

Mother Tongue



Mother Tongue

XXVI (2025)

In Memory of
Pierre Bancel, Vitaly Shevoroshkin,
Christopher Ehret and Colin Renfrew

Edited by
Gregory Haynes

Gorgias Press LLC, 954 River Road, Piscataway, NJ, 08854, USA

www.gorgiaspress.com

2026 Copyright © by Gorgias Press LLC

All rights reserved under International and Pan-American Copyright Conventions. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise without the prior written permission of Gorgias Press LLC.

2026

α



ISBN

Paperback 978-1-4632-5026-3

eBook 978-1-4632-5027-0

ISSN 1087-0326

Cover art incorporates material from the following images:

(1.) Venus of Brassempouy, 25th millennium BC, mammoth ivory, discovered 1894 in Grotte du Pape, public domain. (2.) Interior of the grotto of Lascaux, photo by JoJan, screened, used under the Creative Commons Attribution 4.0 International license (see below). (3.) Aboriginal rock art on the Barnett River, Mount Elizabeth Station, photo by Graeme Churchard in 2013, screened, used under the Creative Commons Attribution 2.0 Generic license (free to share and adapt with attribution <https://creativecommons.org/licenses/by/2.0/deed.en>). (4.) Giant Deer, replica of depiction at the cave of Lascaux, public domain. (5.) Hands, Cuevas de las Manos, Argentina, screened, used under the Creative Commons Attribution-Share Alike 2.5 Generic license, free to share and adapt with attribution (<https://creativecommons.org/licenses/by/2.5/deed.en>). (6.) Inscription from the tomb of Darius, Naghsh-e rostam, Iran. Extracted and screened from a photo created by Diego Delso in 2016, used under the Creative Commons Attribution-Share Alike 4.0 International license: (<https://creativecommons.org/licenses/by-sa/4.0/deed.en>).

Library of Congress Cataloging-in-Publication Data

A Cataloging-in-Publication Record is available at the Library of Congress.

Printed in the United States of America

The authorized representative in the EU for product safety and compliance is EU eucomply OÜ,
Pärnu mnt 139b-14,11317, Tallinn, Estonia, hello@eucompliancepartner.com, +3375690241

CONTENTS

MEMORIALS

Pierre Bancel (1958–2025)	1
<i>Magali Bancel</i>	
Vitaly V. Shevoroshkin (1932–2023)	4
<i>Irén Hegedűs and Václav Blažek</i>	
In Memory of Christopher Ehret (July 27, 1941–March 25, 2025)	15
<i>Patrick Manning</i>	
Colin Renfrew: a note on his interests outside archaeology	19
<i>Roger Blench</i>	

ARTICLES

Genetics and the Prehistory of Dravidian Languages	23
<i>Sundar Narayan</i>	
Etruscan–Hurro-Urartian Relations	51
<i>Václav Blažek</i>	
The Puzzle of Pan-African Roots	73
<i>Roger Blench</i>	
Remarks on Spanish <i>conmigo</i> , <i>contigo</i> , <i>consigo</i>	101
<i>Jan Henrik Holst</i>	
Language Origins and Group Behavior: Proto-speech, Syntactic Language, and Spread of Language	113
<i>Patrick Manning</i>	
The Pre-Indo-European Substratum in the British Isles and the Pre-Uralic Substratum in Fennoscandia: Traces of Common Origin	145
<i>Václav Blažek</i>	
Revisiting Greek ἄχθος	149
<i>Gregory Haynes</i>	

BOOK REVIEW

Review of Jan Henrik Holst, <i>Proposals for the diachronic interpretation of Tocharian</i>	161
<i>Carsten Peust</i>	

IN MEMORIAM

Pierre Bancel (1958–2025)

MAGALI BANCEL

Pierre Jean Bancel, French linguist, translator, and editor, was born in Paris on 17 August 1958 and passed away in June 2025. For four decades, he pursued one of humanity’s oldest questions: how language began. His work combined scientific discipline with philosophical depth and a quiet conviction that understanding language is inseparable from understanding what it means to be human.

Education and Early Career

Bancel studied at the Université Lumière Lyon II, where he earned a Master’s degree in Language Sciences with a focus on comparative linguistics, instrumental phonetics, and Bantu languages. Trained in both the humanities and empirical science, he sought to bridge precision and imagination—a balance he would later describe as “the only path capable of revealing new truths, though constantly threatened from both sides.”

Fluent in several languages and deeply versed in the classics (Latin, Greek, Old French), he reflected a lifelong fascination with the diversity and unity of human expression. After teaching in Paris, he worked for leading French publishers such as *Éditions Le Robert* and *Larousse*, contributing to major dictionaries and editorial projects. He later served as an editor and proofreader for *Le Monde*, *Courrier International*, *Pour la Science*, and *La Recherche*.

Research on the Origins of Language

In the 1990s, Bancel joined the Association for the Study of Language in Prehistory (ASLIP) at Harvard University and began publishing in *Mother Tongue Journal*. There he met lifelong collaborators, Merritt Ruhlen, Alain Matthey de l’Etang, and John D. Bengtson, with whom he would develop a rigorous yet daring body of comparative linguistic work.

Their collaboration took off after Bancel’s French translation of Merritt Ruhlen’s *The Origin of Language* (Belin, 1997), a book that profoundly influenced him. An entry in Ruhlen’s global etymology list, the kin term *kaka*, sparked their joint research. As Matthey later recalled, “with *kaka* it was the jackpot: the term appeared across hundreds of languages worldwide with the meaning ‘maternal uncle.’” This discovery became the foundation of a two-decade investigation into

what they called *Proto-Sapiens*, a hypothesized ancestral system underlying all human kinship and pronoun structures.

Their most cited studies include:

- *Tracing the Ancestral Kinship System: The Global Etymon kaka* (2002)
- *Kin Tongue: Nursery Terms in Relation to Language Acquisition* (2005)
- *Where Do Personal Pronouns Come From?* (2010)
- *Brave New Words* (in Lefebvre et al., *New Perspectives on the Origins of Language*, 2013)
- *The Proto-Sapiens Prohibitive/Negative Particle ma* (2019)
- “On the Pronoun Roots *n* (1 sg) and *m* (2 sg)” in *The Native Languages of the Americas* (2021)

Together, these works sought to illuminate how the first human words emerged from biological, cognitive, and social evolution—a bold endeavor conducted with scholarly discipline and intellectual courage.

Translator and Mediator of Ideas

Pierre Bancel’s translations helped shape linguistic discourse in France. He brought into French two major works by Stanford linguists: Merritt Ruhlen’s *The Origin of Language* and Joseph H. Greenberg’s *Indo-European and the Eurasian Language Family*, thus giving French readers access to the American school of macro-comparison. His meticulous, lucid translations were praised for both scientific accuracy and literary elegance.

Later Years and United Nations Service

From 2010 onward, Bancel worked as a translator for the United Nations in New York, Geneva, and Vienna, while continuing his linguistic research. He served as co-editor of *Mother Tongue Journal* from 2021 until his death. His colleagues valued his sharp editorial eye, exacting standards, and generosity with younger researchers.

The Species Pump Hypothesis

At the end of his life, Bancel completed a major synthesis: *Steps in the Evolution of Human Language: The Species Pump Hypothesis*. This book, the result of more than twenty years of research, proposed that the emergence of language was not a gradual cultural accident but an evolutionary transformation—a “species pump” that converted pre-linguistic hominins into *Homo sapiens*. It reflects both his scientific imagination and his conviction that language was the decisive leap that made humanity possible.

Intellectual Influences and Convictions

In his final reflections, Bancel cited the thinkers who shaped his worldview: Greenberg, Darwin, Dawkins, Marx, Freud, and above all the naturalist E. O. Wilson, whose integrative vision of biology and ethics he deeply admired. He wrote that he had always tried to “marry scientific rigor

and creative imagination,” following the example of mentors who combined experiment and intuition.

He saw his work as a continuation of this lineage, a search for order and meaning within the apparent chaos of languages, always tempered by humility: “If we were wrong, at least we closed one dead end. But I leave confident that we have most probably opened a new avenue.”

Character and Legacy

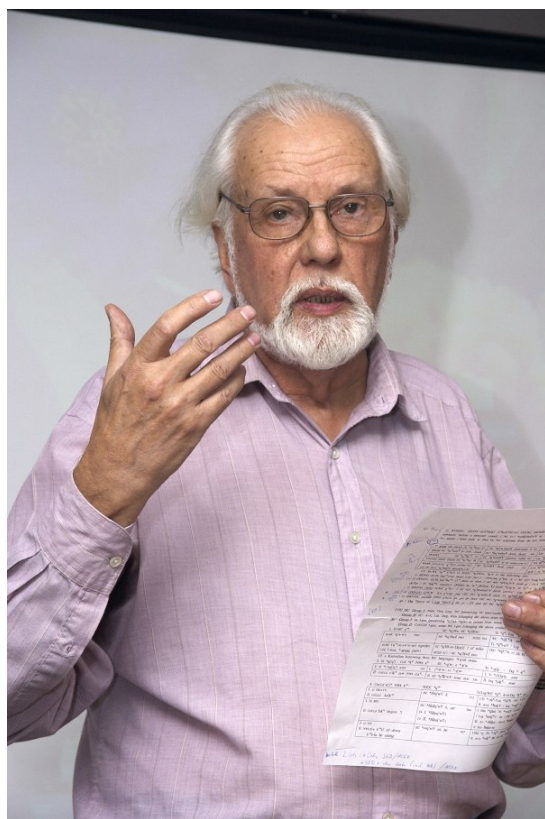
Those who worked with him remember a formidable intellect and a demanding editor who never accepted imprecision, but also a generous and disinterested collaborator who believed ideas belonged to everyone. As Alain Matthey wrote, “You were an extraordinary and formidable companion of intelligence ... certainly the most generous and selfless of men.”

Pierre Bancel leaves a legacy of intellectual independence, cross-disciplinary curiosity, and a body of work that continues to inspire linguists exploring the origins of speech. His writings, translations, and editorial contributions remain a testament to a mind that sought, against all odds, to glimpse the first human words—and through them, the birth of humanity itself.

Magali Bancel
November, 2025

Vitaly V. Shevoroshkin (1932–2023)

IRÉN HEGEDŰS AND VÁCLAV BLAŽEK



Vitaly V. Shevoroshkin was born on 12 January 1932 in the small historical settlement of New Athos on the Black Sea coast (Abkhazia, Georgian Soviet Republic of the USSR). After 1951, he lived and studied in Moscow. Having graduated from Moscow State University in 1956, he started working at the Institute of Russian Language, Academy of Science of USSR in 1962. He obtained his academic degree (“Candidate of linguistics”/Doctor of Philology) at the Russian Academy of Science in 1965.

In 1974 he emigrated to the USA and began teaching first at the University of Maryland (1974–1975), then held a position of associate professor at Yale University (1975–1978). In 1978 he became visiting associate professor of linguistics at the University of Michigan, Ann Arbor, where a year later he was appointed professor of linguistics at the Department of Slavic Languages and Literatures. It is from this same department that he retired after 30 years of teaching on May 31, 2009.

His teaching activity covered a vast area of linguistic studies. He taught courses on Anatolian, and other Indo-European languages, which included synchronic and diachronic approaches to Slavic languages, as well as indigenous languages of the Americas. He was member of linguistic organizations such as *Societas Linguistica Europaea* and the Michigan Linguistic Society (of which he was president in 1986–1987). His publication activity continued for several years after retirement, though abated after 2014, when – after almost 50 years of marriage – he lost his wife, Galina Barinova, also a linguist (married in 1965).

He died on 22 December 2023 in Zürich, Switzerland, where his son, Aleksei lives and works.

Vitaly Shevoroshkin’s oeuvre covered three major areas of linguistic research: A) theoretical aspects of phonetics and phonology, B) ancient Anatolian languages, and C) the hypothesis of distant genetic relationship between language families (with emphasis on the Nostratic hypothesis).

A) Theoretical aspects of phonetics and phonology

Shevoroshkin's early period of linguistic research focused on theoretical issues of sound-chain patterns. This was the topic of his doctoral dissertation,¹ in which his goal was to identify common features of sound sequences shared by languages of the world and to find typological patterns that would facilitate the solution of theoretical issues as well as the decipherment of ancient texts. His monograph of 1965 demonstrated the applicability of the method in the decipherment of the Carian script. He carried on the synchronic line of sound-chain research in a wider scope of languages, and this led to the publication of his monograph of 1969,² in which he argued that since sound chains in all languages are governed by similar principles (regularity in the sequence of vowels and consonants, etc.), the method is capable of identifying classes and subclasses of various sounds (characters) in a text written in any language.

He also co-authored some articles on Russian phonetics and phonology with his wife, Galina Barinova (1972b; 1988d).

B) His activity in the field of Anatolian studies

Vitaly V. Shevoroshkin belongs to the most productive and at the same time to the most respected scholars in the field of the ancient Anatolian languages. Let us recount his Anatolian linguistic interests in some detail. He published 9 contributions about the Anatolian (= Hittite-Luwian) languages in general (1965d; 1967; 1968c; 1970; 1978a; 1979; 1981; 1982; 1988; 1998; 2015b), and together with Paul Sidwell they edited a collective monograph (2002b). Shevoroshkin devoted individual studies to Hittite (1964a), Luwian (1965f) and Lydian (1968f; plus, the monograph of 1967). He concentrated especially on two Anatolian languages of the first millennium BCE:

- Lycian, including Milyan (1963a; 1966c; 1968b, e, f; 1971b; 1977b; 1978b; 1993; 2000a, b; 2002a, b, c; 2004; 2008a; 2010; 2011a, b, c, d; 2012a, b; 2013c; 2014; 2015b), and
- Carian (1962; 1963b, d, e; 1964b, c, d, e, f, g; 1965a, b; 1968e, f; 1977a; 1982–83; 1991a, b; 1992a; 1994a), plus the first monograph about Carian (1965).

Although Shevoroshkin analysed the onomastic data practically in all publications about Anatolian languages, he devoted three special studies focusing exclusively on Anatolian onomastics (1969b; 1978a; 1988a).

Vitaly Shevoroshkin was also a first-class specialist on the scripts of the East Mediterranean and Asia Minor. Besides the standard contributions (1962b; 1965c; 1966a; 1968d; 1972a; 1975) he was also the author of a series of knowledgeable comments on the Russian translations of important studies devoted to the Phaistos Disc, Linear A, Lydian, Carian and Sidetic scripts (1976a, b, c, d, e). It was during the preparation of this volume (edited by Igor M. Djakonov) that Vitaly Shevoroshkin emigrated to the USA. His emigration automatically meant that he became *persona*

¹ «К исследованию некоторых закономерностей строения звуковых цепей. (В связи с проблемами идентификации лингвистических единиц)» [On the investigation of some regularities in the structure of sound-chains. (On the connection with the problems of identifying linguistic units)]

² 2nd edition published in 2004.

non grata, which entailed the prohibition of his publications. For this reason, his name in the volume *Тайны древних писмен: проблемы дешифровки* [Secrets of old scripts: problems of decipherment] was replaced by fictitious А.Ф. Деянов, and this way the volume could be published, although everyone among the linguists of the East Block knew about the real identity of the author.

C) Promoting the Nostratic hypothesis and the study of language in prehistory

In 1962 Shevoroshkin met Aharon Dolgopolsky, they both worked at the Department of Structural Linguistics, Russian Language Institute of the Academy. Dolgopolsky had already been working on his Sibiuro-European (later called Borealic and then Nostratic) comparisons, and since Shevoroshkin himself noted relevant non-Indo-European correspondences to his Indo-European root list, their discussions frequently concerned the Nostratic hypothesis. The development of the hypothesis gained significant momentum, when Vladimir A. Dybo – aware of the fact that beside Dolgopolsky, another young linguist, Illič-Svityč was investigating the same topic – introduced the two scholars to each other. With the early demise of Illič-Svityč in 1966, Shevoroshkin's emigration to the US in 1974 and Dolgopolsky's to Israel in 1976, the Moscow school of Nostratic linguistic research obtained representatives beyond the Russian Academy.

In 1988, Vitaly Shevoroshkin and Ben Stolz organized the *First Interdisciplinary Symposium on Language and Prehistory* at the University of Michigan, Ann Arbor. This event brought together linguists focusing on the Nostratic hypothesis with those investigating other hypotheses of linguistic relationship at a deeper time-depth. This was the first opportunity for American experts to meet a group of Russian and East European linguists (including Aharon Dolgopolsky, then already living in Israel) who were specialising in the reconstruction of the Nostratic protolanguage. This symposium has a special relevance for the historiography of linguistics because its aim was to establish communication between linguists and human geneticists in order to approach the complex issue of linguistic prehistory from an interdisciplinary viewpoint. This was quite a revolutionary momentum at that time, and given the present-day heightened attention to the correlation of population genetics and the historical migration routes of speakers of different languages, it is easy to see that Vitaly intended to steer the development of historical-comparative linguistic research in the right direction. This, however, does not mean that the prestige of Nostratic or long-range linguistic comparison has won general recognition, but the symposium and the subsequent period showed an increased interest in these contested issues of language in prehistory. The proceedings and related materials were published in a sequence of several volumes in Germany (1989a, 1989b; 1990; 1991; 1992).

Vitaly's professional support and friendship is gratefully acknowledged by both authors of this obituary. Vitaly and his wife Galina Barinova were the kindest of hosts and welcomed us to their home in Ann Arbor, Michigan. We also treasure the pleasant memories of meeting them during their professional or private visits in Europe.

Irén Hegedűs (University of Pécs, Hungary) and
Václav Blažek (Masaryk University, Brno, Czech Republic)

Vitaly V. Shevoroshkin's list of publications

Books

- 1965 *Исследования по дешифровке карийских надписей*. [Investigations in the decipherment of Carian inscriptions] Москва: Наука. 358pp.
- 1967 *Лидийский язык* [The Lydian Language] Москва: Наука. 71pp.
- 1969 *Звуковые цепи в языках мира*. [Sound-chains in the world's languages] Москва: Наука. 187pp.
- 1970 (with A. M. Кондратов) *Когда молчат письма. Загадки древней Эгеиды*. [When the writings are silent. Mysteries of the ancient Aegean] Москва: Наука. 227pp.
- 2004 *Звуковые цепи в языках мира*. [Sound-chains in the world's languages] (2nd, rev. edition) Москва: USSR. 184pp.

Edited books

- 1986 (with T.L. Markey) *Typology, Relationship and Time*. Ann Arbor: Karoma.
- 1989a *Reconstructing Languages and Cultures*. (Abstracts and Materials from the First Interdisciplinary Symposium on Language and Prehistory, Ann Arbor, 8–12 November 1988). Bochum: Brockmeyer.
- 1989b *Explorations in Language Macrofamilies*. (Materials from the First Interdisciplinary Symposium on Language and Prehistory). Bochum: Brockmeyer.
- 1990 *Proto-Languages and Proto-Cultures*. (Materials from the First Interdisciplinary Symposium on Language and Prehistory). Bochum: Brockmeyer.
- 1991 *Dene-Sino-Caucasian Languages*. (Materials from the First Interdisciplinary Symposium on Language and Prehistory). Bochum: Brockmeyer.
- 1992 *Nostratic, Dene-Caucasian, Austric and Amerind*. (Materials from the First Interdisciplinary Symposium on Language and Prehistory). Bochum: Brockmeyer.
- 1999 (with Paul Sidwell) *Historical linguistics & lexicostatistics*. Melbourne: Association for the History of Language.
- 2002a (with Paul Sidwell) *Languages and their Speakers in Ancient Eurasia. Dedicated to Professor Aharon Dolgopolsky on his 70th Birthday* (AHL Studies in the Science and History of Language 5) Canberra: Association for the History of Language.
- 2002b (with Paul Sidwell) *Anatolian Languages*. (AHL Studies in the Science and History of Language 6) Canberra: Association for the History of Language.
- 2007 Vladimir Orel's *Russian Etymological Dictionary* vols. 3–4. Theophania Publishing.
- 2009 (with Harald U. Sverdrup) *Bygone voices reconstructed. On language origins and their relationships. In honor of Aharon Dolgopolski*. Copenhagen: Underskoven Publishers ApS.

Articles

- 1957 К истории индоевропейского генетива. [On the history of the Indo-European genitive] *Вопросы языкознания* 6, № 6, 89–90.
- 1962a Карийский вопрос [The Carian problem] *Вопросы языкознания* 11, № 5, 93–100.
- 1962b О начальном этапе дешифровки буквенных письменностей. [On the initial stage of deciphering alphabetic scripts] In: *Симпозиум по структурному изучению знаковых систем*. Москва: Издательство Академии наук СССР, 57–61.
- 1963a Linear A and Lycian. *Nestor* 66, 258–259.

- 1963b Linear A and Karian. *Nestor* 67, 263–264.
- 1963c How Herodotos' father spelt his name. *Nestor* 70, 282–283.
- 1963d Karer auf Kreta. *Nestor* 72, 293–295.
- 1963e О хетто-лувийском характере карийского языка. [On the Hittite-Luwian character of the Carian Language] *Вопросы языкознания* 12, № 3, 83–84.
- 1963f О структуре звуковых цепей. [On the structure of sound-chains] In: *Проблемы структурной лингвистики*. Отв. ред. С.К. Шаумян. Москва: Издательство Академии наук СССР, 164–181.
- 1964a Новые исследования по хеттологии. [New investigations in Hittite studies] *Вопросы языкознания* 13, № 3, 124–133.
- 1964b О новых результатах исследований карийских надписей. [On new findings in the study of Carian inscriptions] *Вопросы языкознания* 13, № 4, 108–111.
- 1964c Статуэтка Исиды с египетско-карийской надписью. [Statue of Isis with Egyptian-Carian inscription] *Вестник древней истории* 1964, № 2, 128–134.
- 1964d On Karian. *Revue Hittite et Asianique* XXII, 74, 1–55.
- 1964e Aegyptisch-karische Inschrift am Sockel einer Isisstatuette (Leningrader Staatsermitage). *Revue Hittite et Asianique* XXII 74, 57–65 and pl. I–IV.
- 1964f Zur karischen Schrift und Sprache. *Kadmos* 3, 72–87.
- 1964g Карийский язык. Современное состояние дешифровки и изучения. [The Carian language. Current status of deciphering and study] In: *Проблемы индоевропейского языкознания. Этюды по сравнительно-исторической грамматике индоевропейских языков*, 18–39. Москва: Наука.³
- 1965a О хетто-лидийском характере карийского языка. [On the Hittite character of the Carian language] *Вестник Московского Гос. Университета. Филология-журналистика* 1965, № 5, 46–57.
- 1965b Опыт дешифровки карийских надписей. [An attempt at deciphering Carian inscriptions] *Вестник древней истории* 91, № 1, 31–50.
- 1965c К происхождению малоазийских и некоторых других буквенных письменностей. [On the origin of Asia Minor and some other alphabetic scripts] *Вестник древней истории* 93, № 3, 132–138.
- 1965d Малоазийские языковые параллели [Asia Minor linguistic parallels]. *Этимология* 1965, 142–159.
- 1965e Этимологические заметки. [Etymological notes] *Этимология* 1965, 235–247.
- 1965f Лувийские племена. [Luwian tribes] In: *Советская историческая энциклопедия. В 16 томах. М.: Советская энциклопедия. 1973–1982. Том 8: КОШАЛА–МАЛЬТА*. Издательство «Советская энциклопедия», 799.
- 1966a Недешифрованные письменности Крита. [Undeciphered scripts of Crete] *Вопросы истории* 1966, № 12, 209–212.
- 1966b К проблеме соотношения звуковых и фонемных цепей. [On the problem of the relationship between sound-chains and phonemic chains] *Исследования по фонологии*. Москва: Наука, 160–165.
- 1966c (with Andrej Korolëv) Милийские этимологии [Milyan etymologies] In: *Проблемы славянских этимологических исследований в связи с общей проблематикой современной этимологии*. Москва: Наука, 42–44.

³ In the same volume the opening chapter was V. M. Illič-Svityč's Древнейшие индоевропейско-семитские языковые контакты. [The earliest Indo-European–Semitic language contacts]

- 1967 Этимологические заметки. [Etymological notes] *Этимология* 1965, 235–247.
- 1968a Тайны письмен Крита. [Secrets of writings from Crete] In: *Будущее науки: Междунар. ежегодник 2: Перспективы, гипотезы, нерешенные проблемы*. Москва: Знание, 305–318.
- 1968b К проблеме ликийского языка. [On the problem of the Lycian language] *Вопросы языкознания* 1968, № 6, 66–80.
- 1968c Zur hethitisch-luwischen Lexik. *Orbis* 17/2, 467–491.
- 1968d Zur Entstehung und Entwicklung der kleinasiatischen Buchstabenschriften. *Kadmos* 7, 150–173.
- 1968e Karisch und Lykisch. In: *Atti e Memorie del I Congresso Internazionale di Micenologia* (1967). Roma: Edizioni dell'Ateneo, 462–472.
- 1968f Karisch, Lydisch, Lykisch. *Klio* 50, 53–69.
- 1969a Zu den 'späthethitischen' Sprachen. *XVII. Deutscher Orientalistentag – 1968, Würzburg*, ed. by Wolfgang Voigt (= ZDMG Suppl. I) Wiesbaden: Steiner, 250–271.
- 1969b Zur Erforschung der kleinasiatischen Onomastik. In: *Disputationes ad montium vocabula aliorumque nominum significationes pertinentes. 10. Internationaler Kongress für Namenforschung II*, ed. by Herwig H. Hornung. Wien: Verlag der Wiener Medizinischen Akademien, 341–350.
- 1969c (with Andrej Korolëv) Lykische Wörter und Namen. *Archiv orientální* 37, 523–542.
- 1970 К сравнительно-исторической характеристике хетто-лувийских языков. [On the comparative-historical characteristics of the Hittite-Luwian languages] In: *Кузнецовские чтения. Тезисы докладов конференции по фонологии и морфонологии*, ред. А. А. Реформатский. Москва: Ин-т языкознания АН СССР, 54–55.
- 1971a О двух [в] в русском языке. [About two /v/ sounds in the Russian language] In: *Развитие фонетики современного русского языка: Фонологические подсистемы*, ред. Высотский С.С. Москва: Наука, 279–286.
- 1971b Zur Erforschung des Milyischen. *Mitteilungen des Instituts für Orientforschung* 17, 206–236.
- 1972a К происхождению и взаимовлиянию алфавитов Балканского полуострова, островов Эгейского моря и Малой Азии. [On the origin and mutual influence of the alphabets of the Balkan Peninsula, the Aegean Islands and Asia Minor] In: *I. симпозиум по балканскому языкознанию. Античная балканистика (23–24 мая 1972 года). Предварительные материалы: Тезисы докладов. Сообщения. Аннотации*. Москва: Институт славяноведения и балканистики АН СССР, 39–41.
- 1972b (with Galina A. Barinova) Регулярные и нерегулярные сочетания согласных в русском языке. [Regular and irregular consonant combinations in Russian] In: *Проблемы структурной лингвистики 1971*, ред. С. К. Шаумян. Москва: Наука, 334–341.
- 1973 (with G.M. Bogomazov and R.F. Paufošima) О некоторых способах реализации консонантных сочетаний в речи. [On some ways of the realisation of consonant combinations in speech] In: *Проблемы теоретической и прикладной фонетики и обучение произношению*. Москва: Издательство УДН им. Патриса Лумумбы, 110–113.
- 1974a Фригийская культура. [Phrygian culture] In: *Советская историческая энциклопедия. В 16 томах*. Москва: Советская энциклопедия. 1973–1982. Том 15: ФЕЛЛАХИ–ЧЖАЛАЙНОР, 426–427.

- 1974b Фригийцы. [Phrygians] In: *Советская историческая энциклопедия. В 16 томах.* Москва: *Советская энциклопедия.* 1973–1982. Том 15: *ФЕЛЛАХИ–ЧЖАЛАЙ–НОР*, 427.
- 1974c Фригия. [Phrygia] In: *Советская историческая энциклопедия. В 16 томах.* Москва: *Советская энциклопедия.* 1973–1982. Том 15: *ФЕЛЛАХИ–ЧЖАЛАЙ–НОР*, 427–428.
- 1975 Zur sidetischen Schrift. *Kadmos* 14, 154–166.
- 1976a Надпись на диске из Феста. [The inscription on the Phaistos Disk] In: *Тайны древних писмен: проблемы дешифровки* (перевод с английского, немецкого, французского и итальянского языка сост., ред., предисл. И. М. Дьяконов). Москва: Прогресс, 31. {Under the name А.Ф. Деянов}
- 1976b Линейное письмо А; Комментарий. [Linear A script; Commentary.] In: *Тайны древних писмен: проблемы дешифровки* (перевод с английского, немецкого, французского и итальянского языка сост., ред., предисл. И. М. Дьяконов). Москва: Прогресс, 83–84; 99–100. {Under the name А.Ф. Деянов}
- 1976c Паракарыйская письменность; Комментарий. [The Para-Carian script; Commentary.] In: *Тайны древних писмен: проблемы дешифровки* (перевод с английского, немецкого, французского и итальянского языка сост., ред., предисл. И. М. Дьяконов). Москва: Прогресс, 304–308; 316–317. {Under the name А.Ф. Деянов}
- 1976d О ‘паралидийском’ письме; Комментарий. [On the Para-Lyidian script; Commentary.] In: *Тайны древних писмен: проблемы дешифровки* (перевод с английского, немецкого, французского и итальянского языка сост., ред., предисл. И. М. Дьяконов). Москва: Прогресс, 318–319; 326–327. {Under the name А.Ф. Деянов}
- 1976e Сидетское письмо; Комментарий. [The Sidetic script; Commentary.] In: *Тайны древних писмен: проблемы дешифровки* (перевод с английского, немецкого, французского и итальянского языка сост., ред., предисл. И. М. Дьяконов). Москва: Прогресс, 328–330; 333–334. {Under the name А.Ф. Деянов}
- 1977a Zu einigen karischen Wörtern. *Münchener Studien zur Sprachwissenschaft* 35, 117–130.
- 1977b Zu einigen Verwandtschaftsbeziehungen im Lykischen und Milyischen. *Münchener Studien zur Sprachwissenschaft* 36, 131–144.
- 1978a Studies in Hittite-Luwian Names. *Names: A Journal of Onomastics* 26.3, 231–257.
- 1978b Sull’ interpretazione delle righe 20-21 della trilingue di Xanthos. *Incontri Linguistici* 4, 238-239.
- 1979 On the Hittite-Luwian numerals. *Journal of Indo-European Studies* 7, 177–198.
- 1981 On Hittite-Luwian and Indo-European etymologies. In: *Bono homini donum. Essays in Historical Linguistics in Memory of J. Alexander Kerns*, ed. by Yoël L. Arbeitman and Allan R. Bomhard. Amsterdam: Benjamins, 259–261.
- 1982a Zu den hethitisch-luwischen Konsonanten. In: *Investigationes philologicae et comparativae. Gedenkschrift für Heinz Kronasser*, ed. by Erich Neu. Wiesbaden: Harrassowitz, 210–214.
- 1982b Research Report: Penutian labial stop correspondences. *California-Oregon Languages Newsletter* (I) 5(2), 5–6.
- 1982–83 Über den Lautwert des karischen Buchstaben θ. *Incontri Linguistici* 8, 71–78.
- 1984[86] Verbesserte Lesungen von karischen Wörtern. *Incontri Linguistici* 9, 199–200.

- 1985 (with Mark Kaiser) On Indo-European laryngeals and vowels. *Journal of Indo-European Studies* 13/3–4: 377–413.
- 1986 (with Mark Kaiser) Inheritance vs. borrowing in Indo-European, Kartvelian, and Semitic. *Journal of Indo-European Studies* 14, 365–378.
- 1987a Indo-European Homeland and Migrations. *Folia Linguistica Historica* 7.2, 227–250.
- 1987b (with Mark Kaiser) On Recent Comparisons Between Language Families: The Case of Indo-European and Afroasiatic. *General Linguistics* 27.1, 34–46.
- 1988a Carian proper names. *Onomata* 12, 497–505.
- 1988b Indo-European consonants in Anatolian. In: *A Linguistic Happening in Memory of Ben Schwartz, Studies in Anatolian, Italic and other Indo-European Languages*, ed. by Yoël L. Arbeitman. (Bibliothèque des cahiers de l'institut de linguistique de Louvain 42) Louvain-La-Neuve: Peeters, 283–303.
- 1988c (with Mark Kaiser) Nostratic. *Annual Review of Anthropology* 17, 309–329.
- 1988d (with Galina Barinova) Some observations on the role of ослышки in linguistic research. *Russkij jazyk / Russian Language Journal* 42, № 141–143, 35–39.
- 1988e On laryngeals. In: *Die Laryngaltheorie und die Rekonstruktion des indogermanischen Laut- und Formensystems*, ed. by Alfred Bammesberger. Heidelberg: Carl Winter Universitätsverlag, 527–546.
- 1989a Methods in interphyletic comparisons. *Ural-Altaische Jahrbücher* 61, 1–26.
- 1989b (with Mark Kaiser) Nostratic. *Anthropology Today* 5, № 6 (Dec., 1989), 22–23.
- 1990a Uralic vocalism and long-range comparison. *Uralo-Indogermanica* 2, 85–94.
- 1990b The Mother Tongue: How linguists have reconstructed the ancestor of all living languages. *The Sciences* 30.3, 20–27.
- 1990c Introduction. In: *Proto-Languages and Proto-Cultures*. (Materials from the First Interdisciplinary Symposium on Language and Prehistory), ed. by Vitaly Shevoroshkin. Bochum: Brockmeyer, 8–12.
- 1991a Лингвистический анализ новых карийских надписей (Находки 1985 г.) [Linguistic analysis of new Carian inscriptions (Findings from 1985)] *Вопросы языкознания* 1991, № 3, 66–75.
- 1991b On Carian Language and Writing. *Perspectives on Indo-European Language Culture and Religion. Studies in Honor of Edgard C. Polomé. vol. I* (= JIES Monograph 7), 116–135.
- 1991c (with A. Manaster Ramer) Some recent work on the remote relations of languages. In: *Sprung from Some Common Source: Investigations into the Prehistory of Languages*, ed. by S.M. Lamb & E. D. Mitchell. Stanford, Cal.: Stanford University Press, 178–199.
- 1991d Introduction. In: *Dene-Sino-Caucasian Languages*. (Materials from the First Interdisciplinary Symposium on Language and Prehistory), ed. by Vitaly Shevoroshkin. Bochum: Brockmeyer, 6–9.
- 1992a On Carian language and writing. *Newsletter for Anatolian Studies* Vol. 8/1, 3.
- 1992b A symposium on the deep reconstruction of languages and cultures. In: *Reconstructing Languages and Cultures*. (Abstracts and Materials from the First Interdisciplinary Symposium on Language and Prehistory, Ann Arbor, 8–12 November 1988), ed. by Vitaly Shevoroshkin. Bochum: Brockmeyer, 6–8.

