. ч. — **me-s* или **-mo-s* (или, если считать *-s* инновационным показателем мн. ч. из субстантивной парадигмы, как и греко-хеттское **-n-*), все не так просто. Встречающиеся в славянских диалектах окончания *-me* и *-mo* — предмет бурной полемики славистов, поскольку они не являются закономерными фонетическими наследниками **-mos* и к тому же встречаются в основном в диалектах, для которых зафиксировано нестандартное поведение слабого ера. Соответственно одни слависты считают их диалектными фонетическими вариантами ст.-слав. *-mъ* < **-mos*, другие — потомками других форм индоевропейских окончаний. Автор никак не выражает своего отношения к этой проблеме и, по-видимому, считает *-mo* потомком **-mos*, *-me* потомком **-mes*, а про глагольные окончания *-mъ* и *-мы* вообще ничего не пишет.

Что касается обоснования того, что «агглютинативная» форма **me-s* — архаическая для индоевропейских языков форма личного местоимения множ. ч., то для этого требуется доказать вторичный характер другой возможной прямой (номинативной) формы мн. ч., **yei* (такая форма восстанавливается по германскому и индоиранскому материалу и не противоречит хеттскому и тохарскому). Это доказательство подробно проведено еще в книге О. СЕМЕРЕНЬИ ([СЕМЕРЕНЬИ 1980: 232–233]); в рецензируемой работе оно повторено и несколько расширено за счет добавления разбора тохарского материала и случаев употребления клитической версии этой основы в глагольных окончаниях. Предполагается, что это — первоначально аналитическая форма двойственного числа, происходящая из числительного **dye-*. Сильным аргументом против вторичности этого местоимения кажется все же его наличие именно как формы множественного числа в анатолийской и тохарской группах, как наиболее рано отделившихся. К. В. БАБАЕВ пытается оспорить этот аргумент следующим образом (с. 163): «...Единственным языком, где **we* не засвидетельствовано в двойственном числе, является хеттский, где этого числа не существовало вовсе. Так как существует множество языков, где **we* функционирует в дуалисе и плюралисе, или же только в дуалисе (на самом деле такова только балто-славянская группа — А. Д.⁸), но нет ни одного, где он был бы только в плюралисе, можно считать доказанным, что первоначальным значением основы было именно значение двойственности, позже перенесенное на множественное число. И возражение о том, что морфологическое двойственное число, как полагают, развилось в индоевропейском только после выделения анатолийских диалектов, здесь не играет роли; значение двойственности может выражаться в языке синтаксически, причем очень тривиально: с помощью числительного *два*». Но задумаемся над тем, каким образом местоимение со значением двойственности может начать выражать множественное число — причем в наиболее далеко разошедшихся между собой диалектах. Для этого нужна либо хорошо развитая морфологическая категория двойственности, подпарадигма двойственного, прочно встроенная в систему именных парадигм и потому получившая возможность влиять аналогическим путем на подпарадигму множественного

⁸ Косвенным указанием на вторичность такого положения в балтославянском может служить как раз литовское *vėdu* < **ye-* + числительное «два», т. е. в дуалисе потребовалась специальная маркировка двойственности.

числа: тогда приходится предполагать, что все-таки в праиндохеттском периода распада было грамматическое двойственное число (хотя, конечно, на более раннем этапе оно могло возникнуть из числительного — но этот этап можно получить только с помощью внутренней реконструкции и надежно подтверждался бы он только в случае абсолютной невозможности получить внешние, ностратические параллели для **uei* как местоимения 1-го л., каковая невозможность — забегая вперед — все же осталась недоказанной в книге К. В. БАБАЕВА); либо надо предполагать для праиндохеттского состояния периода распада развитую категорию инклюзивности, которую также в принципе может выражать форма, образованная от числительного *два* (т. е., практически как предложено в [ГАМКРЕЛИДЗЕ, ИВАНОВ 1984: 254]), затем падение этой категории в одних языках (в анатолийской группе — при этом вытесняется старое эксклюзивное — при такой конструкции — **mes*) и преобразование ее в категорию двойственного числа в других языках (в праобщиндоевропейском?). Но эта конструкция у К. В. Бабаева не рассматривается, а на с. 159 (по поводу семантики местоимения **ne*) прямо утверждается: «В индоевропейских языках категория эксклюзивности/инклюзивности не имеет доказательной базы». В общем, для того, чтобы устранить из восстанавливаемых для праиндоевропейского местоименных корней 1-го л. **uei*, автор должен был бы проделать значительно более подробное построение, демонстрирующее не только тот набор, который восстанавливается для сравнения с ностратическими параллелями, но и как из этого набора получают позднейшие системы. Наличие же фонемы **u* в составе основы местоимения 2-го л. мн. ч. (которое автор рассматривает, вслед за О. СЕМЕРЕНЬИ, в качестве аргумента на с. 164) совершенно не мешает архаичности **uei* — ср. среднекорейское *nà* ‘я’ и *nə* ‘ты’, происшедшие, судя по всему, из разных местоимений с разными начальными согласными. Так что при разрешении проблемы трех потенциальных праиндоевропейских местоимений 1-го л. мн. ч. — основной проблемы, разрешаемой в главе 3 — достаточно очевидным образом К. В. БАБАЕВУ удается продемонстрировать только, что основная зона употребления индоевропейского **ne-* — это не специфическая эксклюзивность или инклюзивность, а косвенные формы местоимения 1-го л. мн. ч., что как раз очень хорошо согласуется с его внешними соответствиями (с. 147–159). Продемонстрировано также, что гипотеза о связи **nV* с инклюзивностью/эксклюзивностью в рамках ностратического сравнения не подтверждается, зато довольно хорошо видно, что оно тяготеет к косвенным формам. Что же касается **uei*, то его исключение из местоимений 1-го л. осталось не до конца убедительным. Сопоставления А. Б. Долгопольского из области семито-хамитских местоимений, действительно, не выглядят слишком надежными; спорно и собственно местоименное сопоставление с картвельским **čwen-*, которое не очень пока понятно, как членится морфологически, но общекартвельский субъектный префикс 1-го л. ед. и мн. чисел **w-*, почему-то не попавший в этимологию А. Б. Долгопольского, очевидным образом напрашивается как параллель.

К числу достоинств этой главы следует также отнести более доскональную, чем у предшествующих авторов, проработку этимологии ностр. **qV*, в которой значительно увеличен привлекаемый материал и семантические обоснования. Не вполне завершенной, тем не менее, оказывается реконструкция системы праностратических местоимений 1-го л. в отношении их синтаксических/семантических свойств; и здесь также, на мой взгляд, сказывается пренебрежение «обратным ходом» реконструкции — построением вывода наличных систем из реконструированной, что, по крайней мере, для обоснования морфологических реконструкций, по-видимому, совершенно необходимо. На с. 92 книги об индоевропейском независимом личном местоимении **me* сначала говорится: «На основании сравнения глагольной системы, где оно выступает в неперфектной парадигме, и местоименной системы, где в ед. ч. оно фигурирует в косвенных падежах, его праязыковым синтаксическим значением должно являться значение первого лица субъекта переходного глагола действия». Вывод был бы совершенно непонятен, если не вспомнить, что К. В. БАБАЕВ допускает первичную активность/эргативность праиндоевропейского. В эргативном языке «косвенная» основа может выражать субъект переходного глагола, приняв форму эргатива. А в следующем абзаце мы видим: «Таким образом, на индоевропейском праязыковом уровне показатель 1-го л. **m-* мог функционировать в качестве независимого полнозначного объектного местоимения». Конструкция, в которой одна и та же форма местоимения является выразителем субъекта переходного глагола и его же объекта, совершенно определенно не согласована. К тому же получается, что по глагольным показателям неперфектная (т. е. предположительно не-стативная) серия совпадает с косвенной основой местоимения, а перфектная (стативная) серия — с прямой основой местоимения (ларингальной). Действительно, в известных мне эргативных языках личное согласование глагола обычно происходит по субъекту, т. е. у переходного

глагола с эргативом, у непереходного — с номинативом. Но на с. 100 говорится: «На пракарт-вельском уровне просматривается аналогичное индоевропейскому, уральскому и алтайскому распределение основ: *me* в функции прямого падежа и **me-n(V)* в функции косвенного падежа». Что касается уральского и алтайского, это, пожалуй, действительно так, но для индоевропейского **me-* — косвенная основа, а **mene* — генитив. При этом автор решительно и, по-видимому, совершенно справедливо утверждает, что для праностратического эргативный строй не восстанавливается, а восстанавливается аккузативный (с. 237–238). Тогда приходится предполагать, что по пути от ностратического к индоевропейскому язык дважды поменял синтаксический строй, сначала развил эргативность, а затем из нее — опять номинативность/аккузативность; это возможно, но кажется маловероятным, чтобы при этом отношения прямых и косвенных форм местоимений остались прежними. При этом для праностратического состояния К. В. БАБАЕВ восстанавливает местоимение первого лица **qV* со значением показателя субъекта при интранзитивном глаголе, а местоимение **me* характеризует как показатель субъекта при транзитивном глаголе (с. 169–172). Но ведь именно наличие подобных специфических маркировок — субъект интранзитивного глагола, противопоставленный субъекту транзитивного глагола — и определяет эргативный или активный строй языка — значит, праностратический все же эргативен? Налицо явное противоречие или недостратифицированность реконструкции, которой автор не уделяет никакого внимания, что не на пользу общей фундированности концепции.

Частные замечания. С. 99. О личных показателях в глагольных формах части монгольских и тунгусо-маньчжурских языков, как восходящих к именительным падежам личных местоимений: автор ссылается на [SINOR 1988]; лучше было бы здесь отнестись к [PORRE 1955] и [BENZING 1955], где действительно доказательно продемонстрировано происхождение этих аффиксов.

С. 103: говорится про ПА **män-* и его роли для реконструкции ПН **mä-* или **me*. Но **män-* — только праторкская форма косвенной основы, прямая основа по тюрк., монг. и ТМ материалу дает **bi*, вообще же восстанавливается чередование **bi / *mi-ne-* (sing.) ; **ba / *mju-n-* (plur.) (см. [EDAL]).

С. 123. О восстановлении ностратического показателя 1-го л. **q-*: «Несмотря на отсутствие соответствий такого рода в классических трудах по ностратике ([Иллич-Свитыч 1971: 147–150], [DOLGOROLSKY 1998: 115]), в последнее время исследователями все чаще приводятся данные в пользу подтверждения системного соответствия между индоевропейскими ларингалами и глухой велярной фонемой **k* других ностратических языков». Такое соответствие есть в [Иллич-Свитыч 1971: 149], это, например, ПН **-q-*, которое К. В. БАБАЕВ и восстанавливает для этого местоимения, вслед за Е. А. Хелимским ([Хелимский 1979: 17–18]). С. 135, об этом же соответствии: «Фонетический переход ларингального в велярный и обратно — частое типологическое явление в языках мира». Но здесь нет речи о переходе ларингального в велярный или наоборот, речь о развитии увулярного (а не велярного) в одних языках в ларингал, в других — в велярный; тоже вполне обычное развитие.

С. 141: о реконструкции ПИЕ местоимения 1-го л., прямой формы: «...Основной согласный местоимения **egho(m)*, который может быть восстановлен ... как **gʰ* на основе индоиранских и балтославянских форм, либо как **gʷ* на основе форм германского, латинского и греческого». Никак нельзя восстанавливать **gʰ* на основании как минимум славянских форм, так как там наблюдается удлинение гласного по закону Винтера (слав. **āzъ*), позиция действия которого — перед индоевропейским звонким непридыхательным. Автор не может не знать этого, т. к. абзацем ниже ссылается на статью В. А. Дыбо о законе Винтера ([Дыбо 2002]).

Реконструкция показателей 2-го л., проведенная в гл. 4, сама по себе менее проблематична, чем реконструкция 1-го л., и, возможно, в связи с этим раздел практически не вызывает существенных возражений. Единственное — к уже упоминавшемуся выше возведению местоимения 2-го л. мн. ч. **ye-* к числительному «два»: оно тем более рискованно, что автор не пытается никак объяснить в этом случае, как могла получиться форма прямого падежа с вокализированным **i*: при этом трудно ожидать отпадения **d-*, в отличие от ситуации с **dye-*.

Частности. С. 183: автор говорит о внешнем сходстве древнеяпонских местоимения 2-го л. *si* и указательного местоимения = местоимения 3-го л. *si*; говорит, что это — совпадение, родственное индоевропейскому соотношению, и в качестве примера индоевропейского соотношения приводит слав. **-si* 2-го л. в глаголе и **si-* — ближайшее дейктическое местоимение. Если бы не было употреблено слово «родственное», а имелась бы в виду возможность случайного совпадения, то все было бы в порядке с этим примером (заметим при этом, что интерпретация соответствующих контекстов из *Маньёсю*, сама по себе не является легко решаемой проблемой), за исключением

того, что, приводя пример на еще одно точно такое же случайное совпадение, как в случае *X*, мы заставляем собеседника слегка усомниться в случайности этого совпадения в случае *X*, так что лучше бы мы этого примера не приводили. Если же речь идет о том, что в ностратических языках наблюдается сходство местоимения 2-го л. и указательного местоимения, поскольку они восходят к ПНостр **si* ‘местоимение 2-го л.’ и **šV* (так в [Иллич-Свитыч 1971], ср. [Дыбо 2004: 95, № 31]: **šV*) ‘дейктическое местоимение’, то славянский пример неправильный, поскольку слав. *сь* ‘этот (ближайший)’ восходит к ПИЕ **kī-s*, представленное в лит. *šis* «этот», лат. *cis* «по эту сторону», хетт. *kā-*, *ki-* ‘этот’ и пр.; чтобы не вводить читателя в недоумение, следовало бы приводить, например, санскр. *-sī* в глаголе и *sa-s*, прямую основу указательного местоимения м. р.

С. 193: гипотеза о том, что *-ti* в ПТМ инклюзивном местоимении 1-го л. **munti* является старым местоимением 2-го л., не представленным самостоятельно в ТМ языках, но сохранившимся в монгольском (**mun-*, соответственно, косвенная основа местоимения 1-го л. мн. ч.), высказанная, например, у Рамстедта и повторенная Редди, цитируется автором в слишком, на мой взгляд, уверенной модальности. Это, действительно, вполне вероятно, но возможно и какое-то другое происхождение показателя **ti* (тем более, что **munti* не восстанавливается для ПТМ до первого распада: в маньчжурском *muse* из той же косвенной основы **mun* и обычного местоимения 2-го л. мн. ч.); а для монг. инклюзива **bida*, который на с. 194 вслед за Рамстедтом объясняется из **bi* ‘я’ и **ta* ‘вы’, такое объяснение фонетически незакономерно, интервокальных озвончений в монгольском мы не знаем. См., например, комментарий к этому месту в русском издании Рамстедта, со ссылкой на действительно информативную работу [Kotwicz 1926].

Пятая глава, «Опыт реконструкции парадигмы показателей лица в ностратическом праязыке», начинается с обзора достижений предшественников и краткого методологического манифеста касательно морфологической реконструкции, под которым я готова подписаться обеими руками. В частности, там проговаривается необходимость верификации морфологической реконструкции путем построения обратного прохода от реконструированной системы к рефлексным. Автор действительно делает в этой главе, пусть очень краткую, попытку такого прохода. Как мы уже говорили выше, у него восстанавливается для праностратического (причем для «широкого» праностратического, включающего афразийский и палеоазиатские языки) системное противопоставление местоимения — субъекта переходного глагола и местоимения — субъекта непереходного глагола, как в 1-м, так и во 2-м л. Реконструкция такого противопоставления опирается прежде всего на индоевропейские факты, а именно, на тяготение ларингального показателя 1-го л. и показателя 2-го л. **t-* к перфектной и медиальной подпарадигмам глагола (для перфекта типологически естественно предполагать первоначально стативное значение, отсюда связь с непереходностью). Значения соответствующих показателей в других ностратических языках в случае наличия противопоставления удается вполне успешно проинтерпретировать в этом же направлении. Что касается направления семантических изменений, автор предполагает, что местоименный маркер субъекта непереходного глагола в части случаев обобщился как показатель при стативе, а кое-где стал выполнять роль эмфатического вынесенного в абсолютную позицию личного местоимения (*Who is his father? — Me*). А местоименный маркер субъекта переходного глагола становится универсальным маркером субъекта любого глагола не в стативе. Это все — действительно вполне мыслимые развития. В то же время при полиперсональном спряжении этот маркер субъекта переходного глагола, по мнению автора, логично переходит на роль маркера объекта переходного глагола. Он ссылается при этом на многочисленные типологические параллели такого развития в книге [Corbett & al.: 99–100]. Не имея под рукой книги Корбетта, я не могу оценить точность этих параллелей, но само развитие отнюдь не кажется логичным — предлагается чуть ли не энантиосемический переход без каких-либо промежуточных этапов. Невозможно представить себе синтаксическое развитие, при котором могла бы произойти такая перемаркировка актантов. С другой стороны, такой механизм мог бы объяснить часто предполагаемые диахроническими типологами (но, впрочем, кажется, ни разу не предьявленные) случаи перехода от эргативной маркировки актантов к аккузативной. В общем, логика предполагаемого перехода неясна, а подтверждающие примеры недоступны прямому наблюдению. Дальнейшее развитие основы объектного маркера как общекосвенной основы, впоследствии вытесняющей кое-где из парадигмы супплетивную «прямую» основу субъекта там, где она возникла из бывшего интранзитивного субъектного местоимения, опять же, вполне правдоподобно. Последующие построения по поводу общего происхождения косвенной основы местоимения 1-го л. **nV* вместе с неясного характера местоимением 2-го л. **nV* и показателем косвенной/атрибутивной формы **nV* от пол-

нозначного слова со значением 'сторона' возможны, но совершенно не обязательны, во всяком случае, до тех пор, пока мы не убедимся в том, что *V* во всех этих случаях — одна и та же гласная. Пока «омонимия» всех этих морфем основана на совпадении одного элемента.

Параграф о процессах трансформации парадигм местоимений в ностратических языках и параграф о формах множественного числа местоимений написаны совершенно реалистически и содержат много типологически интересных примеров.