- 1992c Introduction. In: *Proto-Languages and Proto-Cultures*. (Materials from the First Interdisciplinary Symposium on Language and Prehistory), ed. by Vitaly Shevoroshkin. Bochum: Brockmeyer, 7–8.
- 1993 Lykisch und andere späthanatolische Sprachen. In: *Akten des II. Internationalen Lykien-Symposions; Wien, 6-12. Mai 1990*, ed. by Jürgen Borchhardt and Gerhard Dobesch. (Ergänzungsbände zu den Tituli Asiae Minoris, 17) Wien: Verlag der Österreichischen Akademie der Wissenschaften, 39–52.
- 1994a Carian – three decades later. *Decifrazione del cario. Atti del “Primo Simposio internazionale sulla Decifrazione del Cario”*, Roma 1993, ed. by M. Giannotta et al. Roma: Consiglio Nazionale Ricerche, 131–166.
- 1994b Deep reconstruction of languages and semantics. *Origins of Semiosis: Sign Evolution in Nature and Culture*, ed. by Winfried Nöth. Berlin, New York: De Gruyter Mouton, 293–308.
- 1998a Анатолийские и славянские сходжения в индоевропейском лексическом фонде. [Anatolian and Slavic similarities in the Indo-European lexical fund.] In: *Язык: изменчивость и постоянство. К 70-летию Л. Л. Касаткина*, ред. М. Л. Каленчук. Москва: Институт русского языка РАН.
- 1998b 1998 Symposium on Nostratic at Cambridge. *Mother Tongue* 31, 28–32.
- 1999a Nostratic languages: internal and external relationship. In: *Nostratic: Examining a Linguistic Macrofamily*, ed. by Colin Renfrew and Daniel Nettle. Cambridge: The McDonald Institute for Archaeological Research, 75–91.
- 1999b Nostratic and Sino-Caucasian: two ancient language phyla. In: *From Neanderthal to Easter Island (Festschrift W. W. Schuhmacher)*, ed. by N. A. Kirk & P. J. Sidwell. Melbourne: Association for the History of Language, 44–74.
- 2000a Notes on Lycian and Milyan. In: *The Asia Minor Connexion. Studies on the Pre-Greek Languages in Memory of Charles Carter*, ed. by Yoël L. Arbeitman. (Orbis Supplementa 13). Leuven: Peeters, 205–211.
- 2000b Symmetric construction in Milyan inscriptions. In: *Слово в тексте и в словаре. Сборник статей к семидесятилетию академика Ю. Д. Апресяна*. Москва: Языки русской культуры, 324–344.
- 2002a Milyan passages with *neu* and *ni(-ke)*. In: *Anatolian Languages*, ed. by Vitaly Shevoroshkin and Paul Sidwell. (AHL Studies in the Science and History of Language 6) Canberra: Association for the History of Language, 210–232.
- 2002b Word combinations in Milyan and Lycian inscriptions. In *Studia Linguarum 3/1-2. De omnibus linguae rebus scibilibus et quibusdam aliis. Memoriae A.A. Korolëv dicata*, ed. by A. S. Kassian and A. V. Sidel'tsev. (= *Studia linguarum* 3) Moscow: Languages of Slavonic Culture, 117–189.
- 2002c Gods and priests in Milyan and Lycian inscriptions. In: *The Linguist's Linguist: A Collection of Papers in Honour of Alexis Manaster Ramer*, ed. by Fabrice Cavoto. München: LINCOM, 405–420.
- 2003 Salishan and North Caucasian. *Mother Tongue* 8, 39–64.
- 2004a Proto-Salishan and Proto-North-Caucasian consonants: A few cognate sets. In *Nostratic Centennial Conference: the Pécs Papers*, ed. by Irén Hegedűs and Paul Sidwell. Pécs: Lingua Franca Group, 181–191.
- 2004b Topics in Milyan. In: *Per Aspera ad Asteriscos. Studia indogermanica in honorem Jens Elmegard Rasmussen sexagenarii indibus martiis anno MMIV*, ed. by Adam

- Hyllested et al. (Innsbrucker Beiträge zur Sprachwissenschaft 112) Innsbruck: Institut für Sprachen und Literaturen, 513–526.
- 2008a Introduction to Milyan. *Mother Tongue* 13, 63–96.
- 2008b On the Origin of Salish, Wakashan, and North Caucasian Languages. *International Journal of Modern Anthropology* 1, 84–121.
- 2009 On basic words in Salish, Wakashan and Lezghian. *Bygone voices reconstructed. On language origins and their relationships. In honor of Aharon Dolgopolski*, ed. by V. V. Shevoroshkin and Harald U. Sverdrup). Copenhagen: Underskoven Publishers ApS, 237–241.
- 2010 Four notes on Milyan. In *Исследования по лингвистике и семиотике. Сборник статей к юбилею Вяч. Вс. Иванова*. Москва: ЯСК, 156–167.
- 2011a Misadventures of Zeus in the Lycian Kingdom 2500 Years Ago. *Aramazd* 6:1, 24–42.
- 2011b Why are the nymphs grieving in Xanthos. *Aramazd* 6:2, 141–156.
- 2011c К пониманию милийских текстов. [On understanding Milyan texts.] *Слово и язык. Сборник статей к восьмидесятилетию академика Ю. Д. Апресяна*. Москва: ЯСК, 588–619.
- 2011d On the origin of Milyan nouns. *Mother Tongue* 16, 63–101.
- 2012a Хиазмы и сходные структуры в милийских текстах. [Chiasmuses and similar structures in Milyan texts] *Армянский гуманитарный вестник* 4, 98–120.
- 2012b Anatolian laryngeals in Milyan. In: *The Sound of Indo-European: Phonetics, Phonemics, and Morphophonemics*, ed. by Benedicte Nielsen Whitehead. Museum Tusculanum Press, 456–483.
- 2013a Salishan and North-Caucasian. *Global Journal of Human Social Science Research* 13, No 10-G.
- 2013b Милыйский язык. [The Milyan language] In: *Языки мира: Реликтовые индоевропейские языки Передней и Центральной Азии*. Ред. колл.: Ю.Б. Коряков, А.А. Кибрик. Москва: Институт языкознания РАН, 154–165.
- 2013c On the origin of Milyan verbs. *Mother Tongue* 18, 189–210.
- 2014 Milyan *trija*. *Aramazd* 8.1-2, 229–261.
- 2015a Milyan accusative constructions *lijeiz lupeliz* and *pleliz lijaiz*. *Historische Sprachwissenschaft* 128, 193–204.
- 2015b Notes on Anatolian languages. *Mother Tongue* 20, 113–148.

Reviews

Shevoroshkin regularly published reviews in Russian journals, such as *Вопросы языкознания* [Questions of Linguistics] and *Вестник древней истории* [Herald of Ancient History], as well as abroad. These reviews covered books on general linguistics, Anatolian languages, onomastics, and writing systems. Here is a just a sample of his numerous reviews:

- 1962 Martinet, A.: *Elements de linguistique generale*. *Вопросы языкознания* 11–12, № 6, 131–135.
- 1964 Palmer, L.R.: *Mycenaeans and Minoans. Aegean Prehistory in the Light of the Linear B Tablets*. L., 1961; & Heubeck, A.: *Praegraeca. Sprachliche Untersuchungen zum vorgriechisch-indogermanischen Substrat*. Erlangen, 1961. *Вестник древней истории* 1964, № 2, 192–194.
- 1965 Gusmani, R.: *Lydisches Wörterbuch*. *Вопросы языкознания* 1965, №2, 111–116.

- 1966a Zgusta, L.: Kleinasiatische Personennamen, Prag, 1964. *Вестник древней истории* 1966, № 2, 235.
- 1966b Zgusta, L.: Anatolische Personennamensippen, Teil I. Texte. Teil II. Beilagen, Prag, 1964. *Вестник древней истории* 1966, № 2, 235–237.
- 1969 Гиндин, Л. А.: Язык древнейшего населения юга Балканского полуострова (Фрагмент индоевропейской ономастики) . *Вестник древней истории* 1969, № 4, 146–152.
- 1971 Jensen, Hans: Die Schrift in Vergangenheit und Gegenwart, B., 1969. *Вестник древней истории* 1971, № 2, 136–138.
- 1972a (with G.M. Kallerman) Об изучении палайских текстов (O. Carruba: Das Palaische. Texte, Grammatik, Lexikon. Wiesbaden, 1970). *Вестник древней истории* 1972 № 3, 195–203.
- 1972b (with Z. Shalyapina) Venezky, R. L.: The structure of English orthography (The Hague–Paris: Mouton 1972). *Linguistics* 84, 89–100.
- 1972c Jensen, H.: Sign, symbol and script: An account of man's efforts to write (London, Allen & Unwin 1970). *Linguistics* 92, 110–114.

In Memory of Christopher Ehret (July 27, 1941–March 25, 2025)

PATRICK MANNING

Christopher Ehret, one of the most distinguished of historical linguists working on African and global materials, died on March 25, 2025, of the after-effects of a cardiac arrest in April 2022. His scholarship maintained high standards in studies of African languages and advanced the early history of Africa through language. His research brought high points in classification and reconstruction of African languages, archaeological studies that clarified early changes in African technology, economy, and society. He was employed at UCLA from the award of his PhD in 1968, became a full professor in 1978, and then held an emeritus Distinguished Professorship from his retirement in 2011.

His principal books were works on Southern Nilotic and Southern Cushitic history, and volumes on the reconstruction of Afroasiatic and Nilo-Saharan languages. These include:

- *An African Classical Age*
- *The Archaeological and Linguistic Reconstruction of African History*
- *History and the Testimony of Language*
- *Ancient Africa: A Global History, to 300 CE*
- *Civilizations of Africa.*

Ehret spent most of his life in Southern California, particularly Ventura County. He and I first met in 1959, when we were both 18 and part of a cohort of twenty male freshmen at Caltech. Because we were local Southern Californians, we had to live off campus; we became friends and often had dinner with a small group at a nearby drugstore. Ehret was an immensely lively and gifted student, but the scientific curriculum proved not to be his preference. After the first year he transferred to the nearby University of Redlands, where he earned a BA in history that included work in African history and linguistics. He later pursued graduate study in historical linguistics at Northwestern University. We met again in early 1966, this time as students in a multidisciplinary seminar for dissertation researchers preparing for fieldwork in Africa, taught by the noted Jan Vansina at the University of Wisconsin. By then I too had moved into African history and was preparing a Wisconsin dissertation in economic history. We remained in touch throughout his life.

He conducted dissertation research in Kenya for two years, where he established a life-long tie with historian Bethwell A. Ogot, a founding historian of East Africa who served as Chancellor of Moi University until 2013. Ehret completed dissertation work and was employed in 1968 as one of four young historians of Africa at UCLA, along with Edward Alpers, Robert Griffiths, and Boniface Obichere. In 1976, archaeologist Merrick Posnansky joined the History Department as the senior Africanist, and he and Ehret rapidly collaborated on a co-edited, multidisciplinary book.⁴

Ehret's early publications, in 1967 and 1968, were adventurous: linguistic analyses of the spread of cattle and milking and then of sheep in Africa, arguing that livestock moved from Central

⁴ Posnansky and Ehret, *The Archaeological and Linguistic Reconstruction of African History* (1982)

Sudanic-speaking areas of the Upper Nile to regions far to the south in a time before Bantu-speaking people had occupied Eastern and Southern Africa. While he was to revise certain of these arguments in time, his early commitment to such analysis enabled him to sharpen his thinking for another half-century.

Viewed over a longer time frame, Ehret created his own path but also echoed the work of Joseph H. Greenberg, who classified the languages of Africa in publications from 1948 to 1963. Greenberg further demonstrated that Bantu languages originated in the Cameroon grasslands, adjoining similar language groups, in contrast to the existing interpretation of Malcolm Guthrie. Ehret took his PhD at Northwestern, where Greenberg had studied in the 1930s. Like Greenberg, he too published research on all four African language phyla, working at multiple levels. He published detailed books on reconstruction of Afroasiatic languages (1995) and Nilo-Saharan languages (2001), where reconstruction is the detailed follow-up of classification. His studies of Khoesan addressed both classification and reconstruction.⁵

For the Kongo-Kordofanian language family, Ehret focused primarily on Bantu languages. In the generation following Greenberg's classification, numerous scholars elaborated the details of Bantu linguistic history and migration, among them Ehret, who contributed several important articles. In 2001 a special issue of a journal centered on his lead article reviewing these advances; it included fourteen scholarly commentaries and Ehret's response.⁶ In 2015, three of Ehret's former graduate students published a concise survey of 5500 years of Bantu social history, arguably the best social history of a large region over such an extended time span.⁷

In Afroasiatic languages, Ehret's studies added much new knowledge, though debates continued on the location of the Afroasiatic homeland. Greenberg had coined the term "Afroasiatic" to replace the limited and biased "Hamito-Semitic" classification that gave primacy to the place of Semitic. Ehret approved of Igor Diakonoff's 1975 proposed revision of the title to "Afrasian"; in practice Ehret used both "Afroasiatic" and "Afrasian". Classification within this phylum was strengthened by Harold Fleming's 1969 argument that Omotic languages of Ethiopia should be considered as a separate group, parallel to all other Afroasiatic languages. Ehret adopted Fleming's discovery and treated Omotic as one of two original Afroasiatic subgroups, the other being Erythraic. This new framework added a stage to the evolution of Afroasiatic and likely expanded its time-frame.⁸

But the thesis of Hamito-Semitic as languages arising in Asia, effectively challenged by Greenberg, returned every generation in a new form, using civilizational categorization as an implicit or explicit factor rather than using language only as a basis of classification. Pederson's initial 1903 proposal for Nostratic languages comprised languages of Eurasia, including Hamito-

⁵ Ehret's reconstructions of Afroasiatic and Nilo-Saharan do not give full accounts of sound changes among the languages but take a substantial step in that direction. His studies of Khoesan included a dictionary of Sandawe (co-authored with Patricia Ehret) and a study of extinct Khoesan languages of East Africa.

⁶ C. Ehret, "Bantu Expansions: Re-envisioning a Central Problem of Early African History," and "Christopher Ehret Responds," *International Journal of African Historical Studies* 34, 1 (2001): 5–41 and 82–87. (Pp. 42–81: responses from 14 scholars).

⁷ Catherine Cymone Fourshey, Rhonda M. Gonzales, and Christine Saidi. *Bantu Africa: 3500 BCE to Present*. Oxford University Press, 2017

⁸ Ehret later analyzed classification and reconstruction of Omotic languages.

Semitic in Asia and Africa. Post-war Russian linguists expanded their view of Nostratic to include Caucasian and Dravidian languages. Supporters of an expanded Nostratic phylum occasionally treated Indo-European and Semitic as founding groups rather than as subordinates in the proposed macro-family. This had the effect of underestimating the historical depth of Afroasiatic and Nostratic. In contrast, Ehret emphasized Semitic as a subordinate group in Afroasiatic, while Greenberg's 2000 publication proposed a Eurasiatic phylum that included Indo-European as one of six principal subgroups in northeastern Eurasia.

The underestimation of the time depth of Afroasiatic was especially evident in an influential *Science* article by Jared Diamond and Peter Bellwood. Citing Diakonoff, they argued that Afroasiatic originated with the rise of food production in the Levant around 9500 BCE and later spread into Egypt and North Africa with agricultural expansions around 5000 BCE. In their published response, Ehret, Keita, and Newman demonstrated that Diamond and Bellwood had misrepresented Diakonoff's views, and they further argued that Afroasiatic originated in Africa at least ten millennia earlier.⁹

In a dozen articles written after 2000, Ehret continued to offer details on the North Erythraean subgroup of Afroasiatic, based in northern Ethiopia, and its subdivisions to the north (Egyptian), east (Semitic), and west (Chadic and Amazigh). While details are yet to be worked out, especially in the timing, geography, and classification of Semitic languages, Ehret left a consistent vision of the structure and timing of Afroasiatic, and of its interactions with adjoining phyla in Africa and Eurasia.

Meanwhile, much of Ehret's work addressed localized issues alongside broader phylum-wide structures. His separate analyses of Nilo-Saharan, Kongo-Kordofanian, and Afroasiatic languages traced the gradual westward movement of languages from the Nile Valley within each family, each associated with its own characteristic economic and social patterns. In this work, Ehret engaged with a rich and expanding literature in African history, archaeology, and oral tradition (see his *Testimony of Language*), including a recent proposed linkage between the languages of the Kordofan Hills and Kordofanian groups in West Africa, themselves related to Niger-Congo. Likewise, another line of his research followed the gradual southward and westward movements from East Africa of Khoesan, Nilo-Saharan, and Kongo-Kordofanian languages.

Through the wide range of Ehret's studies in geography and social scale, he developed broad-based knowledge of early African history. I think he may have known more about the early history of Africa and the implications of that history than any other person. He expressed that knowledge in his 2002 *Civilizations of Africa*, which expanded the geographical and temporal scope for teaching African history up to 1800, in a book widely used in undergraduate courses. This text relied on the balance of bottom-up history and broad interpretation in his earlier books at language-group and regional scale. The second edition, in 2016, was confirmed as the most successful textbook on early Africa.

⁹ C. Ehret, S. O. Y. Keita, and Paul Newman. "The Origins of Afroasiatic," *Science* 306 (3 December 2004): 1680–1681; Jared Diamond and Peter Bellwood, "Farmers and Their Languages: The First Expansions," *Science* 597 (2003): 597–603.

To list just a few of Ehret's many other accomplishments were his co-authorship of two chapters on East Africa in the UNESCO *General History of Africa*; his remarkable article on "Reconstructing Ancient Kinship in Africa";¹⁰ his co-authorship in four major articles on genetic analysis of African history; his methodological analysis, *History and the Testimony of Language* (2011); and the Huggins Lectures on early African history that he delivered at Harvard on November 5, 6, and 7, of 2021.¹¹

The delivery of the Huggins Lectures gave Ehret the opportunity and encouragement to write a concise survey of African history. What appeared was a brilliant work, *Ancient Africa: A Global History to 300 AD*, published posthumously in 2023. He chose such "firsts" in Africa as ceramics, early agriculture, iron production, commercial towns, all in the savanna and Sahara. In a chapter on "The Africanity of Egypt," Ehret traced the movements of Afroasiatic-speakers from 20,000 BC in Ethiopia toward the lower Nile, settling among previous inhabitants and maintaining Afroasiatic-descended languages, with subgroups moving on to east and west. He joins this story to the rise of the Middle Nile Culture Area among speakers of Nilo-Saharan languages, especially 6000–3100 BP, where Qustul rose as a state with its Nabta Playa ritual center, the southern neighbor to the Afroasiatic-speaking Naqada state. As drought weakened the savanna and Qustul, Naqada thrived and conquered Qustul, forming the Old Kingdom of Egypt but retaining its ritual ties to Qustul. The concluding chapter of *Ancient Africa* is a narrative of the world history of speaking humans, showing the prominent place of Africa in the totality. Ehret notes the parallels in developments among world regions yet shows that African developments often preceded those in other regions. Thus the book addresses many of the specific issues on which Ehret has written over the years, with his updated statements on language and history; yet it is also a model of a way to write a continental history in a global context.

Shortly after submitting his manuscript to Princeton University Press, he suffered a severe cardiac arrest. He survived and lived comfortably but was no longer able to work on the book. Princeton University Press and Patricia Ehret were able to arrange a plan for editorial corrections that brought the book into print in 2023.

Chris Ehret put a great amount of energy into interaction and care for his family, including children, grandchildren, and close friends. His family interactions supported his scholarly activities. The Ehret family organized a memorial service on July 27, 2025, with a Zoom connection that included numerous participants in addition to the many more at the California site. Family members and colleagues spoke briefly but movingly, with heartfelt expressions of his kindness, his modesty and brilliance. He was tenacious in pursuing his many interests but was never combative. The proceedings of the memorial service were taped, including a wonderful photo montage of Chris's life.

¹⁰ In Nicholas J. Allen, Hilary Callan, Robin Dunbar, and Wendy James (eds.), *Early Human Kinship: From Sex to Social Reproduction*, pp. 200–231, 259–269. Oxford: Blackwell, 2008.

¹¹ <https://hutchinscenter.fas.harvard.edu/event/huggins-lectures-christopher-ehret-1-3>

Colin Renfrew: a note on his interests outside archaeology

ROGER BLENCH

Colin Renfrew, who died 24th November 2024, has been the subject of a wide variety of obituaries, reminiscences and memorials. These can easily be found online and need not be repeated here. However, I want to draw attention to his interests outside archaeology proper, which have hardly featured in these pieces. Almost all have been written by archaeologists and tend to feature his archaeological monographs, especially concerning the Aegean. But later in his lengthy career he turned to a variety of other topics, reflecting wide-ranging interests. These included global linguistics, but also the illicit antiquities trade and the tracing of stolen archaeological items, archaeogenetics, archaeoacoustics, and even the parallels between archaeology and the creativity of modern artists.

Colin Renfrew was the first director of the McDonald Institute (1990 onwards) as well as the Disney Professor of Archaeology at Cambridge University until 2004. This gave him considerable freedom to appoint individuals to follow up on his various interests and to initiate projects that would usually be considered outside the purview of a traditional archaeology department. This has now effectively been written out of the Cambridge narrative of archaeology. Although his portrait still hangs in the Institute, effectively all of the units he started have been closed down and the books relating to these divergent topics have been taken off display. The McDonald itself has been incorporated into the Department of Archaeology.

I do not propose to cover all these areas, merely to focus on his contributions to the interdisciplinary study of linguistics and archaeology. His interest in this type of long-range linguistics was reflected in his membership of the ASLIP council of fellows from 1993 until his death. His first major publication in this area was in 1987, when he published *Archaeology and Language: The Puzzle of the Indo-European Origins*. This represented a major challenge to the conventional model of Indo-European, the Kurgan hypothesis, originally championed by Marija Gimbutas, which had the earliest speakers as horse-riders erupting out of Central Asia. The ‘Anatolian hypothesis’ argued that Proto-Indo-European originated approximately 9,000 years ago in Anatolia and spread westwards via farming throughout the Mediterranean and into central and northern Europe. This was, paradoxically, not really a linguistic hypothesis, since it did not examine reconstructions of agricultural terms in any detail. This was the subject of much controversy and it was claimed by later proponents that Bayesian analysis of Indo-European languages supported this view. While it is certainly true that farming did gradually spread out from Anatolia during the time frame posited, it now seems this was not necessarily a linguistic expansion, or at least we cannot now associate it with particular languages. From 2015 onwards, new genetic data, especially concerning the Yamnaya hypothesis, gave further strong support to the Central Asian origin model (Lazaridis et al. 2025).

However, the reception of the Indo-European book encouraged Colin Renfrew to expand his model to global language phyla. He first floated these ideas in a paper to the Philological Society

(Renfrew 1989), but in an influential journal article, *Archaeology, genetics and linguistic diversity* he extended them to apply various types of dispersal to different language phyla (Renfrew 1992). Unfortunately (at least from my point of view) he adopted the Greenberg/Ruhlen classification, so that Amerind and Indo-Pacific made their appearances. Starting with a hunting/foraging/aquatic dispersal (Khoisan etc.), he assigned some phyla to agricultural dispersals and thence to a rather ill-defined 'northern dispersal' and to 'elite dominance' (which was not matched to language phyla). This was certainly an extremely bold and undoubtedly different approach to the conjunction of linguistics and archaeology. His contacts with Peter Bellwood, a frequent visitor to Cambridge at that time, convinced him of the importance of farming dispersals as a major element in language dispersal. A conference was held in Cambridge on this topic and an edited book was published, exploring the link (or otherwise) in the language phyla of the world (Bellwood & Renfrew 2002). This book is missing from the bibliography on Wikipedia, which is an interesting omission.

However, in some ways his most significant contribution was one where his name did not appear as an author. In the early 1990s, the third World Archaeological Congress was planned for New Delhi, taking place 4th–11th December 1994. Colin Renfrew urged Peter Ucko, the prime mover, to include a session on archaeology and language. Colin was too busy to be more than minimally involved in its organisation and, at his suggestion, Peter Ucko brought together myself and archaeologist Matthew Spriggs, and we in turn invited a large number of attendees to what turned out to be one of the most successful sessions of an otherwise somewhat chaotic conference. Selected papers plus several commissioned papers were eventually published in four volumes (Blench & Spriggs 1997, 1998, 1999a,b). It is fair to say this represented something of the state of the art at the time. Among other things, the first international presentation of the new mathematical models of historical linguistics surfaced at these sessions, in a paper by Russell Gray not included in the published proceedings.

Many linguists of a historical persuasion were convinced that this was the beginning of a great marriage of the two disciplines, and that interdisciplinary thinking, also bringing in the emerging science of historical human genetics, was about to revolutionise our understanding of the past. Regrettably, things took altogether a rather different turn. With Colin Renfrew stepping down from the Cambridge professorship in 2004, there was a reversion back to a more hardnosed 'scientific' archaeology. An emphasis on excavation returned, along with laboratory sciences, and collaboration with all other types of human sciences was largely discontinued. Colin Renfrew's interest in the topic remained, however, and in 2019, when I gave a presentation in the Leverhulme Centre for Human Evolutionary Studies, Cambridge, which presented a new model of world language phyla dispersals, he attended and gave comments as a discussant.

I cannot say I agreed with many of Colin Renfrew's specific hypotheses, but I absolutely appreciated the energy and enthusiasm with which he pursued these ideas. Although interest in these topics among a new generation of archaeologists is very restricted, honourable mention should go to the University of Gent and its Bantu studies which have continued to seek funding to

pursue and publish interdisciplinary work. Those of us who were inspired at the time continue researching these topics, which is, in part, the legacy of Colin Renfrew.

REFERENCES

- Bellwood, P. & C. Renfrew, eds. 2002. *Examining the farming/language dispersal hypothesis*. Cambridge: McDonald Institute for Archaeological Research.
- Blench, Roger M. & M. Spriggs. 1997. eds. *Archaeology and Language, I: theoretical and methodological orientations*. London: Routledge.
- Blench, Roger M. & M. Spriggs. 1998. eds. *Archaeology and Language, II*. London: Routledge.
- Blench, Roger M. & M. Spriggs. 1999a. eds. *Archaeology and Language, III*. London: Routledge.
- Blench, Roger M. & M. Spriggs. 1999b. eds. *Archaeology and Language, IV*. London: Routledge.
- Lazaridis, I., Patterson, N., Anthony, D., Vyazov, L., Fournier, R., Ringbauer, H., Olalde, I., Khokhlov, A.A., Kitov, E.P., Shishlina, N.I. and Ailincăi, S.C., 2025. "The genetic origin of the Indo-Europeans." *Nature*, pp.1-11.
- Renfrew, C. 1987. *Archaeology and Language: The Puzzle of Indo-European Origins*. London: Pimlico.
- Renfrew, C. 1989. "Models of change in language and archaeology." *Transactions of the Philological Society*, 87(2): 103–155.
- Renfrew, C. 1992. "Archaeology, genetics and linguistic diversity." *Man*, 27: 445–478.

GENETICS AND THE PREHISTORY OF DRAVIDIAN LANGUAGES

SUNDAR NARAYAN
LAMBTON COLLEGE, SARNIA, ONTARIO, CANADA

Dedicated to David McAlpin (1945-2023)

ABSTRACT

Recent genetic studies, which are perhaps not yet familiar to linguists, have cast a new light on the origins of the Dravidian language family of India, so this review paper compares Dravidian historical linguistics and archaeology with recent archaeogenetics publications and concludes that they support the Proto-Elamo-Dravidian (Proto-Zagrosian) model of McAlpin and Southworth.

Keywords: Genetics, Elamo-Dravidian, Zagrosian, Tamil, India, Indus Valley, Aryans, Iran, Elamite.

1. INTRODUCTION

Dravidian is the world's sixth largest family of languages, comprising over 24 languages (Steever, 2017) that are spoken by 250 million people living mainly in South Asia. See Figure 1, which shows the linguistic map of India and a few other countries. In southern India, the 4 major Dravidian languages, i.e., Telugu, Tamil, Kannada and Malayalam, have a total of 240 million speakers, a number that has just started to decrease due to low birth rates, a first in overpopulated India. About 20% (Misra, 2024) of India's people speak the Dravidian languages whose speakers have included Nobel laureates and world chess champions. Nearly all the rest of India's population, who live mainly in North India, speak various Sanskrit-derived Indo-Aryan (IA) languages, e.g., Hindi, Punjabi, Gujarati, Marathi, Bengali, etc., which are a subset of the larger Indo-European family (Cardona, 2025). The two other language families found in India are Tibeto-Burman and Austro-Asiatic (e.g., Munda), but they are spoken only by small, remote tribal populations (Southworth, 2005: 51–52).

A total of 9 million Tamils live in northern Sri Lanka and Malaysia. North Dravidian languages such as Kurukh (Kurux) and Malto are spoken by small populations of tribal peoples in eastern India and Bangladesh. Pakistan has over 2 million speakers of Brahui, which is regarded as a North Dravidian language by some linguists (Krishnamurti, 2003: 20–22). In recent decades,

over 3 million Dravidian speakers, many of them educated professionals, have settled in the U.S., U.K., Canada and Australia.

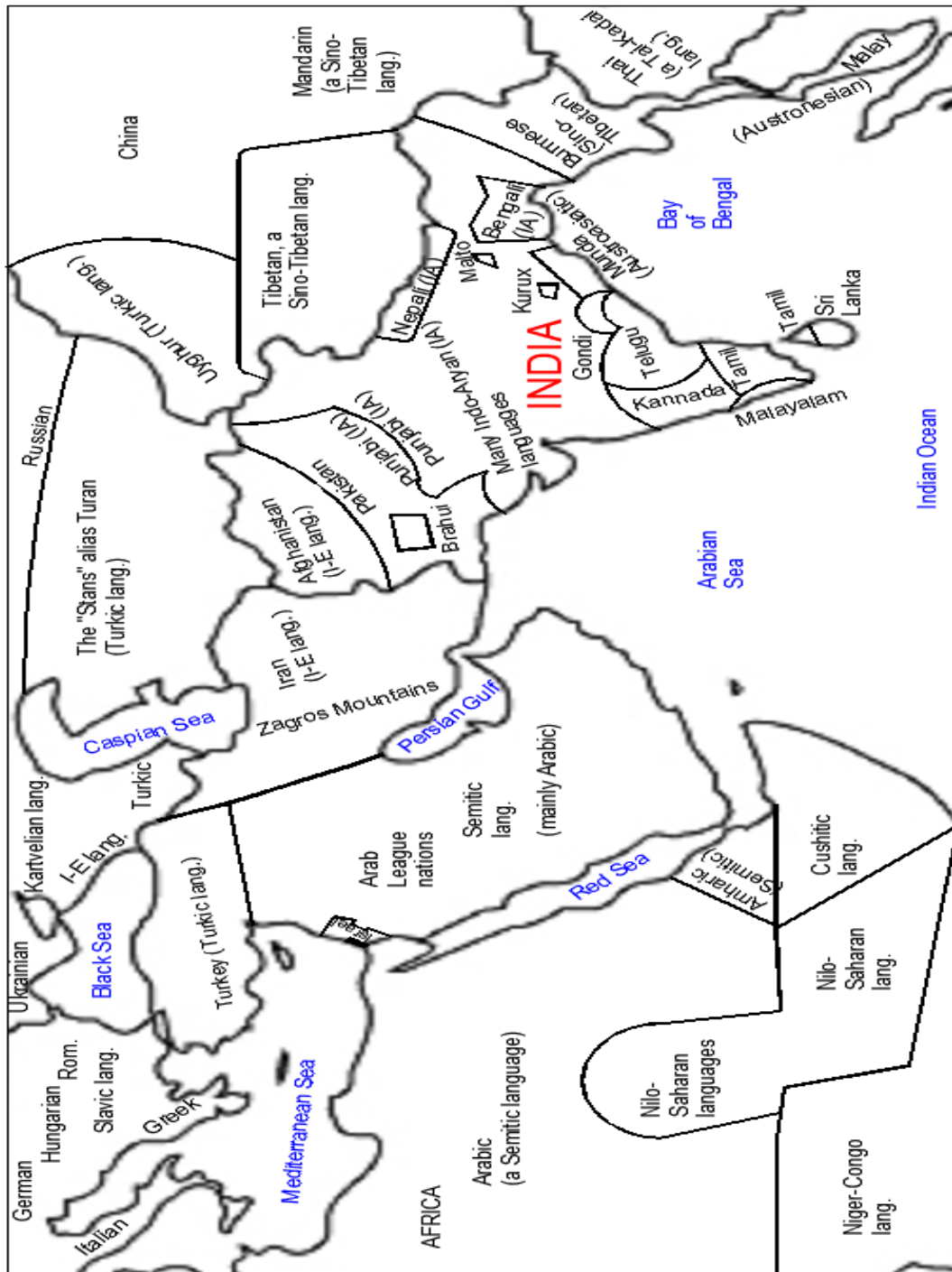


Figure 1: Contemporary Linguistic Map of India and Other Countries

1.1 A Very Brief History of Dravidian Linguistics and Literature

In 1816, Francis Whyte Ellis, a British official working in India, was the first to write that southern India's languages form a distinct family of languages which had no connection to Sanskrit except in the form of the many Sanskrit loanwords in them (Krishnamurti, 2003: 16–17).

Between 1856 and 1913, the various editions of Robert Caldwell's *Comparative Grammar of Dravidian* convinced scholars that Ellis had been right. The modern era of Dravidian linguistics began in 1961 when Krishnamurti wrote his book on Telugu verbal bases. In the same year, Burrow and Emeneau published the *Dravidian Etymological Dictionary* (DED), which is still a standard reference work, especially its 2nd edition in 1984. Proto-Peninsular Dravidian reconstructions are given in Krishnamurti (2003: 523) and the Starling Database. The books and/or papers, listed in the references section of this paper, of Krishnamurti, Subrahmanyam, Zvelebil, McAlpin, Parpola, Steever, Andronov and Southworth form the literature of Dravidian linguistics. The Dravidian Linguistics Association of India sells useful papers and books. See also this paper's additional references section written by an anonymous referee.

Only 4 Dravidian languages are literary languages. The literatures of Kannada, Telugu and Malayalam began in medieval times, but Tamil literature and inscriptions date back to the 3rd century BCE (Steever, 1998). The claim that the potsherds bearing Tamil graffiti found in Keezhadi (also spelled Keeladi) are from the 6th century BCE is not accepted by non-Tamil archaeologists (Charuchandra, 2019). The earliest extant books in Tamil, called the Sangam poems, were written by bards who supposedly were members of the third Sangam (meaning academy, from the Sanskrit *sangha*). Claims by some Tamilians about the 'lost' poems from the 'first' and 'second' Sangams, supposedly founded around 10000 BCE and 5500 BCE (Sastri, 1897), which got submerged by rising sea-levels along with the geologically impossible vast Kumari Kandam continent, are not credible (Ramaswamy, 2004 and Zvelebil, 1992).

Recent genetics papers, for instance, Narasimhan et al. (2019), suggest the possibility that Ancient Ancestral South Indians (AASI) may have been speaking Proto-Dravidian (PD) as their mother tongue, so the next section of this paper, which is on archaeogenetics, examines this possibility.

2. ARCHAEOGENETICS AND THE PREHISTORY OF INDIA

2.1 The Recent African Origin of Humans

It is now well-accepted by paleontologists that the human species, *Homo sapiens*, originated in East Africa 300 kya (kiloyears ago) and started spreading across the world 60 kya to 50 kya (Handwerk, 2021). See Figure 2, which shows only a few migration routes (continuous lines with arrows) from 300 kya to 40 kya and a few other relevant migration routes (short segments with arrows) from 10 kya to 4 kya. The xOoA migration from Africa to Australia, shown as a dashed line in Figure 2, took place much earlier than the southern migration route, which is shown—as previously mentioned—by continuous lines with arrows. The genome from the xOoA migration is now found only in Australia and Papua New Guinea; No trace of it or earlier human or hominin migrations to India is seen genetically in South Asians today (Culotta, 2016). So, it is highly unlikely

that these earlier migrations contributed anything to later human speech in India or elsewhere (Cullotta, 2016), hence they will be ignored in this paper.

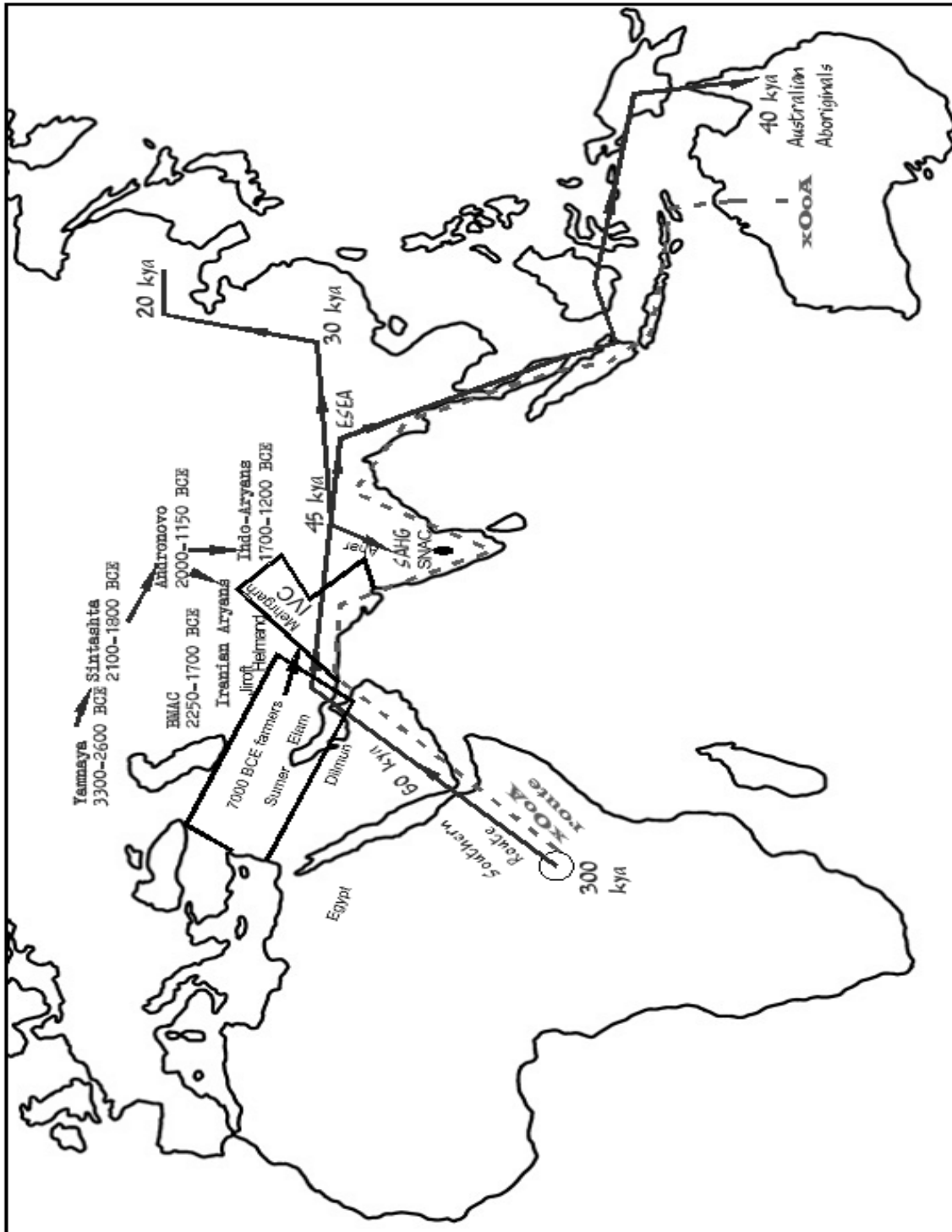


Figure 2: A few selected migration routes from the Paleolithic and Neolithic ages. The Yamnaya-to-India route is not known even approximately, which explains why no archaeological evidence has been found.