А вот к разделу о падежах в ностратическом праязыке у меня находятся существенные возражения. Во-первых, трудно согласиться с тем, что мы имеем какие-то реальные свидетельства того, что праностратический язык был изолирующим. Этого, во всяком случае, нельзя вывести из того, что некоторые морфемы, дошедшие до нас в виде аффиксов, раньше были отдельными словами. Это — результат уже неоднократно упоминавшегося автором процесса грамматикализации, свойственного всем языкам, а не только изолирующим. Просто при настоящем уровне обработки ностратической морфологии мы умеем восстановить только такие случаи. Вполне возможно, что в праностратическом состоянии имелась развитая флективная морфология, окончания которой впоследствии затерлись процессами конца слова и заменились новыми, а грамматические чередования в корнях затерлись процессами парадигматической унификации и сохраняются только в нерегулярных случаях соответствия гласных, которые мы на настоящем уровне реконструкции ностратического вокализма либо не видим как нерегулярные, либо относим из осторожности к разным корням. Примеры: английский язык, в котором бывшая флективность видна только в вокализме глагольного корня и других остаточных формах; когда все *took* заменяется на ⁺*taked*, *oxen* на ⁺*oxes*, а (уже) вторичные *feet* на ⁺*foots*, *ee* не будет видно вовсе. Романские языки, если над ними провести морфологическую реконструкцию, что касается именного словоизменения, позволят только — в ряде тонких случаев и в основном с применением старофранцузских форм — понять, что у некоторых слов были прямая и косвенная формы, но никак не восстановить латинскую систему склонения.

Затем, у автора, насколько я могу понять, происходит какая-то путаница с понятием эргативности. Он правильно пишет, что в эргативном языке одинаково маркируются субъект непереходного глагола и объект переходного, и иначе — субъект переходного глагола (с. 237–238). Но почему-то дальше он замечает: «И субъект, и объект всегда маркируются», и рассуждает о том, что, поскольку полиперсональное спряжение для ностратических языков вторично, то праностратический нельзя назвать эргативным языком. Во-первых, множество эргативных языков не имеют полиперсонального спряжения. Во-вторых, они могут и вовсе не иметь личного спряжения, и не обязаны маркировать тождество/разницу актантов внутри глагольной словоформы. Важна именно одинаковость выражения субъекта непереходного глагола и объекта переходного при отличии выражения субъекта переходного — а его-то сам же автор и восстановил для праностратического, и значит, его праностратический либо эргативен, либо активен (в активном языке субъект непереходного глагола маркируется по-разному в зависимости от того, агенс он или пациенс, в первом случае — как субъект переходного, во втором — как объект переходного). Но эта реконструкция, как он сам замечает, вступает в противоречие с довольно хорошей восстановимостью для праностратического показателя прямого объекта (т. е. аккузатива) в отличие от номинатива — эта форма неизвестна эргативным языкам.

Чтобы разрешить противоречие, можно было бы предложить попытаться рассмотреть еще одну возможность реконструкции употребления личных показателей в индоевропейском глаголе. Если предположить, что одна из серий является по происхождению приименными показателями, т. е. посессивными по происхождению, т. е. восходящими к косвенным формам местоимений, а другая — собственно глагольными показателями, восходящими к прямым основам местоимений, то для двух серий ностратических местоимений можно восстанавливать, соответственно, различие прямой и косвенной основ. Возможно, имеет смысл рассмотреть подробнее такую конструкцию.

Книга содержит также краткую библиографию по проблемам дальнего родства и грамматической типологии и — очень удобное для читателя — приложение с парадигмами местоимений и личных показателей в ностратических и пара-ностратических языках (с. 257–295).

В целом следует признать, что, хотя, как я уже писала в начале, нельзя считать, что эта работа окончательно разрешила проблему ностратических местоимений, книга получилась полезная и интересная, написана достаточно профессионально, что выгодно выделяет ее среди многих работ по дальнему родству, а дискуссионность ряда положений только добавляет ей интересности и противодействует застою научной мысли.

- ГАМКРЕЛИДЗЕ, ИВАНОВ 1984 — Тамаз В. ГАМКРЕЛИДЗЕ, Вячеслав Вс. ИВАНОВ. [თ. ბაჰყრელიძე, ზ. ივანოვიძე ~ Thomas V. GAMKRELIDZE & VJACHESLAV VS. IVANOV]. *Индоевропейский язык и индоевропейцы. Реконструкция и историко-типологический анализ праязыка и протокультуры*. Тбилиси: Издательство Тбилисского Университета [*ინდოევროპული ენა და ინდოევროპელები. წინარეებისა და წინარეკულტურის რეკონსტრუქცია და ისტორიულტიპოლოგიური ანალიზი*. თბილისი: თბილისის უნივერსიტეტის გამომცემლობა ~ *Indo-European and the Indo-Europeans. A Reconstruction and Historical Typological Analysis of a Protolanguage and a Proto-Culture*. Tbilisi: Publishing House of the Tbilisi State University].
- ДЫБО 1971 — В. А. ДЫБО. Закон Васильева–Долобоко и акцентуация форм глагола в древнерусском и среднеболгарском // *Вопросы языкознания*, 2; стр. 93–114.
- ДЫБО 2007 — А. В. ДЫБО. Реконструкция праогузского спряжения // *Аспекты компаративистики*. Вып. 2 (= *Orientalia et classica XI: Труды Института восточных культур и античности*). М.: Издательство РГГУ; стр. 259–280
- Иллич-Свитыч 1971 — В. М. Иллич-Свитыч. *Опыт сравнения ностратических языков*. Т. 1 (b–K). М.: «Наука».
- МОРЕВ и др. 1961 — Л. Н. МОРЕВ, Ю. Я. ПЛАМ, М. Ф. ФОМИЧЕВА. *Тайский язык (Языки зарубежного Востока и Африки / Под общ. ред. проф. Г. П. Сердюченко)*. АН СССР, Ин-т народов Азии. М.: Изд-во Восточной литературы.
- ОРЕЛ 1990 — В. Э. ОРЕЛ. К происхождению личных местоимений в семито-хамитском // *Сравнительно-историческое языкознание на современном этапе. Конференция памяти В. М. Иллич-Свитыча*. М.; стр. 54.
- РАМСТЕДТ ГКЯ — Г. РАМСТЕДТ. *Грамматика корейского языка*. М.: «Прогресс». 1951.
- СЕМЕРЕНЬИ 1980 — О. СЕМЕРЕНЬИ. *Введение в сравнительное языкознание*. М.
- СТАРОСТИН 1993 — С. А. СТАРОСТИН. Заметки о древнекитайском языке // *Знак*. Сборник статей по лингвистике, семиотике и поэтике. М.: Русский учебный центр; стр. 93–126.
- СТАРОСТИН 2007 — С. А. СТАРОСТИН. Заметки о древнекитайском языке // С. А. СТАРОСТИН. *Труды по языкознанию*. М.: «Языки Славянской Культуры»; стр. 499–528.
- СУНИК 1978 — О. СУНИК. Местоимения «сам», «свой» и их морфологические дериваты в алтайских языках // *Очерки сравнительной морфологии алтайских языков*. М.; стр. 232–268.
- ХАЙДУ 1985 — П. ХАЙДУ. *Уральские языки и народы*. М.: «Прогресс».
- ХЕЛИМСКИЙ 1982 — Евгений А. ХЕЛИМСКИЙ. *Древнейшие венгерско-самодийские языковые параллели*. М.: «Наука».
- ЭДЕЛЬМАН 1990 — Д. И. ЭДЕЛЬМАН. *Сравнительная грамматика восточноиранских языков. Морфология. Элементы синтаксиса*. М.
- BENZING 1955 — J. BENZING. *Die tungussischen Sprachen. Versuch einer vergleichender Grammatik*. Wiesbaden.
- CORBETT & al. — *Heads in Grammatical Theory / Ed. by Greville G. CORBETT, Norman M. FRASER, Scott McGLASHAN*. Cambridge: Cambridge University Press (July 30, 1993).
- DOLGOPOLSKY 1998 — Aharon DOLGOPOLSKY. *The Nostratic Macrofamily & Linguistic Paleontology (Papers in the Prehistory of Languages)*. The McDONALD Institute for Archaeological Research.
- DYBO 2002 — V. A. DYBO. Balto-Slavic Accentology and WINTER's Law // *Studia linguarum*, 3/2, М.; pp. 295–515.
- DYBO 2004 — A. V. DYBO. Some Peculiarities of Altaic Reflexes of Nostratic Sibilants // *Nostratic Centennial Conference: the Pécs Papers / Ed. by Irén HEGEDŰS and Paul SIDWELL*. Pécs: *Lingua Franca* Group; pp. 85–114
- HEINE–KUTEVA 2007 — Heine B., Kuteva T. *The Genesis of Grammar: A Reconstruction*. Oxford, 2007.
- ITABASHI 1998 — ITABASHI Yoshizo [板橋 義三]. The Old Japanese Personal Pronouns as an Etymological Problem // *UAI*, Vol. 70.
- KOTWICZ 1926 — W. KOTWICZ. *Les pronoms dans les langues altaïques*. Kraków.
- POPPE 1955 — N. POPPE. Introduction to Comparative Mongolian Studies // *Mémoires de la Société finno-ougrienne*, 110.
- SIEWIERSKA 2004 — A. SIEWIERSKA. *Person*. Cambridge.
- SINOR 1988 — Sinor D. The Problem of the Ural-Altaic Relationship // *The Uralic Languages. Description, History and Foreign Influences*. Leiden; pp. 706–741.

ABC Etymological Dictionary of Old Chinese
By Axel SCHUESSLER
Honolulu: University of Hawai'i Press. 2007. xx + 668 pp.

George STAROSTIN
(Russian State University for the Humanities)

Appearing somewhat surreptitiously and certainly lacking the proper publicity, Axel SCHUESSLER's *Etymological Dictionary of Old Chinese* is, in fact, arguably the biggest thing that has so far happened to Chinese historical linguistics studies in the 21st century. Finally, the scholarly world has come into possession of a worthy reference source on the history and origins of a significant portion of Old Chinese vocabulary — the first such source ever published. Regardless, then, of whether Prof. SCHUESSLER's work is flawless or not, it constitutes a major landmark in the history of Sinology, and will serve as the benchmark by which many, if not most, further works dealing with the history of the Chinese language will have to be judged.

With an immense baggage of knowledge on both the internal developments and the external connections of Chinese accumulated over the last century by dozens of specialists, Chinese and Western alike, the appearance of such a dictionary was, of course, a historical inevitability. Today, the Sino-Tibetan affiliation of Chinese is no longer disputed as it once used to be, and, with the acknowledged status of Austronesian, Austro-Asiatic, Hmong-Mien (Miao-Yao), and Tai-Kadai language families as forming strong areal rather than genetic connections with Chinese, it has become generally clear how to treat these connections from a methodological point of view. Moreover, the fact that the idea of finally compiling and publishing a proper etymological dictionary of Old Chinese was carried out by none other than Prof. SCHUESSLER is quite auspicious, since he is one of the biggest Western authorities on Chinese historical linguistics; his earlier "Dictionary of Early Zhou Chinese" is easily the finest resource in existence on this archaic and, even today, least understood and well-studied (but most important for various philological studies) period of the Chinese language, and, over the past several decades, he has made numerous valuable contributions to the study of Chinese historical phonology, grammar, and external linguistic relations.

Several important and useful features of the dictionary should be stressed at once. First, it is preceded by a detailed, 150-page long introduction that carefully explains the structure of the work, the methodological principles underlying it, and the general linguistic situation in and around Ancient China as assumed by the author for the times when Old Chinese was still a spoken language. On an equally significant note, we get a thorough description of the author's view on the phonological and morphological structure of Old Chinese, as well as the origins of this structure, reconstructed through comparison with other branches of the Sino-Tibetan family. This is particularly meaningful inasmuch as no general consensus has so far been reached by linguists on many of the raised issues, and it is very important for specialists and casual readers alike to learn more about the motivation behind many (if not most) of Prof. SCHUESSLER's etymologies offered in the main body of the work.

Concerning this main body, primary emphasis is placed on making the work a word-based rather than character-based dictionary — a decision with which most Sinologists, no doubt, will heartily concur. Relying too heavily on the graphic form of the etyma would make the reader miss many obvious (and some not so obvious) connections between words that, already from the earliest times, were written with different Chinese characters, yet were nevertheless etymologically related — or, vice versa, assume false connections between words of entirely different etymological origin, whose only similarity was in that the same character used for one word would also be used for a different one due to phonetic similarity between the two.

The dictionary is, therefore, structured according to the phonetic principle, with all the words arranged in the order of their modern *pinyin* readings. A separate character index would have been nice, but is not obligatory provided one knows the primary reading of the character; characters denoting derived words whose etymologies are inseparable from etymologies of other ('primary')

words simply refer to these words. E. g., a search for the etymology of 豹 *bào* ‘leopard’ will lead the reader to the entry *bào* 豹 → *bó* 駮, which should be read: “*bào* ‘leopard’ is originally derived from *bó* ‘variegated’, see under *bó*”. This is a fairly reasonable approach.

The overall structure of the entry, besides the expected reconstruction for Middle Chinese and Old Chinese (usually several distinct stages of the latter) and the etymological section, usually includes other valuable information, such as attested dialect forms and/or transcriptions of the etymon in question in such medieval sources as the ’Phags-pa transcribed *Měnggǔ Ziyùn* [蒙古字韻]. The latter information, as the author correctly remarks, may occasionally shed light on the word’s later development. However, it is provided somewhat inconsistently.

Obviously, no matter how ample our knowledge on Chinese etymology may be at present, multiple Old Chinese stems still lack any kind of historical explanation. Such words are simply omitted from the dictionary altogether, so the reader should not expect to open it in the hopes of finding at least something about any character ever recorded in attested Old Chinese sources.

Aside from a somewhat clumsy system of abbreviations (a problem exacerbated by the fact that different types of abbreviations are decoded in fairly distant parts of the book, so that it takes a fair amount of time to properly understand the contents of each entry), the general layout of the dictionary is, therefore, almost impeccable. Much more complicated, but, of course, also much more important is the issue of properly evaluating its scientific content.

We should begin with the obligatory remark that the creation of an etymological dictionary for Old Chinese runs into a set of problems quite different from the ones involved in the creation of, for instance, an etymological dictionary for any given Indo-European language. First, Old Chinese itself is, essentially, a reconstructed language, at least in its phonological (and partially morphological) aspects; therefore, before comparing Old Chinese forms to their hypothetical cognate forms in other Sino-Tibetan languages or their hypothetical ancestors in non-Sino-Tibetan languages from which they are supposed to have been borrowed, it is necessary to have a solid historical justification for the Chinese form itself.

Second, unlike Indo-European, Proto-Sino-Tibetan has yet to be properly reconstructed. Despite significant progress in this field, achieved through the work of scholars such as P. BENEDICT, R. SCHAFER, N. C. BODMAN, W. S. COBLIN, S. STAROSTIN, I. PEIROS and others, unresolved issues in Tibeto-Burman, let alone Sino-Tibetan, linguistics still abound, with several competing versions of Sino-Tibetan phonology, multiple languages and whole subgroups still linked to the main “bulk” of Sino-Tibetan only by a few suggested cognates, and no definitive corpus of Sino-Tibetan etymologies; the only extensive comparative dictionary of Sino-Tibetan languages in existence is I. PEIROS’ and S. STAROSTIN’S “*Comparative Vocabulary of Five Sino-Tibetan Languages*”, which is a very serious piece of work, but, with only five languages involved, is obviously not enough. Thus, any etymological dictionary of Old Chinese will, by definition, rest on a somewhat shaky foundation, since it will be attempting to link the Chinese forms to a family that has not yet been properly reconstructed itself.

Third, there is still nothing like a universally accepted scenario for both prehistorical and historical contacts between speakers of Old Chinese and speakers of other language families. At best, everyone recognizes the reality of such a contact between the Chinese (including “proto-Chinese”) and various members of the hypothetical “Austic” macro-family, but the exact chronological, geographical, and quantitative scale of most of these contacts is far from being established. Numerous resemblances between Old Chinese and Tai-Kadai (or Hmong-Mien, or Mon-Khmer, &c. &c.) etyma have been spotted by researchers, and much of this evidence has been painstakingly documented in Prof. SCHUESSLER’S work, but for quite a few of them we still cannot definitely establish, for instance, the direction of borrowing, or even an approximate time and place in which the borrowing occurred (provided it is indeed a borrowing, and not a chance resemblance, or a trace of some older, deep-level relationship).

All of these complex factors have to be kept in mind by whoever wishes to use Prof. SCHUESSLER’S book as a reliable reference source. The biggest limitation (one for which Prof. SCHUESSLER certainly cannot be held responsible), is that this is, indeed, the first etymological dictionary of Old Chinese — and, like all first etymological dictionaries of any language or language family, it is only the first step in the direction of producing a thorough deciphering of the prehistory of the Chinese lexicon. Most of the etymologies have to be taken with a grain of salt, not because they are unrealistic or seriously violate

any principles of comparative Sinology or comparative linguistics as a whole (such cases are quite rare), but simply because there is a high chance that better, more convincing etymologies will replace them in the future, as our knowledge of the linguistic situation in South-East Asia steadily increases.

In order to clarify this statement, it is perhaps reasonable to voice some specific concerns on each of the three issues raised above.

The basic foundation for all of the etymologies offered in the dictionary is the reconstructed Old Chinese (OC) form. At present, the two most comprehensive reconstructions for Old Chinese, available in published form, are BAXTER ([1992]) and STAROSTIN ([1989]); both are remarkably similar in many respects, especially concerning the treatment of OC finals, but significantly differ in others, such as proposed subsystems of OC initials. For the most part, SCHUESSLER's version of OC is compatible with BAXTER's, although he proposes quite a few of his own modifications to the system, at the same time keeping in touch with more recent hypotheses by L. SAGART ([1999]) and others. STAROSTIN's work is mostly ignored (odd enough, STAROSTIN ([1989]) is not even mentioned in the bibliography!), which is a pity, since many of the hypotheses presented there make a lot of sense from a general Sino-Tibetan perspective (e. g. the reconstruction of a phonological opposition between simple and palatalized affricates and fricatives, which, according to STAROSTIN, is a good match for the same opposition in Tibetan and Lolo-Burmese).