2.2 South Asian Hunter-Gatherers (SAHG), the ancestors of the AASI

As Figure 2 shows, humans crossed into Arabia 60 kya. Some of them reached the present India-Nepal border 45 kya, i.e., close to the start of the Upper Paleolithic Age. There, this population trifurcated into South Asian Hunter-Gatherers (SAHG, who later became the AASI), East and Southeast Asians (ESEA) and Australasians (AA), according to Yang (2022), but other references (e.g. Bennett et al., 2024) show the trifurcation happening in Pakistan or Burma. The SAHG entered India some 45 kya and slowly spread across much of that sub-continent, including its west coast, but South Asia's archaeological record does not reveal much about the SAHG. Cave paintings, which began to appear along with the Mesolithic, i.e., 10,000 BCE (Clottes, 2004), may be associated with the SAHG, as can the Southern Neolithic Archaeological Complex (SNAC) of 2500 BCE onwards (Thapar, 2025), which is discussed later.

It is impossible to know what language the three aforesaid groups spoke 45,000 years ago. The only language choice that is available from a few brave linguists (Fleming et al., 2013) is a language superfamily called Borean, which was spoken 45 kya. Borean was proposed as the ancestor of many of today's language families or proposed superfamilies, including Nostratic. Today, the descendants of the SAHG in South India, ESEA, and Australasians speak Dravidian, Austric, and Australian languages, all probable descendants of Borean. The Austric superfamily proposal combines several existing recognized language families of East and Southeast Asia and is not accepted by most linguists (van Driem, 2001). The Australian language superfamily combines Pama-Nyungan and non-Pama-Nyungan, so it includes all Aboriginal languages (Harvey and Mailhammer, 2024).

When, if at all, did the SAHG start speaking Dravidian languages? Since genetics has proven (Yang, 2022) that the SAHG, ESEA and AA diverged from each other, it is reasonable to suppose that Dravidian, Austric and Australian may be remotely related to one another, as discussed below.

2.2.1 *Is there a Relationship between Australian and Dravidian?*

At present, standard textbooks on historical linguistics state that 90% of the 250 remaining Aboriginal languages of Australia belong to the Pama-Nyungan language family. The remaining 10% belong to what is called the non-Pama-Nyungan family. Harvey and Mailhammer (2024) have combined the 2 language families and have reconstructed Proto-Australian.

Eminent Australianist Dixon (1980: 19) admits that there is no way of telling whether proto-Australian was spoken 30,000 years or longer ago, or more recently, such as 4000 years ago, but it seems clear that nearly all the languages of Australia form one genetic family going back to Proto-Australian (Dixon, 1980: 228). O'Grady and Hale (2004) write that there was a great extinction of languages in Australia 5 kya when the Pama-Nyungan languages replaced existing languages. This view is now widely accepted (Blazek, 2006), making it impossible to prove any remote connections between Dravidian and Australian. Harvey and Mailhammer (2024) conclude that Proto-Australian was most likely spoken 6000 years ago in the Northern Territory of Australia before spreading all over that continent. Bergstrom et al. (2016) point out that the overall evidence is consistent with a lack of gene flow from India to Australia in the Holocene, so connections

between Pama-Nyungan and India's languages are unlikely. Dixon (1980: 236–237) points out that a thorough search of Dravidian and Australian yields no systematic correspondences of form and meaning which could even hint at a genetic connection between the two. Thus, the conclusion is that any connection between Dravidian and Australian remains unproven.

2.2.2 Is there a Relationship between Austric languages and Dravidian?

The Austric superfamily of languages is a proposal, not accepted by most linguists, that combines several recognized language families of East and Southeast Asia, such as Austroasiatic, Austronesian, Tai-Kadai and Hmong-Mien into Greater Austric (van Driem, 2001). No systematic comparisons of these families with Dravidian seem to have been done by linguists. A comparison, using only wordlists, of Dravidian with Greater Austric or Sino-Tibetan reveals no cognates, though Caldwell (1913: 75) did find that Dravidian and Tibetan had somewhat similar personal pronouns. In eastern central India, some Dravidian languages have long been in contact with the Munda languages of the Austroasiatic family, but there is no genetic relationship (Krishnamurti, 2003: 26). Due to the above reasons, it is almost certain that there are no deep connections between Dravidian and any Austric language.

Thus, archaeology and linguistics are unable to prove that the Dravidian is related to Austric and/or Australian and/or that the SAHG/AASI of southern India spoke Dravidian languages. So, it is possible that in 10,000 BCE, the SAHG/AASI could have been speaking unknown languages which had no relationship to Australian or Austric. Both statements are consistent with the possible birth of Proto-Dravidian 1 (PDr1) somewhere between Iran and the Indus Valley Civilization (IVC) of Pakistan and north-west India, as discussed in the next section.

3. GENETICS AND THE ELAMO-DRAVIDIAN HYPOTHESIS

In contemporary Iran, over 70% of the people speak one of several Indo-Iranian languages; nearly all the rest speak various Turkic languages. But in Bronze Age Iran which had only non-Aryan and non-Turkic people living in it, Stolper (2004) says that at its maximum extent, the Elamite linguistic area could have stretched from Azerbaijan in the Northwest to the eastern shore of the Persian Gulf and all the way east to Balochistan, i.e., at the western border of the IVC. This corresponds to the first quadrant of the rectangle in Figure 2, i.e., from the present Iran-Pakistan border to the shores of the Caspian Sea.

Bellwood (2013) points out that languages in pre-state societies usually spread through human migrations. Pathak et al. (2024) trace the creation of the L1-M22 Y-chromosome haplogroup 20.6 kya somewhere in the Middle East, its expansion there till 10 kya, and its arrival near the Indus River 8 kya. The authors, one of whom (Phillip Endicott) is a linguist, state that their study supports McAlpin's Elamo-Dravidian language family proposal (which a few linguists actively oppose) by suggesting a connection between the roots of all Y-chromosomal L1-M22 lineages and the ancestry of Caucasus Iranian Hunter-Gatherers (CIHG or CHG). These were closely related to the Iranian farmers mentioned earlier, and the CIHG are the best candidate for disseminating both the Elamite and Dravidian languages to south-western Iran (Elam) and the IVC, respectively, with

both these locations being the final areas of migration from yet another region inhabited by a population with CIHG ancestry. Their results do not support the suggestion that L1-M22's geographical expansion started in Elam, but do not rule it out either.

McAlpin (2015) renamed Proto-Elamo-Dravidian as Proto-Zagrosian and Proto-Dravidian 2 as Proto-Peninsular Dravidian (PPD) and updated his original proposal (McAlpin 1981). McAlpin's Zagrosian family tree is shown in Figure 3 below. His original term PDr1 is retained in it and in this paper. Note that authors other than McAlpin and Southworth refer to PPD as Proto-Dravidian (PD).

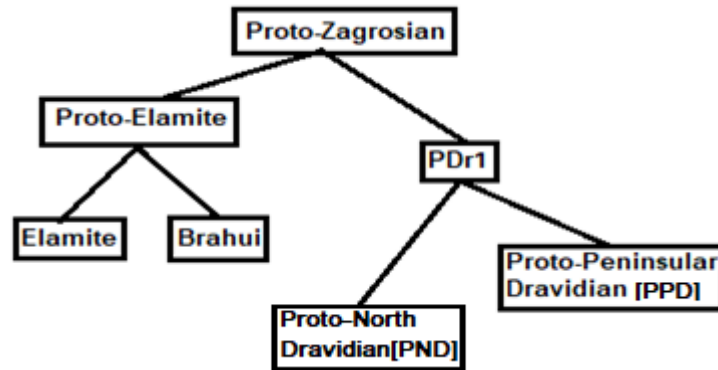


Figure 3: Proto-Zagrosian and its descendants; PND = Proto-Kurux-Malto (McAlpin, 2015).

In Figure 3 above, PDr1 is McAlpin's newly proposed proto-language that was intermediate between Proto-Zagrosian and PPD. McAlpin (2015) states his reasons for moving Brahui from the Dravidian language family, where most linguists place it, to the Elamite 'family'. Genetic studies by Pagani et al. (2017) of the Brahui people have conclusively proven that they have no genetic affinity with South Indians and are no different from their Balochi neighbors. Brahui today has many Balochi loanwords. Pagani et al. (2017) conclude that the Brahui are a remnant of what was once a more widespread population of Brahui speakers. This strengthens the case for McAlpin's Proto-Zagrosian model. Among Dravidianists, only Georgiy Starostin (2002) is skeptical about Elamo-Dravidian, but Zvelebil (1990: 113), Southworth (2011), Bomhard (1996) and Krishnamurti (van Driem, 2023) support it, though the latter initially had opposed it (Krishnamurti, 2003: 44–45). Blazek (1999) tried to link Elamite with Afroasiatic but does not view his theory as being opposed to McAlpin's and sees no obstacle to bringing all three language families together (Starostin, 2002). Most non-Dravidianists consider Elamite to be a language-isolate (Stolper, 2004). Unlike Starostin, they do not actively oppose its membership in the Zagrosian family.

According to Pathak et al. (2024), the L1-M22 haplogroup could have arisen 20.6 kya anywhere within the tilted rectangular area shown in Figure 2, south-east of Turkey; then, this haplogroup reached the Indus River in 6000 BCE, having already spread all over the Middle East by 8000 BCE. Bomhard (1996: 111–112) states that Proto-Elamo-Dravidian (PED, i.e., McAlpin's Proto-Zagrosian) came into being between 8000-7000 BCE, which seems to agree with the findings of Pathak et al. (2024). Pejros and Shnirelman (1988) said that PED's homeland was the

western or central Iran of today. The latter is east of the Zagros mountain range. The two authors proposed that PED split into Proto-Elamite in the west (Iran) and Proto-Dravidian, i.e., the PDr1 of McAlpin (2015) in the east (Pakistan), with this split occurring in 5000 BCE according to Bomhard (1996: 111–112), though Andronov (1964) obtained 4000 BCE as the date according to lexicostatistics. Thus, the aforesaid linguistic dates accord well with the Pathak et al. and other genetics papers, and the Indus Valley Civilization, as discussed in the next two sections.

4. INDUS VALLEY CIVILIZATION AND ARCHAEOGENETICS

In 1925, British and Indian archaeologists announced the discovery of the Indus Valley Civilization (IVC) in what is now Pakistan, though later archaeological digs in post-independence India would find similar settlements. Excavations in Harappa, Punjab and Mohenjo-Daro, Sind, revealed well-planned cities with brick houses and, amazingly, planned sewerage systems, the first in the ancient world. By the 1970s, radiocarbon dating had conclusively proven that the IVC began around 3300 BCE, was in its mature form from 2600 BCE to 1900 BCE and declined from 1900 BCE till 1300 BCE due to climatic, economic and other less well-understood reasons (Wright, 2010).

Only in the 21st century has it become possible to know the ethnic make-up of the IVC's inhabitants. Ancient DNA (aDNA) analyses of 12th to 1st millennium BCE samples, 523 of them, gathered from Iran and the 'Stans' by Narasimhan et al. (2019) showed that Indus Valley people had an ancestry profile that was 45–82% Iranian farmer-related and 11–50% from Ancient Ancestral South Indians (AASI). The Iranian farmers had up to 10% Western Siberian Hunter Gatherer (WSHG) but very little Anatolian genome. The ancestors of the IVC people were formed by admixture between 5400 to 3700 BCE. Narasimhan et al. used the DNA from modern Andamanese hunter-gatherers (AHG) as a proxy for the now-extinct AASI, but this follows well-accepted archaeogenetics practice (Reich et al., 2009; Debortoli et al., 2020). Yelmen et al. (2019) used South Indian Irula DNA as the proxy and reached the same conclusions as Narasimhan et al.

During the IVC's decline from 1900–1300 BCE, some of its people migrated south and east, mixed with the AASI living there to form the Ancestral South Indian (ASI) population. The remaining IVC population interbred with one or more waves of newcomers, the Indo-Aryan-speaking mainly male settlers, many having the R1a Y-chromosome, who had completed their centuries-long migration from the Yamnaya Steppe in present-day Ukraine and Russia, to form the Ancestral North Indians (ANI) (Narasimhan et al., 2019).

Today, most North Indians and South Indians have varying amounts of ANI and ASI ancestry, called the Modern Indian Cline. Indo-Aryan-speaking North Indians today have more ANI, IVC Iranian farmer, and Steppe ancestry and less ASI or AASI ancestry than the Dravidian-speaking South Indians have now (Narasimhan et al., 2019).

Narasimhan et al. also found that Yamnaya Steppe ancestry first arrived by 2100 BCE in the Bactria-Margiana Archaeological Complex (BMAC), which lay between present-day Afghanistan and the Caspian Sea, but the BMAC people, whose descendants are most of today's Iranians, are not closely related to Indians. That paper and Figure 2 show that the Old Indo-Aryan speakers of

Yamnaya ancestry either went to Iran or northern India in the early 2nd millennium BCE, a fact that Hindus subscribing to the Indigenous Aryans theory deny (Shahane, 2019). The Indo-Aryans arrived in the IVC region when it was in decline.

Perhaps the most well-known fact about the IVC is the undeciphered IVC ‘script’ which is discussed in the next section.

5. IVC SEALS AND SCRIPT

The first Indus seal containing what seemed to be writing was found in 1875. By 1929, thousands of seals and seal impressions had been found. They bore inscribed images and/or glyph-like signs; see Figure 4a. Perhaps understandably, scholars simply assumed that the Indus seals contained unreadable writing in an unknown language. It would not be till the 21st century that such a casual assumption would be questioned by Farmer, Sproat and Witzel (2004). However, many aspects of that paper were criticized by Parpola (2005, 2007). Rao et al. (2009) claimed to have disproven the not-a-script thesis, eliciting a sharp response from Farmer, Sproat, and Witzel (2009). Despite these three authors, most scholars still believe that the Indus seals do bear written inscriptions.

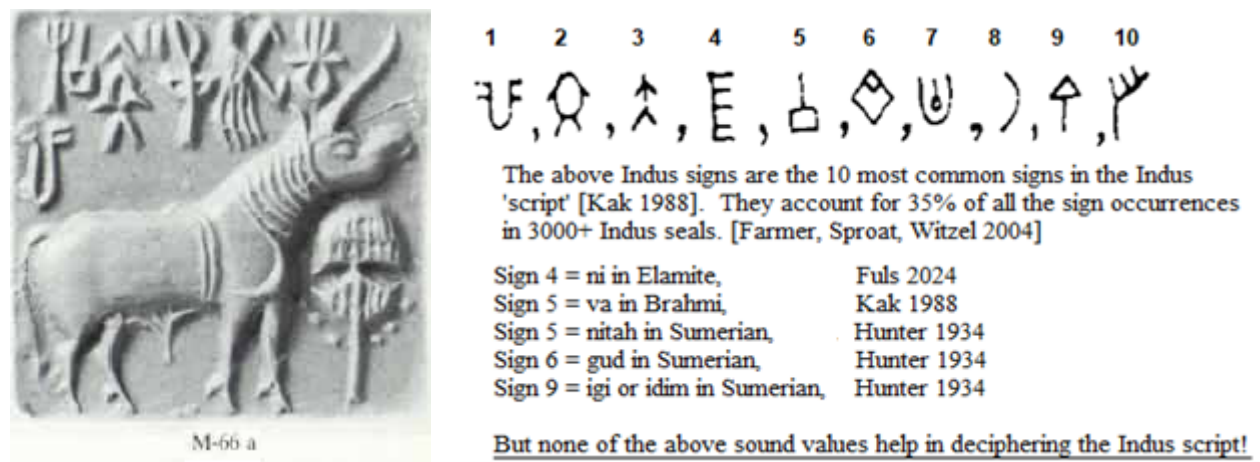


Figure 4a: Seal impression with right-to-left writing, Figure 4b: Ten most common signs of Indus script.

Since the first detailed study of the Indus script by Hunter (1934), there have been nearly 70 Indus script decipherment attempts. The books or papers, now available online, of Heras (1953), Mahadevan (2002, 1982), and Parpola (1994) explain why the more famous decipherments were not accepted by a majority of scholars. Most decipherment attempts have assumed that the script encodes Sanskrit or Dravidian. The Sanskrit-based decipherments were dismissed for linguistics-related reasons, but now they can all be dismissed in light of the new genetics-based evidence, which clearly shows that Indo-Aryans, hence Sanskrit, first entered the IVC region only during the 2nd millennium BCE, when the IVC was in its decline and seal production had stopped. As stated earlier, languages in pre-state societies usually spread through human migrations. The orthodox opinion among scholars for over 60 years has been that the IVC seal writing is in PPD (which they refer to as Proto-Dravidian or just Dravidian), even though, as Farmer, Sproat and Witzel (2004)

have pointed out, such an assumption has failed to produce a single, verifiable reading. However, Dravidian place names ending in ‘paLLi’, ‘vali’ or ‘wari’ can be found even today from the Indus valley to central India (Witzel, 1999). Such names may predate 300 BCE, since Jain literature mentioning ‘paLLi’ is from that date (Sankalia, 1949).

The Indus script was contemporary (2800–1900 BCE) with the Linear Elamite script. In Elamite Iran, the first Proto-Elamite tablets with writing in the Proto-Elamite script first appeared in 3100 BCE; they bear a clear resemblance to the proto-cuneiform Sumerian tablets of the Mesopotamian Uruk IV period, which antedated the tablets by 100 years. A few still question Desset’s recent full decipherment claim of the Proto-Elamite and Linear Elamite scripts, but Linear Elamite is considered partially deciphered, due to the earlier efforts of Mader (Fuls, 2024). Only one of the 10 most common Indus signs (see Figure 4b) has an equivalent in Linear Elamite (Fuls, 2024), so an impending decipherment of the IVC script is not likely. Bronze Age Jiroft, near the IVC, had an entirely different script from the IVC (Basello, 2006).

Though Mahadevan (1998) held that the IVC was monolingual (and Dravidian), it is possible that the writing on different seals may encode different languages, greatly complicating any decipherment attempt. The case for a multilingual IVC is strong and is discussed in the next section.

6. POSSIBLE IVC LANGUAGES

6.1 PDr1 as an IVC language

Narasimhan et al. (2019) showed that the ancestors of the IVC people were formed by admixture between Iranian farmers and SAHG/AASI, during 5400-3700 BCE, i.e., before the IVC began, in north-western India and Pakistan. The work of Bomhard (1996), Pejros and Shnirelman (1988) and McAlpin (2015) suggests that the eastern dialect of Proto-Zagrosian became PDr1 in 5000 BCE or 4000 BCE (Andronov, 1964). So, it is likely that IVC’s forefathers spoke PZ and, later, PDr1, which probably accepted loanwords from the language/s of the SAHG/AASI when they interbred with the Iranian farmers. The oldest settlement in the Iran-Pakistan area is Mehrgarh, Pakistan, which dates back to 7000 BCE; PDr1 speakers may have lived in Mehrgarh’s Periods 2 and 3 from 5500-3500 BCE. The nearby Helmand Culture cities, Shahr-i-Sokhta and Mundigak, had trading links with the IVC, so they too may have hosted PDr1 speakers. PDr1 may have been the major language of the IVC during much of its existence, except perhaps during its decline. PDr1 was probably linguistically closer to PZ and Elamite than to PPD. If PDr1 was the language of the Indus seals, that may explain why scholars have been unable to read PPD into the seals.

If PDr1 got replaced by PPD, that cannot be discovered from the still-undeciphered Indus seals. However, Witzel (1999) points out that Hinduism’s oldest book, the Rigveda, composed in Vedic Sanskrit by Indo-Aryan immigrants to the IVC area during the declining centuries of the IVC, contains important clues. The Rigveda has been chanted unchanged and memorized by India’s Brahmins since its composition in the 2nd millennium BCE, so it is like a tape-recording, according to Witzel who bases his quest for Dravidian and foreign loanwords in the Rigveda on Kuiper’s list of 380 such words which form only 4% of the total word count of the Rigveda. Witzel

writes that the Rigveda reflects the Punjab province of Pakistan and India from 1700-1200 BCE. This date range agrees very well with various genetics studies that show that one or more waves of Indo-Aryans entered India in the 2nd millennium BCE. Witzel assigns each of the 10 books, except book 9, of the Rigveda to the following three layers: i) the early Rigvedic period from c.1700-1500 BCE, ii) the middle Rigvedic period from c.1500-1350 BCE, and iii) the late Rigvedic period from c.1350-1200 BCE.

Witzel reports that the early layer of the Rigveda contains no Dravidian words, but they suddenly appear in the middle layer and the late layer of the Rigveda. It is exciting to imagine that the tape-recording that is the Rigveda might have caught PPD soon after its birth during the middle Rigvedic period, but that is too good to be true! As van Driem (1999) points out, it is more likely that Indo-Aryans met Dravidian speakers for the first time in post-IVC Punjab of the 15th century BCE. PPD may date back to around 1800 BCE (Southworth, 2011), i.e., after the start of IVC's decline, as discussed later. By that time, seal production had stopped, so the writing on the seals cannot be in PPD.

6.2 Para-Munda as an IVC language?

Witzel (1999) also claims that, in addition to PPD words, the foreign words in the Rigveda also include “Para-Munda” words of apparently Austroasiatic origin. Witzel tries to show that the Para-Munda substrate was the predominant language of the IVC. As van Driem (1999) points out, the case for a Dravidian Indus is still intact, with or without the Para-Munda substrate in the Rigveda or the IVC. In his paper, Witzel admits that he is not a Dravidian specialist. Krishnamurti (2003: 38–39), who was one, points out that Nikita Gurov has shown that several of Witzel's Para-Munda words have Dravidian etymologies based on compounding and not prefixing and comments that it would have been better if Witzel had said that he did not know the true source of the 300+ borrowings in the Rigveda.

Probably because Para-Munda was confused with Proto-Munda, Witzel renamed Para-Munda as the Kubha-Vipas (K-V) substrate (Witzel, 2009), which may or may not have been related to the Munda languages. Because Southworth's standard work (2005: 68) writes that Old Indo-Aryan had Munda/AA words, Witzel's Para-Munda or K-V hypothesis is further discussed below.

The Munda peoples, classified as Scheduled Tribes by the Indian government, still live in eastern central India (see figure 1), except the Munda-speaking Korku tribe, who live in central India. The Munda people speak the eponymous Munda languages, which belong to the Austroasiatic (AA) family, which today is found in Southeast Asia, especially Cambodia, Laos and Vietnam, where AA speakers form the majority. The Munda's mostly male AA ancestors migrated from south-east Asia to east India 4.3 kya according to Arunkumar et al. (2015) and 2 to 3.8 kya according to Tattē et al. (2019). These dates make it impossible for Munda speakers to be in the IVC, 2000 km away. An earlier genetic study by Zhang et al. (2015) had the Munda arriving in East India 10 kya, but this study seems to have been superseded by Tattē et al. (2019), who cite it. Most speakers of Ho, a Munda language, live in East India, but a few live near Punjab. So, the latter must be recent migrants to the area and could not have interacted with Indo-Aryans. Also,

the Modern Indian Cline on which most present-day North and South Indians fall (Narasimhan et al., 2019) neither needs nor includes the Austroasiatic speakers' genome.

If the Rigveda indeed contains K-V words, they need to be from an unknown non-Munda AA source, which somehow did not contribute to the Modern Indian Cline, but this is special pleading. It is more probable that there are no Para-Munda or K-V words in the Rigveda (Krishnamurti, 2003: 38–39; van Driem, 1999).

6.3 Language X and Meluhhan as IVC languages?

Witzel (1999) also says that the Rigveda's Para-Munda substrate overlays a still older local level, which may be identical to Masica's "language X" of the pre-Aryan Gangetic plains. Language X had been proposed by Masica (1979), who found that 80% of the agricultural terms in modern Hindi cannot be traced to any other language, extinct or extant. Masica (1979) did not examine if these agricultural terms were part of the vocabularies of languages ancestral to modern Hindi, but Southworth (2005: 66–69) says that they are of general occurrence in Indo-Aryan, so he says that it is reasonable to posit the "Indus" languages as the source of this material.

Meluhhan was proposed as an IVC language by Southworth (2005: 64–66) and Witzel (1999). Meluhhan is needed to explain the 40 non-Dravidian loanwords in Sumerian that refer to goods from Meluhha, which is identified by scholars as the IVC. Meluhhan might have been spoken only in the Balochistan-IVC border region, according to Southworth and Witzel.

As Narasimhan et al. (2019) found, the Iranian farmers who were the source population of the IVC, along with the SAHG, had up to 10% Western Siberian Hunter-Gatherers (WSHG) genome due to admixture during their stay in eastern Iran. So, the unknown WSHG language might have influenced the IVC's languages.

Thus, the languages of the IVC's early and mature phases may have been PDr1, Meluhhan and Language X, all of which may have had SAHG and WSHG loanwords. Another possible but very unlikely IVC language may have been an unknown Austroasiatic language like Witzel's Kubha-Vipas. During the late 2nd millennium BCE, in what is now North India and Pakistan, all Bronze Age languages (except perhaps Brahui and Burushaski, both spoken only in mountainous areas) got replaced by Old-Indo-Aryan (OIA) dialects, i.e., Sanskrit dialects, while multiple waves of Indo-Aryans mingled with the IVC population. Their descendants, the ANI who had some ASI genome, spread across north India and Pakistan, along with Vedic Hinduism. In the next millennium, OIA dialects gave way to Middle Indo-Aryan, alias Prakrit languages, whose daughter languages are still widely spoken today (Southworth, 2005: 44–47).

6.4 Bronze Age Language/s of the SAHG in South India

The question as to which languages were spoken by the SAHG living in southern India during the IVC period before contact with southbound IVC migrants cannot be answered at present. However, Southworth (2005: 66, 329) suggests Proto-Nilgiri as one of the non-Dravidian languages of South India, which preceded the arrival of PPD. Zvelebil (1990: 63–70) listed a few words from Southworth's Proto-Nilgiri which are now found in the substrate in Irula, a Dravidian language closely

related to Tamil. It is important to note that Southworth uses the term ‘Proto-Nilgiri’ to refer to a non-Dravidian language, whereas the Starling Database uses that term to refer to the Dravidian ancestor of the Irula language. The Irula scheduled tribe was once mistakenly bracketed with Australia’s Aboriginals as part of the ‘Australoid race’, because of their phenotypic resemblance. Today, genetics has proven that the Irula are genetically the closest South Indian population to the AASI and have no Australian genome (Yelmen, 2019). As Witzel (1999) points out, more research on the pre-Dravidian substrate in Tamil, Telugu, etc. is needed. Besides Southworth’s Proto-Nilgiri, no other non-Dravidian language has been inferred as a language of the South Indian AASI. It does not seem impossible that two endangered language isolates in southern South Asia, i.e., Nihali (Nahali) and the Vedda language, now spoken by small tribal populations, may be relics or descendants of the original languages of the AASI people before they started speaking Dravidian.

7. ZAGROSIAN AND DRAVIDIAN FAMILY TREE

In southern India, the Southern Neolithic Archaeological Complex (SNAC) and other neolithic sites are from 2500 BCE onwards. The claim that the Iron Age started in Tamil country in 3345 BCE is almost certainly false (Krishnan, 2025). In his book, Southworth (2005: 49, 245) repeatedly states that SNAC (see figure 2) should be associated with Proto-Peninsular Dravidian. But in a later paper (Southworth, 2011) written after he started subscribing to the Proto-Zagrosian theory, Southworth admits that other languages may have been spoken in the SNAC area before they were completely replaced by PPD, perhaps sometime between 1800-1200 BCE, i.e., during phase III of the southern Neolithic age. So, the dates that Southworth (2005) suggests in his book for PPD and PSD are some 500 years too early, and his 2011 paper has more accurate dates. Southworth’s (2011) suggested timeframe (1800-1200 BCE) for PPD in the Southern Neolithic seems to be now confirmed by the new archaeogenetic evidence of Narasimhan et al. and other such papers, which show that the ASI and ANI populations formed between 2000-1000 BCE, and by archaeological evidence. The 1800-1200 BCE timeframe also agrees with Dravidian words appearing in the Rigveda, which was composed during this period in Northwest India.

Based on the new genetics studies and the earlier linguistic studies of several previously mentioned authors, the Zagrosian superfamily’s linguistic tree is shown in Figure 5, which merges the Proto-Zagrosian model of McAlpin and Southworth with the relevant part of the Dravidian family tree (Krishnamurti, 2003: 493). The language splits and dates from various references are given in Figure 5, and the 3900-year time-depth of PPD in Figure 5 broadly agrees with the Bayesian phylogenetic study of Kolipakam et al. (2018), which concludes that the Dravidian language family is 4.5 kya old.

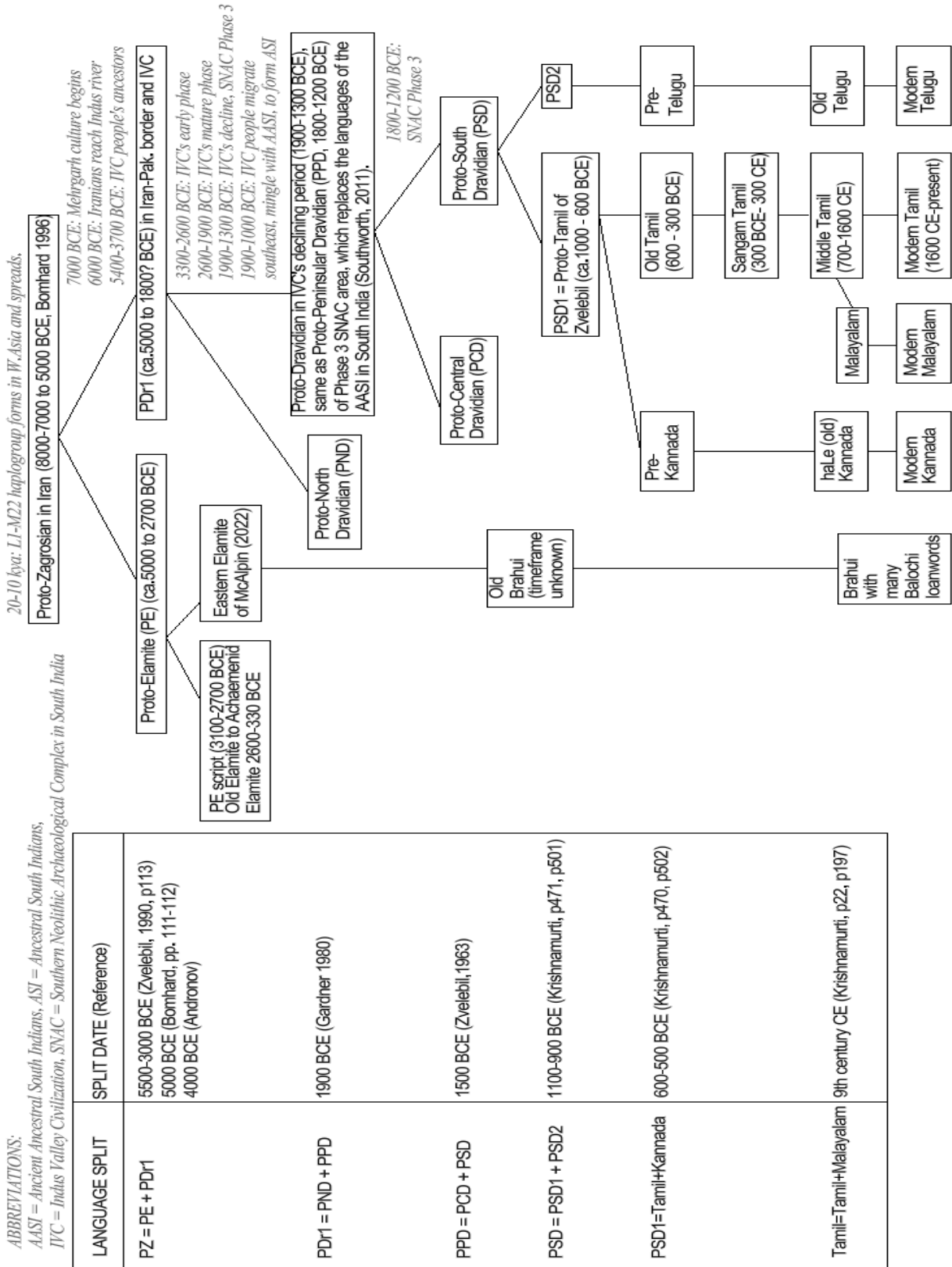


Figure 5: Zagrosian and Dravidian Family Tree with Dates of Language Splits

Pejros and Shnirelman (1988) used linguistic archaeology, i.e., the reconstructed PPD words (available in Southworth (2005b)) for some food crops, especially wheat and barley, and paleobotanical finds of these in archaeological sites to suggest that PPD originated in a dry deciduous forest region where wheat and barley were grown along with some rice, sorghum and African millet. The only archaeological region where these can all be found together is in the southeast corner of the IVC (see figure 2), where the post-IVC Kayatha and Ahar chalcolithic sites are, 1000 km north of SNAC. The authors claim that from this region, agricultural expansion occurred eastward and southward. This agrees with the finding of Narasimhan et al. (2019) that IVC people migrated in these directions before the formation of the ASI. So, it is likely that the PPD-speaking newcomers to Phase III of SNAC originated in Kayatha and Ahar, and agriculture may have played a role in the replacement by PPD of the original languages of the SNAC. Southworth (2005b) writes that PPD was spoken in eastern central India because the daughter languages of PND, PCD and PSD are now spoken around this area. His assertion is compatible with the papers of Pejros and Shnirelman, and Narasimhan et al., if we take eastern central India as the second home of PPD before its split into PCD and PSD in 1500 BCE (Zvelebil, 1963), as shown in Figure 5. Perhaps, a few PPD speakers migrated from Kayatha and Ahar to Punjab also, explaining the presence of PPD words in the middle layer of the Rigveda.

8. THE ORIGINS OF PROTO-ZAGROSIAN?

Proto-Zagrosian's possible origin in the Fertile Crescent near Anatolia is thought-provoking after one reads Heggarty et al. (2023), who describe their use of linguistics and genetics data to reconcile the Anatolian theory of the origin of Proto-Indo-European (PIE) with the rival Kurgan theory, which holds that PIE's original home was the Steppe. They suggest that PIE originated 8 kya in Anatolia, and its speakers migrated to a new secondary home in the Yamnaya Steppe and then over much of Eurasia.

PZ and PIE are members of the proposed Nostratic superfamily of languages, a concept which is widely rejected by historical linguists except long-rangers like Kerns, who wrote that the Fertile Crescent may have been the original home of Nostratic (Bomhard, 1996: 128–129). However, genetic studies reveal no Anatolian genome in modern Indians (Pathak et al. 2024), weakening this hypothesis.

Pagel et al. (2013) include Dravidian in their Eurasiatic superfamily along with PIE, but this paper attracted much criticism on linguistic grounds (Heggarty, 2013 and Thomason, 2013). Pagel et al. have Dravidian originating in Siberia, which is now contradicted by the new genetics papers mentioned earlier. The Uralo-Dravidian connection studied by Pudas-Marlowe (1974) in her dissertation cannot be dismissed, according to Southworth (2005: 49). Susumu Ohno tried to prove a relationship between Tamil and faraway Japanese. Both Zvelebil (1990: 116–118) and Krishnamurti (2005: 45) refuted his claim. History records that Tamils interacted with the Japanese for the first time only during the 20th century CE. Dravidian's possible connections with other languages or families have been investigated, but Austerlitz (Krishnamurti, 2005: 43) concludes that such studies, often done by amateurs, are unsystematic and wrong.