An even worse problem is that, despite the length of the introduction, the reader does not get a very coherent picture of the phonological structure of Old Chinese as a whole. There is no general inventory table for OC, nor do we get a strict and compact set of rules to let us easily understand the transition from OC to Middle and Modern Chinese. In a rather chaotic way, chapters that deal with OC phonology take, as their point of departure, either OC itself ("Initial consonants", "Final consonants"), or Middle Chinese ("MC tones and their Old Chinese equivalents", "Initial and medial *j* and the MC divisions"), so that the relevant information sometimes gets doubled or even tripled in several locations. Discussion can easily branch from OC to either its descendants (MC), ancestors (Sino-Tibetan) or contacts (Austro-Asiatic, Tai, &c.); it would have probably been a better choice to treat all these connections separately and more systematically, especially considering that many of the statements made in these sections are debatable, and a clearer presentation of them is essential for further research.

The methodology of the reconstruction itself occasionally suffers from relying too heavily on isolated pieces of potential evidence that do not seem to form a coherent system. This is a very important point, since Old Chinese phonology has, for the most part, been established through a very specific brand of internal reconstruction, and one of the perils of internal reconstruction is that it generally relies much more on subjective judgement than external reconstruction, based on identifying regular phonetic correspondences between related languages. In reconstructing the phonology of OC, our most reliable sources are those whose statistic significance is unquestionable, namely, data gained from the analysis of OC rhymes and well-represented *xiesheng* (phonoideographic) series of Chinese characters. Unsurprisingly, it is the conclusions gained from these types of data that are among the less controversial among specialists on OC reconstruction.

Unfortunately, recent research on OC tends to focus on less reliable data — such as various isolated "oddities" of the *xiesheng* series; approximate phonetic glosses in Han-time dictionaries of Chinese characters; and, worst of all, on the analysis of so-called "word families" — etymological groupings of characters based on similarities in their pronunciation and meaning, even when these similarities do not seem to form easily explainable patterns. This dense mixture of subjective judgement on the part of both ancient Chinese and modern Western philologists inevitably leads to multiple complications of the existing reconstruction that do not have sufficient backing, and vary greatly from scholar to scholar, at the same time confusing the student and making him/her question the more solid parts of OC reconstruction as well, even up to rejecting the possibility of such a reconstruction altogether.

Fortunately for students and scholars alike, Prof. SCHUESSLER's approach to OC is far more "moderate" than, e. g., Laurent SAGART's ([1999]), or, in certain respects, even that of B. KARLGREN (good examples of KARLGREN's exaggerated reliance on "word-families" are adduced on p. 8 of the dictionary). Nevertheless, I would still consider it way too dependent on subjective judgement. Thus, on the same p. 8, we meet with the following statement:

"For example, it seems obvious that the words *jiàn* 監 **krâms* 'look at' and *lǎn* 覽 **râm?* 'to see' are related, but what the difference in later tones and the presence / absence of an initial **k*-

might have entailed is so far a matter of speculation. On the other hand, we can confidently state that *zhì* 織 **təkh* or **təks*, literally ‘something that has been woven’, is a regular exopassive derivation from *zhī* 織 **tək* ‘to weave’. We consider both *jiàn* and *lǎn*, and *zhī* and *zhì* to be allofams in their respective word families”.

The last sentence, in my opinion, constitutes a grave methodological error. It is clear, both to the reader and to the author, that the two examples described above are of an entirely different nature — the latter being one example in a general pattern that can be illustrated by multiple additional examples, the former being merely a subjective similarity that does not form part of any subsystem; how much more ‘obvious’ does it really seem that *jiàn* and *lǎn* are related than, for example, English *rod* and *p-rod* (the former is, of course, something you prod with), or *drain* and *s-train*, two words whose meanings seriously overlap in some contexts, but which nevertheless have entirely different etymologies? We should also keep in mind that, although OC was unquestionably much richer in phonemic inventory than its modern descendants, and more permissible in its phonotactics, the number of allowed syllables was still limited, and the overall degree of homonymy in it (including homonyms with similar meanings) must have been greater than in the European languages we are used to.

From a methodological standpoint, then, it seems more reasonable for the moment to keep words like *jiàn* and *lǎn* apart rather than place them under the same etymon (as the author does on p. 305) — thus creating the illusion that there exists some kind of true scientific evidence to consider them “allofams”. The same recommendation can be applied to dozens, if not hundreds, of etymologizations offered in the main body of the dictionary. An extreme example is *KA* ‘solid, hard’ (p. 261), which, according to the author, serves as the basis of a whole bunch of stems like *ka-k*, *ka-ng*, *ka-r*, *ka-n*, &c., with meanings like ‘hard; solid; strong; dry; withered; freeze’, scattered throughout Chinese, Tibeto-Burman, and even Mon-Khmer languages — but for most of the quoted examples, there is no real evidence other than a purely intuitive feeling that they really descend from a single ancestor **KA*, and, as of now, I see no other possibility than to regard all of them as separate roots.

Sometimes this liberal policy on word-families leads to a fair amount of confusion in particular etymological entries. Thus, although there is a connection proposed between *bīng* 冰 (OC **prəŋ*) ‘ice’ and *líng* 凌 (OC **rəŋ*) id., with both placed under the same entry (p. 168), it is impossible to reconcile this connection with the proposed etymology: the author reasonably relates OC **prəŋ* to PTB **pam* ‘snow, ice’, but this means that *p-* in OC **prəŋ* is part of the root, not a prefix (*contra* [SAGART 1999]); what, then, does the *p*-less form **rəŋ* represent? To answer this question, the author has to substitute his own reconstruction for W. BAXTER’S, respectively 冰 **prjəŋ* and 凌 **b-rjəŋ*; however, since BAXTER’S clusters like *b-r-*, &c., are also fairly controversial, he then reverts back to his own reconstruction, tentatively suggesting a semantically remote Austro-Asiatic origin for the forms and stating that ‘the unusual initial alternation *p ~ r* also speaks for a non-ST origin’. So which etymology is preferable — the ST one or the Austro-Asiatic one? And if it is the latter, do the two forms represent independent borrowings from different Austro-Asiatic sources and / or into different OC dialects? Under these conditions, it would seem more reasonable to separate the two words.

Another problem is drawing questionable conclusions from several disconnected pieces of hypothetical evidence, which also happen to contradict other evidence. A case in point here is the author’s reconstruction of an initial consonant cluster **ml-* in OC (pp. 89–90). It has long been recognized that the initial lateral resonant in OC gives bifurcated reflexes in MC: OC **l-* > MC *j-* in most cases, but also *dž-* in a smaller group of examples. Prof. SCHUESSLER attempts to explain the difference by setting up an OC cluster **ml-* (or, rather, **m-l-*, which he separates from the unhyphenated **ml-*; the phonetic difference between the two is left unexplained), for which MC *dž-* is the regular reflexation, and gives various examples of both proposed Tibeto-Burman cognates and Miao-Yao borrowing sources that seem to confirm this reconstruction.

The problem, of course, is that for most of these cognates, we are not really sure they have to be reconstructed with an initial **m-* for Proto-Sino-Tibetan. Thus, 食 **m-lək* ‘eat’ is compared with PTB **m-lyak* ‘to lick’, but there is almost no evidence, except for a few scattered Kuki-Chin forms, that the PTB form indeed contained an initial **m-*. The same goes for PTB **(m-)loŋ* ‘canoe’ (the PTB reconstruction here and in most other places is taken directly from [BENEDICT 1972]) > Kuki-Naga **m-launŋ*,

but no *m-* is found in any other TB language. About the only good example in this group is PTB **m-lay* ‘tongue’ (where sufficient evidence is found in more than one Tibeto-Burmese subgroup to assert that *m-* really reflects an old state of affairs rather than a local innovation) = OC 舐 **m-le?* ‘to lick’, but it goes without saying that a hypothesis cannot be solidly based on but one example of such a kind.

Two more examples show how MC *dž-* < OC **m-l-* is connected with words with initial prenasalized consonants in Miao-Yao: MC 稊 *džjuet* ‘glutinous millet’ < MY **nblut* ‘glutinous, sticky’; MC 舌 *džjät* ‘tongue’ < OC **m-let* < Yao **byel⁶*, Miao **nplai⁶* id. However, the two Miao-Yao parallels adduced here are actually quite different. The former, according to [PEIROS 1998: 141], is PMY **mblVt* > PM **mblo-Q* (Hmu *nə*, Hmong *mplou*), PY **mblut* (Mien *bjut*, Kimmun *blət*). Borrowing from an early ancestor of PMY is likely here, with OC **lut* or **(m-)lut* having no better etymology. On the other hand, the PMY form for ‘tongue’, also according to PEIROS, is rather **mbret* > PM **mblə-Q* (Hmu *ni*, Hmong *mplai*), but PY **mbjet* (Mien *bjet*, Kimmun *bjet*). The two types of correspondences are not distinguished in [WANG & MAO: 1995] (in both cases, **mbl-* is reconstructed for PMY), but they do exist, and a reconstruction like **mbret* is in good agreement with other types of correspondences where PM **-l-* or **-r-* is the equivalent of PY **-j-* (see [PEIROS 1998: 136–137] for the full set). Furthermore, Prof. SCHUESSLER seems to agree with this, as he quotes the PMY reconstruction for ‘tongue’ as **nbret* rather than **nblet* (p. 467).