9. CONCLUSIONS

- 1) Because of the new archaeogenetics evidence, linguists should take a second look at the Proto-Elamo-Dravidian or Proto-Zagrosian family trees of McAlpin and Southworth, whose timelines, based on older linguistic studies, agree very well with the new genetic evidence.
- 2) There is more evidence supporting an origin in the Indus Valley Civilization (IVC) than for a more ancient origin within India, for Dravidian, which seems to be a daughter family of the Zagrosian superfamily of Elamite Iran, as shown in Figure 5. That figure summarizes the genetic, linguistic and archaeological evidence discussed in this paper.
- 3) PDr1 (the mother of Proto-Peninsular Dravidian (PPD)), Meluhhan, and Language X may have been the languages of the IVC. Since these are unreconstructed languages, deciphering the IVC seals would be very difficult, explaining why 70+ decipherment attempts have failed.
- 4) Proto-Peninsular Dravidian (PPD) emerged during the IVC's decline in the 2nd millennium BCE and was taken by migrating IVC people to South India, where it replaced the languages spoken by the Ancient Ancestral South Indians (AASI) who interbred with the newcomers. How and why this language replacement happened is not clear, but it is relevant that Sanskrit dialects replaced the Bronze Age languages of North India. Proto-Peninsular Dravidian (PPD) may have borrowed AASI words, and there may be a hitherto undiscovered substrate in today's Dravidian languages whose descent from PPD is shown in Figure 5.

ACKNOWLEDGEMENT

The author thanks the anonymous referee who gave him the list of additional references, which can be found below the next section, and recommended its inclusion.

REFERENCES

- Andronov, Mikhail, S. 1964. Lexicostatistic Analysis of the Chronology of Disintegration of Proto-Dravidian, *Indo-Iranian Journal*, 7, 2-3(1964), 170-186.
- Anonymous. 2025. Official Use of Tamil by Country 2025, *World Population Review*, article URL: <https://worldpopulationreview.com/country-rankings/official-use-of-tamil-by-country>
- Arunkumar, GaneshPrasad, et al. 2015. A late Neolithic expansion of Y chromosomal haplogroup O2a1-M95 from east to west, *Journal of Systematics and Evolution*, volume 53, issue 6, 546–560. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/jse.12147>
- Ayyangar, R. Raghava. 1952. *tamizhar varalaaru* (book in tamil), Thapovanam Press, 218–223.
- Basello, Gian Pietro. 2006. The Tablet from Konar Sandal B (Jiroft), URL: <http://www.elamit.net/elam/jiroft.pdf>
- Bellwood, P. 2013. *The Encyclopaedia of Global Human Migration*, John Wiley & Sons.
- Bennett, E. Andrew, et al. 2024. Reconstructing the Human Population History of East Asia through Ancient Genomics, article URL: <https://www.cambridge.org/core/elements/reconstructing-the-human-population-history-of-east-asia-through-ancient-genomics/0524D629660B5E43FC7094C043D54C6A>

- Bergstrom et al. 2016. Deep Roots for Aboriginal Australian Y Chromosomes, *Current Biology*, 26(6): 809-913, URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC4819516/>
- Blazek, Vaclav. 1999. Elam: a bridge between Ancient Near East and Dravidian India?, article in *Archaeology and Language IV*, Edited by Roger Blench and Matthew Spriggs, URL: <https://archive.org/details/ElamABridgeBetweenAncientNearEastAndDravidianIndia1999/page/n1/mode/2up?view=theater>
- Bomhard, Allan R. 1996. *Indo-European and the Nostratic Hypothesis*, Signum, p111. URL: <https://archive.org/details/BomhardIndoEuropeanAndTheNostraticHypothesis1996>
- Caldwell, Robert, 1913. *A Comparative Grammar of the Dravidian*, Kegan Paul, Trench, Trubner.
- Cardona, George. 2025. Indo-Aryan Languages, *Encyclopaedia Britannica* article, URL: <https://www.britannica.com/topic/Indo-Aryan-languages>
- Charuchandra, Sukanya. 2019. Experts question the dates of script in Tamil Nadu's Keeladi Excavation Report, *The Wire*, science article, URL: <https://science.thewire.in/society/history/keeladi-settlement-tamil-nadu-department-of-archaeology-tamil-brahmi-script-indus-valley-civilisation/>
- Clottes, Jean. 2025. Exploring the Rock Art of Central India, Bradshaw Foundation article URL: https://www.bradshawfoundation.com/india/central_india/index.php
- Culotta, Elizabeth. 2016. Almost all living people outside Africa trace back to a single migration more than 50,000 years ago, *Science*, article URL: <https://www.science.org/content/article/almost-all-living-people-outside-africa-trace-back-single-migration-more-50000-years>
- Debortoli, Guilherme. 2020. Novel insights on demographic history of tribal and caste groups, *Nature Scientific Reports*, (2020) 10:10075, URL: <https://pubmed.ncbi.nlm.nih.gov/32572090/>
- Dixon, Robert M.W. 1980. *The Languages of Australia*, Cambridge University Press, URL: <https://archive.org/search?query=title%3A%28the%20languages%20of%20australia%29%20AND%20creator%3A%28dixon%29>
- Dixon, Robert M.W. 1997. *The rise and fall of languages*, Cambridge University Press, 36–37, p91, URL: <https://dokumen.pub/the-rise-and-fall-of-languages-9780521626545.html>
- Farmer, Steve; Richard Sproat, and Michael Witzel. 2004. The Collapse of the Indus-Script Thesis, *Electronic Journal of Vedic Studies (EJVS)* 11-2 (2004), 19–57. URL: <https://safarmer.com/fsw2.pdf>
- Farmer, Steve; Richard Sproat, and Michael Witzel. 2009. A Refutation of the Claimed Refutation of the Nonlinguistic Nature of Indus Symbols, URL: <https://safarmer.com/Refutation3.pdf>
- Fleming, Harold et al. 2013. The Early Dispersions of Homo Sapiens sapiens and proto-Human from Africa." In *Mother Tongue*, issue XVIII, 143–188.
- Fuls, Andreas. 2024. Comparison of Linear Elamite and Indus Writing Systems, *Iranian Journal of Archaeological Studies*, 2024, Vol. 14, No. 1. URL: https://ijas.usb.ac.ir/article_8463_6b02e06a8bea25c88b2a2d228fb1a4af.pdf
- Gardner, P. 1980. Lexicostatistics and Dravidian Differentiation in situ. *Indian Linguistics* 41, 1980.
- Handwerk, Brian. 2021. An Evolutionary Timeline of Homo sapiens, *Smithsonian Magazine*, article URL: <https://www.smithsonianmag.com/science-nature/essential-timeline-understanding-evolution-homo-sapiens-180976807/>
- Harvey, Mark and Mailhammer, Robert. 2024. *Proto-Australian: Reconstruction of a Common Ancestor Language*, De Gruyter Mouton, Summary Article by the authors is in the URL:

- <https://theconversation.com/new-evidence-confirms-our-indigenous-languages-have-a-common-source-but-how-they-spread-remains-a-mystery-242576>
- Heggarty, Paul. 2013. Ultraconserved Words and Eurasiatic?, *PNAS*, 110(35) e3254. URL: <https://www.pnas.org/doi/10.1073/pnas.1309114110>
- Heggarty, Paul, et al. 2023. Language trees with sampled ancestors, *Science*, vol. 381, issue 6656, URL: <https://www.science.org/doi/10.1126/science.abg0818>
- Heras, Henry. 1953. *Studies in Proto-Indo-Mediterranean Culture*, book published by Indian Historical Research Institute, 29–61. URL: <https://archive.org/details/in.ernet.dli.2015.84973/mode/2up?q=fairservis>
- Hunter, G.R. 1934. The Script of Harappa and Mohenjodaro, London: Kegan Paul, Trench, Trubner & Co., 203–210. URL: <https://archive.org/details/in.ernet.dli.2015.283741/mode/2up>
- Kak, Subhash. 1988. A Frequency Analysis of the Indus Script, *Cryptologia*, July 1988, volume XII Number 3, 129–143.
- Kolipakam V, et al. 2018. A Bayesian phylogenetic study of the Dravidian language family. *Royal Society Open Science* 5: 171504. URL: <http://dx.doi.org/10.1098/rsos.171504>
- Krishnamurti, Bhadriraju. 2003. *The Dravidian Languages*, Cambridge University Press, URL: https://tamilnavarasam.in/books/others/the_dravidian_languages.pdf
- Krishnan, P.A. 2025. Tamil Nadu CM's Iron Age Claims, *Outlook India* (magazine), URL: <https://www.outlookindia.com/national/tamil-nadu-cms-iron-age-claims-omitting-inconvenient-questions>
- Mahadevan, Iravatham. 1982. S. R. Rao's Decipherment of the Indus Script, *The Indian Historical Review*, Volume VIII, Number 1-2, 58–73. URL: https://rmrl.in/en/dl/research-papers/mahadevan/ebook?id=S.R.%20Roas%20decipherment%20of%20Indus%20script_1982b
- Mahadevan, Iravatham. 1998. The Complete Interview by Harappa.com, URL: <https://www.harappa.com/content/iravatham-mahadevan-complete-interview>
- Mahadevan, Iravatham. 2002. Aryan or Dravidian or Neither?, *Electronic Journal of Vedic Studies (EJVS)* 8-1 (2002) 1–19. URL: <https://hasp.ub.uni-heidelberg.de/journals/ejvs/article/view/833/920>
- Masica, Colin. 1979. Aryan and Non-Aryan Elements in North Indian Agriculture, chapter in *Aryan and Non-Aryan in India*, a book edited by Madhava N. Deshpande and Peter Edwin Hook, Center for South and Southeast Asian Studies, Ann Arbor, 55–151.
- McAlpin, David. 1981. Proto-Elamo-Dravidian: The Evidence and its Implications, *Transactions of the American Philosophical Society*, Volume 71, Part 3. URL: https://www.researchgate.net/publication/270069377_Proto-Elamo-Dravidian_The_Evidence_and_Its_Implications
- McAlpin, David. 2015. Brahui and the Zagrosian Hypothesis, *Journal of the American Oriental Society*, 135(3):551. URL: https://www.researchgate.net/publication/285628453_Brahui_and_the_Zagrosian_Hypothesis
- McAlpin, David. 2022. Modern Colloquial Eastern Elamite, *Al-Burz*, 2022, 14(1), 64–123. ISSN (online) 2521-408X. URL: <http://alburz.uob.edu.pk/journal/index.php/alburz/index>
- Mishra, Udit. 2024. How should South India deal with its ageing population, *Indian Express* newspaper, article, URL: <https://indianexpress.com/article/explained/explained-economics/how-should-south-india-deal-with-its-ageing-population-9632112/>
- Narasimhan, Vagheesh et al. 2019. The Formation of Human Populations in South and Central Asia, *Science*, Volume 365, Issue 6457, URL: <https://www.science.org/doi/10.1126/science.aat7487>

- O'Grady, Geoff and Hale, Ken. 2004. The Coherence and Distinctiveness of the Pama-Nyungan Family, article in *Australian Languages*, a book by Claire Bowern and Harold Koch, John Benjamins Publishing. URL: https://books.google.ca/books?id=jSU6AAAAQBAJ&pg=PA69&redir_esc=y#v=onepage&q&f=false
- Pagani, Luca et al. 2017. An Ethnolinguistic and Genetic Perspective on the Origins of the Dravidian-Speaking Brahui in Pakistan, *Man India*, 2017;97(1):267-278. URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC5378296/>
- Pagel, Mark, et al. 2013. Ultraconserved words point to deep language ancestry across Eurasia, *PNAS*, 110(21) 8471-8476. URL: <https://www.pnas.org/doi/10.1073/pnas.1218726110>
- Parpola, Asko. 1994. *Deciphering the Indus script*, Cambridge University Press, p59. URL: <https://sindhilanguagelibrary.com/assets/documents/pdf/647f088f1fff3.pdf>
- Parpola, Asko. 2005. Study of the Indus Script, *Paper read at the 50th ICES Tokyo Session*, 19 May 2005, URL: <https://www.harappa.com/sites/default/files/pdf/indusscript%282%29.pdf>
- Parpola, Asko. 2008. Is the Indus Script indeed not a writing system?, *Airavati, Felicitation volume in honour of Iravatham Mahadevan*, varalaaru.com publishers, URL: <https://www.harappa.com/sites/default/files/pdf/indus-writing.pdf>
- Pathak, Ajai Kumar, et al. 2024. Human Y chromosome haplogroup L1-M22 traces Neolithic expansion in West Asia and supports the Elamite and Dravidian connection, *iScience*, Volume 27, Issue 6, 21 June 2024, 110016, URL: <https://www.sciencedirect.com/science/article/pii/S2589004224012410#bib30>
- Pejros, Ilja and Shnirelman, Viktor. 1988. "In Search of the Dravidian Homeland", paper presented in the first *International Interdisciplinary Symposium on Language and Prehistory*, Ann Arbor, 8-12 November, 1988. A 2-page version of this paper is in *Reconstructing Languages and Cultures*, edited by V.V. Shevoroshkin, Bochum: N. Brockmeyer, 1989, 70–71, a reference that is easier to find.
- Pudas-Marlow, Elli Johanna. 1974. *More on the Uralo-Dravidian Relationship*, Dissertation, University of Texas at Austin, January 1974. URL: <https://archive.org/details/PudasMarlowMoreOnTheUraloDravidianRelationship>
- Ramaswamy, Sumathi. 2004. *The Lost Land of Lemuria*, University of California Press.
- Rao, Rajesh P.N. et al. 2009. Entropic Evidence for Linguistic Structure in the Indus Script, *Science* 324(5931):1165, URL: <https://www.science.org/cms/asset/f142c776-4f40-4b10-9d09-a4bd790a628a/pap.pdf>
- Reich, David et al. 2009. Reconstructing Indian Population History, *Nature*, 2009, Sept 24; 461(7623):489-494, URL: <https://www.science.org/doi/10.1126/science.aat7487>
- Sankalia, H. D. 1949. Studies in the historical and cultural geography and ethnography of Gujarat (places and peoples in inscriptions of Gujarat: 300 B.C. - 1300 A.D.) Poona: Deccan, p53.
- Sastri, M. Seshagiri. 1897. *Essay on Tamil Literature*, S.P.C.K Press, Madras, India.
- Shahane, Girish. 2019. Why Hindutva supporters love to hate the discredited Aryan Invasion Theory, Article in *Scroll.in*, a digital news service, <https://scroll.in/article/937043/why-hindutva-supporters-love-to-hate-the-discredited-aryan-invasion-theory>
- Southworth, Franklin. 2005. *Linguistic Archaeology of South Asia*, RoutledgeCurzon. URL: https://archive.org/details/linguisticarchaeologyofsouthasiafranklinsouthworthc.routledge_599_K

- Southworth, Franklin. 2005b. Proto-Dravidian Agriculture, 7th ESCA Round Table Conference, Kyoto, June 2005. URL: https://www.sas.upenn.edu/~fsouth/from_ccat/Proto-DravidianAgriculture.pdf
- Southworth, Franklin, 2011. *Rice* (2011) 4:142–148, DOI 10.1007/s12284-011-9076-9
- Starostin, George, 2002. On the genetic affiliation of the Elamite language, *Mother Tongue* (ISSN: 1087-0326), vol. VII, 2002: p147. URL: <https://starlingdb.org/Texts/elam.pdf>
- Steever, Sanford B. 2017. The Dravidian Language Family, Article in Linguistic Typology, Cambridge University Press, p887. URL: <https://www.cambridge.org/core/books/abs/cambridge-handbook-of-linguistic-typology/dravidian-language-family/7E945B94A2692E0CBD1EDC5BF593E60C>
- Steever, Sanford. 1998. Introduction to the Dravidian Languages, chapter in *The Dravidian Languages*, edited by Sanford B. Steever. URL: <https://www.cambridge.org/core/books/abs/cambridge-handbook-of-linguistic-typology/dravidian-language-family/7E945B94A2692E0CBD1EDC5BF593E60C>
- Stolper, Matthew Wolfgang. 2004. *The Cambridge Encyclopaedia of The World's Ancient Languages*, Cambridge University Press, 60–61.
- Tatte, Kai, et al. 2019. The genetic legacy of continental scale admixture, *Nature Scientific Reports*, (2019) 9:3818. URL: <https://www.nature.com/articles/s41598-019-40399-8>
- Thapar, Romila. 2025. Developments in the Ganges Basin, *Encyclopaedia Britannica*, article URL: <https://www.britannica.com/place/India/Developments-in-the-Ganges-basin#ref484967>
- Thomason, Sally. 2013. Ultraconserved words? Really?, *Language Log*, May 8, 2013. URL: <https://languagelog.ldc.upenn.edu/nll/?p=4612>
- Van Driem, George. 1999. On the Austroasiatic Indus Theory, *Mother Tongue Journal*, Special Issue, October 1999. URL: <https://archive.org/details/MotherTongueSpecialIssueOctober1999>
- Van Driem, George. 2001. *Languages of the Himalayas*, Brill, p298, URL: https://books.google.ca/books?id=fiavPYCz4dYC&pg=PA298&redir_esc=y#v=onepage&q&f=false
- Van Driem, George. 2023. The Father Tongues L, R and P, *Mother Tongue Journal*, XXIV, 45–56.
- Witzel, Michael. 1999. Early Sources for South Asian Substrate Languages, *Mother Tongue Journal*, Special Issue, October 1999, 1–70. URL: <https://archive.org/details/MotherTongueSpecialIssueOctober1999>
- Witzel, Michael. 2019. Early ‘Aryans’ and their neighbors outside and inside India, *Journal of Biosciences*, (2019)44:58, published by the Indian Academy of Sciences.
- Witzel, Michael. 2009. The linguistic history of some Indian domestic plants. *Journal of BioSciences* 34(6): 829-833. URL: <https://dash.harvard.edu/server/api/core/bitstreams/7312037c-99bf-6bd4-e053-0100007fdf3b/content>
- Wright, Rita. 2010. *The Ancient Indus: Urbanism, Economy and Society*, Cambridge University Press, 1, 43, 115–125, 308–314.
- Yang M. 2022. A genetic history of migration, diversification, and admixture in Asia. *Human Population Genetics and Genomics*, 2022; 2(1):0001. URL: <https://doi.org/10.47248/hpgg2202010001>

- Yelmen, Burak. 2019. Ancestry-Specific Analyses Reveal Differential Demographic Histories, *Molecular Biology and Evolution*, Volume 36, Issue 8, August 2019, Pages 1628–1642, URL: <https://doi.org/10.1093/molbev/msz037>
- Zhang, X, et al. 2015. Y-chromosome diversity suggests southern origin and Paleolithic backwave migration, *Science Reports*, 2015, Oct 20; 5:15486. URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC4611482/>
- Zvelebil, Kamil. 1990. *Dravidian Linguistics*, Pondicherry Institute of Linguistics and Culture, p113. URL: <https://archive.org/details/zvelebildravidianlinguisticsanintroduction1990>
- Zvelebil, Kamil. 1992. *Companion Studies to the History of Tamil Literature*, E. J. Brill, Leiden, p246.

ADDITIONAL REFERENCES

The additional references given below by a referee are very useful to Dravidian scholars. The paragraph below classifies the bibliography given in this section, into various topics.

- Dravidian language prehistory:** Starostin 1997, 1998, 2000a, 2000b. (See the bibliography below for more detailed citations)
- Dravidian-Uralic:** Andronov 1961, 1968, 1969, 1971; Austerlitz 1971; Bouda 1953, 1956; Burrow 1943, 1968; Eronen 1977; Hevesy 1932; Larsson 1982; Marlow 1980; Schrader 1924, 1936; Sebeok 1945; Tuttle n.d.; Tyler 1968, 1990.
- Dravidian-Altaic:** Bouda 1953, 1956; Fujiwara 1975, 1981; Hulbert 1905; Meile 1949; Menges 1964, 1969, 1975, 1989; Miller 1981, 1983, 1986; Ohno 1970, 1980, 1981, 1983a, 1983b, 2001; Vacek 1978-2013c; Zvelebil 1985b, 1991a, 1991b.
- Dravidian-Kartvelian:** Fähnrich 1965, 1991, 2016.
- Dravidian-Elamite:** Blažek 2002a; Zvelebil 1974, 1985a.
- Dravidian-Afroasiatic:** Blažek 2002a, 2002b; Homburger 1955, 1957; Shamsuddin 1973-1975.
- Dravidian** as a member of the **Nostratic** macrofamily: Blažek 2003, 2007, 2009; Bomhard 2023; Dolgopolsky 2012; Illič-Svityč 1971-1984.
- Pre-Dravidian** substratum with **Australian** affinities: Blažek 2006.
- Proto-Indus script:** Al'bedil' 1986; Knorozov 1981; Knorozov, Al'bedil', Volčok 1981; Vacek 1970; Zide & Zvelebil 1970, 1976; Zvelebil 1973, 1985c.

BIBLIOGRAPHY

- Aalto, Pentti. 1972. The Alleged Affinity of Dravidian and Fenno-Ugrian. In: *Proceedings of the II. International Conference-Seminar of Tamil Studies, 1968*. Madras: International Association of Tamil Research, 262-266.
- Al'bedil', M.F. 1986. Tipy protoindijskix nadpisej. In: *Drevnie sistemy pis'ma – étničeskaja semiotika*, ed. by Ju.V. Knorozov. Moskva: Nauka, 36-68.
- Andronov, Mikhail, S. 1961. New Evidence of Possible Linguistic Ties Between the Deccan and the Urals. In: *Dr. R.P. Sethu Pillai Silver Jubilee Commemoration Volume*. Madras: University of Madras, 137-140.
- Andronov, Mikhail, S. 1968. Dravidian and Uralian: A peep into the prehistory of language families. In: *Two Lectures on the Historicity of Language Families*. Annamalainagar: Annamalai University, 14-35.

- Andronov, Mixail, S. 1969. O xaraktere dekkano-ural'skix analogij. [On the Nature of Deccano-Uralic Analogies]. In: *Jazykovye universalii i lingvističeskaja tipologija*. Moskva: Nauka, 308-321.
- Andronov, Mikhail, S. 1971. Comparative studies on the nature of Dravido-Uralian parallels (A peep into the pre-history of language families). In: *Proceedings of the Second International Conference-Seminar of Tamil Studies*, vol. 1, ed. by Ronald E. Asher. Madras: Christian Literature Society, 267-77.
- Austerlitz, Robert. 1971. Long-range comparisons of Tamil and Dravidian with other language families in Eurasia. In: *Proceedings of the Second International Conference-Seminar of Tamil Studies*, vol. 1, ed. by Ronald E. Asher. Madras: Christian Literature Society. 254-261.
- Blažek, Václav. 2002a. Elam: a Bridge between the Ancient Near East and Dravidian India? *Mother Tongue* 7, 123-146.
- Blažek, Václav. 2002b. Some New Dravidian - Afroasiatic Parallels. *Mother Tongue* 7, 171-199.
- Blažek, Václav. 2003. Lexica Nostratica Addenda. *Mother Tongue* 8, 11-22.
- Blažek, Václav. 2006. Was there an Australian Substratum in Dravidian? *Mother Tongue* 11, 275-294.
- Blažek, Václav. 2007. Dravidian Numerals. *Mother Tongue* 12, 203-215.
- Blažek, Václav. 2009. Dravidian Numerals. *Journal of Language Relationship* 1, 69-80 (revisited and enlarged version of Blažek 2007)
- Blažek, Václav. 2016. Hydronymia Ṛgvedica. *Linguistica Brunensia* 64/2, 7-54.
- Bomhard, Allan R. 2023. *A Comprehensive Introduction to Nostratic Comparative Linguistics*. Florence (SC).
- Bouda, Karl. 1953. Dravidisch und Uralaltaisch. *Ural-Altäische Jahrbücher* 25, 161-173.
- Bouda, Karl. 1956. Dravidian and Uralaltaic. *Lingua* 5, 129-144.
- Burrow, Thomas. 1943. Dravidian Studies IV: The body in Dravidian and Uralian. *Bulletin of the School of Oriental and African Studies* 11, 328-356.
- Burrow, Thomas. 1968. *Collected Papers on Dravidian Linguistics*. Annamalainagar: Annamalai University.
- Dolgopolsky, Aharon. 2012. *Nostratic Dictionary*³. Cambridge: The McDonald Institute for Archaeological Research.
- Eronen, Jarmo. 1977. Uralo-dravidalaisista yhteyksistä. [On Uralic-Dravidian correspondences] In: *Suomen antropologi* 1977/1, 33-38. (Helsinki).
- Fane, Hannah. 1980. Sumerian-Dravidian Interconnections: the Linguistic, Archeological and Textual Evidence. *Journal of Dravidian Studies* 9, 286-305.
- Fähnrich, Heinz. 1965. Iberokaukasisch und Drawidisch. *Bedi Kartlisa / Revue de Kartvilologie* 19-20, 136-158.
- Fähnrich, Heinz. 1991. Kartwelisch-drawidische Sprachparallele. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 44, Nr. 3, 339-343.
- Fähnrich, Heinz. 2016. *Die Kartwelier. Grundsprache – Kultur – Lebensraum*. Wiesbaden: Reichert.
- Fujiwara, Akira. 1975. The body in Japanese and Dravidian. *Gengo Kenkyu* 68, 119-121.
- Fujiwara, Akira. 1981. The Japanese-Dravidian Vocabulary of Flora and Fauna. *Bulletin of the International Institute for Linguistic Sciences* 2/4, 73-97. (Kyoto Sangyo University)
- Hevesy, Wilhelm von. 1932. *Finnisch-Ugrisches aus Indien*. Wien: Manzsche Verlags- und. Universitäts-Buchhandlung.
- Homburger, Liliás. 1955. L'Inde et l'Afrique. *Journal de la Société des africanistes* 25, 13-18.

- Homburger, Liliás. 1957. De quelques éléments communs à l'égyptien et aux langues dravidiennes. *Kêmi* 14, 26-33.
- Hulbert, Homer B. 1905. *A comparative grammar of the Korean language and the Dravidian languages of South India*. Seoul.
- Illič-Svityč, Vladislav M. 1971-1984. *Opyt sravnenija nostratičeskix jazykov*, I-III. Moskva: Nauka.
- Knorozov, Jurij V. 1981. Protoindijskie nadpisi (k probleme dešifrovki). *Sovetskaja étnografija* 1981, n. 5, 47-71.
- Knorozov, Jurij V., Al'bedil', M.F., Volčok, B.J. 1981. *Proto-Indica 1979: Report on the investigation of the Proto-Indian texts*. Moskva: Nauka.
- Larsson, Lars-Gunnar. 1982. Some Remarks on the Hypothesis of an Uralo-Dravidian Genetic Linguistic Relationship. *Fenno-Ugrica Suecana* 5, 169-184.
- Lee, Ki-Moon. 1973. Linguistic Ties Between Korea and Japan. *Korea Journal* 1973, 37-42.
- Marlow, Elli Johanna Pudas. 1980. Uralic-Dravidian Hypothesis: What Evidence? In: *Congressus Internationalis Fenno-Ugristarum* 5. Pars 2. Turku: Suomen Kielen Saura, 245-250.
- Meile, Pierre. 1949. Observations sur quelques caracteres communs des langues dravidiennes et des langues altaïques. *Actes du XXI Congrès International des Orientalistes*. Paris: Imprimerie Nationale, 207-209.
- Menges, Karl H. 1964. Altajisch und Drävidisch. *Orbis* 13, 66-103.
- Menges, Karl H. 1969. The Dravido-Altaic Relationship. *Journal of Tamil Studies* 1, 35-39.
- Menges, Karl H. 1975. Dravidian and Altaic. *Central Asiatic Journal* 19, 202-205.
- Menges, Karl H. 1977. Dravidian and Altaic. *Anthropos* 72, 129-179.
- Menges, Karl H. 1989. East-Nostratic: Altaic and Dravidian. In: *Reconstructing Languages and Cultures. Abstracts and Materials from the First International Interdisciplinary Symposium on Language and Prehistory* (Ann Arbor, 1988), ed. by Vitaly V. Shevoroshkin. Bochum: Universitätsverlag Brockmeyer (Bochum Publications in Evolutionary Cultural Semiotics Vol. 20.), 59-62.
- Miller, Roy A. 1981. Review of S. Ohno's *Nihongo to Tamirugo*. *Language* 59, 207-211.
- Miller, Roy A. 1983. Tamil and Japanese? *Bulletin of the School of Oriental and African Studies* 12/2, 366-396.
- Miller, Roy A. 1986. Tamil and Japanese? *Bulletin of the School of Oriental and African Studies* 49/3, 557-560.
- Ohno, Susumo. 1970. *The Origin of the Japanese Language*. Tokyo: Japan Cultural Society.
- Ohno, Susumo. 1980. *Sound Correspondences Between Tamil and Japanese*. Tokyo: Gakushuin University.
- Ohno, Susumo. 1981. *Nihongo to tamirugo*. [Japanese and Tamil] Tokyo: Shinchosha.
- Ohno, Susumo. 1983a. A Study of the Relationship between Tamil and Japanese. *Journal of Dravidian Studies* 12, 366-396.
- Ohno, Susumo. 1983b. The Loss of Initial *c* in Tamil and *s* in Japanese. *Uy arayvu* 1: 16-21.
- Ohno, Susumo. 2001. *Sound correspondences between Tamil and Japanese*. Chennai: International Institute of Tamil Studies.
- Schrader, Otto. 1924. Dravidisch und Uralisch. *Zeitschrift für Indologie und Iranistik* 3, 81-112.
- Schrader, Otto. 1936. On the "Uralian" Element in the Drāviḍa and the Muṇḍā Languages *Bulletin of the School of Oriental Studies* 8, No. 2/3, 751-762.
- Sebeok, Thomas A. 1945. Finno-Ugric and the languages of India. *Journal of American Oriental Society* 65, 59-62.

- Shamsuddin, K.O. 1973-1975. Origin of the Dravidian and Semitic languages. In: *Proceedings of the First All India Conference of Dravidian Linguists* (Trivandrum, June 3-4 1971), Vol. I, ed. by V.E. Subramoniam, E. Valentine. Trivandrum, 3-12.
- Starostin, George. 1997. On the reconstruction of velar phonemes in Proto-Dravidian. In: *Studia linguarum*, 1. Moskva: Izdatel'stvo RGGU, 190-212.
- Starostin, George. 1998. Alveolar Consonants in Proto-Dravidian: One or More? In: *Proceedings of the International Conference on South Asian Languages* (July 1-4, 1997). Moscow, 183-194.
- Starostin, George. 2000a. Dravidian roots with initial *j- and possible Indo-European cognates. In: *Problemy izučeniya dal'nego rodstva jazykov na rubeže tretjego tysjačiletija. Doklady i tezisny meždunarodnoj konferencii*. Moskva: Izdatel'stvo RGGU, 219-221.
- Starostin, Georgij S. 2000b. *Rekonstrukcija fonologičeskoj sistemy pradravidijskogo jazyka*. Moskva: Ph.D. Thesis.
- Tuttle, E.H. n.d. *Finnic and Dravidian*. New Haven.
- Tyler, Stephen A. 1968. Dravidian and Uralian: the Lexical Evidence. *Language* 44, 798-812.
- Tyler, Stephen A. 1990. Summary of Noun and Verb Inflectional Correspondences in Proto-Dravidian and Proto-Uralic. In: *Proto-Languages and Proto-Cultures. Materials from the First International Interdisciplinary Symposium on Language and Prehistory* (Ann Arbor, 1988), ed. by Vitaly V. Shevoroshkin. Bochum: Universitätsverlag Brockmeyer (Bochum Publications in Evolutionary Cultural Semiotics Vol. 25), 68-76.
- Vacek, Jaroslav. 1970. The Problem of the Indus Script. *Archiv Orientální* 38, 198-212.
- Vacek, Jaroslav. 1978. The Problem of the Genetic Relationship of the Mongolian and Dravidian Languages. *Archiv Orientální* 46, 141-151.
- Vacek, Jaroslav. 1981. The Dravido-Altaic Relationship - Lexical and Sound Correspondences. In: *Proceedings of the 5th International Conference-Seminar of Tamil Studies*, I. Madras, 159-170.
- Vacek, Jaroslav. 1983. Dravido-Altaic: The Mongolian and Dravidian Verbal Bases. *Journal of Tamil Studies* 23, 1-17.
- Vacek, Jaroslav. 1987. The Dravido-Altaic Relationship. *Archiv Orientální* 55, 134-149.
- Vacek, Jaroslav. 2002a. Emphasizing and Interrogative Enclitic Particles in Dravidian and Altaic. In: *MONGOLICA PRAGENSIA '02, Ethnolinguistics and Sociolinguistics in Synchrony and Diachrony*. Edited by Jaroslav Vacek and Alena Oberfalzerova. Prague: Triton, 151-185.
- Vacek, Jaroslav. 2002b. Dravidian and Altaic "Deer – Antelope – Bull". *Indologica Taurinensia* XVIII, 251-266.
- Vacek, Jaroslav. 2002c. *Flowers and Formulas. Nature as Symbolic Code in Old Tamil Love Poetry*. Prague: Charles University (Studia Orientalia Pragensia XXV).
- Vacek, Jaroslav. 2003. *Verba dicendi* and related etyma in Dravidian and Altaic (Etyma with initial sibilants of affricates). In: *MONGOLICA PRAGENSIA '03, Ethnolinguistics and Sociolinguistics in Synchrony and Diachrony*. Edited by Jaroslav Vacek and Alena Oberfalzerova. Prague, Triton, 175-205.
- Vacek, Jaroslav. 2004a. Dravidian and Altaic "Sheep – Deer – Cattle". In: *South-Indian Horizons (Felicitation Volume for François Gros on the occasion of his 70th birthday)*. Edited by Jean-Luc Chevillard (Editor) and Eva Wilden (Associate Editor) with the collaboration of A. Murugaiyan. Paris: Institut Français de Pondichéry - École Française d'Extrême-Orient, 333-345.

- Vacek, Jaroslav. 2004b. Dravidian and Altaic – In Search of a New Paradigm. *Archiv Orientální*, Vol. 72, 384–453. Reprinted in: *International Journal of Dravidian Linguistics* (Trivandrum, India), Vol. XXXV, No.1, January 2006, 29–96.
- Vacek, Jaroslav. 2004c. *Verba dicendi* and related etyma in Dravidian and Altaic 2.1. Etyma with initial velars. In: *MONGOLICA PRAGENSIA '04, Ethnolinguistics and Sociolinguistics in Synchrony and Diachrony*. Edited by Jaroslav Vacek and Alena Oberfalzerová. Prague: Triton, 195–230.
- Vacek, Jaroslav. 2005a. Dravidian and Altaic ‘to break, to beat, to crush, to rub’ (Ta. *kavi* – Mo. *xabir-* / *xayar-* – MT. *KAPU-* / *KAVBL-* / *KAIḌTARĀ-* – OT. *kam-* / *kak-*). *Archiv Orientální*, Vol. 73, 325–334.
- Vacek, Jaroslav. 2005b. *Verba dicendi* and related etyma in Dravidian and Altaic 2.2. Etyma with initial velars and final liquids. In: *MONGOLICA PRAGENSIA '05, Ethnolinguistics and Sociolinguistics in Synchrony and Diachrony*. Edited by J. Vacek and A. Oberfalzerova. Prague: Triton, 171–197.
- Vacek, Jaroslav. 2005c. Dravidian and Altaic parts of the body 1. Breast. In: *Acta Mongolica*, Vol. 5 (246). Dedicated to the 100th Birthday of Professor Yo'nsiebu' Biyambyn Rinc'en (1905–1977). Ulaanbatar 2005, 65–70.
- Vacek, Jaroslav. 2006a. Dravidian and Altaic – In search of a New Paradigm. *International Journal of Dravidian Linguistics* (Trivandrum, India), Vol. XXXV, No.1, 29–96. (Reprint of Vacek 2004b)
- Vacek, Jaroslav. 2006b. Dravidian and Altaic ‘to bow, bend, stoop, incline, curve’. 1. (C)VC- roots with initial labials and medial velars or labials. *Archiv Orientální*, Vol. 74, 183–202.
- Vacek, Jaroslav. 2006c. Dravidian and Altaic parts of the body 2. Hair, feather. In: *Acta Mongolica*, Vol. 6 (267). Dedicated to the 90th Birthday of professor Denis Sinor. Edited by Ts. Shagdarsuren. Ulaanbaatar: Centre for Mongol Studies, National University of Mongolia, 79–82.
- Vacek, Jaroslav. 2006d. *Verba dicendi* and related etyma in Dravidian and Altaic 3.1. Etyma with initial labials (*p-*, *b-*, *f-*, *v-*, *m-*). In: *MONGOLICA PRAGENSIA '06, Ethnolinguistics and Sociolinguistics in Synchrony and Diachrony*. Edited by Jaroslav Vacek and Alena Oberfalzerová. Prague: Triton, 163–190.
- Vacek, Jaroslav. 2007a. Dravidian and Altaic parts of the body – 3. Heart, chest, inside, mind, thought. In: *Mongolian Studies*, Vol. 22, 2 (Seoul: The Korean Association for Mongolian Studies), 27–45.
- Vacek, Jaroslav. 2007b. Dravidian and Altaic to bend / to bow – elbow / knee / ankle. Initial dental stop or nasal, medial velar or labial. *Archiv Orientální* 75, 3, 395–410.
- Vacek, Jaroslav. 2007c. *Verba dicendi* and related etyma in Dravidian and Altaic 3.2. Etyma with initial labials (*p-*, *b-*, *f-*, *v-*, *m-*) and root-final liquids and retroflex stops. In: *MONGOLICA PRAGENSIA '07 sive Folia linguarum Orientis selecta (FLOS), Ethnolinguistics, Sociolinguistics and Culture*, Vol. 1. Edited by Jaroslav Vacek and Alena Oberfalzerová. Prague: Charles University and Triton, 85–111.
- Vacek, Jaroslav. 2008a. *Verba dicendi* and related etyma in Dravidian and Altaic 4.1. Etyma with initial dentals (*t-*, *d-*, *n-*). In: *MONGOLO-TIBETICA PRAGENSIA '08, Linguistics, Ethnolinguistics, Religion and Culture*. Vol. 1/1. Edited by Jaroslav Vacek and Alena Oberfalzerová. Prague: Charles University and Triton, 99–133.