Thus, if PMY ‘tongue’ is really **mbret* (or **nbret*), there is no solid ground for assuming that the word has anything to do with OC 舌 **(m-)let* ‘tongue’; it is not clear why OC **m-l-* should have been borrowed as PMY **mbr-* instead of **m(b)l-* proper.

A further scrap of “evidence” for OC *m-l-* is allegedly found within Chinese itself: since OC 繩 **ləŋ* ‘string’ > MC *džjəŋ* uses the character 𪛗 *mrəŋ?* ‘toad’ > MC *məŋ* as phonetic, it has to be reconstructed as **m-ləŋ* (and this is in agreement with MC *dž-* as a proper reflexion of OC *m-l-*). However, it is not mentioned that the same phonetic series also includes such a well-known character as MC 蠅 *jiəŋ* ‘fly’, which, according to the author’s own rules, could not have contained an initial **m-* in OC, since it is MC *jiəŋ*, not *džjəŋ*. So this character has to be reconstructed as OC **ləŋ* or **jəŋ* (p. 576) — which makes the preceding argument hollow, since it proves that not every character in this phonetic series has to contain an initial *m-*, and that OC 繩 **ləŋ* rather than **m-ləŋ* would be perfectly acceptable for it. In fact, the connection between **C-r-* / **C-l-* and **r-* / **l-*-type syllables within the same phonetic series need not necessarily presuppose that the **r-* / **l-*-type syllables necessarily or almost necessarily go back to **C-r-* / **C-l-*-type syllables as well; this idea, upheld by W. BAXTER ([1992]) and L. SAGART ([1999]), occasionally has some interesting applications, but cannot function as a universal rule.

Additional concerns: if it is indeed true that MC *dž-* < OC **m-l-*, whereas MC *j-* < OC **l-*, how do we explain numerous “doublets” attested in MC dictionaries which, for the exact same characters in the exact same meanings, give both *j-* and *dž-* (a good list can be found in [STAROSTIN 1989: 514])? Do they all represent OC alternations between *l-* and *m-l-*? And also, what do we do with various potential TB cognates for OC words with **(m-)l-* > MC *dž-* that have no traces of an initial nasal whatsoever? E. g., the author gives OC 贖 **m-luk* > MC *džjwok* ‘to ransom’ = Written Tibetan *blu-ba*, *blu-s* ‘to buy off, ransom’, saying that “the WT forms can theoretically derive from earlier **mlu(t)*”; but he does not mention the quite possibly related Kachin *gə-lo* ‘to settle, pay an indemnity’ or Lushai *hlo?* ‘wages, salary’, or Kiranti **lək* ‘lend’, with no nasals anywhere; also, if Tibetan *bl-* < *ml-*, why doesn’t any other example in SCHUESSLER’s group contain Tibetan *bl-* to confirm this? Finally, some very reasonable etymologies seem to be rejected simply because they do not fit in with the rule, e. g. OC 乘 **ləŋ* ‘to mount, ascend’ > MC *džjəŋ* = Tibetan *laŋ* ‘rise, get up’, Kachin *luŋ* ‘ascend’ ([PEIROS & STAROSTIN 1996: III: 19], not quoted in the dictionary).

Overall, with virtually no internal evidence for OC **m-l-*, and the external evidence being inconsistent, controversial, and contradictory, it feels more prudent to simply scrap the whole thing — for the moment, at least, until it becomes possible to present a more convincing structure. Currently, however, it is more reasonable to simply admit a dialectal variation reason for the bifurcation of OC **l-* into *dž-* and *j-* in MC.

Speaking of dialectal variation, another minor flaw of the dictionary is that occasionally, theoretical statements are made that do not seem to have sufficient basis in fact and may give a somewhat misleading picture of the historical development of Chinese. Thus, on p. 7 the author introduces the concept of “rural dialects”, saying that “words with rare and unusual features typically

have meanings with a rustic or vulgar flavour”. He then lists the following words as “rural” or “vulgar” (identified on the basis of having initial *x*- < OC voiceless resonants instead of expected regular reflexes): “beard, to face / toward, ribs (of a horse), to know, to vomit, to rear animals, stupid, to roar, tiger, pig”. While the general explanation of *x*-containing words as dialectal varieties may be true, it is absolutely not clear what exactly is so “rural” or “vulgar” about notions like “beard”, “to face”, “to know”, or “tiger”. Further below, we also learn that words like “old man”, “to fall” (!), and “wrist” also “conjure up a rural sphere” (although other quoted words, like “farm” and “sickle”, certainly could do that). It is simply hard to understand why it is necessary to set up and insist on the “rural” vs. “non-rural” (urban?) distinction when the data do not confirm it.

Already in our discussion of the hypothetical OC **m-l-*, we have indicated certain problems with TB etymologies for supposed OC cognates that crop up in the dictionary. Fortunately, they are not ubiquitous, and in most cases, TB cognates are reasonable both from the phonetic and semantic sides. Not all the known etymologies are quoted; where formerly proposed etymologies contradict the author’s ideas on phonetic correspondences between OC and TB, they are usually omitted. This is understandable, but, again, one should keep in mind that the development from PST to OC as traced (in a rather chaotic form) in the introduction is not the final word on Sino-Tibetan, and that various alternative hypotheses to some of the proposed solutions are quite probable. Thus, the idea that aspirated consonants in OC are of a secondary origin (and, therefore, we should not look for specific TB correspondences to OC aspirates) is mostly based on “word family” hypotheses and ideas of sound symbolism that are controversial and unprovable (pp. 58–61); BENEDICT’s earlier explanation of OC unaspirated stops as reflecting traces of lost prefixes ([BENEDICT 1972: 165]) or STAROSTIN’s even more complex set of rules to explain the interaction between voiced, voiceless, and aspirated articulation in ST ([STAROSTIN 1989: 136–146, &c.) both seem more preferable.

Concerning those etymologies of Chinese words for which an external, non-ST source has been suggested, some of the methodological problems that inevitably crop up here are already well described by the author himself on pp. 9–10 (“Identification of cognates”) and need not be addressed further in this review. One thing, however, that must be mentioned is that the author’s historical scenario for contacts between Chinese speakers and speakers of other language families seems to be seriously incomplete. Throughout the dictionary, he most prominently quotes such potential sources or targets of borrowing as Austro-Asiatic, Tai-Kadai, and Miao-Yao languages. These are, indeed, almost universally recognized as donors to / recipients from OC, although nobody has so far presented a coherent chronological model of these contacts. However, while almost overstating the importance of these contacts, he insists that “other language families” had “little contact with Chinese” (p. 133). This can hardly be so, at least in respect to two other prominent families:

(a) Austronesian. Numerous similarities between Austronesian and Chinese have been spotted by L. SAGART, who has even ventured so far as to postulate a genetic relationship between the two based on these connections (the so-called ‘Sino-Austronesian’ theory). In reality, most of the connections can be explained in contact and areal terms (with a few ones possibly indicating some ultra-deep level relation), but they are so numerous that ignoring them altogether is hardly a good solution. Thus, for OC \pm **thâ?* ‘earth’ we are told that “there is no obvious cognate and etymology” (p. 502), apart from a few phonetically dubious Austro-Asiatic parallels, but, in fact, there is a perfectly fine Austronesian connection: cf. PAN **-taq* ‘earth, soil’ [SAGART 2002: 6]. Apart from SAGART’s works, numerous examples on Chinese-Austronesian or, wider, Sino-Tibetan-Austronesian connections are given in [PEIROS 1998] and other sources.

(b) Altaic. The idea that speakers of Altaic languages (or, if one does not subscribe to the Altaic theory, Turkic, Mongolic, Tungus-Manchu, and “pre-Korean” languages) had little or no contact with the earliest Chinese speakers is very odd, considering that nearly all the lands to the north of the Yellow River must have been populated with predominantly Altaic speakers long before the Chinese actually settled in its basin, and that epigraphic and historical records in China are full of references to contacts with nomadic tribes of the North that, in all probability, spoke Altaic languages.

It is true that the issue of Altaic-Chinese language contacts has been relatively unexplored, due to a variety of factors such as:

— typological dissimilarities between Altaic languages and Chinese, making such comparison less comfortable than the one between Chinese and more “compatible” languages such as Tai-Kadai or Mon-Khmer;

— emphasis on later, historical period contacts between Han-time Chinese and Turkic / Mongolic tribes, obscuring earlier contacts;

— lack of direct prehistorical knowledge (earlier than the first centuries A. D.) about the ancestors of Turkic, Mongolic, &c., and the never-ending debate about the status of the Altaic connection.