- Vacek, Jaroslav. 2008b. Dravidian and Altaic parts of the body 4. Liver, spleen. In: *Acta Mongolica*, Vol. 8 (306). Dedicated to the 70th Birthday of Professor D. Tumurtogoo. Ulaanbaatar, 7–14.
- Vacek, Jaroslav. 2009a. Dravidian and Altaic – problems and solutions. In: *Oriental Studies. Proceedings of the International Conference of Oriental Studies. 55 years of the Committee of Oriental Studies of the Polish Academy of Sciences (1952-2007)*. Edited by Agata Bareja-Starzyńska and Marek Mejer. *Rocznik Orientalistyczny* 62, 1, 228–239.
- Vacek, Jaroslav. 2009b. Dravidian and Altaic ‘to bend’ – Ta. *kavi-* / *kavił-*, Mo. *kebüi-* / *köbüre-*, MT. *KAMPI-* / *KEBILE-*. In: *MONGOL STUDIES* (Research Papers), Volume XXX (318). Editors-in-chief D. Zayabaatar, M. Uuganbayar. Ulaanbaatar, School of Mongolian Language and Culture, National University of Mongolia, 157–168.
- Vacek, Jaroslav. 2009c. Dravidijskij i altajskij – jazyki drevnix intensivnyx kontaktov. (Dravidian and Altaic – ancient high contact languages). In: *Problemy mongolovednyx i altaističeskix issledovanij. Materialy meždunarodnoj konferencii, posvjaščennoj 70-letiju professora V.I. Rassadina*. Elista, 50–58.
- Vacek, Jaroslav. 2009d. Dravidian and Altaic ‘to wave / to tremble – arm / shoulder-(blade) / wing’. In: *Acta Mongolica*, Vol. 9 (320). Dedicated to the 80th Birthday of Professor Rachewiltz. Ulaanbaatar 2009, 7–14.
- Vacek, Jaroslav. 2009e. *Verba dicendi* and related etyma in Dravidian and Altaic 4.2. Etyma with initial dentals (*t-*, *d-*, *n-*) and root-final liquids and retroflex stops. In: *MONGOLO-TIBETICA PRAGENSIA '09, Linguistics, Ethnolinguistics, Religion and Culture*. Vol. 2/1, 2009. Edited by J. Vacek and A. Oberfalzerová. Charles University and Triton, Praha, 123–143.
- Vacek, Jaroslav. 2009f. Dravidian and Altaic – two layers in Dravidian due to ancient high contact? In: *MONGOLO-TIBETICA PRAGENSIA '09, Linguistics, Ethnolinguistics, Religion and Culture*. Vol. 2/2, 2009. Edited by J. Vacek and A. Oberfalzerová. Praha: Charles University and Triton, 77–109.
- Vacek, Jaroslav. 2010a. *Verba dicendi* and related etyma in Dravidian and Altaic 5.1. Etyma with initial vowels (*i-*, *e-*, *u-*, *o-*, *a-*) and root-final stops, affricates, nasals and glides. In: *MONGOLO-TIBETICA PRAGENSIA '10. Linguistics, Ethnolinguistics, Religion and Culture*. Vol. 3/1. Edited by Jaroslav Vacek and Alena Oberfalzerová. Praha: Charles University and Triton, 123–158.
- Vacek, Jaroslav. 2010c. Dravidian and Altaic parts of the body 5. ‘Extremities’ – head, finger, foot/leg. *Rocznik Orientalistyczny* 63, 1, 271–283.
- Vacek, Jaroslav. 2010d. Dravidian and Altaic – ancient high contact languages? *International Journal of Dravidian Linguistic* (International School of Dravidian Linguistics, University of Trivandrum), Vol. 39, No. 2, 117–138.
- Vacek, Jaroslav. 2011g. *Verba dicendi* and related etyma in Dravidian and Altaic 5.2. Etyma with initial vowels (*i-*, *e-*, *u-*, *o-*, *a-*) and root-final liquids. In: *MONGOLO-TIBETICA PRAGENSIA '11. Linguistics, Ethnolinguistics, Religion and Culture*. Vol. 4/1. Edited by Jaroslav Vacek and Alena Oberfalzerová. Praha: Charles University and Triton, 135–160.
- Vacek, Jaroslav. 2012. Dravidian and Altaic ‘fear, timidity, worry’ 1. In: *MONGOLO-TIBETICA PRAGENSIA '11. Linguistics, Ethnolinguistics, Religion and Culture*. Vol. 5/1. Edited by Jaroslav Vacek and Alena Oberfalzerová. Praha: Charles University and Triton, 91–108.
- Vacek, Jaroslav. 2013a. Dravidian and Altaic – implications of the established parallels. In: *Mongolian Studies*, Vol. 34 (Seoul, The Korean Association for Mongolian Studies), 209–234.

- Vacek, Jaroslav. 2013b. Dravidian and Altaic – a new beginning. In: *The X International Congress of Mongolists*. August 9–13, 2011, Ulaanbaatar. In: *Proceedings of Congress*. Олон улсын монголч эрдэмтдийн Х их хурлын илтгэлүүд, IV боть, Улаанбаатар, 2013, 205–209.
- Vacek, Jaroslav. 2013c. Three Dravidian and Altaic etymologies – ‘to shrink, contract; to leave, abandon; to grind, pound’. In: *A Window onto the Other. Contributions on the Study of the Mongolian, Turkic and Manchu-Tungusic Peoples, Languages and Cultures Dedicated to Jerzy Tulisow on the Occasion of His Seventieth Birthday*. Edited by Agata Bareja-Starzyńska, Jan Rogala and Filip Majkowski. Warszawa: Dom Wydawniczy Elipsa, Wydział Orientalistyczny Uniwersytetu Warszawskiego, 248–257.
- Zide, Arlene R.K. & Zvelebil, Kamil V. 1970. Review of Parpola et alii and Knorozov et alii. *Language* 46, 952–968.
- Zide, Arlene R.K. & Zvelebil, Kamil V. 1976. *The Soviet Decipherment of the Indus Valley Script: translation and Critique*. The Hague – Paris: Mouton (Janua Linguarum series Practica 156).
- Zvelebil, Kamil V. 1973. The so-called „Dravidian“ of the Indus Inscriptions. *Compte-rendu de la Troisième Conférence Internationale*. Pondichéry, 32–41.
- Zvelebil, Kamil V. 1974. Dravidian and Elamite – a real breakthrough? *Journal of American Oriental Society* 93, 384–385.
- Zvelebil, Kamil V. 1985a. Review of D.W. McAlpin 1981. *Journal of American Oriental Society* 105, 364–372.
- Zvelebil, Kamil V. 1985b. Tamil and Japanese – are they related? The hypothesis of Susumu Ohno. *Bulletin of the School of Oriental and African Studies* 48, n. 1, 116–120.
- Zvelebil, Kamil V. 1985c. Recent Attempts at the Decipherment of the Indus Valley Script and Languages (1965–1980): A Critique. In: *Indus Valley to Mekong Delta, Explorations in Epigraphy*, ed. by N. Karashima. Madras, 151–87.
- Zvelebil, Kamil V. 1991a. Long-Range Language Comparison in New Models of Language Development: The Case of Dravidian. *Journal of Dravidian Studies* 1/1, 21–31.
- Zvelebil, Kamil V. 1991b. *Long-Range Linguistic Comparisons: The Case of Dravidian*. Utrecht: Utrecht University.
- Zvelebil, Kamil V. 1991c. Dravidian and Japanese Once Again. *Archiv Orientální* 59, 73–77.

ETRUSCAN–HURRO-URARTIAN RELATIONS

VÁCLAV BLAŽEK

ABSTRACT

This paper reviews earlier comparisons of Etruscan with North Caucasian, Hurro-Urartian with North Caucasian, and Etruscan with Hurro-Urartian. After critical evaluation, the most promising correspondences are retained and supplemented with new comparanda, forming a synthesis that supports a genealogical Etruscan–Hurro-Urartian relationship. The comparative minicorpus includes 47 lexical items; morphology is excluded (reserved for future study). Though modest in size, this sample is representative given the limited corpus of Etruscan words with known or probable meaning and the divergent semantic profiles of Hurrian and Urartian.

1. MOTIVATION

This study evaluates the hypothesis of an Anatolian origin for Etruscan and its relatives via a genetic link with Hurro-Urartian, as an alternative to proposed ties with Indo-European languages of Anatolia (e.g., Adrados 1989, 1994, 2005; Georgiev 1962, 1974, 1979; Woudhuizen 2019; critically, Neu 1991; Simon 2021).¹

2. HISTORY OF RESEARCH

The Etruscan language is the best-documented member of the Tyrsenian language family, which also includes two other epigraphic languages: Raetic and Lemnian. The family was named by Helmut Rix (1998). The close relationship between Lemnian and Etruscan was first demonstrated by Brandenstein (1948, cols. 1925–1929) and subsequently refined by Eichner (2012, 2013, 2019).

The Hurro-Urartian language family is a well-established genealogical unit, first proposed by Hrozný (1915: 42) and definitively demonstrated by Friedrich (1935; cf. also Kapancjan 1936: 61–63). It comprises two languages attested in cuneiform: Hurrian (2nd millennium BCE) and Urartian (first half of the 1st millennium BCE). A third member, Kassite, has been suggested (Schneider 2003), but is not considered here. Early comparisons with North Caucasian, particularly the Nakh branch, were pioneered by Bork (1909) and summarized by Trombetti (1922–23: 105–108).

¹ First recorded by Herodotus (Hist. 1.94) ca. 440 BCE. See Appendix.

The Etruscan–North Caucasian relationship was first anticipated by Thomsen (1899: 375–381). Hrozný (1917: 57; 1928: 172–173) discussed and expanded these grammatical comparisons, though he attributed them to Indo-European influence. A new era in Hurro-Urartian diachronic studies began with Igor M. Diakonoff, a leading specialist in cuneiform languages. His comparative grammar (1961) marked the first milestone, followed by studies (1967; 1979) that noted ties to East Caucasian without specific evidence. In later works (1971; 1978), he incorporated East Caucasian comparanda into his Hurro-Urartian grammar sketch. Collaborating with Sergei Starostin, Diakonoff (1986; Russian 1988) advanced the field by integrating preliminary East Caucasian reconstructions into Hurro-Urartian comparisons. This lexical material was included in *A North Caucasian Etymological Dictionary*, where Nikolaev and Starostin (1994) presented their Proto-North Caucasian model. Meanwhile, Ivanov (1988) explored Etruscan parallels with Hurrian, Urartian, and early reconstructions of North Caucasian branches. Independently, Orel and Starostin (1990) compared Etruscan with reconstructed East Caucasian, first at the 1988 Ann Arbor Symposium on Language and Prehistory. Their study offered 59 Etruscan–East Caucasian correspondences, including 13 involving Hurro-Urartian. Similar tripartite comparisons appear in Robertson (2006) and Blažek (2010). Exceptions include Facchetti (2002; 2018) and Akulov (2024), which focus directly on Etruscan–Hurrian relations.

The Etruscan data in this study draw primarily from Steinbauer (1999), supplemented by Wallace (2008), Facchetti (2008), Rix (2004), Bonfante (2002), and d’Aversa (1994). Lemnian follows Eichner (2019). Hurrian relies mainly on Richter (2012, 2016), with additional material from Laroche (1980), Wilhelm (2004), Wegner (2004), and Nozadze (2007). Urartian is based chiefly on Arutjunjan (2001), supplemented by Wegner (2004b), Salvini and Wegner (2014), and Diakonoff (1963).

3. COMPARATIVE WORDLIST

1. Etruscan *ac-* ‘machen’ (Steinbauer 1999, 391)

// Hurrian *ag-* ‘to lead, bring, take up’ (Steinbauer 1999, 4-5); Urartian *ag(u)-* ‘to lead’ (Arutjunjan 2001, 431).

Note 1: Concerning semantics, cf. Latin *agere* ‘to lead, drive, conduct’ and ‘to do, act, labor’.

Note 2: In Etruscan there were no voiced stops.

Note 3: Diakonoff & Starostin (1986: 57, n. 142) connected HU **ag-* with North Caucasian **ʔVrkV-* ‘to drive’ = **=ər̥kĕw* (NCED 267).

Note 4: Orel & Starostin (1990: 63, n. 47) compared Etruscan *ac-* ‘to have; to do’ with North Caucasian **ʔāGGV* ‘to hold, take’ = **-āGĀr* (NCED 254–55).

2. Etruscan *ali-* ‘geben, schenken, weihen’ (Steinbauer 1999: 396)

// Hurrian *al(e)-* ‘nahe heranbringen, annähern; heranreichen; apportare, avvicinare’ (Richter 2012: 11).

Note 1: Orel & Starostin (1990: 61, n. 1) compared Etruscan *al-* ‘to give’ with HU **ar-* ‘to give.’

Note 2: Cf. East Caucasian **-iĭV* ‘to give’ > Čečen *-al-a* id. etc. (NCED 640–41).

3. Etruscan *am-* ‘to be’ (Steinbauer 1999, 397)

// Hurrian *amm-* ‘to reach, arrive at; ankommen, gelangen, erreichen; atteindre, arriver, faire parvenir’ (Wegner & Bomhard 2020/2023, 99; Richter 2012, 23).

Note 1: The semantic difference is compatible, cf. Czech *přibýt* ‘to arrive’ vs. *být* ‘to be’.

Note 2: Orel & Starostin (1990, 61, 4) compared Etruscan *am-* ‘to be’ with Hurrian *mann-*, Urartian *man-* and Proto-Lezgian **ʔi-ma(n)-* ‘to be’.

4. Etruscan *an* ‘he, she; which’; *in* ‘it’ (Pfiffig, while Rix preferred the relative function; see Steinbauer (1999: 93, 397)

// Hurrian *anni* ‘this’ vs. *ani* ‘that’ (Kassian 2010 [2011]: 418–19), Urartian *ini* ‘this,’ anaph. *ina* (Arutjunjan 2001: 448; Wilhelm 2004: 128; Salvini & Wegner 2014: 37).

5. Etruscan *ara-* ‘to put, place; make’ (Wallace 2008: 247)

// Hurrian *ar-*, Urartian *ar(u)-* ‘geben’ (Richter 2012: 41–43; Arutjunjan 2001: 436).

Note 1: Diakonoff & Starostin (1986: 51, n. 125) compared HU **ar-* ‘to give’ with East Caucasian **ʔi-ʒʒV* ‘to give’ > Čečen *-al-a* id. etc. = **-iLV* (NCED 640–41).

Note 2: Orel & Starostin (1990: 61, n. 1) compared HU **ar-* ‘to give’ with Etruscan *al-* ‘to give.’

6. Etruscan *atia-* ‘Mutter’ (Steinbauer 1999: 400)

// Hurrian *ašte/i* ‘(Ehe)Frau,’ *atte* ‘Frau’ (Richter 2012: 59–60, 69).

Note 1: Cf. Lithuanian *môtė* ‘wife,’ *móteris* ‘woman’ vs. Latvian *māte* ‘mother,’ Latin *māter* id.

Note 2: Independently identified by Akulov 2024: 42.

Note 3: Orel & Starostin (1990: 63, n. 42) compared Etruscan *ati* ‘mother’ with East Caucasian **ada(j)-* ‘mother’ = **ʔadaj* ‘mother / father’ (NCED 397–98).

Note 4: Diakonoff & Starostin (1986: 39, n. 86) compared Hurrian *ašti* ‘woman, wife’ with East Caucasian **ć(V)dV* > Dargwa *cade* ‘female’; Čečen *stē* id. = **cVjdV* (NCED 375).

7. Etruscan *avil* ‘year’, Lemnian *aviz* ~ Etruscan gen. sg. *avils* (Steinbauer 1999: 401–402)

// Hurrian *°arbu* ‘year’ in such compounds as *šin=arbu* ‘two years old,’ *kig=arbu* ‘three years old,’ (Nuzi) *tumun=arbe* ‘four years old,’ etc. (Wilhelm 2004: 115; Richter 2012: 201, 381, 468).

Note 1: The comparison is acceptable only in the case of the compatibility of Etruscan *l* and Hurrian *r* and the metathesis in one of these languages. Alternatively, it is possible to speculate about the Etruscan proto-form **arvil*.

Note 2: Orel & Starostin (1990: 62, n. 30) compared Etruscan *avil-* ‘year’ with Hurrian *šawali* ‘year’, Urartian *šāli* id., but for the correspondence of HU **š-* vs. Etruscan *ø* there are no supporting arguments. That is why it is possible to offer an alternative solution: the HU compound **ša(w)-* + **(a)wali*, ‘full year’ vel sim., cf. Hurrian *ša(w)-* ‘vollmachen’ (Richter 2012: 340).

8. Etruscan *calusu-* ‘ausgezeichnet, blühend, best-’ (Steinbauer 1999: 403)

// Hurrian *kel-* ‘beneficare, zufrieden/wohl sein, to be/make well, please’ (Richter 2012: 202).

9. Etruscan *cap(i)*- ‘take’ (Wallace 2008: 248)

// Hurrian *kab-*, *kapp-* ‘(aus)füllen, aufschütten’ (Richter 2012: 186).

Note: Compared by Facchetti (2018: 14).

10. Etruscan *caru* ‘gebaut’, *cerixunce* ‘er baute, ließ bauen’, *cerinu* ‘errichtet’ (Steinbauer 1999: 407–08) ~ *car-*, *cer-* ‘to make, build’ (Bonfante 2002: 215)

// Hurrian *kur-* ‘herstellen’ (Richter 2012: 227).

11. Etruscan *cel-* ‘Erde’ (Steinbauer 1999: 407)

// Urartian *kaquli-* ‘stone’ (Arutjunjan 2001: 451).

Note: Cf. North Caucasian **qwiłə* ‘rock, cliff, stone’ (NCED 939).

12. Etruscan *ci-* ‘three’, *ciem* ‘minus three’, *citz*, *ciz(i)* ‘three times’, *cis-um* ‘aber dreier’, *ciemzaθrum-* ‘17’ = ‘20 – 3’, *cialχ-/cealχ-* ‘30’, besides *cezp-* ‘8’, *cezpz* ‘8 times’, *cezpālχ* acc. ‘80’ (Steinbauer 1999: 96, 100, 408–09), if *cezp-* originates from **ce-z=p(i)* ‘three times on {five}’ (Robertson 2006: 6).

// Hurrian *kige* ‘three’, *kigae* ‘dreifach’, *kiški* ‘three times’ or ‘third’ < **kikši* < **kig(a/e)=ži*, *kikmani* ‘13’ or ‘30’ (Richter 2012: 201–202).

Note 1: Gluhak (1978: 26) compared the Etruscan and Hurrian numerals, the latter he mistakenly designated as Urartian.

Note 2: Ivanov (1988: 214) and Orel & Starostin (1990: 61, n. 10) already compared Etruscan *ci-* ‘three’ with Hurrian *kig-* ‘three’ and Proto-Nakh **qo* id.

13. Etruscan *clan* ‘Sohn’, gen. *clens/ś* (Steinbauer 1999: 409), pertinentive pl. *celeniarasi*, besides *clenaraši*, gen. pl. *cliniaras* (Wallace 2008: 248)

// Hurrian *kalge* ‘small, weak’ (Richter 2012: 181).

Note: The semantic dispersion corresponds to Sumerian *tur* ‘small, little, young; child’, *tur₅* ‘newborn; weak, sick’ (Halloran 2006: 282).

14. Etruscan *ei* ‘not’ (Bonfante 2002: 215; cf. Steinbauer 1999: 416)

// Hurrian *oia* (*u-ia-*) ‘nein’ (Wilhelm 2018: 573) ~ *ua* / *uja* ‘nein’ (Richter 2012: 481), Urartian *ui* ‘not’ (Arutjunjan 2001: 47) ~ *ú-i* ‘nein’ (Salvini & Wegner 2014: 114).

Note: Cf. Robertson (2006: 18).

15. Etruscan *-em* ‘minus’ (in subtraction), e.g. *θunem* ‘minus one’, *eslem* ‘minus two’, *ciem* ‘minus three’ (Steinbauer 1999: 416, 489; Bonfante 2002: 215)

// Hurrian *-ma* ~ *-mma* ‘negative forms of the 3rd person’ (Kassian 2010 [2011]: 408), Urartian *mi* ‘negative (prohibitive) proclitic in the optative’ (Diakonoff & Starostin 1986: 94) ~ *me(i)* ‘пусть не; ни’ (Arutjunjan 2001: 454).

Note 1: Cf. Robertson (2006: 18).

Note 2: Diakonoff & Starostin (1986: 94) compared HU with North Caucasian **ma* ‘a prohibitive and negative particle’ (NCED 797).

16. Etruscan *ena-* ‘jetzt, heute, gegenwartig’ (Steinbauer 1999: 417)

// Hurrian *henne/i* ‘jetzt, damals, nun’, Urartian *hini(e)* ‘jetzt’ (Richter 2014: 151; Arutjunjan 2001: 446). In HU, it is probably a compound of the deictic element **hi*, which is identified in Urartian *ikukaḫini* ‘the same’ < **iki=uka=hi=ni* vs. *ikukani* ‘the same’ < **iki=uka=ni* (Salvini & Wegner 2014: 37, 110), and **eni/*ini*; cf. Hurrian *inna* ‘now; if’ or temporal ‘wenn, als’ (Richter 2012: 86).

Note: Diakonoff & Starostin (1986: 32, n. 61) compared HU with North Caucasian **hVnV* ‘now’ = **h[ä]nV* (NCED 487).

17. Etruscan *esvi-* ‘später, nachher’ (Steinbauer 1999: 418)

// Hurrian *ešawa* ‘on the far side’ (Richter 2012: 102–103).

18. Etruscan *eši* ‘vom selben’, *sa-* ‘dasselbe’ (Steinbauer 1999: 418, 462)

// Urartian *isi* ‘всякий’, i.e. ‘every’ (Arutjunjan 2001: 449) or *aišei* ‘irgendjemand’ (Salvini & Wegner 2014: 107).

19. Etruscan *heva* ‘all, jeder’ (Steinbauer 1999: 418)

// Hurrian *heyari*, pl. *heyarunna* / *hé-pa-ru-un-na* ‘alle, every’ (Nozadze 2007: 153).

Note: Cf. North Caucasian **xwE* ‘together, close to’ (NCED 1063–1064).

20. Etruscan pl. *hušur* & *husiur* ‘Kind, Nachkomme’ (Steinbauer 1999: 430)

// Hurrian *hišši* ‘Sproß’ (Richter 2016: 418).

Note: Orel & Starostin (1990: 62, n. 19) compared Etruscan *hus-*, *huš-* ‘boy, son’ with Urartian *aršə* ‘adolescents, boys’ and East Caucasian **u-išwV* = **išwĖ* ‘son, daughter’ (NCED 671–672).

21. Etruscan *ica* ‘this’, later *eca*, acc. *ican*, *ikan*, later *ec(u)n*, loc. *ec(-s)*, with postp. *ecθ(i)*; proclitic nom. *ca*, acc. *cn*, gen. *cla*, loc. *ce(-š)*, *cei(-š)*, *c(-š)*, with postp. *calti*, *clθ(i)* (Steinbauer 1999: 92–93, 430; Bonfante 2002: 92)

// Urartian **iki* in *ikukaḫini* ‘the same’ < **iki=uka=hi=ni* & *ikukani* ‘the same’ < **iki=uka=ni* (Arutjunjan 2001: 448; Salvini & Wegner 2014: 37, 110).

22. Etruscan *ilacva-* ‘Monat, Fest’ with the suffix *-cva-* (Steinbauer 1999: 431, 87)

// Hurrian *eli* ‘Fest’ (Richter 2012: 82).

Note: Independently identified by Facchetti (2018: 14).

23. Etruscan *ipa-*, later *epa-* ‘who, which’ (Bonfante 2002: 94)

// Hurrian *abi* ~ *awi* ‘who’ (Kassian 2010 [2011]: 425, #98; Richter 2012: 38; Wegner & Bomhard 2020/2023: 55).

24. Etruscan *ita* ‘this’, later *eta*, acc. *itan*, *itun*, later *et(a)n*, gen. *etula*, loc. *etule*; proclitic nom. *ta*, acc. *tn*, loc. *tei*, *tei(-ś)* (Steinbauer 1999: 92; Bonfante 2002: 92)

// Urartian *ute* ‘nobody’, analyzable as *u-te*; the first component can be identified with Urartian *ui* ‘not’, *ui aini* ‘nobody’, where the second component means ‘another’ (Arutjunjan 2001: 471; Salvini & Wegner 2014: 107, 114).

Note: Cf. East Caucasian **tV* ‘that’ > Dargwa: *Akušā it*, *Čirag it*, *taś* etc. (NCED 993).

25. Etruscan *lautn* ‘family’, *lautni* ‘freedman’, *lautnita* & *lautniθa* ‘freedwoman’ (Wallace 2008: 250)

// Urartian *lutu*, attested in several variants: ^{SAL}*lu-tú*, ^{LU.SAL}*lu-tú-ni*, ^{SAL}*lu-tú-ni-e* ‘woman’ (with the same suffix as the synonym ^{LU/SAL}*uediani*), ^{SAL}*lu-tú*^{MEŠ} ‘women’ (Arutjunjan 2001: 453).

Note: Robertson (2006: 12) made the same comparison, adding Avar *ḷ:ádi*, pl. *ḷ:ud-bí* ‘wife’ (NCED 764–65).

26. Etruscan *max* ‘five’, *muvalx* ‘fifty’, Lemnian *mav* ‘five’ (Steinbauer 1999: 439; Wallace 2008: 220), if the final *-x* was the ordinal suffix as in *θunx-* ‘first’ (Jacemirskij 2007: 190)

// Hurrian *eman* ‘ten’ (Richter 2012: 83–85), if the final *-n* can be identified with the Hurrian plural relator *-na* (cf. Wilhelm 2004a: 106–108).

Note 1: First compared by Blažek (2010a: 121).

Note 2: If the final *-x* was the enclitic conjunction (Wallace 2008: 116), it is legitimate to interpret Tyrsenian **mu(va)-x* as ‘plus X’, where X = ‘4’ with regard to the interpretation of Etruscan > Latin *mundus* as **“four cardinal points”* (see below). In this case the Etruscan/Tyrsenian numeral ‘5’ would be formed as the sum **θu-mu(va)-x* ‘1+4’. This hypothesis supports the existence of the common numeral ‘4’ in Tyrsenian **mu(va)-* & *mundus* and Hurrian **(ta-)muni-* (see #29). Accepting it, this isogloss conspicuously resembles the Anatolian numerals ‘4’: Hittite *me(i)u-/meia-*; Cuneiform Luwian *māuṽa-* ‘four’, Hieroglyphic Luwian *miwa-*, reconstructed on the basis of the sign IIII *mi* and the record nom.-acc.pl.c. IIII-*wa/i-zi*; Lycian *mupm̃m-* ‘fourfold’ (Kloekhorst 2008: 571), and it is legitimate to ask, who borrowed from whom.

27. Etruscan *men-* ‘errichten, setzen, weihen’ (Steinbauer 1999: 443)

// Hurrian *min-* ‘geben, nehmen’ (Richter 2012: 250).

28. Etruscan *mi* ‘I’, acc. *mini* ‘me’ (Steinbauer 1999: 91)

// Urartian *-me* enclitic dat. ‘me’ (Arutjunjan 2001: 454; Salvini & Wegner 2014: 37).

29. Etruscan > Latin *mundus* ‘heavens, sky; world, earth; subterranean vault’ [Plautus+], originally perhaps ‘four cardinal points’ with regard to the probable equivalence *Roma quadrata* = *mundus* (cf. Carter Platner 1901: 420–25; 1908: 180; Toporov 1983: 123–24), which is applicable to both ‘sky’ and ‘world’

// Hurrian *tumni* ‘fourth’, *tumunzi* ‘fourth’, *tumnadi* ‘consisting of four; Vierheit’ (Richter 2012: 468–69). With regard to the ordinal *tumunzi* ‘4th’ and cardinal *tamri* ‘9’ (Richter 2012: 436), the numeral ‘4’ is derivable from **tamuni* and ‘9’ from **tamun-nari-* ‘4 + 5’ (cf. Blažek 2010: 118–

19). It is tempting to identify in the first syllable the Hurrian form *tea*, which has been interpreted as ‘much, many, very, numerous, big, completely’. The hypothetical compound **tea-muni-* would mean ‘completely four’, perhaps about fingers of a hand without the thumb. Without the first component, but with the derivational suffix *-adi* we have got *munadi* ‘consisting of four’, which perfectly corresponds to Latinized Etruscan *mundus*.

Note: Cf. Quechua *tahuantinsuyu* /*tawantinsuyul* ‘four regions of Andes’ (*tawa* ‘four’; *Anti* ‘Andes’; *suyu* ‘region’ – DQEQ 16–17; 595, 617) = ‘the Inca empire’.

30. Etruscan *netei* ‘Schwiegermutter’ (Steinbauer 1999: 449)

// Hurrian *nera* ‘Mutter’ (Richter 2012: 275) ~ ‘жена деверя’ (Nozadze 2007: 257).

Note 1: Perhaps a primary compound of the type **ner-atia* ‘man’s mother’; cf. Urartian *nara* ‘people, nation’.

Note 2: Kapanejan (apud Martirosyan 2010: 505) connected the Hurrian kinship term *nera* with Armenian *nēr*, gen.-dat.sg. *niri* ‘husband’s brother’s wife; husband’s the other wife’. The latter has been derived from IE **(H)jēnh₂-ter-ih₂-*.

31. Etruscan *nu-θe* ‘to listen to’ (Wallace 2008: 251); the suffix *-θ-* is productive in the Etruscan verbal morphology (Steinbauer 1999: 142)

// Hurrian *nui* ‘ear’ (Richter 2012: 281).

32. Etruscan *nun-* ‘to bring, involve’, *nuna* ‘offering’ (Bonfante 2002: 217)

// Urartian *nun(a)-* ‘to come, arrive’ (Arutjunan 2001: 456); cf. Hurrian *un-* ‘kommen’, trans. ‘bringen’ (Richter 2012: 490).

Note: Connected by Robertson (2006: 9).

33. Etruscan *rasna* ‘etruskisch; national’ (Steinbauer 1999: 459) = *Ῥασέννα* by Dionysius Halicarnassus (*Antiquitates Romanae*, 1.30.3: ‘Their own name ... is the same as that of one of their leaders, *Ῥασέννα*’)

// Urartian *arše* ‘child, boy, lad, youth’ (Arutjunan 2001: 436). The Hurrian counterpart may be identified in the Ugaritic alphabetic record *aršm* ‘young boy’? (Richter 2012: 48).

Note 1: The ethnonyms motivated by the meaning ‘child, boy, youth’ are relatively frequent in Indo-European traditions: e.g. *Čech* can be etymologized on the basis of the appellative preserved in Slovenian *čèh* ‘boy from 10 to 15 years; herdsman’, Kašubian *čex* ‘youngster, child’, all from Common Slavic **čexъ*; *Goth* of the same origin as Norwegian *gut*, Swedish *gutt* ‘boy’; the Germanic tribe of *Buri* [Tacitus: *Germania* 43] of the same origin as Old Norse *burr*, Old English *byre* ‘son’ etc. (Blažek 2010b: 17).

Note 2: Diakonoff & Starostin (1986: 36, n. 71) compared the Urartian word with such the North Caucasian forms as Lak *ars* ‘son’, Dargi *urši* ‘son, boy’ etc. (NCED 671–672: North Caucasian **=išwĒ*, where the forms with the medial *-r-* are explained through the metathesis of the *r*-suffix of the oblique stem).

34. Etruscan *rene-θi* ‘in der Hand; zur Verfügung’ (Facchetti 2008: 231; cf. Steinbauer 1999: 460 without any translation)

// Hurrian *nariya* ‘five’ (Wilhelm 2004: 115).

Note 1: Facchetti (2018: 20) connected these forms, accepting metathesis in one of the compared languages.

Note 2: The probable external cognates from other branches of the Sino-Caucasian macrofamily indicate that the sequence *r...n* is primary: Burušaski of Hunza *-rín* & *-rín̄*, pl. *rín̄cín̄*, Nagir pl. in *očan̄*, Yasin *-rén̄*, (together with the root): pl. *-rén̄(čin̄)* ‘hand’ (Berger 1998: 364–65); Yeniseian **r̄ɔŋ* ‘hand’ > Ket *l’añ*, Yug *l’ɔŋ* ‘hand’, Arin *lan-t’ūŋ* ‘wing’ (Starostin 1995: 178). In North Caucasian a promising cognate can be identified in the numeral **ʔrān̄λE* ‘6’ (NCED 219–220), if it is analyzable as a compound of **ʔrān̄* ‘5’ & **λE*. The latter component is derivable from the North Caucasian verb **=ālĖw* ‘to lie, put, lead’ > Nakh **=ill-* ‘to lie, put upon (something)’, **t-ill-* ‘to put (from above)’, Čamalal *=al-* ‘to begin’; Tsezian *ɔL* ‘to be’; Bežta *=ol-*, Gunzib *=ol-* ‘to finish’; Lezgian *ʔel̄:wi-* ‘to put, lie’; West Caucasian **λ’ə-* ‘to lie’ (NCED 278–79). The primary semantics could be ‘six’ = ‘(one) put upon five’ or ‘beginning the (new) five’. A similar structure is assumed for Indo-European **(K)syeḱs* ‘6’, namely **ḡhes-* ‘hand’ & **yeḱs-* ‘to grow, rise’ (**-ḱ-* is confirmed by Lithuanian *vešėti* ‘to grow vigorously, thrive; prosper, flourish’), i.e. ‘overgrowing the hand’ (see Blažek 1999: 239–41).

35. Etruscan *ruva* ‘Bruder’ (Steinbauer 1999: 462)

// Hurrian *ewuru* / *ewiru* ‘heir, Erbe’ (Nozadze 2007: 131).

Note 1: These forms are compatible only in the case of metathesis in one of the compared languages. It seems, it was Hurrian, where the initial *r-* was probably impossible in the inherited lexicon (see n. 24).

Note 2: The semantic difference has analogy e.g. in Old Irish *orb* ‘heir’, *com-arbe* ‘successor’, Gothic *arbi* ‘heir’, vs. Russian *rebēnok* ‘child’, *rebjáta* ‘children’ (Pokorny 1959: 781–82).

36. Etruscan *sex* ‘Tochter’ (Steinbauer 1999: 463), if it is derivable from *selχ*; cf. *mex* ‘Herrin, Königin’, if it represents an adaptation of Phoenician *mlkt* ‘queen’ (Krahmalkov 2000: 289); cf. Hebrew *malḫah* ‘wife of the king’ (HAL 592)

// Hurrian *šala* ‘Tochter’ (Richter 2012: 344), Urartian *silā* ‘wife’ (Arutjunan 2001: 460).

Note 1: Cf. Ivanov (1988: 212). Independently identified by Akulov (2024: 42).

Note 2: Related can be Common Burušaski **salé-n* > Yasin *salén*, *selén* ‘sister(s) and daughter(s) of a man’, Hunza & Nagir *silāžín* ‘woman, female relatives’ (Berger 1998: 373, 378).

Note 3: Orel & Starostin (1990: 63, n. 43) compared Etruscan *sec*, *sex* ‘daughter’ with East Caucasian **CVḳwā* ‘female’.

Note 4: Diakonoff & Starostin (1986: 42, n. 94) connected Hurrian *šala* ‘daughter’ and Urartian *silā* ‘wife’ with East Caucasian **-ilćéwi* ‘female relative’ = **rīlćwī* (NCED 952).

37. Etruscan *snaenθ* ‘Genossin, Zofe, Magd’ (Steinbauer 1999: 467) ~ ‘maid, companion’ (Bonfante 2002: 219)

// Hurrian *šena* ‘brother’ (Richter 2012: 384–85).

Note 1: Concerning semantics, cf. Egyptian *sn* ‘brother; companion’ (Erman & Grapow IV: 150–151) ~ *sn* ‘brother’ vs. *snw* ‘companion, fellow’, *snt* ‘sister’ vs. dual *snwt(y)* ‘female companions’ (Faulkner 1962/1991: 230).

Note 2: Independently identified by Akulov (2024: 42).

Note 3: Cf. East Caucasian: Avar-Andi **ccVnV* ‘second cousin’ (NCED 322).

38. Etruscan *sval(a)-* ‘lebend(ig)’ (Steinbauer 1999: 469) ~ *sval* ‘alive, living’ (Wallace 2008: 252)

// Hurrian *šawali* ‘year’ (Richter 2012: 351), Urartian *šāli* id. (Arutjunan 2001: 462).

Note 1: Cf. Vedic *āyu-* ‘life, lifetime’ vs. Latin *aevum* ‘period of time’.

Note 2: Orel & Starostin (1986: 30, n. 55) compared HU **šawal-* ‘year’ with North Caucasian **swērHo* ‘year; old’ (NCED 968–69).

Note 3: Orel & Starostin (1990: 62, n. 30) compared Hurrian *šawali* ‘year’ with Etruscan *avil-* ‘year’, while Etruscan *sval-* ‘to live’ they combined with Hurrian *zurgi* ‘blood’ and East Caucasian **čāžwV* ‘alive; blood’ = **čāLwV* (NCED 376).

39. Etruscan *ten-, tn-* ‘to act as magistrate’ (Bonfante 2002: 219) ~ ‘ableisten’ (Steinbauer 1999: 476)

// Hurrian *tan-*, dial. (Ugarit) *tin-* ‘tun, machen’ (Richter 2012: 436–438), Urartian *tan-* ‘machen’ (Salvini & Wegner 2014: 114) ~ *tanu-* ‘to found’ (Arutjunan 2001: 466).

40. Etruscan *tinia-* ‘Tag; Zeus, Juppiter’ (Steinbauer 1999: 478)

// Hurrian *tueni* ‘Tag’ (Richter 2012: 478).

Note: Orel & Starostin (1990: 62, n. 31) compared Etruscan *tin-* ‘day’ with PNakh **dē*, obl. **dēni-* ‘day’ < East Caucasian **rīhV* (NCED 952).

41. Etruscan *trepu-* ‘Handwerker, Zimmermann’ (Steinbauer 1999: 481); Lemnian *tover* ‘Schmiedegott’, reconstructed on the basis of the Lemnian city-name *Toverona*, corresponding to ancient Ἡφαίστεια [Herodotus, 6.140] according to Eichner (2019: 120)

// Hurrian *tabiri* ‘smith’ > Sumerian *TABIRA/TIBIRA* ‘coppersmith’ > Assyrian *tabīru* id. (Richter 2012: 439).

42. Etruscan *trutnuθ* ‘Blitzdeuter’ = Latin *fulguriator* (Steinbauer 1999: 482), if it is analyzable as **t(V)ru-* ‘fire’ & *ten-/tn-* ‘to act as magistrate’ ~ ‘ableisten’ (see above), together perhaps ‘that one which attends to {celestial} fires’

// Hurrian *tari-* ‘fire’ (Richter 2012: 444).

Note: The semantic development from ‘fire’ to ‘lightning’ has analogy e.g. in Old Norse *elding* ‘lightning’ vs. *eldr* ‘fire’ or in Irish *teintreach* ‘lightning’ vs. *teine* ‘fire’ (Buck 1949: #1.57).

43. Etruscan *Tursikina* ‘gens’ from Chiusi, 7th cent. BCE; cf. Greek Τῦρσενοί, later Τῦρρενοί, Umbrian *Tursko-*, Latin *Tusci* (d’Aversa 1994: 55; Steinbauer 1999: 483)

// Hurrian *tarše* ‘Mensch’, *taršuwan(n)i* ‘Menschheit’ (Richter 2012: 447); Urartian ^{LÚ}*taršuani(e)* ‘man, human being’, ^{LÚ}*taršuanarani* ‘people, settlement’ vs. *nara* ‘people, nation’ (Arutjunan 2001: 456, 467).

Note: Diakonoff & Starostin (1986: 39, n. 87) reconstructed the HU compound **tar-s(u)wa-*, where the first component was connected with Hurrian *turohḫa* ‘male’ and the second one with North Caucasian **čwījo* ‘man, male’ (NCED 336).

44. Etruscan *θi* ‘Wasser; Bach, Fluß’ (Steinbauer 1999: 488)

// Hurrian *šije* ‘eau’ ~ *ī* in the Ugaritic transcription (Richter 2012: 366, 544).

Note 1: Independently identified by Facchetti (2018: 14) and Akulov (2024: 43).

Note 2: Diakonoff & Starostin (1986: 49, n. 118) connected Hurrian *šije* ‘water’ with East Caucasian **ʔä-šV* ‘to become wet’; alternatively with East Caucasian **xxänšy* ‘water’ > Karata *ššēji* id. etc. = **xǎñhĩ* (NCED 1060).

45. Etruscan *θu(n)-* ‘eins’, later *θu*, further *θunem* ‘weniger eins’, *θunz* ‘einmal’ (Steinbauer 1999: 100, 489)

// Hurrian *šukki* ~ *šugi* ‘one’, *šukši* ‘erstens’ (Richter 2012: 406–407), Urartian *šuki* ‘once’, *šusi(ni)* ‘ein ... einzig’ (Arutjunan 2001: 466; Salvini & Wegner 2014: 114).

Note 1: The first one who compared the Etruscan and Urartian numerals ‘1’ was probably Stoltenberg (1943: 243).

Note 2: Orel & Starostin (1990: 62, n. 20) compared Etruscan *θu* ‘one’ with Urartian *š(V)-usə* ‘first’ and North Caucasian **cHš* ‘one’ (NCED 323–324).

46. Etruscan *un* acc. sg., *une* dat. sg.; *unu* acc. pl. ‘you’ (Rix 1991; Bonfante 2002: 91, 220; cf. Wallace 2001: 58–59, 254)

// Hurrian *we* ‘selbständige Person der 2Ps. Sg.; *wella* 2Ps. Pl. (Richter 2012: 309–310).

Note 1: The comparison with Etruscan nom. *mi*, acc. *mini* ‘I’ indicates the unattested nom. sg. *u* vel sim.

Note 2: Diakonoff & Starostin (1986: 82) compared Hurrian *we* with East Caucasian **w_lV* ‘thou’ > Lezgin *wun*, Arči *un* id. etc. < North Caucasian *uō* id. (NCED 1014).

Note 3: Related are also Burušaski *un* ‘thou’ (Berger 1998: 456); Yeniseian **ʔaw/ʔu* id. > Ket *ū*, Yug *u*; Kott *au*, pl. *auoŋ*; Arin *au*; Pumpokol *úe*.

47. Etruscan *zal* ‘zwei’, loc. *esals*, subtracted *eslem*, multiplied *eslz* ‘zweimal’, *zelur* ‘two at the time’ (Bonfante 2002: 97); *zaθrum* ‘zwanzig’ (Steinbauer 1999: 96–100, 499); Raetic *pal* ‘zwei’ (Rix 1998: 57; Salomon 2020: 283); Lemnian *azlas* ‘den zweien’ (Eichner 2019: 96).

// Hurrian *ošella* in *šinišella*, which should correspond to Hittite *2-an-ki* ‘twice’ by Kammenhuber (see Laroche 1980: 234; Richter 2012: 382), but a more fitting interpretation could be ‘two both’² or German ‘zwei beide.’³ Friedrich (1934: 314–316), followed by Laroche (1980: 171, 288) also mentioned the Hurrian suffix *-ššil* appearing in the theonyms of Mitanni-Aryan origin, namely ^{DINGIRmeš}*Mi-it-ra-aš-ši-il* and ^{DINGIRmeš}*ú-ru-wa-na-aš-ši-il*. They apparently correspond to the Vedic divine names in dual *Mitrā* ‘both Mitras’, i.e. ‘Mitra and Varuṇa’, and *Váruṇā* ‘both Varuṇas’, i.e. ‘Mitra and Varuṇa’. Both the theonyms are also attested as the copulative compound of the type dvandva, with the common dual ending: nom. *Mitra-Váruṇau*, gen.-loc. *Mitra-Váruṇayoḥ*, dat.-abl.-instr. *Mitra-Váruṇābhyām*. Laroche (1980: 171) interpreted the suffix *-ššil* as ‘les deux’.

Note 1: The Etruscan-Hurrian comparison was identified by Ivanov (1988: 214).

Note 2: Orel & Starostin (1990: 61, n. 12) compared Etruscan *zal* ‘2’ with HU counterparts: according to more recent sources Hurrian *šina*, *šine/i* ‘2’ (Richter 2012: 380), Urartian *šinamdini* ‘вдвое’ (Arutjunan 2001: 464), plus Nakh **ši?*, obl. **šina-* ‘2’ (NCED 845–846). But HU **šina-* ‘2’ is apparently of Semitic origin, cf. Akkadian m. *šina* (status absolutus), Eblaite *šina* ‘2’ (Blažek 2010: 118; Robertson 2006: 9; Facchetti 2018: 18, following p.c. with Mauro Giorgieri in 2011), while Nakh **ši?*, obl. **šina-* ‘2’ looks as a result of the areal influence of Hurrian or Urartian.

Note 3: The comparison of Etruscan *zal* ‘2’ with HU **šina-* ‘2’ would imply the correspondence of Etruscan *-l-* and HU *-n-*, but for it are missing other supporting arguments.

4. CONCLUSION

Of the 47 proposed comparisons, 20–21 are nouns (substantival or adjectival), 11 are pronouns, adverbs, or grammatical words, 6 are numerals or numeral-derived, and 9–10 are verbs. Given the small size and low semantic overlap of the attested Etruscan, Hurrian, and Urartian lexicons, this distribution supports—tentatively—a genetic relationship between Etruscan (or more broadly Tyrsenian) and Hurro-Urartian. The connection appears stronger than any with North Caucasian, though the latter awaits systematic quantification.

Additional evidence includes Hurro-Urartian etymologies for Greek mythological names (*Apollo*, *Artemis*, *Leto*, *Athene*, etc.; cf. Blažek 2007, 2014, 2016, 2017, 2022) and archaeological-linguistic traces of Hurro-Urartian dialect(s) in western Anatolia, the Aegean, and possibly Cyprus. Northwest Anatolian origins for pre-Etruscan Tyrsenians are also defended (Beekes 2003; Kloekhorst 2021: 58–72). The present data allow the hypothesis that these two population movements reflect a single ethno-linguistic continuum.

² E.g. two both of our own.

³ E.g. zwei beide von unseren Selbst.

APPENDIX A₁: CORRESPONDENCES AMONG CONSONANTS

Etruscan	Hurrian	etymologies (##)	Etruscan	Uartian	etymologies (##)
<i>c-</i>	<i>k-</i>	8, 9, 10, 12, 13	<i>c-</i>	<i>-q-</i>	11
<i>-c-</i>	<i>-g-</i>	1	<i>-c-</i>	<i>-g-</i>	1
			<i>-c-</i>	<i>-k-</i>	21
<i>h-</i>	<i>ḫ-</i>	19, 20			
<i>-i- (= ṡi)</i>	<i>-j-</i>	14	<i>-i- (= ṡi)</i>	<i>-ṡi-</i>	14
<i>-l-</i>	<i>-l-</i>	2, 11, 13, 22	<i>l-</i>	<i>l-</i>	25
<i>m-</i>	<i>m-</i>	27			
<i>m-</i>	<i>-m-</i>	26, 29	<i>m-</i>	<i>-m-</i>	28
<i>-m-</i>	<i>-m(m)-</i>	3, 15			
<i>n-</i>	<i>n-</i>	30, 31	<i>n-</i>	<i>n-</i>	32
<i>-n-</i>	<i>n-</i>	34			
<i>-n-</i>	<i>-n(n)-</i>	4, 16, 37, 39	<i>-n-</i>	<i>-n-</i>	4, 16, 32
<i>-p-</i>	<i>-b-</i>	41			
<i>-p-</i>	<i>-b-/-pp-</i>	9			
<i>-p-</i>	<i>-b-/-w-</i>	23			
<i>r-</i>	<i>-r-</i>	34, 35	<i>r-</i>	<i>-r-</i>	33
<i>-r-</i>	<i>-r-</i>	5, 10, 41, 42, 43	<i>-r-</i>	<i>-r-</i>	5
<i>s-</i>	<i>š-</i>	36, 37	<i>s-</i>	<i>s-</i>	36
<i>-s-</i>	<i>-š-</i>	43	<i>-s-</i>	<i>-š-</i>	33, 34
<i>-s-/-ś-</i>	<i>-š-</i>	20	<i>-ś-</i>	<i>-s- or -š-</i>	18
<i>sv-</i>	<i>š(a)w-</i>	38			
<i>-sv-</i>	<i>-š(a)w-</i>	17			
<i>t-</i>	<i>t-</i>	39, 40, 41, 42, 43	<i>t-</i>	<i>t-</i>	39, 43
			<i>-t-</i>	<i>-t-</i>	24, 25
<i>-ti- (= tṡi)</i>	<i>-št-/-tt-</i>	6			
<i>θ-</i>	<i>š-</i>	44, 46	<i>θ-</i>	<i>š-</i>	45
<i>-v-</i>	<i>-y-/-p-</i>	19			
<i>-v-</i>	<i>-rb-</i>	7			
<i>z-/-s-</i>	<i>-š-</i>	47			

APPENDIX A₂: CORRESPONDENCES AMONG VOWELS OF THE FIRST SYLLABLE

Etruscan	Hurrian	etymologies (##)	Etruscan	Uartian	etymologies (##)
<i>a</i>	<i>a</i>	1, 2, 3, 4, 5, 6, 7, 9, 26, 33	<i>a</i>	<i>a</i>	1, 5, 33
<i>a/e</i>	<i>e/i</i>	8, 47			
<i>a/e</i>	<i>u</i>	10	<i>e</i>	<i>u</i>	11, 14
<i>e</i>	<i>a</i>	15, 36			
<i>ø/e</i>	<i>a</i>	13			
<i>ø/e</i>	<i>a/i</i>	39	<i>e</i>	<i>i</i>	16, 18, 36
<i>ø</i>	<i>a</i>	42			
<i>e</i>	<i>e</i>	16, 17, 19, 30			
<i>i</i>	<i>e</i>	22	<i>i</i>	<i>e</i>	28
<i>i</i>	<i>i</i>	12, 44	<i>i</i>	<i>i</i>	4, 21
<i>i</i>	<i>a</i>	23			
<i>i</i>	<i>ue</i>	40	<i>i</i>	<i>u</i>	24
<i>u</i>	<i>u</i>	29, 31, 45	<i>u</i>	<i>u</i>	45
<i>u</i>	<i>u/i</i>	35			
<i>u</i>	<i>a</i>	43	<i>u</i>	<i>a</i>	43
			<i>au</i>	<i>u</i>	25

APPENDIX B: TEXT PASSAGES

Dionysius Halicarnassus: *Antiquitates Romanae*, 1.27

1.27.1. οἱ δὲ μετανάστας μυθολογοῦντες αὐτοὺς εἶναι Τυρρηγνὸν ἀποφαίνουσιν ἡγεμόνα τῆς ἀποικίας γενόμενον ἀφ’ ἐαυτοῦ θέσθαι τῷ ἔθνει τοῦνομα: τοῦτον δὲ Λυδὸν εἶναι τὸ γένος ἐκ τῆς πρότερον Μηονίας καλουμένης, παλαιὸν δὴ τινα μετανάστην ὄντα: εἶναι δ’ αὐτὸν πέμπτον ἀπὸ Διὸς, λέγοντες ἐκ Διὸς καὶ Γῆς Μάνην γενέσθαι πρῶτον ἐν τῇ γῇ ταύτῃ βασιλέα: τούτου δὲ καὶ Καλλιρόης τῆς Ὠκεανοῦ θυγατρὸς γεννηθῆναι Κότυν: τῷ δὲ Κότυϊ γήμαντι θυγατέρα Τύλλου τοῦ γηγενοῦς Ἀλίην δύο γενέσθαι παῖδας Ἀσίην καὶ Ἄτυν:

2. ἐκ δὲ Ἄτυος καὶ Καλλιθέας τῆς Χωραίου Λυδὸν φῦναι καὶ Τυρρηγνόν: καὶ τὸν μὲν Λυδὸν αὐτοῦ καταμείναντα τὴν πατρίαν ἀρχὴν παραλαβεῖν καὶ ἀπ’ αὐτοῦ Λυδίαν τὴν γῆν ὀνομασθῆναι: Τυρρηγνὸν δὲ τῆς ἀποικίας ἡγησάμενον πολλὴν κτήσασθαι τῆς Ἰταλίας καὶ τοῖς συναραμένοις τοῦ στόλου ταύτην θέσθαι τὴν ἐπωνυμίαν.

3. Ἡροδότῳ δὲ εἴρηνται Ἄτυος τοῦ Μάνεω παῖδες οἱ περὶ Τυρρηγνόν, καὶ ἡ μετανάστασις τῶν Μηόνων εἰς Ἰταλίαν οὐχ ἐκούσιος. φησὶ γὰρ ἐπὶ τῆς Ἄτυος ἀρχῆς ἀφορίαν καρπῶν ἐν τῇ γῇ Μηόνων γενέσθαι, τοὺς δὲ ἀνθρώπους τέως μὲν ὑπὸ τῆς φιλοχωρίας κρατουμένους πολλὰ διαμηχανήσασθαι πρὸς τὴν συμφορὰν ἀλεξήτρια, τῇ μὲν ἑτέρα τῶν ἡμερῶν μέτρια σιτία προσφερομένους, τῇ δ’ ἑτέρα διακαρτεροῦντας: χρονίζοντος δὲ τοῦ δεινοῦ διανείμαντας ἅπαντα τὸν δῆμον διχῇ κλήρους ταῖς μοίραις

ἐπιβαλεῖν: τὸν μὲν ἐπ' ἐξόδῳ τῆς χώρας, τὸν δ' ἐπὶ μονῇ, καὶ τῶν Ἄτυος παίδων τὸν μὲν τῇ προσνεῖμαι, τὸν δὲ τῇ.

4. λαχούσης δὲ τῆς ἅμα Λυδῶ μοίρας τὴν τοῦ μεῖναι τύχην, ἐκχωρῆσαι τὴν ἐτέραν ἀπολαχοῦσαν τῶν χρημάτων τὰ μέρη, ὀρμισαμένην δ' ἐπὶ τοῖς ἐσπερίοις μέρεσι τῆς Ἰταλίας, ἔνθα ἦν Ὀμβρικοῖς ἡ οἴκησις, αὐτοῦ καταμείνασαν ἰδρύσασθαι πόλεις τὰς ἔτι καὶ κατ' αὐτὸν ἐκεῖνον οὔσας.

1.27.1. 'But those who relate a legendary tale about their having come from a foreign land say that Tyrrhenus, who was the leader of the colony, gave his name to the nation, and that he was a Lydian by birth, from the district formerly called Maeonia, and migrated in ancient times. They add that he was the fifth in descent from Zeus; for they say that the son of Zeus and Gê was Manes, the first king of that country, and his son by Callirrhoê, the daughter of Oceanus, was Cotys, who by Haliê, the daughter of earth-born Tyllus, had two sons, Asies and Atys:

2. from the latter of whom by Callithea, the daughter of Choraëus, came Lydus and Tyrrhenus. Lydus, they continue, remaining there, inherited his father's kingdom, and from him the country was called Lydia; but Tyrrhenus, who was the leader of the colony, conquered a large portion of Italy and gave his name to those who had taken part in the expedition.

3. Herodotus, however, says that Tyrrhenus and his brother were the sons of Atys, the son of Manes, and that the migration of the Maeonians to Italy was not voluntary. For they say that in the reign of Atys there was a dearth in the country of the Maeonians and that the inhabitants, inspired by love of their native land, for a time contrived a great many methods to resist this calamity, one day permitting themselves but a moderate allowance of food and the next day fasting. But, as the mischief continued, they divided the people into two groups and cast lots to determine which should go out of the country and which should stay in it; of the sons of Atys one was assigned to the one group the other to the other.

4. And when the lot fell to that part of the people which was with Lydus to remain in the country, the other group departed after receiving their share of the common possessions; and landing in the western parts of Italy where the Umbrians dwelt, they remained there and built the cities that still existed even in his time.'

1.30.3. Ῥωμαῖοι μέντοι ἄλλαις αὐτὸ προσαγορεύουσιν ὀνομασίαις: καὶ γὰρ ἐπὶ τῆς χώρας, ἐν ᾗ ποτε ὤκησαν, Ἑτρούριαν προσαγορευομένης Ἑτρούσκους καλοῦσι τοὺς ἀνθρώπους: καὶ ἐπὶ τῆς ἐμπειρίας τῶν περὶ τὰ θεῖα σεβάσματα λειτουργιῶν, διαφέροντας εἰς αὐτὴν ἐτέρων, νῦν μὲν Τούσκους ἀσαφέστερον, πρότερον δ' ἀκριβοῦντες τοῦνομα ὥσπερ Ἕλληνες Θυοσκοὺς ἐκάλουν: αὐτοὶ μέντοι σφαῖς αὐτοὺς ἐπὶ τῶν ἡγεμόνων τινὸς Ῥασέννα τὸν αὐτὸν ἐκεῖνω

1.30.3. 'The Romans, however, give them other names: from the country they once inhabited, named Etruria, they call them Etruscans, and from their knowledge of the ceremonies relating to divine worship, in which they excel others, they now call them, rather inaccurately, Tusci, but formerly, with the same accuracy as the Greeks, they called them Thyoscoi. Their own name for themselves, however, is the same as that of one of their leaders, Rasenna.'

Dionysius Halicarnassus: *Antiquitates Romanae*, Vol I-IV, ed. Karl Jacoby. Leipzig: Teubner 1885.

Dionysius of Halicarnassus: *The Roman Antiquities*, Vol. I, with an English translation by Earnest Gary, on the basis of the version of Edward Spelman Cambridge (Mass.): Harvard University Press – London: Heinemann 1937.

Herodotus: *Historiae* 1.94

[3] ἐπὶ Ἄτυος τοῦ Μάνεω βασιλέος σιτοδείην ἰσχυρὴν ἀνὰ τὴν Λυδίην πᾶσαν γενέσθαι, καὶ τοὺς Λυδοὺς τέως μὲν διάγειν λιπαρέοντας, μετὰ δὲ ὡς οὐ παύεσθαι, ἅκεα δίζησθαι, ἄλλον δὲ ἄλλο ἐπιμηχανᾶσθαι αὐτῶν. ἐξευρεθῆναι δὴ ὧν τότε καὶ τῶν κύβων καὶ τῶν ἀστραγάλων καὶ τῆς σφαίρης καὶ τῶν ἀλλέων πασέων παιγνιέων τὰ εἶδεα, πλὴν πεσσῶν τούτων γὰρ ὧν τὴν ἐξεύρεσιν οὐκ οἰκηιοῦνται Λυδοί. [4] ποιεῖν δὲ ὥδε πρὸς τὸν λιμὸν ἐξευρόντας, τὴν μὲν ἐτέρην τῶν ἡμερέων παίζειν πᾶσαν, ἵνα δὴ μὴ ζητέοιεν σιτία, τὴν δὲ ἐτέρην σιτέεσθαι παυομένους τῶν παιγνιέων. τοιοῦτῳ τρόπῳ διάγειν ἐπ' ἕτεα δυῶν δέοντα εἴκοσι. [5] ἐπεῖτε δὲ οὐκ ἀνιέναι τὸ κακὸν ἀλλ' ἔτι ἐπὶ μᾶλλον βιάζεσθαι οὕτω δὴ τὸν βασιλέα αὐτῶν δύο μοίρας διελόντα Λυδῶν πάντων κληρῶσαι τὴν μὲν ἐπὶ μόνῃ τὴν δὲ ἐπὶ ἐξόδῳ ἐκ τῆς χώρας, καὶ ἐπὶ μὲν τῇ μένειν αὐτοῦ λαγχανούσῃ τῶν μοιρέων ἑωυτὸν τὸν βασιλέα προστάσσειν, ἐπὶ δὲ τῇ ἀπαλλασσομένῃ τὸν ἑωυτοῦ παῖδα, τῷ οὐνομα εἶναι Τυρσηνόν. [6] λαχόντας δὲ αὐτῶν τοὺς ἐτέρους ἐξίεναι ἐκ τῆς χώρας καταβῆναι ἐς Σμύρνην καὶ μηχανήσασθαι πλοῖα, ἐς τὰ ἐσθεμένους τὰ πάντα ὅσα σφι ἦν χρηστὰ ἐπίπλοα, ἀποπλέειν κατὰ βίου τε καὶ γῆς ζήτησιν, ἐς ὃ ἔθνεα πολλὰ παραμειψαμένους ἀπικέσθαι ἐς Ὀμβρικούς, ἔνθα σφέας ἐνιδρύσασθαι πόλιας καὶ οἰκέειν τὸ μέχρι τοῦδε. [7] ἀντὶ δὲ Λυδῶν μετονομασθῆναι αὐτοὺς ἐπὶ τοῦ βασιλέος τοῦ παιδός, ὅς σφεας ἀνήγαγε, ἐπὶ τούτου τὴν ἐπωνυμίην ποιευμένους ὀνομασθῆναι Τυρσηνοὺς. Λυδοὶ μὲν δὴ ὑπὸ Πέρσῃσι ἐδεδούλωντο.

<<https://sacred-texts.com/cla/hh/hh1090.htm>>

[3] 'In the reign of Atys son of Manes there was great scarcity of food in all Lydia. For a while the Lydians bore this with what patience they could; presently, when the famine did not abate, they looked for remedies, and different plans were devised by different men. Then it was that they invented the games of dice and knuckle-bones and ball and all other forms of game except dice, which the Lydians do not claim to have discovered. [4] Then, using their discovery to lighten the famine, every other day they would play for the whole day, so that they would not have to look for food, and the next day they quit their play and ate. This was their way of life for eighteen years. [5] But the famine did not cease to trouble them, and instead afflicted them even more. At last their king divided the people into two groups, and made them draw lots, so that the one group should remain and the other leave the country; he himself was to be the head of those who drew the lot to remain there, and his son, whose name was Tyrrhenus, of those who departed. [6] Then the one group, having drawn the lot, left the country and came down to Smyrna and built ships, in which they loaded all their goods that could be transported aboard ship, and sailed away to seek a livelihood and a country; until at last, after sojourning with one people after another, they came to the Ombrici, where they founded cities and have lived ever since. [7] They no longer called themselves Lydians, but Tyrrhenians, after the name of the king's son who had led them there.'

Herodotus: *The Histories*, with an English translation by A. D. Godley. Cambridge: Harvard University Press 1920.

‘In the reign of Atys the son of Manes their king there came to be a grievous dearth over the whole of Lydia; and the Lydians for a time continued to endure it, but afterwards, as it did not cease, they sought for remedies; and one devised one thing and another of them devised another thing. And then were discovered, they say, the ways of playing with the dice and the knucklebones and the ball, and all the other games excepting draughts (for the discovery of this last is not claimed by the Lydians). These games they invented as a resource against the famine, and thus they used to do: on one of the days they would play games all the time in order that they might not feel the want of food, and on the next they ceased from their games and had food: and thus they went on for eighteen years. As however the evil did not slacken but pressed upon them ever more and more, therefore their king divided the whole Lydian people into two parts, and he appointed by lot one part to remain and the other to go forth from the land; and the king appointed himself to be over that one of the parts which had the lot to stay in the land, and his son to be over that which was departing; and the name of his son was Tyrsenos. So the one party of them, having obtained the lot to go forth from the land, went down to the sea at Smyrna and built ships for themselves, wherein they placed all the movable goods which they had and sailed away to seek for means of living and a land to dwell in; until after passing by many nations they came at last to the land of the Ombricans, and there they founded cities and dwell up to the present time: and changing their name they were called after the king's son who led them out from home, not Lydians but Tyrsenians, taking the name from him.’

The History of Herodotus: translated into English by George Campbell Macaulay. London: Macmillan and Co. 1904.

REFERENCES

- Adrados, Francisco R. 1989. “Etruscan as an IE Anatolian (but not Hittite) Language”. *Journal of Indo-European Studies* 17, 363–383.
- Adrados, Francisco R. 1994. “More on Etruscan as an IE-Anatolian Language”. *Historische Sprachforschung* 107, 54–76.
- Adrados, Francisco R. 2005. “El etrusco como indoeuropeo anatolio: viejos y nuevos argumentos”. *Emerita* 73, 45–56.
- Akulov, Alexander. 2024. “Etruscan is related to Hurro-Urartian”. *Cultural Anthropology and Ethnosemiotics* 10/1, 40–44.
- Arutjunjan, N.V. 2001. *Korpus urartskix klinoobraznyx nadpisej*. Erevan: Izdatel'stvo Gitutjun – Nacional'naja akademija nauk Respubliki Armenii.
- d'Aversa, Arnaldo. 1994. *Dizionario della lingua etrusca*. Brescia: Paideia Editrice.
- Beekes, Robert S.P. 2003. *The Origin of the Etruscans*. Amsterdam: Koninklijke Nederlandse Akademie van Wetenschappen.
- Berger, Hermann. 1998. *Die Burushaski-Sprache von Hunza und Nagir, Teil III: Wörterbuch Burushaski-Deutsch, Deutsch-Burushaski*. Wiesbaden: Harrassowitz (Neuindische Studien, Bd. 13).
- Blažek, Václav. 1999. *Numerals. Comparative-Etymological Analyses and their Implications*. Brno: Masarykova univerzita.

- Blažek, Václav. 2007. “Theonymica Helleno-Semitica II: Pallas Athānā / Athānā Potnia Virgin or Lady?”. *DO-SO-MO. Fascicula Mycenologica Polona* 7, 161–176.
- Blažek, Václav. 2010a. “Hurrian Numerals”. In: *Issledovanija po lingvistike i semiotike. Sbornik statej k jubileju Vjač. Vs. Ivanova*, red. T.M. Mixajlova. Moskva: Jazyki slavjanskix kutur, 117–123.
- Blažek, Václav. 2010b. “Czechs = Young Men?”. In: *Sprache und Leben der frühmittelalterlichen Slaven. Festschrift für Radoslav Katičić zum 80. Geburtstag*, hrg. von Elena Stadnik-Holzer & Georg Holzer. Frankfurt am Main: Peter Lang (Schriften über Sprachen und Texte 10, 224) 2010, 13–28.
- Blažek, Václav. 2014. “Etymologizing ‘unetymologizable’ Greek dendronyms”. *Graeco-Latina Brunensia* 19/1, 41–50.
- Blažek, Václav. 2016. “Artemis and her family”. *Graeco-Latina Brunensia* 21/2, 37–48.
- Blažek, Václav. 2017. “Apollo the Archer”. In: *Ancient Greek Linguistics: New Approaches, Insights, Perspectives*, ed. by Felicia Logozzo & Paolo Poccetti. Berlin – Boston: Walter de Gruyter, 643–662.
- Blažek, Václav. 2022. “Greek γέφυρα ”dam; bridge“”. In: *Greek – Latin – Slavic. Aspects of Linguistics and Grammatography*, edited by Barbora Machajdíkova & Ľudmila Eliášová Buzássyová. Tübingen: Narr, 141–160.
- Bonfante, Giuliano & Larissa. 2002. *The Etruscan Language: An Introduction*². Manchester – New York: Manchester University Press.
- Bork, Ferdinand. 1909. “Die Mitanni-Sprache”. *Mitteilungen der Vorderasiatische Gesellschaft*, 14 (Berlin: Peiser).
- Brandenstein, Wilhelm. 1948. “Tyrrhener”. In: *Paulys Realencyclopädie der classischen Altertumswissenschaft VIIA*. 2nd ed. by Konrat Ziegler. Stuttgart: Metzler, cc. 1909–1938.
- Brugnatelli, Vermondo. 1982. “Questioni di morfologia e sintassi dei numerali cardinali semitici”. *Firenze: La nuova Italia editrice*.
- Buck, Carl D. 1949. *A Dictionary of Selected Synonyms in the Principal Indo-European Languages*. Chicago – London: The University of Chicago Press.
- Bugge, Sophus. 1883. *Beiträge zur Erforschung der etruskischen Sprache*. Stuttgart: Heitz.
- Carter, Jesse B. 1908. “Roma Quadrata and the Septimontium”. *American School of Classical Studies in Rome* 12/2, 172–183.
- Diakonoff, Igor M. 1971. *Hurrisch und Urtäisch*. München: Kitzinger (Münchener Studien zur Sprachwissenschaft, Beiheft Neue Folge 6, translated from Russian by Karl Sdrembek).
- Diakonoff, Igor M. & Starostin, Sergei A. 1986. *Hurro-Urartian as an Eastern Caucasian Language*. München: Kitzinger (Münchener Studien zur Sprachwissenschaft, Beiheft 12. Neue Folge).
- D’jakonov, Igoř M. 1961. “Sravnitel’no-grammatičeskij obzor xurritskogo i urartskogo jazykov”. In: *Peredneaziatskij sbornik – Voprosy xettologii i xurritologii*, red. I.M. Djakonov & G.V. Cereteli. Moskva: Izdatel’stvo vostočnoj literatury, 369–423.
- D’jakonov, Igoř M. 1963. *Urartskie pišma i dokumenty*. Moskva – Leningrad: Izdatel’stvo Akademii nauk SSSR.
- D’jakonov, Igoř M. 1967. “Xurritskij i urartskij jazyki”. In: *Jazyki drevnej Perednej Azii*, ed. by Igoř M. D’jakonov. Moskva: Nauka, 111–165.
- D’jakonov, Igoř M. 1978. “Xurrito-urartskij i vostočno-kavkazskie jazyki”. In: *Hin arevelk’ / Drevnij Vostok III*. Erevan: Izdatel’stvo Akademii nauk Armjanskoj SSR, 25–38.