Nevertheless, all of these problems are purely technical and psychological, and many of them can now be overcome with the appearance of *An Etymological Dictionary of the Altaic Languages* ([EDAL]); material from this source can be used for productive comparison with reconstructed OC forms even if one is skeptical about Altaic as a whole, since it also includes reconstructions for uncontested branches. A few examples of why exploring the Altaic connection is a must for further Sinologic studies, taken from [STAROSTIN 2007], will suffice:

(a) For 兔 *tù* ‘hare’ SCHUESSLER offers an OC reconstruction **lhâh*, saying that “the OC initial **lh-* is revealed by the graph’s use as phonetic in a word for ‘tiger’ in *Zuǒzhuàn*” (p. 502). However, there is no such character as 𪛗 in attested versions of the *Zuǒzhuàn*. The character in question is found in the *Shuōwén*, where it is defined as part of the Chǔ word for ‘tiger’, 烏𪛗. The *Zuǒzhuàn* quotes obviously the same word as 於菟, whereas the *Hànshū* (not the *Hòu Hànshū*, as stated on p. 282 of the dictionary) quotes it as 於釋; for the latter SCHUESSLER gives the modern reading *yú-shì* < MC *ʔjwo-śjäk*, but in this particular use, the *Jiyùn* apparently gives the MC reading *ʔjwo-duo* > modern *yú-tú*. What this all leads to is to state that there is no obvious connection between the regular OC form for tiger, 虎 **hlâʔ*, and the above dialectal word for ‘tiger’, which could have easily been OC **ʔa-dâ(k)* instead of the suggested **ʔa-lhak*.

This, in turn, means that the word for ‘hare’, which SCHUESSLER reconstructs as OC **lhâh*, may easily have been **thâh* or even an earlier **thâks*. If so, then the connection with Middle Korean *twos(ki)* ‘hare’ is more complex than a borrowing ‘from Chinese in connection with the animal cycle’. The Korean word finds a good Altaic parallel in Proto-Tungus-Manchu **tuksa-* ‘hare’ ([EDAL: 1451]), which, in turn, is in that language derived from a verbal stem with the meaning ‘to run’. Thus, it is, in fact, the Chinese word that is more likely to represent an old borrowing from an Altaic source rather than vice versa.

(b) OC 麥 **mrək* ‘wheat’ is compared to Tibetan *bra-bo* ‘buckwheat’ and possibly also to Proto-Lolo-Burmese **g-ra* id. (p. 374); this etymology, despite the obvious phonetic issues (final *-k*, *m-* / *b-* alternation, &c.), could be acceptable if no other were available, but cf. Tungus-Manchu **murgi* ‘barley’, Middle Korean *mír* ‘wheat’, Old Japanese *mugji* ‘wheat, barley’ < Altaic **mǰurgu* ([EDAL: 935]). With three consonantal segments matching, this seems a more likely solution.

(c) OC 敵 **dêk* ‘enemy’ is compared to PTB **m-ta:y* ‘avenge, retaliate’ (p. 209), but cf. rather Turkic **jagi* ‘enemy, war’, Mongolic **dajin* ‘war’ < Altaic **dâgi* ([EDAL: 457]).

Additionally, in [STAROSTIN 2007] one can find reasonable Altaic etymologies for several important items in Chinese lexicon that are altogether lacking in Prof. SCHUESSLER’s dictionary — such as 琴 *qín* ‘*qín*, musical instrument’, 鵲 *què* ‘magpie’, and even the famous “Chinese unicorn” 麒麟 *qílín* (traced back to an Altaic word for ‘deer’). Not all the etymologies offered there will find acceptance, but cases like these, nevertheless, seem to disprove, once and for all, the idea that it is fruitless to include Altaic in the list of potential “donors” to the Chinese language.

Concluding this review, I would once again like to stress that all of the issues raised here do not in any way diminish the general importance of the work. On the contrary, even where one might feel the author has missed some valid data, or has jumped to unwarranted conclusions, this should still constitute a stimulus for discussion and improvement rather than a pretext to dismiss the work altogether. My only concern is that scholars and students that are not well-versed in historical Chinese and Sino-Tibetan studies may make the mistake of taking this dictionary for the *definitive* word on Chinese etymology, which it most certainly is not. I would not, therefore, recommend it for “blind” usage to any such student or scholar. However, if one approaches it with due caution, keeping in mind most of the *caveats* that have been voiced above, it will certainly be a valuable aid in any kind of research that has to do with the history of Chinese language, be it linguistics, philology, or history as such.

- BAXTER 1992 — William H. BAXTER. *A Handbook of Old Chinese Phonology*. MOUTON–DE GRUYTER, Berlin–New York, 1992.
- BENEDICT 1972 — P. K. BENEDICT. *Sino-Tibetan: A Conspectus*. Cambridge University Printing House, 1972.
- EDAL — S. STAROSTIN, A. DYBO, O. MUDRAK. *An Etymological Dictionary of Altaic Languages*. HdO Vol. 8-1, 8-2, 8-3. Leiden: BRILL, 2003.
- PEIROS 1998 — I. PEIROS. *Comparative Linguistics in Southeast Asia // Pacific Linguistics Series*, Vol. 142. Australian National University, Canberra, 1998.
- PEIROS & STAROSTIN 1996 — I. PEIROS, S. STAROSTIN. *A Comparative Vocabulary of Five Sino-Tibetan Languages* (6 vols.). Melbourne, 1996.
- SAGART 1999 — L. SAGART. *The Roots of Old Chinese // Current Issues in Linguistic Theory*, Vol. 184., Amsterdam–Philadelphia: John BENJAMINS Publishing Company.
- SAGART 2002 — L. SAGART. *Sino-Tibeto-Austronesian: an Updated and Improved Argument*. Paper presented at *Ninth International Conference on Austronesian Linguistics (ICAL9)*. 8–11 January 2002. Canberra, Australia.
- STAROSTIN 1989 — С. А. СТАРОСТИН. *Реконструкция древнекитайской фонологической системы [A Reconstruction of the Old Chinese Phonological System]*. Moscow, “Nauka”, 1989.
- STAROSTIN 2007 — S. A. STAROSTIN. *Altaic and Chinese // С. А. СТАРОСТИН. Труды по языкознанию*. М.: «Языки Славянской Культуры» [*Works on Linguistics*. Moscow: “Languages of Slavic Culture”]; pp. 850–853.