- D'jakonov, Igoř M. 1979. "Xurritskij i urartskij jazyki". In: *Jazyki Azii i Afriki III: Jazyki drevnej Perednej Azii (nesemitskie). Iberijskokavkazskie jazyky. Paleoaziatskie jazyki*, red. G.D. Sanžeev, A.N. Kononov, I.F. Vardul'. Moskva: Glavnaja redakcija vostočnoj literatury, 50–78.
- D'jakonov, Igoř M. & Starostin, Sergej A. 1988. "Xurrito-uratskie i vostočnokavkazskie jazyki". In: *Drevnij Vostok: Ėtnokul'turnye svjazi*, red. G.M. Bongard-Levin & V.G. Ardzinba. Moskva: Glavnaja redakcija vostočnoj literatury izdatel'stva 'Nauka', 164–207.
- DQEQ = *Diccionario quechua - español – quechua*. Cusco: Gobierno regional Cusco 2005.
- Eichner, Heiner. 2012. "Neues zur Sprache der Stele von Lemnos (Erster Teil)". *Journal of Language Relationship / Voprosy jazykovogo rodstva* 7, 9–32.
- Eichner, Heiner. 2013 "Neues zur Sprache der Stele von Lemnos (Zweiter Teil)". *Journal of Language Relationship / Voprosy jazykovogo rodstva* 10, 1–42.
- Eichner, Heiner. 2019. "Die Stele Lemnia: Vorstellung ihrer neuen Interpretation samt angestrebter Beweisführung". In: *'And I Knew Twelve Languages'". A Tribute to Massimo Poetto on the Occasion of His 70th Birthday*, edited by Natalia Bolatti Guzzo and Piotr Taracha. Warsaw: Agade Bis - University of Warsaw - Faculty of Oriental Studies, 91–133.
- Erman, Adolf & Grapow, Hermann. 1971. *Wörterbuch der ägyptischen Sprache*, I–V. Berlin: Akademie-Verlag.
- Facchetti, Giulio M. 2002. *Appunti di morfologia etrusca. Con un'appendice sulle questioni delle affinità genetiche dell'etrusco*. Rome: Olschki.
- Facchetti, Giulio M. 2005. "The Interpretation of Etruscan texts and its Limits". *Journal of Indo-European Studies* 33/3–4, 359–388.
- Facchetti, Giulio M. 2008. "Etruskisch". In: *Wieser Enzyklopädie: Sprachen des europäischen Westens*, I. Band (A–I), herausgegeben von Ulrich Ammon & Harald Haarmann. Klagenfurt: Wieser Verlag, 221–235.
- Facchetti, Giulio M. 2018. "I numerali hurrici: note etimologiche e interlinguistiche". In: *EXPRESSIO. Rivista di Linguistica, Letteratura e Comunicazione* 2, 13–22.
- Faulkner, Raymond O. 1962/1991. *A Concise Dictionary of Middle Egyptian*. Oxford: Griffith Institute – Ashmolean Museum.
- Festus, Sextus Pompeius: *De verborum significatione quae supersunt cum Pauli epitome*, emendata et annotata a Carolo Odofredo Muellero. Leipzig: Sumptibus Simmelii eiusque socii 1880.
- Festus, Sextus Pompeius: *De verborum significatu quae supersunt cum Pauli epitome*, edidit Æmilius Thewrewk de Ponor. Budapest: Sumptibus academiae litterarum hungaricae 1889.
- Fiesel, Eva. 1936. "Bemerkungen und Berichtigungen". *Studi Etruschi* 10, 321–325.
- Francfort, H.-P., Tremblay, X. 2010. "Marhaši et la civilisation de l'Oxus". *Iranica Antiqua* 45, 51–224.
- Friedrich, Johannes. 1934. "Mitraššil Uruṇanaššel". *Orientalia* 12, 311–317.
- Friedrich, Johannes. 1935. "Zum Subaräischen und Urartäischen". *Analecta Orientalia* 12, 122–135.
- Georgiev, Vladimir I. 1962. *Hethitisch und Etruskisch. Die hethitische Herkunft der etruskischen Sprache*. Sofia: Academie Bulgare des Sciences (*Linguistique Balkanique* 5.1).
- Georgiev, Vladimir I. 1974. "Etruskisch und Hethitisch. Ein Vergleich der bekannten Tatsachen der etruskischen Grammatik." *Balkansko ezikoznanie* 17/1, 5–40.
- Georgiev, Vladimir I. 1979. *La lingua e l'origine degli Etruschi*. Roma: Editrice Nagard.
- Gluhak, Alemko. 1978. "Etruscan Numerals". *Linguistica* 17/1, 25–32.

- Hadas-Lebel, Jean. 2017. “Le nombre ‘cent’ en étrusque. À propos de étr. *sran*.” *Rasenna* 5, 1–7.
- HAL = *The Hebrew and Aramaic Lexicon of the Old Testament*, I-II, by Ludwig Koehler & Walter Baumgartner, translated by M.E.J. Richardson. Leiden – Boston – Köln: Brill, 2001.
- Halloran, John Alan. 2006. *Sumerian Lexicon*. Los Angeles: Logogram Publishing.
- Hoenigswald, Henry M. 1941. “On Etruscan and Latin Month-Names”. *The American Journal of Philology* 62/2, 199–206.
- Hrozný, Bedřich. 1915. “Die Lösung des hethitischen Problems”. *Mitteilungen der Deutschen Orient-Gesellschaft* 56, 17–50.
- Hrozný, Bedřich. 1917. *Die Sprache der Hethiter – ihr Bau und ihre Zugehörigkeit zum indogermanischen Sprachstamm*. Leipzig: Hinrich.
- Hrozný, Bedřich. 1928. “Etruskisch und die ‘hethitischen’ Sprachen”. *Zeitschrift für Assyriologie* 38, 171–180.
- Ivanov, Vjačeslav V. 1988. “Drevnevostočnye svjazi étrusskogo jazyka”. In: *Drevnij Vostok: Étnokul’turnye svjazi*, red. G.M. Bongard-Levin & V.G. Ardzinba. Moskva: Glavnaja redakcija vostočnoj literatury izdatel’sstva ‘Nauka’, 208–218.
- Jacemirskij, Sergej. 2007. “Étrusskie čislitel’nye: problemy i itogi issledovanija”. In: *Aspekty komparativistiki 2: Orientalia et Classica. Trudy Instituta vostočnyx kul’tur i antičnosti XI*. Moskva: Rossijskij gosudarstvennyj humanitarnyj universite, 187–196.
- Kapancjan, Gr. 1936. *Obščie élementy meždu urartskim i xettskim jazykami*. Èrevaň: Gosudarstvennoe Izdatel’stvo SSRA.
- Kassian, Alexei. 2010 [2011]. “Hurro-Urartian from the lexicostatistical viewpoint”. *Ugarit-Forschungen. Internationales Jahrbuch für die Altertumskunde Syrien-Palästinas* 42, 383–452.
- Klimov, Georgij A. 1967. “Zaimstvovannye čislitel’nye v obščekartvel’skom?” *Étimologija* 1965, 307–310.
- Kloekhorst, Alwin. 2008. *Etymological Dictionary of the Hittite Inherited Lexicon*. Leiden – Boston: Brill.
- Kloekhorst, Alwin. 2021. “Luwians, Lydians, Etruscans, and Troy: The linguistic landscape of Northwestern Anatolia in the pre-classical period”. In: *The Political Geography of Western Anatolia in the Late Bronze Age*, ed. Ivo Hajnal, Eberhard Zangger, Jorrit Kelder. Budapest: Archaeolingua (Series Minor, 45), 49–75.
- Krahmalkov, Charles R. 2000. *Phoenician-Punic Dictionary*. Leuven: Peeters en Department Oosterse Studies.
- Kretschmer, Paul. 1921. “Pelasger und Etrusker”. *Glotta* 11, 276–285.
- Kretschmer, Paul. 1929. “Υπηρία”. *Glotta* 18, 110–111.
- Lejeune, Michel. 2009. “Les termes de la numération. Notice rédigée par M. Lejeune pour le ThLE II”. In: *Autour de Michel Lejeune*. Actes des Journées d’étude organisées à l’Université Lumière Lyon 2 – Maison de l’orient et de la Méditerranée (2–3 février 2006), ed. Frédérique Biville et Isabelle Boehm. Lyon: Maison de l’orient et de la Méditerranée – Jean Pouilloux (Collection de la Maison de l’orient, 43), 308–317.
- Liddell, Charlton T. & Scott, Charles. 1958. *A Latin Dictionary*. Oxford: Clarendon Press.
- Laroche, Emmanuel. 1980. *Glossaire de la langue hourrite*. Paris: Klincksieck.
- Mańczak, Witold. 1983. “Das etruskische Numerale *ša*”. *Glotta* 61/1–2, 103–105.
- Martirosyan, Hrach K. 2010. *Etymological Dictionary of the Armenian Inherited Lexicon*. Leiden – Boston: Brill.
- NCED = *A North Caucasian Etymological Dictionary*. Sergei L. Nikolayev & Sergei A. Starostin. Moscow: Asterisk Publishers 1994.

- Neu, Erich. 1991. "Etruskisch – eine indogermanische Sprache Altanatoliens?". *Historische Sprachforschung* 104, 9–28.
- Nozadze, Nana. 2007. *Leksika xurritskogo jazyka / Vocabulary of the Hurrian Language*. Tbilisi: Tbilisi State University.
- Orel, Vladimir & Starostin, Sergei. 1990. "The position of Etruscan". In: *Proto-Languages and Proto-Cultures*, ed. by Vitaly Shevoroshkin. Bochum: Brockmeyer, 60–66.
- Oštir, Karl. 1921. *Beiträge zur alarodischen Sprachwissenschaft I*. Wien: Beyer.
- Pallotino, Maximo. 1968. *Testimonia linguae etruscae*². Firenze: La nuova Italia.
- Patri, Sylvain. 2009. "Aspects de la position génétique du hurrite". In: *General Studies and Excavations at Nuzi 11/2 in Honor of David I. Owen on the Occasion of his 65th Birthday*. (October 28, 2005), ed. by Gernot Wilhelm. Bethesda: CDL Press, 341–354.
- Pauli, C. 1882. "Die etruskischen Zahlwörter". *Etruskische Forschungen und Studien III*, ed. W. Deecke & C. Pauli. Stuttgart: Heitz.
- Platner, Samuel Ball. 1901. "The Pomerium and Roma Quadrata". *The American Journal of Philology* 22/4, 420–425.
- Richter, Thomas. 2012. *Bibliographisches Glossar des Hurritischen*. Wiesbaden: Harrassowitz.
- Richter, Thomas. 2016. *Vorarbeiten zu einem hurritischen Namenbuch*, Erster Teil: *Personennamen altbabylonischer Überlieferung vom Mittleren Euphrat und aus dem nördlichen Mesopotamien*. Wiesbaden: Harrassowitz.
- Rix, Helmut. 1991. "Etrusco un, une, unu «te, tibi, vos» e le preghiere dei rituali paralleli nel Liber Linteus". *Archeologia Classica*, 43/1: *Miscellanea etrusca e italica in onore di Massimo Pallottino*, 665–691.
- Rix, Helmut. 1998. *Rätisch und Etruskisch*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck (Innsbrucker Beiträge zur Sprachwissenschaft. Vorträge und kleinere Schriften, 68).
- Rix, Helmut. 2004. "Etruscan". In: *The Cambridge Encyclopedia of the World's Ancient Languages*, by Roger Woodard. Cambridge: University Press, 943–966.
- Robertson, Ed. 2006. *Etruscan's genealogical linguistic relationship with Nakh-Daghestanian: A preliminary evaluation*. Ms.
<<https://www.theelen.info/%5B20151101%5D%20Etruscan%20numerals%20-2-.pdf>>
- Salomon, Corinna. 2020. "Raetic". *Palaeohispanica* 20, 263–298.
- Salvini, Mirjo & Wegner, Ilse. 2014. *Einführung in die urartäische Sprache*. Wiesbaden: Harrassowitz.
- Schneider, Thomas. 2003. "Kassitisch und Hurro-Urartäisch: Ein Diskussionsbeitrag zu möglichen lexikalischen Isoglossen". *Altorientalische Forschungen* 30, 372–381.
- Simon, Zsolt. 2021. "The alleged Anatolian loanwords in Etruscan: A reconsideration". In: *Studies in the languages and language contact in Pre-Hellenistic Anatolia*. Federico Giusfredi & Zsolt Simon, eds., with the editorial assistance of Elena Martínez Rodríguez.
- Skutsch, F. 1895. "Zu den etruskischen Zahlwörtern". *Indogermanische Forschungen* 5, 256–265.
- Slotty, Friedrich. 1937. "Die etruskischen Zahlwörter". *Archiv orientální* 9, 379–404.
- Slunečko, Vladimír. 2003. "Etruské číslovky {status quaestionis}". *Auriga* 45, 25–34.
- Smeets, Rieks. 1989. "On Hurro-Urartian as an Eastern Caucasian Language". *Bibliotheca Orientalis* 46, No 3/4, cols. 259–279.
- Starostin, Sergei A. 1995. "Sravnitel'nyj slovar' enisejskix jazykov". In: *Ketskij sbornik: Lingvistika*. Sergei Starostin, ed. Moskva: Vostočnaja literatura, 176–316.
- Steinbauer, Dieter H. 1999. *Neues Handbuch des Etruskischen*. Scripta Mercaturae Verlag.

- Stephani Byzantii. *Ethnica*, volumen iv: π-υ, recensuerunt Germanice verterunt adnotationibus indicibusque instruxerunt Margarethe Billerbeck et Arlette Neumann-Hartmann. Berlin – Boston: De Gruyter 2016.
- Stephanos Byzantini. 1849/1992. *EΘNIKΩΝ. A geographical lexicon on ancient cities, peoples, tribes and toponyms*. A. Meineke, ed. Berlin: Reimer / reprint Chicago: Ares Publishers.
- Stoltenberg, Hans L. 1943. “Die Bedeutung der etruskischen Zahlenamen”. *Glotta* 30, 234–244.
- Thomsen, Vilhelm. 1899. “Remarques sur la parenté de la langue étrusque”. *Det Kongelige Danske Videnskabernes Selskabs* 1899, 373–398.
- Toporov, Vladimir N. 1983. “K semantike četveričnosti (anatolijskoe **meu-* i dr.)”. *Ėtimologija* 1981, 108–130.
- Trask, Robert Lawrence (Larry). 2008. *Etymological Dictionary of Basque*. Edited for web publication by Max W. Wheeler. Falmer (Brighton): University of Sussex.
- Trombetti, Alfredo. 1922–23. *Elementi di glottologia*. Bologna: Zanichelli.
- Trombetti, Alfredo. 1928. *La lingua etrusca. Grammatica. Testi con commento. Saggi di traduzione interlineare. Lessico*. Firenze: Rinascimento del libro.
- Urbanová, Daniela & Blažek, Václav. 2008. *Národy starověké Itálie, jejich jazyky a písma*. Brno: Host.
- Wallace, Rex E. 2008. *A Manual of Etruscan*. Ann Arbor – New York: Beech Stave Press.
- Wegner, Ilse. 2007. *Einführung in die hurritische Sprache*². Wiesbaden: Harrassowitz.
- Wilhelm, Gernot. 2004a. “Hurrian”. In: *The Cambridge Encyclopedia of the World's Ancient Languages* by Roger Woodard. Cambridge: University Press, 95–118.
- Wilhelm, Gernot. 2004b. “Urartian”. In: *The Cambridge Encyclopedia of the World's Ancient Languages*. Roger Woodard, ed. Cambridge: University Press, 119–137.
- Woudhuizen, Fred C. 2008. *Etruscan as a Colonial Luwian Language: The Comprehensive Version*. Amsterdam: Dutch Archaeological and Historical Society.

Václav Blažek
 Department of Linguistics & Baltic Studies
 Masaryk University, Brno – Czech Republic
blazek@phil.muni.cz

THE PUZZLE OF PAN-AFRICAN ROOTS

ROGER BLENCH

Department of Indigenous Affairs,
Government of Arunachal Pradesh
McDonald Institute for Archaeological Research
University of Cambridge
Department of History, University of Jos
Kay Williamson Educational Foundation
8, Guest Road
Cambridge CB1 2AL
United Kingdom
Voice/ Ans (00-44)-(0)1223-560687
Mobile worldwide (00-44)-(0)
E-mail rogerblench@yahoo.co.uk
<http://www.rogerblench.info/RBOP.htm>

Jos, 1 April, 2026

TABLE OF CONTENTS

1. Introduction	74
2. Data tables	75
3. Conclusions	91
Abbreviations	92
References	94

TABLES

Table 1: Acronyms for language-specific sources	92
---	----

ACRONYMS AND CONVENTIONS

#	quasi-reconstruction
C	consonant
IPA	International Phonetic Association
N	nasal
V	vowel

1. INTRODUCTION

African language phyla are generally ‘well-behaved’ in the sense that it is relatively easy to assign a given language to a phylum, based on its core lexicon and morphology. Africa has a relatively small number of isolates (Blench 2017), given that it is the centre of dispersal for modern humans. Until recently, it was deemed to have four phyla, Khoisan being counted as a unit. More recently, Khoisan has been split into three independent phyla. Nilo-Saharan remains the most problematic phylum and not all researchers accept either a) that it is a unit or b) that all purported branches actually fit together (Blench 2026). On the other hand, there has been a persistent thread of argument for uniting Nilo-Saharan and Niger-Congo into a single macrophylum, Niger-Saharan which would join the two families (see Blench 1995 and references up to that date). Westermann (1911) in his first version of the *Sudan Sprachen* put together Nilotic and languages we would regard as Niger-Congo. However, he later seems to have dropped this idea, since he makes no reference to it in his later *Sudansprachen* (Westermann 1927).

In the last few decades there have been various further attempts to bring together Niger-Congo and Nilo-Saharan, including Gregersen (1972), Creissels (1981), Boyd (1999). Hall & Hall (1974) were the first authors to draw attention to the striking similarities of the ATR vowel systems in Niger-Congo and Nilo-Saharan. Unfortunately, all this is rather inconclusive, despite the massive expansion of data over this period. Personally, although my earlier papers argued for a genetic

connection between the two phyla, I now consider the similarities to be the outcome of early borrowing. One reason for that is that the homelands of the two phyla seem to be at opposite sides of Africa, Nilo-Saharan in the Ethio-Sudan borderlands and Niger-Congo in the west of West Africa.

The existence of macrophyla, such as Nostratic or Transeurasian, will always be controversial, in part because it is difficult to disentangle ancient loanwords from true cognates beyond a certain time-depth. So it is with Niger-Saharan; there are certainly common lexical items and strong evidence for typological commonalities, such as ATR vowels. When I wrote my original papers these seemed to me convincing evidence for genetic affiliation. However, I have subsequently come to see that many of the lexical items are also attested in Afroasiatic and more rarely in Khoisan and even the isolates. If so, then some or all of these similarities may be due to ancient loans, although when and where this transmission took place will remain uncertain.

In a first approach to this topic, I focused on crabs, turtles and frogs, at first sight rather unlikely candidates for Pan-Africanisms (Blench 1997). The argument was that, although these are marginal foods at present, this type of aquatic fauna was highly salient at the time when foraging groups were dispersing, and when human beings had yet to learn to fish in open water. This may have something going for it, but subsequent research has revealed some widespread roots which also have Pan-African distributions, but which seem largely arbitrary. Why, for example, should ‘knee’ be so widely attested and not ‘arm’ or ‘leg’?

The purpose of this paper is simply to compile datasets for the most common Pan-African roots and to see what conclusions can be drawn. The data for crabs, turtles and frogs, appropriately updated, is reprised here. §2 consists of the data tables, and the conclusions in §3 ask what emerges from this exercise.

2. DATA TABLES

The data tables have been gradually compiled over many years, as part of a broader exercise to identify common Niger-Congo roots. The attestations are divided by phylum, with isolates cited first. The sources are abbreviated, and the compressed references are keyed in the references at the end. Major proposals for Niger-Congo reconstruction, namely Westermann (1927) [Proto-Western Sudanic] and Mukarovsky (1976-77) [Proto-Western Nigritic] are given although I would not now propose the forms they present. Among other things, they were unaware of the Dogon family of languages, which plausibly provide evidence for the earliest stages of Niger-Congo.

1. come I #6woN-

Ph	Family	Group	Language	Attestation	Comment	Source
AA	Chadic	West	Ywom	bəl	come	JI
AA	Chadic	West	Kulere	bo	come	JI
AA	Chadic	West	Tsagu	bàà	come	JI
AA	Chadic	Central	Tera	ḡa	come	JI
AA	Chadic	East	Mawa	oobon	come	Roberts (p.c.)
NS	ES		proto-Daju	*ḡoN	come	RCS
NS	ES		proto-Nilotic	*ḡun pl. *pə	come	D
NS	ES	Nilotic	Nuer	ben	come	RCS

Ph	Family	Group	Language	Attestation	Comment	Source
NS	ES	Nilotic	Padang Dinka	bɔ	come	RCS
NS	ES		Mabaan	bɛɛd	come	RCS
NS	ES		Nara	wɔ	come	Ed91
NS	Fur		Fur	bo-	reach	Ja90
NC			PWS	bía, bá	<i>kommen</i>	W
NC			PWN	pík- (bhik-?)	arrive, come	M
NC	Dogon		Walo	wá	<i>venir</i>	RMB
NC	Dogon		Toro So	vè	<i>venir</i>	C-G
NC	Ijò		Proto-Ijò	*ḡó	come	KW
NC	Mande		Guro	ḡo	<i>venir</i>	ALMCI
NC	Kru		Seme	bɛ	<i>venir</i>	ALKrCI
NC	Atlantic		Fulfulde	ḡad-	come	No89
NC	Atlantic		Temne	-fàt	come near, approach	Wi07
NC	Atlantic	North	Mankanya	bi, ban	<i>arriver</i>	Sg
NC	Atlantic	North	Diembereng	ben	come	Wi07
NC	Atlantic	North	Balanta	-bēn	<i>venir</i>	Sg
NC	Gur		Deg	bà	<i>venir</i>	ALGCI
NC	Gur		Bwamu	bwala	<i>venir</i>	ALGCI
NC	Adamawa		Samba Leeko	bādīn	<i>être proche</i>	Fa04
NC	Adamawa		Gimbe	ḡók	<i>arriver</i>	Di16
NC	Adamawa		Wiyaa	ba-	come	KI90
NC	Ubangian		Mba-ne	ḡó-	<i>venir</i>	Mo88
NC	Kwa	Guan	Gonja	bà	come	Sn89
NC	Kwa		Avikam	bá	<i>arriver</i>	JR
NC	Kwa		Twi	bà	come	Chr33
NC	Kwa		Ga	bà	come	Kr73
NC	VN	Gbe	Ewe	vá		JR95
NC	VN	Yoruboid	Yoruba	èbá	proximity	Ab58
NC	VN	Igboid	Ọ̀nị́ca Igbo	-bía	come	Wi72
NC	VN	Nupoid	Nupe	bé	come	Ba14
NC	VN	Idomoid	Etulo	ba	come	Ar83
NC	EBC	Ukaan	P-Ukaan	*wàg	come	Ab99
NC	EBC	Plateau	Pe	bén	come	RMB
NC	EBC	Plateau	Eloyi	bó	come	Ar83
NC	EBC	Jukunoid	Kuteb	bá	come	Robert Koops
NC	Bantoid	Dakoid	Sama Mum	báá	come	BS10
NC	Bantoid	Tivoid	Tiv	vá	come	Ab40
NC	Bantoid	Ekoid	Ejagham	ḡá	come	John Watters
NC	Bantoid	Beboid	Bebe	bwó	come	SIL wordlists
NC	Bantu		PB	bút	come, go back	BLR3

Commentary: Not attested in Afroasiatic outside Chadic and thus almost certainly a borrowing. Dimmendal (1988:35) notes that the irregular plural **pɔ* must be reconstructed to Proto-Nilotic and forms with initial p- appear elsewhere, hinting at a still greater time-depth. The Kru form is

unique and Seme is isolated, so this may be loan, as is the unique citation in Ubangian. The Bantu form makes it only doubtfully cognate. Westermann (1927: 209) noted that this word frequently shows up as a future auxiliary in Niger-Congo languages. Palatalisation is scattered throughout Niger-Congo but nowhere forms a consistent pattern. The *#ba* form is so persistent in Niger-Congo that it seems likely that is where it originated. The appearance in Chadic and only two branches of Nilo-Saharan suggest ancient borrowing. There are also parallels in Indo-European, e.g. French *venir*.

Refs: W. 209; Ar64: 36; M. 435; E. 563;

2. to split, cut, break *#keri*

Ph	Family	Subgroup	Language	Attestation	Comment	Source
KS	Southern		!Xóǀ	t'kx'āla	open meat to flat- ten it	Tr94
KS	Central		Proto-East Khoe	*kade	cut flesh in strips	Vo97
AA	Omoti		proto-Dizoid	k'aat'-	cut	Be03
AA	Agaw		Bilin	kər	break	Ap06
AA	Cushitic	South	Dahalo	k'eer-	chop	To91
AA	Cushitic	South	Iraqw	qeet	be broken	KM03
AA	Cushitic	East	Arbore	k'uur-	cut	Be20
AA	Chadic	Masa	Masa	káná	cut	JI
AA	Chadic	Central	Dghwede	klà	cut	JI
AA	Chadic	West	Warji	kón	cut	JI
AA	Chadic	West	Ngas	can	but	JI
AA	Berber		Kabyle	cerref	<i>couper</i>	Da82
NS	Kuliak		Ik	ka-ukot-	cut	He99
NS	Koman		Uduk	kwal	split	Ot19
NS	Berta		Undu	'kíírà	split hard sub- stance	RCS
NS	ES	Nubian	Dongola	gor	cut grass	RCS
NS	ES		Nyimang	kír	cut	RCS
NS	Fur		Fur	karr-	split	Ja90
NS	Saharan		Beria	kót	<i>couper</i>	JC04
NS	Saharan		Daza	kər	cut	Le50
NS	Songhay		Gao	körtù-	tear, split	Pr77
NC			PWS	ká-	<i>schneiden, ab- brechen</i>	W
NC			PWS	kuà	<i>schnitzen</i>	W
NC			PWN	cet-	cut	M
NC			PWN	kél-	cut	M
NC			PWN	ken-	split	M
NC			PWN	khant-	cut (off)	M
NC			PWN	kwèk-(kwyèk-)	cut	M
NC	Dogon		Ana	keðele	cut	RMB

Ph	Family	Subgroup	Language	Attestation	Comment	Source
NC	Dogon		Toro Tegu	ʃɛ	cut throat	Heath (p.c.)
NC	Ijoid		P-Ijɔ	kãã/kãõ	tear	KW
NC	Ijoid		Defaka	kakara	tear	Connell (p.c.)
NC	Ijoid		Defaka	kée	cut	Connell (p.c.)
NC	Mande		Bambara	kári	cut	Ba96
NC	Kordofanian	Talodi	Dagik	kəra	cut, scratch	Va16
NC	Atlantic	North	Wolof	xar	tear, cut	MG91
NC	Atlantic	South	Bullom	keth-	<i>couper</i>	Guillaume Seegerer
NC	Atlantic	South	Kisi	kəl	cut	Ch00
NC	Bijogo		Bijogo	-kɔp(i)	<i>abattre</i>	Sg02
NC	Bijogo		Bijogo	-kɛnk(i)	<i>couper, tailler</i>	Sg02
NC	Kru		Bakwe	ká	fell, cut down	Ye09
NC	Kru		Nyabwa	cei	cut	ALKrCI
NC	Adamawa		Galke	kəŋ	<i>couper</i>	La62
NC	Adamawa		Mumuye	ka	break	Sh83
NC	Adamawa		Mumuye	ko	cut in two	Sh83
NC	Adamawa	Jen	Jen	kə	cut	N117
NC	Ubangian		Ndunga-le	-kələ-	<i>couper</i>	Mo88
NC	Gur		Kulango	kɔrɪ	<i>couper</i>	Mi07
NC	Kwa		Igo	klô	break pot	JR
NC	Kwa		Ga	kɔ́	break off	Kr73
NC	VN		Nupe	ke	be split, be cut	Ba14
NC	EBC		Tarok		break, cut	RMB
NC	EBC	Jukunoid	Hone	káp	break (stick)	Sh80
NC	EBC	Jukunoid	Wapha	kəm	break	Sh80
NC	Bantoid	Furu	Bikyak	kɔ̀	<i>erbrechen</i>	Kiessling (p.c.)
NC	Bantoid	Mambiloid	Mambila	ker3	cut	PM95
NC	Bantoid	Dakoid	Nnakenyare	kék	cut	BS10
NC	Bantoid	Ekoid	Nkim	kēm	cut	Cr69
NC	Bantoid	Ekoid	Mbe	kɛ́t~ɾ	cut	JP
NC	Bantoid	Tikar	Tikar	kɛti	cut	Ja88
NC	Bantoid	Ring	Oku	kâk	‘to cut loose’	SIL wordlists
NC	Bantoid	Eastern	PEG	*kEk	cut	ELV
NC	Bantu		CB	-kəd	cut	G
NC	Bantu		CB	céc-	cut	G
NC	Bantu		PB	kád	<i>tear, cut</i>	BLR
NC	Bantu	Manenguba	Akɔɔse	-kwəl	cut	He87

Commentary: This root has almost certainly been loaned very widely in Africa. The occurrence in Chadic is most likely a localised loan, but in Cushitic and Omotic these forms are widespread (see examples under *k'er* ‘split’ and *kaal-ta* ‘axe’ in Lamberti & Sottile 1997:411, 435). This lexeme has been proposed as a ‘world etymology’, and the cognate set would presumably then including English ‘cut’ (see Bender 1997:122 for more Nilo-Saharan examples). Temiar (Aslian, Austroasiatic) has *cer* ‘to cut’. Although the first CV seems to be cognate throughout, the diversity

of C2 may be explained by reduplication and compounding. The common Manenguba forms probably are cognate if $k \rightarrow s$, but the original initial is conserved at least in Akɔɔse as an alternative. The labial-velar is only preserved in Mbe, hence its reconstruction remains doubtful, although the labialisation in Akɔɔse also suggests its presence.

Refs: W. 227+243; Ar64. 43; M. 75+202+203+256+22; G. 280+304+1029+1033+1045; Gr:80; G:97,135,154; B:122, 133; Boyd (1994:106)

3. knee I #kulu

Ph	Family	Sub-group	Language	I	II	Gloss	Source
KS	Southern		!Xóð	g xúu)			Tr94
KS	Central		Kxoe- /Anda	kúdù			Vo97
KS	Central		Shua-Cara	(kú)kúdù			Vo97
KS	Northern		Ju 'hoan	g!xàà			Di4
AA	Omoti	North	Wolaytta	gulba-ta			LS
AA	Cushitic		*PC	*gulb-/*gwilb-			E87
AA	S Cushitic		Asax	ngulu-et			KM03
AA	Cushitic	South	Dahalo	gilli			To91
AA	Cushitic	Agaw	Bilin	gərəb			Ap06
AA	Semitic	Ethiopia	Amharic	gulbāt			Hu89
AA	Chadic		Sukur	kırım			JI
AA	Chadic		Tera	xulukti			JI
NS	Kuliak		Ik	kutuŋ			He99
NS	Shabo		Shabo	hutu/kutti			DA19
NS	Koman		Kwama	dugol			Ot19
NS	Berta		Berta	guʃuŋ			Be89
NS	Kunama		Kunama	tùgà		?C	Be01
NS	ES		Kenzi	kur(ti)			RCS
NS	Kadu		Katcha	kúúgέ	pl. nu-gúúgi		Sch94
NS	Maba		Mesalit	kàdíŋó			Ed91
NS	Fur		Fur		kùrù		Ja90
NS	CS		Mangbetu	nè-káátì pl. è-			DD
NS	Saharan	East	Beria		kurru		JC04
NC			PWN		-kwudi-	knee	M
NC			PWN		-gwùl-	shin; (lower) leg	M
NC	Dogon		Tebul Ure	kúndúgó		knee	RMB
NC	Dogon		Tiranige	kúndzúgó		knee	Heath (p.c.)
NC	Ijoid		Biseni	íkóníée		knee	KW
NC	Mande		Soninke	xuruŋi/o pl. -onu		genou	VV
NC	Katla-Tima		Domurik	kuruŋa		knee	RCS
NC	Kordofanian	Talodi	Dagik	k-əŋgù /n-		knee	Va16
NC	Atlantic	North	Bedik	ε-dəkəla /ma-		knee	Wi07
NC	Atlantic	South	Gola	ké-gòlo		foot, leg	We21

Ph	Family	Sub-group	Language	I	II	Gloss	Source
NC	Bijogo		Caravela	kunu		<i>genou</i>	Guillaume Segerer
NC	Kru		Jrewe		kùrù	<i>genou</i>	Ma83
NC	Kru		Bakwe		koto	<i>genou</i>	Ye09
NC	Adamawa		Leko of Mapeo	dun		leg	Bd94
NC	Ubangian		Yakoma	lì-kũrũ		<i>genou</i>	Mo88
NC	Kwa	Guang	Chumburung	kù-ŋúrí?		knee	Sn89
NC	Kwa	GTML	Igo	ì-lùkù /a-		if metath-JR esis	
NC	VN	Gbe	Ewe		kòlí	<i>genou</i>	Ro95
NC	VN	Edoid	Emai	úgú'òè		knee [òè =SE07 leg]	
NC	VN	Igboïd	Igbo of Udi	íkperè		knee	KW
NC	EBC	Kainji	Vori	dì-rú pl. á-rú		knee	RMB
NC	EBC	Plateau	Hyam	kpurú		knee	RMB
NC	Bantoid	Buru	Buru	enú		knee	Robert Koops
NC	Bantoid	Tivoid	Batu Afi	nún		knee	Robert Koops
NC	Bantoid	Ring	Isu	ínǔ		knee	SIL Wordlists
NC	Bantu		Zones C, H etc.		#-koto, godo	-knee	BLR3
NC	Bantu		Zones E, F etc.		-kónò	forearm, leg	BLR3
NC	Bantu		CB		-gùdu	leg	G

Commentary: A preliminary version of this dataset appears in Blench (1997) and this root appears to have extra-African cognates. Gregersen (1972) treats these as two distinct sets for ‘leg’ and ‘knee’ but they are probably to be put together and the more doubtful cognates discarded. A rather different form, *BU(N)KA is proposed as a ‘world etymology’ in Bengtson & Ruhlen (1994). Dimmendaal (1988:46) reconstructs Proto-Nilotic **kɛlɪ* for leg, a semantic change that also took place in Bantoid. BLR3 says the alternation between dúí/núí ‘*est à étudier*’, but if this reconstruction is correct, the two are simply different erosional results of an original –nd-. Why this word should be so similar in so many parts of the continent and beyond is hard to explain. English ‘knee’ might be added.

Bender (1997:133) pursues linkages that include a purported PNC root **khon* for ‘knee’ and brings in Mende *kon* ‘head’ because the ‘knee as head of the leg’¹. This analysis is not used here. Within Niger-Congo, Biseni is isolated within Ijoid and may thus be a loan from a neighbouring language. East Benue-Congo seems to have a distinctive shift, the loss of the *ku-* prefix and the development of *-ru(ŋ)* as the stem. This is retained in Bantoid as least as late as Ejagham. There are two potential cognates in Bantu, the more widespread #-*kónò* which generally means ‘forearm’ and the less common #-*koto/godo* which is knee directly. These do not resemble Bantoid, and the Bantu forms

¹ A metonymy typical of Amazonia (Aikhenvald 2006)

may be cognate with other branches of Niger-Congo. It is probable that #-*kónò* is just a chance resemblance.

Refs: G.:101,123, G. 884; M. 163+304; B:133; B81:.261, Gr.:82,84,

4. knee II #-duŋ-

Ph	Family	Subgroup	Language	Attestation	Gloss	Source
AA	Chadic	A3	Mupun	fùrùm	knee	Fr92
AA	Chadic	Bole-Tangale	Tangale	purum	knee	Jl
AA	Chadic	Ron	Kulere	‘arôm	knee	Jl
NS	CS	SBB	Bongo	đũ	<i>cuisse</i>	Pierre Nougayrol
NC			PWS	-lu-	<i>Knie</i>	W
NC			PWN	-dui-	knee	M
NC			PWN	-l/dũku	joint, knee	M
NC	Katla-Tima		Kaalak	romə	knee	RCS
NC	Katla-Tima		Domurik	ku-runoo /i-	knee	RCS
NC	Atlantic	South	Mmani	i-donk /n-	knee	Wi07
NC	Bijogo		Bijogo	ku-nu /ka-	<i>genou</i>	Guillaume Segerer
NC	Pere		Pere	gbòndũyũ	<i>genou</i>	DC
NC	Gur		Proto-Central	*dun	<i>genou</i>	Ma75
NC	Gur	Oti-Volta	Konni	dũ-ŋ pl. dũné	knee	Mike Cahill
NC	Adamawa		Leko of Mapeo	dun	leg	Bo94
NC	Adamawa	Jen	Munga Doso	đwǎŋ	knee	N117
NC	Gbaya		P-Gbaya	zúgũr	<i>genou</i>	Mo95
NC	Kwa		Kebu	nũnũ-rə	knee	Jacques Rongier
NC	Bijogo		Caravela	ku-nu /ŋa-	<i>genou</i>	Guillaume Segerer
NC	VN	Nupoid	Nupe	edũ	thigh	Ba14
NC	EBC	Plateau	Fyem	đu-rũm /a-	knee	RMB
NC	EBC	Plateau	Yankam	rũŋ /a-	knee	RMB
NC	EBC	Lower Cross	PLC	*é-lóŋ	knee	Co91
NC	EBC	Upper Cross	Ubaghara	irũŋ	knee	JS
NC	Bantoid	Dakoid	Sama Mum	lũũ		BS10
NC	Bantoid	Buru	Buru	e-nũ /a-nu		Robert Koops
NC	Bantoid	Bendi	Bekwara	irĩnũŋ		RS
NC	Bantoid	Nyang	Kenyang	né-nén /má-		Mb98
NC	Bantoid	Tivoid	Batu Afi	ø-nũn /á-		Robert Koops
NC	Bantoid	Beboïd	Bukwe	ńnyũ		SIL Wordlists
NC	Bantoid	Ring	Isu	ĩŋũ		SIL Wordlists
NC	Bantoid		Ngie	ĩ-nũwyĩ /a-		SIL Wordlists
NC	Bantoid	Ekoid	Ejagham	êrũŋ	5/6	Cr69
NC	Bantoid	Mbe	Mbe	lè-lũ /bè-		JP
NC	Bantu		PB	*dũĩ		BLR3
NC	Bantu		CB	-dũ, -dũĩ-dũĩ		G

Commentary: The cognates of Bullom ‘knee’ in South Atlantic show a wide semantic range, for example, Temne ‘leg’ (*a-ləŋk*) and ‘arm’ *a-loŋk*, Baga Koba *a-raŋk* ‘thigh’. This root is widespread

in West Chadic and the #-*rum* element may have been loaned into Plateau and its prefix re-analysed. *Pere* looks like a composite of the root for ‘knee’ with the root for ‘leg’.