Insert 1. Cognate percentage chart for Mande languages, generated by STARLing

Language	DNG	DNB	DNK	TUR	MNO	GUR	YRE	GBA	MWA	WAN	BEN	KPE	KPA	DZU	DUN	SEK	BNK	JOO	BBO	BZJ	BZK	BZT	SON	MAU	JLK	DFN	BMN	MNK	XSN	MDK	KRN	LLE	MGF	KKB	KNO	VAI	SSO	JLN	JGO	JRI	NMU	MEN	LKO	BND	LOM	KPL	KPG	BKB	BOK	BUS	LBR	SAN	KYN	SHN		
DNG	0	0.93	0.9	0.77	0.75	0.72	0.74	0.63	0.72	0.6	0.62	0.35	0.4	0.42	0.35	0.38	0.37	0.42	0.36	0.24	0.27	0.26	0.29	0.36	0.38	0.36	0.35	0.36	0.36	0.34	0.34	0.34	0.34	0.34	0.29	0.28	0.37	0.34	0.36	0.35	0.34	0.33	0.35	0.34	0.37	0.36	0.42	0.43	0.42	0.51	0.5	0.52	0.45			
DNB	0.88	0	0.88	0.74	0.74	0.69	0.74	0.62	0.71	0.62	0.64	0.35	0.43	0.43	0.37	0.42	0.37	0.39	0.35	0.25	0.27	0.26	0.3	0.36	0.38	0.37	0.38	0.38	0.38	0.34	0.34	0.34	0.34	0.33	0.32	0.29	0.33	0.31	0.38	0.36	0.37	0.35	0.34	0.37	0.36	0.4	0.42	0.4	0.52	0.49	0.5	0.44				
DNK	1.03	1.15	0	0.8	0.73	0.72	0.76	0.63	0.69	0.64	0.62	0.31	0.37	0.38	0.33	0.4	0.34	0.38	0.37	0.23	0.25	0.27	0.27	0.36	0.38	0.36	0.36	0.36	0.34	0.34	0.34	0.33	0.34	0.29	0.28	0.34	0.31	0.38	0.36	0.38	0.34	0.32	0.36	0.35	0.37	0.37	0.41	0.41	0.41	0.49	0.47	0.49	0.44			
TUR	1.75	1.85	1.59	0	0.73	0.74	0.78	0.64	0.76	0.67	0.67	0.33	0.38	0.39	0.33	0.41	0.34	0.39	0.36	0.22	0.26	0.26	0.27	0.38	0.39	0.38	0.4	0.39	0.39	0.36	0.36	0.36	0.35	0.35	0.31	0.29	0.32	0.32	0.36	0.34	0.35	0.33	0.29	0.32	0.35	0.35	0.36	0.42	0.43	0.44	0.51	0.47	0.51	0.44		
MNO	1.82	1.88	1.92	1.93	0	0.76	0.77	0.64	0.72	0.65	0.6	0.32	0.39	0.39	0.35	0.45	0.36	0.39	0.37	0.25	0.27	0.27	0.3	0.36	0.37	0.37	0.38	0.37	0.37	0.34	0.36	0.37	0.38	0.34	0.34	0.33	0.34	0.35	0.39	0.35	0.39	0.3	0.29	0.3	0.33	0.37	0.37	0.48	0.49	0.46	0.47	0.45	0.49	0.45		
GUR	1.95	2.09	1.95	1.85	1.77	0	0.87	0.63	0.8	0.72	0.66	0.31	0.38	0.39	0.33	0.39	0.33	0.39	0.36	0.24	0.25	0.26	0.26	0.34	0.37	0.36	0.37	0.36	0.36	0.36	0.33	0.34	0.35	0.36	0.34	0.31	0.3	0.33	0.31	0.38	0.32	0.36	0.31	0.29	0.3	0.33	0.34	0.34	0.43	0.45	0.43	0.47	0.48	0.46	0.42	
YRE	1.86	1.86	1.78	1.7	1.72	1.25	0	0.62	0.81	0.69	0.67	0.34	0.41	0.41	0.37	0.44	0.38	0.39	0.34	0.25	0.27	0.26	0.28	0.38	0.39	0.4	0.4	0.4	0.39	0.36	0.38	0.39	0.38	0.38	0.36	0.33	0.34	0.34	0.38	0.36	0.37	0.33	0.32	0.34	0.33	0.36	0.35	0.47	0.48	0.47	0.51	0.51	0.52	0.48		
GBA	2.41	2.47	2.4	2.36	2.36	2.41	2.46	0	0.64	0.57	0.58	0.3	0.36	0.37	0.31	0.36	0.31	0.32	0.33	0.23	0.24	0.26	0.25	0.32	0.33	0.33	0.34	0.34	0.34	0.3	0.34	0.34	0.31	0.29	0.27	0.33	0.29	0.34	0.33	0.33	0.33	0.3	0.33	0.33	0.33	0.39	0.39	0.39	0.47	0.46	0.44	0.4				
MWA	1.98	2	2.11	1.8	1.98	1.59	1.54	2.36	0	0.74	0.61	0.35	0.39	0.41	0.35	0.43	0.35	0.41	0.33	0.21	0.25	0.24	0.28	0.35	0.36	0.35	0.37	0.37	0.34	0.34	0.33	0.34	0.33	0.34	0.33	0.31	0.28	0.32	0.32	0.36	0.31	0.33	0.33	0.29	0.28	0.31	0.31	0.35	0.35	0.47	0.51	0.48	0.52	0.54	0.55	0.49
WAN	2.56	2.47	2.37	2.21	2.32	1.95	2.14	2.75	1.88	0	0.56	0.26	0.32	0.34	0.28	0.4	0.28	0.35	0.31	0.2	0.25	0.24	0.24	0.34	0.35	0.36	0.35	0.35	0.35	0.31	0.33	0.33	0.33	0.33	0.3	0.29	0.29	0.28	0.38	0.33	0.36	0.31	0.27	0.3	0.32	0.36	0.35	0.42	0.44	0.42	0.44	0.45	0.45	0.41		
BEN	2.45	2.37	2.48	2.21	2.54	2.26	2.19	2.7	2.5	2.78	0	0.35	0.39	0.39	0.34	0.37	0.35	0.36	0.36	0.27	0.29	0.28	0.29	0.35	0.35	0.34	0.37	0.35	0.34	0.35	0.33	0.33	0.34	0.32	0.34	0.29	0.27	0.28	0.31	0.31	0.29	0.3	0.29	0.29	0.28	0.31	0.29	0.4	0.42	0.42	0.48	0.47	0.48	0.41		
KPE	4.24	4.24	4.59	4.36	4.45	4.63	4.32	4.73	4.24	5.09	4.24	0	0.82	0.85	0.78	0.54	0.77	0.49	0.38	0.29	0.32	0.29	0.32	0.43	0.43	0.42	0.43	0.43	0.43	0.42	0.43	0.4	0.41	0.44	0.36	0.35	0.35	0.36	0.38	0.38	0.34	0.29	0.25	0.29	0.25	0.26	0.26	0.24	0.25	0.24	0.27	0.26	0.21	0.23		
KPA	3.79	3.61	4.05	4	3.89	3.95	3.74	4.16	3.9	4.49	3.92	1.48	0	0.97	0.91	0.6	0.89	0.42	0.38	0.28	0.31	0.24	0.34	0.36	0.38	0.37	0.4	0.4	0.39	0.38	0.33	0.33	0.32	0.32	0.32	0.27	0.3	0.31	0.33	0.3	0.28	0.27	0.24	0.25	0.25	0.27	0.26	0.28	0.3	0.28	0.31	0.32	0.26	0.28		
DZU	3.65	3.58	3.93	3.89	3.86	3.92	3.7	4.04	3.76	4.33	3.86	1.34	0.57	0	0.89	0.64	0.88	0.46	0.39	0.29	0.31	0.27	0.36	0.4	0.41	0.41	0.45	0.44	0.43	0.42	0.39	0.39	0.37	0.35	0.34	0.32	0.33	0.34	0.36	0.32	0.3	0.32	0.27	0.3	0.28	0.29	0.29	0.28	0.31	0.28	0.32	0.34	0.26	0.29		
DUN	4.24	4.02	4.36	4.4	4.23	4.43	4.03	4.62	4.18	4.86	4.3	1.69	1.02	1.08	0	0.58	0.84	0.39	0.34	0.28	0.28	0.25	0.29	0.35	0.37	0.37	0.39	0.39	0.39	0.39	0.36	0.33	0.33	0.31	0.29	0.28	0.29	0.3	0.31	0.28	0.26	0.23	0.22	0.24	0.22	0.24	0.23	0.23	0.25	0.23	0.28	0.27	0.22	0.23		
SEK	3.93	3.63	3.79	3.72	3.45	3.9	3.55	4.14	3.59	3.84	4.03	2.9	2.59	2.37	2.66	0	0.86	0.46	0.43	0.32	0.3	0.33	0.4	0.46	0.47	0.46	0.49	0.5	0.5	0.48	0.45	0.44	0.43	0.43	0.38	0.38	0.39	0.47	0.41	0.39	0.37	0.34	0.35	0.33	0.37	0.36	0.31	0.33	0.3	0.3	0.35	0.28	0.26			
BNK	4.08	4.05	4.29	4.43	4.12	4.36	3.95	4.55	4.2	4.91	4.23	1.7	1.09	1.16	1.39	2.81	0	0.39	0.35	0.28	0.3	0.26	0.35	0.37	0.39	0.38	0.4	0.4	0.39	0.37	0.35	0.35	0.34	0.36	0.33	0.3	0.3	0.3	0.3	0.34	0.3	0.28	0.24	0.23	0.25	0.24	0.25	0.26	0.26	0.28	0.26	0.28	0.3	0.24	0.25	
JOO	3.64	3.92	3.96	3.92	3.89	3.87	3.91	4.47	3.75	4.21	4.09	3.22	3.67	3.39	3.92	3.36	3.86	0	0.41	0.29	0.34	0.34	0.4	0.4	0.39	0.37	0.4	0.4	0.4	0.4	0.41	0.41	0.41	0.41	0.39	0.36	0.34	0.35	0.41	0.38	0.37	0.28	0.27	0.26	0.28	0.3	0.3	0.3	0.29	0.3	0.28	0.29	0.24	0.26		
BBO	4.12	4.18	4.05	4.09	4.01	4.12	4.26	4.36	4.43	4.59	4.09	4	4	3.86	4.33	3.59	4.23	3.76	0	0.42	0.36	0.38	0.37	0.4	0.38	0.37	0.39	0.4	0.4	0.41	0.39	0.43	0.44	0.4	0.37	0.39	0.39	0.4	0.39	0.38	0.37	0.33	0.37	0.37	0.35	0.35	0.27	0.29	0.26	0.29	0.29	0.26	0.25			
BZJ	5.33	5.3	5.48	5.73	5.23	5.44	5.3	5.56	5.77	6.07	5.04	4.78	4.89	4.82	4.94	4.51	4.88	4.78	3.68	0	0.74	0.81	0.5	0.45	0.44	0.43	0.46	0.46	0.46	0.46	0.45	0.42	0.39	0.39	0.44	0.38	0.38	0.31	0.3	0.36	0.37	0.32	0.28	0.25	0.3	0.31	0.29	0.29	0.18	0.18	0.15	0.2	0.2	0.16	0.17	
BZK	5	4.96	5.3	5.09	5.01	5.23	5.05	5.45	5.27	5.27	4.76	4.45	4.81	4.54	4.89	4.64	4.7	4.3	4.16	1.88	0	0.84	0.53	0.51	0.52	0.49	0.51	0.5	0.51	0.47	0.46	0.43	0.41	0.44	0.41	0.41	0.29	0.3	0.4	0.39	0.34	0.28	0.26	0.3	0.29	0.27	0.26	0.22	0.23	0.21	0.23	0.2	0.2	0.19		
BZT	5.11	5.2	5.07	5.2	5	5.11	5.16	5.13	5.4	5.36	4.96	4.78	5.33	5.01	5.27	4.43	5.2	4.33	3.98	1.55	1.39	0	0.52	0.44	0.44	0.45	0.45	0.45	0.45	0.45	0.46	0.46	0.44	0.48	0.41																					

Insert 2. Cognate percentage chart for Mande languages, generated by STARLing