Refs: W. 252; M 109; G.722+729

5. tortoise, turtle #kuru

Ph	Family	Language	Attestation		Gloss	Source
—	Sandawe	Sandawe	khú	rú	tortoise	EPT12
—	Hadza	Hadza	k’õ	ló	tortoise	KM
—	Hadza	Hadza	k’úú	tá-	turtle	KM
—	Laal	Laal	kú	Nán	<i>petite tortue</i>	Boyeldieu (p.c.)
AA	S Cushitic	Gorowa	ka	n ke’	tortoise	KM03
AA	E Cushitic	Burji	ko	c’áa	tortoise, turtle	Hu89
AA	Beja	Beja	se ku	ur	tortoise	Hudson (n.d.)
AA	W. Chadic	Hausa	kùŋ	ku ruu	tortoise	Ab49
AA	W. Chadic	Mwaghavul	kú	r	tortoise	JI
AA	C. Chadic	Huba	kwà	kú rù m	tortoise	Kr81
AA	Masa	Lame	gù	rè i	tortoise sp.	Sa82
AA	E. Chadic	Toram	kùn	gù rù	turtle	Jungrauthmayr (p.c.)
AA	Berber	Kabyle	tafe	k ru rt	tortoise	Da82
KS	North	Auen	!gu	ru	tortoise-shell	Bl56
KS	Central	Naro	go	e	tortoise	Tr86
KS	Central	Mohissa	cu	ru	tortoise	Bl56
NS	Kuliak	Ik	ro ki	ro k	tortoise	He99
NS	Koman	Kwama	k’	u kif	turtle	Ot19
NS	Songhay	Songhay	ń	kúú r Á	small tortoise	BWK94
NS	Saharan	Kanuri	kó	ro wú	tortoise	Cy94
NS	Maba	Maba	fa k	ruu n	tortoise	Ed91
NS	Surmic	Didinga	bo-	ko l	tortoise	RCS
NS	ES	Dinka	le-	ku r	tortoise	RCS
NS	CS	Asua	ùn	gú lú	tortoise	DD
NS	CS	Ma’di	o	kù	tortoise	RCS
NS	Kadu	Krongo	-kó	ò ŋ (ní-)	tortoise	Re85
NC		PWN	-kwú	lu	tortoise	M
NC	Dogon	Tebul Ure	aŋguŋ	gu ru	tortoise	RMB
NC	Ijoid	Furupagha	ɔbo	kɔ rɔ	turtle	KW
NC	Ijoid	Nembe	obo	ŋgo ro	turtle	KW
NC	Mande	Yaure	kú	lú	tortoise	ALMCI
NC	Senufo	Nabaj	xu	ru	tortoise	ALGCI
NC	Kru	Bete G	kó	lɔ	<i>tortue</i>	Ma83
NC	Kordofanian	Masakin	(k)ə	rə	tortoise	RCS

Ph	Family	Language	Attestation			Gloss	Source
NC	Atlantic	Sereer		xɔ	m	<i>tortue</i>	Cr72
NC	Isolate	Mbre		k	ru we	tortoise	Denis Creissels (p.c.)
NC	Gur	Tiefo	ba	kū	ò	<i>tortue</i>	HO13
NC	Adamawa	Mumuye		ku	ru	tortoise	Sh83
NC	Ubangian	Geme		kú	lō	<i>tortue</i>	Mo88
NC	Kwa	Mbatto	ómó	k	rō é	tortoise	ALKCI
NC	VN	Gbe-Ewe		k	lo	<i>tortue</i>	Ro95
NC	VN	Işekiri	ólu	kú	rú mè	tortoise	CO91
NC	VN	Nupe	dù	kú		tortoise	Ba14
NC	EBC	Doka [Plat-a- eau]		ku	l	tortoise	BCCW
NC	EBC	Ankwa [Plat-ì- eau]		k	rù	tortoise	RMB
NC	EBC	Koring Ntezi è		kù	l	tortoise	JS
NC	EBC	Proto-Lower ú- Cross		kút	/ ŋ-	tortoise	Co91
NC	EBC	Səgəmək i- [Kainji]		kú	ru	tortoise	RMB
NC	Bantoid	Cambap [Mambiloid]		kpó	ró	tortoise	Connell (p.c.)
NC	Bantoid	Tikar [Tikar]		kpu	lu'	tortue	Stanley (ined.)
NC	Bantoid	Keme- kə zung]E. Be- boid]		ku	lā	tortoise	SIL wordlists
NC	Bantoid	Balep ṅ- [Ekoid]		kú	l	tortoise	Cr69
NC	Bantoid	Ugarə u [Tivoid]		ku	lu lu	tortoise	SIL wordlists
NC	Bantoid	Mbat [Jara- wan]		kò	l	turtle	Green (p.c.)
NC	Bantu	P-Manen- guba		kú	l'	tortoise	He87
NC	Bantu	CB -		kú	dù	tortoise	G

Commentary: An early version of this table was presented in Blench (1997) where it was argued that the importance of turtles and tortoises in the gathering phase of human history had made this word particularly salient. The diversity of forms attested may reflect the fact that different species may have compound names (see the Kanuri and Aiki forms). Extra-African cognates have been recorded;

Phylum	Family	Language	Attestation
Austroasiatic	Munḍā	Sora	kola
Dravidian		Tamil	kurulaḷai

The diversity of the forms attested may reflect the fact that different species may have compound names (see the Kanuri and Aiki forms). Turtle/tortoise is the only item of fauna to be widely attested in Niger-Saharan. Greenberg cites parallels from Kordofanian and also Keiga, now classified as Nilo-Saharan. Hoffmann (1970:15–16) points out that this word was also borrowed into Chadic from Benue-Congo, occurring throughout West Chadic and sporadically in Central Chadic. Ironically, he concluded that this might be evidence that these attestations were unconnected. Since Afroasiatic cognates seem only to be recorded in languages near to Niger-Saharan, it is reasonable to conclude that they are ancient loans.

Refs: M. 306; G. 321+325+1259+1260; C.:321, Gr.:88, G.:159

6. crab #kala

Ph	Family	Language	Attestation	Gloss	Source
Isolate	Hadza	Hadza		goma:	KM
AA	W. Chadic	Hausa	káágwáá		Ab49
AA	C. Chadic	Bana	kwérekwérekiŋ		Richard Gravina
NS	C. Sudanic	Mbay	ká-bàr		Ke97
NC		PWS	-ka-, -ka(l)-		W
NC		PWN	-kála		M
NC	Ijoid	Nembe	à-kàngà		Ka64
NC	Mande	Gban	klá	crab	Valentin Vydrine
NC	Pere	Pere		kamu	crabe DC
NC	Kru	Bakwe	kāā		crab Ye09
NC	Kru	Tepo	ka		crab Ma83
NC	Kru	Koyo		kèmè	crab Ma83
NC	Atlantic	Temne	a-kara		crab Wi07
NC	Atlantic	Bijogo		e-gomba /u-	crabe GS02
NC	Gur	Mõõre	gará-ga		Ca76
NC	Adamawa	Jen	Jen	kōŋ	crab N117
NC	Gbaya	Bodoe	káyá		crabe Mo95
NC	Kwa	Baule	kángâ		crabe ALKwCI
NC	Kwa	Ga	gà'á		lagoon crab Kr73
NC	VN	Gbe	Ewe	à-gálà	crabe Ro95
NC	VN	Nupoid	Nupe	kara	crab Ba14
NC	EBC	Cross River	Obolo	úka	Co91
NC	EBC	Cross River	PUC	*'-kání	crab JS
NC	Bantoid	Mambiloid	Mambila		kaab21 PM95
NC	Bantoid	Beboid	Kemezong		kikambo crab SIL wordlists
NC	Bantoid	Ekoid	Ejagham	ekáé	crab John Watters
NC	Bantu	CB	-káda		crab G

Commentary: African crabs are highly speciated but can be divided into three categories; marine, freshwater and land crabs. The taxonomy of freshwater crabs is given in Cumberlidge (1999). Material on land crabs is not easily available but they are widespread throughout the continent and are frequently culturally important because of their role in divination systems. Crab divination is reviewed in Blench and Zeitlyn (1990) which shows that the words for ‘spider’ and ‘crab’ are etymologically interconnected in the Bantu borderland because of their comparable significance in divination systems.

There appear to be two intertwined roots, one with r/l/d in C₂ position, the second, less common, with –mb-, in the same slot. All those with a lateral have a copy central vowel, generally –a, whereas the final vowel in type II is much more variable. The table presents a sample of crab names. Mukarovsky (1976:144) adds further Niger-Congo cognates. The Niger-Congo roots are discussed in Williamson & Shimizu (1968:92).

‘Crab’ also has widespread Eurasian cognates (Blench 1997). The table below sets out some attestations and reconstructions that have been proposed for ‘crab’ in Old World language phyla.

Phylum	Family	Language	Attestation	Source
Japonic		Modern Japanese	kani	
Altaic		Modern Korean	ke	
Austroasiatic	Proto-Mon-Khmer		*kə(n)taam	Diffloth (1994)
Austroasiatic	Proto-North Bahnaric		*katam	Smith (1972)
Austronesian	Proto-Austronesian		*kaRaŋ	Blust (ined.)
Austronesian	Proto-Nuclear Micronesian		*karika	Jeff Marck (p.c.)
Andamanese	Great Andaman	Aka Biada	kátta-da	Portman (1887:22)
Andamanese	Little Andaman	Onge	tekandue	Dasgupta & Sharma (1982)
Sino-Tibetan	Tibeto-Burman	Tamang	khakre	Rana (2005)
Kusundic		Kusunda	kakchi	Rana (2005)
Dravidian	Common Dravidian		kup(p)i	Burrow & Emeneau (1984:158)
Indo-European		Greek	karkinos	
Vasconic		Basque	karramorro	Trask (p.c.)

Commentary: Clues to the processes at work in African languages can be gained by comparison with Indo-European associations of words for ‘crab’. Indo-European has a root *#kar-* meaning ‘hard’, which has a complex association with words for ‘crab’². Latin *cancer* and Greek *karkinos* are both derived from reduplications of the original root, the image apparently being the hardness of the crab’s shell. A similar association also exists in Niger-Congo; Westermann (1927:240) reconstructs *#kual-* for ‘to be hard’ in Proto-Western Sudanic, and *#-kal-* for ‘crab’.

² Thanks to Václav Blažek for this observation.

Refs: W. 230; M. 193; G. 981;

7. skin, hide #kulu

Ph	Family	Subgroup	Language	Attestation	Gloss	Source
AA	Omotic		Wolayta	galba	skin	Be03
AA	Beja		Beja	kurbe	skin	Be20
AA	Agaw		Kwaräsa	korbe	skin/leather	Ap06
AA	Cushitic	East	Ba'iso	galba	skin	Ha78
AA	Cushitic	South	Alagwa	kaari	hide, skin	KM03
AA	Semitic		Amharic	qurbät	hide	Hu89
AA	Chadic	East	Ndam	géré	skin	JI
AA	Chadic	Central	Kotoko	ṛkèṛṛ	skin	JI
AA	Chadic	West	Tala	kuur	skin	JI
KS	Southern		!Xóõ	gúnu	dried skin	Tra94
KS	Central		Naro	khò	<i>Haut/Fell</i>	Vo97
NS	Kunama		Kunama	agala		Be01
NS	ES	Nilotic	Nuer	kul		RCS
NS	ES	Surmic	Murle	kween		RCS
NS	Saharan		Teda	koro-ta		Le50
NS	Songhay		Zarma	kúurú		DM78
NC			PWS	-ku, -kua (+Nasal)	<i>Haut</i>	W
NC			PWS	-ge, -gel-	<i>Haut</i>	W
NC			PWN	-kwuba	skin, bark, husk	M
NC			PWN	-puk-, -puak-	bark, skin	M
NC	Dogon		Tebul Ure	gudugo	skin	RMB
NC	Dogon		Yanda Dom	gùzù	skin	Heath (p.c.)
NC	Ijoid		P-Ijò	*ɔpaṛãṛ	bark	KW
NC	Ijoid		Defaka	íkpa	skin	BC
NC	Katla-Tima		Kaalak	kpu	bark of tree	RCS
NC	Katla-Tima		Domurik	k-ààrà /y-	skin	GS13
NC	Rashad	Tagoi	Tagoi	k-awár /h-	bark	RCS
NC	Mande		Kpelle	kɔlɔ	<i>peau</i>	Ma83
NC	Kordofanian	Talodi	Ngile	k-íṛi pl. w-úṛi	bark	Sch81b
NC	Kordofanian	Talodi	Nding	kinjini	bark	Sch81b
NC	Atlantic	North	Nalu	ma-kũ/ a-	skin	Wi07
NC	Atlantic	South	Gola	ókɔlɔ	skin	Wi07
NC	Kru	Isolate	Kuwaa	kũ̀	<i>peau</i>	Ma83
NC	Gur	Oti-Volta	Bieri	kwanu	<i>peau</i>	Ma75
NC	Gur	Oti-Volta	Dagaari	gbani	<i>peau</i>	Ma75
NC	Adamawa	Mumuye	Zing	koo	skin	Sh83
NC	Adamawa	Vere	Jango	kɔku	bark	RMB
NC	Adamawa	Jen	Maghdi	kôṛ	bark	NI17
NC	Ubangian	Ngbaka	Mundu	kòṇò	<i>peau, écorce</i>	Mo88
NC	Ubangian		Baka	kòto	<i>peau</i>	Br10
NC	Ubangian	Sereic	Indri	kua	skin	Sa50
NC	Kwa	GTM	Siwu	à-kó /sì-	skin	RMB

Ph	Family	Subgroup	Language	Attestation	Gloss	Source
NC	VN	Nupoid	Nupe	epà	leather, hide, skin	Ba14
NC	VN	Idomoid	Idoma	apu	skin	Ar83
NC	EBC	Upper Cross	Hohumono	èkpá / ì-	skin	JS
NC	Bantu		PB	pó	skin, bark, peel	BLR3 4790

Commentary: A pan-Africanism with apparently related forms in all four major phyla. A preliminary version of this dataset appears in Blench (1997). Greenberg (1963:21) initially identified this root for Niger-Congo and though he quotes Krongo (p. 157), his form does not correspond to that in Reh (1985), which is not evidently cognate. Creissels (1981:316) points out the Songhay cognate and adds further citations for Niger-Congo. Bender (1997:129) gives further examples for Nilo-Saharan, although he includes ‘basket’ in his semantic set. Other commentators include ‘bark’, for example Uduk (Eastern Sudanic) *khur* ‘bark’. Westermann (1927: 220, 237) separates out two forms which are probably part of a single root. Most Kru languages have reduced to CV, but the nasalisation preserved in Kuwaa provides evidence for the original lateral in C₂ position (Marchese 1983). Nasals are attested in Kru, Gur and Adamawa-Ubangian, which argues for an innovation perhaps at the level of Atlantic-Congo.

Refs: W. 220 + 237; M. 302; G. 392; G.:21, Gr.:84, B:129, E:491

8. fat, oil, grease #morV

Ph	Family	Group	Language	Attestation	Gloss	Source
AA	Omotic		Zayse	moora	fat	Be03
AA	Cushitic		Oromo	moora	fat	Hu89
AA	Semitic		Amharic	mora	fat	Hu89
AA	Chadic		Masa	mùl	oil	Jl
AA	Chadic		Gudu	mar	oil	Jl
AA	Chadic	West	Hausa	mai	oil	Ab49
AA	Chadic	West	Tala	mɔl	oil	Jl
NS	ES		Murle	more	oil	Ly71
NS	ES		*PN	*mɔ-r	oil	Di88
NS	ES		Proto-Daju	*mwi-	oil	RT81
NS	Songhay		Zarma	máání	<i>graisse</i>	BWK94
NC			#PWS	-mì	<i>Fett, Öl</i>	W
NC	Ijoid	Ijo	Furupagha	imeɛlɪ	fat	KW
NC	Rashad		Tagom	omá	oil, fat	RCS
NC	Atlantic	South	Temne	marô	oil, fat	Wi07
NC	Ubangian		Ngbaka	mó	<i>graisse, huile</i>	Mo88
NC	Ubangian	Sereic	Feroqe	mú	oil	Sa50
NC	Kwa		Ga	mù	edible oil	Kr73
NC	VN	Gbe	Ewe	àmì	<i>huile, graisse</i>	Ro95
NC	VN	Igboid	Igbo	mmanɔ	oil	KW72
NC	VN	Nupoid	Nupe	emí	oil, fat, grease	Ba14
NC	EBC	Kainji	Gbiri	mani	oil	RMB
NC	EBC	Kainji	Central Kambari	màni'î	oil	Ho65

Ph	Family	Group	Language	Attestation	Gloss	Source
NC	EBC	Kainji	Lopa	mum	oil	RMB
NC	EBC	Plateau	<u>E</u> ten	mos	fat	RMB
NC	EBC	Plateau	Təsu	amene	fat	RMB
NC	EBC	Plateau	Rukul	manai	fat	RMB
NC	EBC	Cross River	Abua	a-monom	oil	Ga80
NC	Bantoid	Tivoid	Batu	mem	oil	Robert Koops
NC	Bantoid	Ekoid-Mbe	Mbe	mwòr	oil	JP
NC	Bantoid	Jarawan	Mama	murū	oil	MW75

Commentary: The assumption here is that the *m-* is originally a class affix, signifying liquids or mass nouns which have become fused to the stem. Forms for mass nouns with *m-* affixes correspond to Kordofanian *η-* classes in other branches of Niger-Congo (Schadeberg 1989). The ‘original’ form (if that has a meaning in this context) would then have been something like *#muri* or *#ɲuri*. The second set of forms appears to be a rather strained correspondence, but there is evidence the *η-* and *m-* do correspond between Kordofanian and the rest of Niger-Congo (Williamson 1989a). Nonetheless, it is possible there are two roots here. Although this form seems well established enough in Plateau and Kainji to imagine that its source is there and gave rise to Hausa *mai*, the Hausa form may well have been borrowed back into some languages giving rise to irregularities.

Refs: D. 40, W. 257; Ehret p. 312

9. fire #(dw)isi

Ph	Group	Language	Attestation	Source
AA	Semitic	Akkadian	ifaat-	
AA	South Cushitic	Iraqw	‘ała	KM03
AA	Chadic	Karekare	ʔèsî	Schuh (p.c.)
AA	Chadic	Miya	osi	Schuh (p.c.)
NS	ES	Meidob	ussi	We93
NS	Kadu	Miri	issi	Sch94
NS	CS	Miza	a(t)si	Be92
NS	CS	Shemyar	dufu-n	Be92
NS	Saharan	Beria	jíé	JC04
NC	Mande	PSM	*síé	Valentin Vydrine
NC	Mande	Dzuun	sà	Valentin Vydrine
NC	Kordofanian	Moro	isia	RMB
NC	Dakoid	Nnakenyare	yísi	BS10

Commentary: An old areal word with dispersed reflexes, it is possible some forms are just chance resemblances. Not attested in Khoisan. Only recorded in some subgroups of West Chadic and thus probably a local loanword. If the Akkadian form is genuinely related, then it is tempting to assume this an old Afroasiatic root loaned into Nilo-Saharan and thence into eastern Niger-Congo. Bender

(1992:43) reconstructs Proto-Central Sudanic *#co*, but as an areal loan, so reconstruction is probably not a meaningful exercise. The Niger-Congo attestations are very scattered and these could be chance or borrowing.

Refs: Bender (1991:5)

10. #-si dog

Ph	Family	Subgroup	Language	Attestation	Source
AA	Semitic		Amharic	wifa	Hu89
AA	Cushitic		PHEC	wafa	Hu89
AA	Cushitic		Beja	yas	Hudson
AA	Chadic		Mwaghavul	as	RMB
AA	Chadic		Kola	hàzà	Jl
NS	Fur		Fur	asà	Ja90
NS	ES		Proto-Daju	*iise	RT81
NS	ES		Nara	wos	RCS
NS	CS		Baka	ísì	RCS
NS	CS		Lugbara	atsí	RCS
NS	CS		Lendu	kaẓẓ	RCS
NS	Koman		Anej	kas	Ot19
NS	Maba		Masalit	wasi	Ed91
NS	Kadu		Katcha	is(s)i	Sch94
NS	Songhay		Kaado	hànsi	DM78
NC	Dogon		Tommo So	isé	Heath (p.c.)
NC	Rashad	Tagoi	Orig	wusu	SE79
NC	Atlantic		Manjaku	û-bûs	Guillaume Segerer
NC	Atlantic	South	Gola	ósu'n	We21
NC	Adamawa		Dza	iicwá	N117
NC	Adamawa	Mumuye	Zing	zaa	Sh83
NC	VN		Nupe	ɛ̃fi	Ba14
NC	EBC		E. Ogbia	isíà	KW
NC	Bantoid	Mambiloid	Ndoro	síe	RMB

Commentary: Although originally cited by Greenberg (1963:120), more complete evidence was marshalled by Bender (1981:258) with attestations in Fur, Sudanic, Kordofanian and possibly Ari [Omoti]. Not attested in Khoesan, Mande and most branches of Afro-Asiatic, which have variants of *kVr/n-*. This root is extremely widespread in Central Africa and yet does not form a convincing pattern. In Central Sudanic, for example, it is attested in almost every language (Bender 1992:40, 48). Bender separates the roots with initial *b-*, but it is likely that the two forms go together.

Refs: G:120

11. to fly, jump *#pur-*

Ph	Family	Subgroup	Language	Attestation	Gloss	Source
AA	Beja		Beja	biir	fly	Richard Hudson
AA	Agaw		Awngi	pərr-	jump	Ap06

Ph	Family	Subgroup	Language	Attestation	Gloss	Source
AA	Cushitic	East	PEC	*bar(ar)	fly	Sa82
AA	Cushitic	East	Burji	burr-	fly	Sa82
AA	Cushitic	South	Gorowa	furuu'	fly	Km03
AA	Chadic	West	Hausa	ɓùráá	jump	Ab49
AA	Chadic	Central	Yedina	fér	jump	JI
AA	Chadic	East	Kwang	bóre	jump	JI
AA	Berber		Kabyle	ferfer	<i>voler</i>	Da82
KS	Khoe		PNKhoe	*pe	<i>springen</i>	Vo97
KS	Khoe		PEKhoe	*pe	<i>springen</i>	Vo97
NS	Maba		Mesalit	fir		Ed91
NS	Berta		Berta	ho'rəŋ		Be89
NS	East Sudanic		Gaam	pərd-	fly	BA80
NS	East Sudanic		*PN	*pär	fly, jump	Di88
NS	Songhay		Djenné Chiini	firri	fly	He98
NS	Saharan		Kanuri	fār	to jump, fly	Cy94
Isolate			Banġeri Mə	dapiru	to fly	RMB
NC			*PWS	*pi, pil-	<i>fliegen</i>	W
NC	Dogon		Bunɔge	pile	to fly	RMB
NC	Ijoid		P-Ijɔ	*ifōĩ	to fly	KW
NC	Ijoid		Nkɔrɔɔ	fĩ	to fly	KW
NC	Mande		Bamana	pá	fly, jump	Valentin Vydrine
NC	Pere		Pere	fiili	<i>voler, se lever</i>	DC
NC	Atlantic N.	Cangin	Saafi-Saafi	pund	fly	B&P16
NC	Atlantic	North	Biafada	fiili	fly	Wi07
NC	Atlantic	South	Temne	fāl	fly	Da66
NC	Kordofanian		Moro	abəro	to fly	RMB
NC	Kordofanian	Talodi	Dagik	piro	to fly, jump	Va16
NC	Kru		Neyo	ɓlií	<i>voler</i>	M83
NC	Kru		Bakwe	plù	<i>sauter</i>	Ye09
NC	Gur		Lobiri	ufir	<i>voler</i>	ALGCI
NC	Gur	Isolate	Kulango	bií	<i>sauter</i>	Mi07
NC	Gbaya		P-Gbaya	bũlũ	<i>voler</i>	Mo95
NC	Kwa		Ega	fè	<i>voler</i>	ALKwCI
NC	Kwa		Ebrie	fè	<i>voler</i>	ALKwCI
NC	Kwa		Nawuri	fùgí	fly, jump	Sn89
NC	Kwa		Ga	filíki	fly	Kr73
NC	VN	Nupoid	Nupe	fù	to fly	Ba14
NC	VN	Nupoid	Nupe	fùrù	jump (n.)	Ba14
NC	EBC		Lokə	fiilá	to fly	JS

Commentary: The meanings of ‘fly’ and ‘jump’ are regularly intertwined throughout African language phyla. Bender (1997:121) lists more Nilo-Saharan cognates in typically elliptical style. This root presents an intriguing methodological problem. The citations are so similar that they must be related in some way. Is this simply a case of ideophonic convergence, where similar languages come to similar conclusions about the sound of beating wings? Not all languages do this,

as many datasets for ‘fly’ are omitted on the grounds of non-cognacy. Similarly, the semantic link with ‘jump’ which seems to be typical for Africa is not generally found elsewhere, where the common association appears to be ‘to flee’. Swadesh (1971) derived a similar form ideophonically as a world gloss, although he spreads the semantic net wider than here. This etymology is reprised in Bengtson & Ruhlen (1994) with a proposed proto-form **par*. Western Kwa seems to have undergone a distinctive reduction to **fe*.

Refs: Gr.:83, D.:42, W.:275, B: 121, E: 382, Ehret (1987:26)

3. CONCLUSIONS

The paper highlights eleven Pan-African roots, which are attested in either three or four phyla and sometimes additionally in isolates. The majority are nouns, although three verbs are included. The most surprising aspect of this is that three words have an almost identical form, *#kuru*, and a fourth, ‘crab’, *#kala*, is extremely similar.

There are two possible explanations; either pervasive sound symbolism or a chain of borrowing. If sound symbolism has driven convergence, it has affected selected concepts, rather like the word families in Southeast Asia or the templatic morphology in Semitic. For some roots, such as ‘fly’ or ‘crab’ their presence outside Africa must suggest phonaesthetic explanations. Alternatively, borrowing is responsible for these similarities at a very early period of contact. It is clear there are few roots which are well attested in Khoesan. These may have been borrowed into Nilo-Saharan at the period when Khoesan was further north and in contact with Nilo-Saharan and Afroasiatic. Similarly, many of the attestations in Afroasiatic look like regional borrowings. They are found in Chadic, Omotic and Cushitic and not in Berber-Egyptian-Semitic. Niger-Congo and Nilo-Saharan seem to have been contact at an early period, which may explain shared lexicon, but also typological features such as ATR vowel harmony, plural verbs and possibly alternating prefixes, such as are found in East Sudanic and are common to most branches of Niger-Congo.

Under no circumstances are these resemblances chance; the similarities are too great, and the comparison with the diverse lexicon for closely allied concepts (compare ‘leg’ and ‘knee’) argues that either historical processes are at work, or sound symbolism has created convergence. Further work may provide more coherent explanations.

ABBREVIATIONS

1. Principles of deriving acronyms for sources;

1. Single-authored works usually take the first two letters of the surname and add the last two numbers of the year of publication. Thus Childs (2000) becomes Ch00. However, where this might lead to ambiguity, a third letter of the author's name is added to discriminate between two authors. Thus;

Cr81 Creissels (1981)

but;

Chr33 Christaller (1933)

2. Two- and three-authored works usually take the initial letter of each author's surname and combine them into a sequece of capital letters, followed by the last two numbers of the year of publication. Thus Jungraithmayr & Ibrizimow (1995) becomes JI95.

3. Single-authored major comparative sources are identified by a single capital letter. These are;

B Bender (1997)

G Guthrie (1967-1971)

M Mukarovsky (1976, 1977)

W Westermann (1927)

Other publications by these authors are abbreviated in the standard way.

4. Multiply-authored major comparative sources, particularly those in series have slightly idiosyncratic abbreviations derived from the initial letters of the main words in the title.

Acronym	Bibliographic reference	Family
ALGCI	Mensah & Tchagbale (1983)	Gur
ALKCI	Hérault (1983)	Kwa
ALKrCI	Marchese (1983)	Kru
ALMCI	Halaoui, Tera and Trabi (1983)	Mande
BCCW	Williamson & Shimizu (1968) & Williamson (1973)	Benue-Congo

5. No distinction between electronic and paper formats is made for unpublished works. In view of the difficulties of assigning a year to these materials, they are numbered in sequence, following the abbreviation of the author(s) name(s) according to the principles above.

2. Specific languages: classification and source(s)

Table 1. Acronyms for language-specific sources

Acronym	Expansion as reference	Language treated
Ab40	Abrahams (1940)	Tiv
Ab49	Abrahams (1949)	Hausa
Ab58	Abrahams (1958)	Yoruba

Acronym	Expansion as reference	Language treated
Ag86	Agheyisi (1986)	Edo
ALGCI	Mensah & Tchagbale (1983)	Gur
ALKCI	Hérault (1983)	Kwa
ALKrCI	Marchese (1983)	Kru
ALMCI	Halaoui, Tera and Trabi (1983)	Mande
Bd94	Boyd (1994)	Chamba Daka
Be	Bertho (1953)	Dogon
Bo87	Boyeldieu (1987)	Fer & Yulu
Bo93	Boyeldieu (1993)	Sara-Bongo-Bagirmi
BW94	Bernard & White-Kaba (1994)	Zarma
C-G	Calame-Griaule (1968)	Dogon
Ch00	Childs (2000)	Kisi
Chr33	Christaller (1933)	Twi
Cr81	Creissels (1981)	Songhay
Cy	Cyffer (1994)	Kanuri
D	Dimmendaal (1988)	Proto-Nilotic
DC	Ducroz & Charles (1978)	Songhay Kaado
Ed91	Edgar (1991)	Maba group
Fa04	Fabre (2004)	Samba Leko
G	Greenberg (1963)	Africa general
Gr	Gregersen (1972)	Kongo-Saharan
Gt	Guthrie (1967-1971)	Bantu
Ja90	Jakobi (1990)	Fur
Jl	Jungraithmayr & Ibriszimow (1995)	Chadic
JR	Jacques Rongier (ined.)	Kwa languages
JRxx	Jacques Rongier ()	Ewe
Kl90	Kleinwillinghöfer (1990)	Wiyaa
Kr16	Kra (2016)	Kulango
Kri99	Kropp-Dakubu (1999)	Ga
KW	Kay Williamson (p.c.)	Ijoid
LS	Lamberti & Sottile (1997)	Cushitic and Omotic
Lu03	Lux (2003)	Noni
M	Mukarovsky (1976/7)	Proto-Western Nigritic
Man	Manessy (1975)	Oti-Volta
Me	Meeussen (1980)	Proto-Bantu
Mo88	Moñino (1988)	Ubangian
Mo95	Moñino (1995)	Gbaya
No89	Noye (1989)	Fulfulde
P	Prost (1953)	South Mande

Acronym	Expansion as reference	Language treated
PWN	Mukarovsky (1976, 1977)	Proto-Western Nigritic
PWS	Westermann (1927)	Proto-Western Sudanic
PM95	Perrin & Mouh (1995)	Mambila
R	Rottland (1982)	Nilotic
RCS	Roland Stevenson mss.	Nilo-Saharan, Kordofanian
RMB	Author's fieldwork	Various languages
Sch81a	Schadeberg (1981a)	Kadu
Sch81b	Schadeberg (1981b)	Heiban Kordofanian
Sch94	Schadeberg (1994)	Kadu
SE07	Schaefer & Egbokhare (2007)	Emai
Sg	Guillaum Segerer	Atlantic languages
Vo82	Voßen (1982)	Eastern Nilotic
Vo88	Voßen (1988)	Maa
Vo97	Voßen (1997)	Khoesan
Wa	Ejagham	Watters ined
W	Westermann (1927)	Western Sudanic
We21	Westermann (1921)	Gola
Wi07	Wilson (2007)	Atlantic
Ye09	Yepi et al. (2009)	Bakwe

REFERENCES

- Abera, Daniel. 2019. *Shabo English Dictionary: With English Shabo Index*. Independently Published. [DA19]
- Abiodun, Michael Ajibola. 1999. *A comparative phonology and morphology of Ukaan dialects of old Akoko Division*. Ph.D. thesis, University of Ilorin, Nigeria. [Ab99]
- Abraham, R. C. 1940. *A dictionary of the Tiv language*. London: Crown Agents. [Ab40]
- Abraham, R.C. 1949. *Dictionary of the Hausa language*. London: University of London Press. [Ab49]
- Abraham, R.C. 1958. *Dictionary of Modern Yoruba*. London: University of London Press. [Ab58]
- Agheyisi, Rebecca N. 1986. *An Edo-English dictionary*. Benin City: Ethiope Publishing Corporation.
- Appleyard, David L. 2006. *A Comparative Dictionary of the Agaw Languages*. (Kuschitische Sprachstudien – Cushitic Language Studies Band 24). Köln: Rüdiger Köppe Verlag. [Ap06]
- Armstrong, R.G. 1983. The Idomoid languages of the Benue and Cross River valleys. *Journal of West African Languages*, XIII,1:91–149.
- Bailleul, Charles. 1996. *Dictionnaire Bambara-Français*. Bamako: Donniya. [Ba96]
- Banfield, A.W. 1914. *Dictionary of the Nupe Language*. Shonga: The Niger Press. [Ba14]
- Bender, M.L. 1983a. (ed.)– *Nilo-Saharan Language Studies*. Michigan State University Press.
- Bender, M.L. 1989. “Berta lexicon.” In: *Topics in Nilo-Saharan Linguistics*. Bender, M.L. ed. 271–304. Hamburg: Helmut Buske Verlag. [Be89]
- Bender, M.L. 1989a (ed.) – *Topics in Nilo-Saharan Linguistics*. Helmut Buske Verlag, Hamburg.

- Bender, M.L. 1992. "Central Sudanic Segmental and Lexical Reconstructions." *Afrikanische Arbeitspapiere*, 29:5–61. [Be92]
- Bender, M.L. 1997. *The Nilo-Saharan languages: a comparative essay*. Munich: Lincom Europa.
- Bender, M.L. 2001. "English-Kunama lexicon." *Afrikanische Arbeitspapiere*, 65:201–253. [Be01]
- Bender, M.L. 2003. *Omotic Lexicon and Phonology*. Carbondale: Southern Illinois University. [Be03]
- Bender, M.L. 2020. *Cushitic Lexicon*. Berlin: Peter Lang. [Be20]
- Bender, M.L. and M.A. Ayre. 1980. *Preliminary Gaam-English-Gaam dictionary*. Available from the author.
- Bendor-Samuel, John (ed.). 1989. *The Niger-Congo languages*. Lanham: University Press of America.
- Bengtson, J.D. and Ruhlen, M. 1994. "Global etymologies." In: *On the origin of languages: studies in linguistic taxonomy*. M. Ruhlen ed. Stanford University Press, 277–336.
- Bennett, Patrick R. 1983. "Adamawa-Eastern: Problems and Prospects." In: *Studies in African Linguistics, Vol. I*. Dihoff, I. (ed.) Holland: Foris Publications.
- Bernard, Yves & Mary White-Kaba. 1994. *Dictionnaire Zarma-Français*. Paris: ACCT. [BWK94]
- Berry, Jack & Greenberg, Joseph H. (eds.) 1971. *Current Trends in Linguistics Vol. 7, Sub-Saharan Africa*. Mouton, The Hague.
- Black, K., & Black, K. 1971. *The Moro language: Grammar and dictionary* (Linguistic Monograph Series, No. 6). Sudan Research Unit, University of Khartoum.
- Bleek, Dorothea. F. 1956. *A Bushman Dictionary*. New Haven: American Oriental Society. [Bl56]
- Blench, Roger M. & D. Zeitlyn. 1990. "A web of words." *Sprache und Geschichte in Afrika*, 10/11:171–186.
- Blench, Roger M. 1995. "Is Niger-Congo simply a branch of Nilo-Saharan?" In: *Proceedings of the Fifth Nilo-Saharan Linguistics Colloquium, Nice, 1992*. ed. R. Nicolai and F. Rottland. 68–118. Köln: Rudiger Köppe.
- Blench, Roger M. 1997. "Crabs, turtles and frogs: linguistic keys to early African subsistence systems." In: *Archaeology and Language, I*. eds. R.M. Blench and M. Spriggs. 166–183. London: Routledge.
- Blench, R.M. 2017. "African language isolates." In: *Routledge Handbook of Language isolates*. Lyle Campbell, Alex Smith, and Thomas Dougherty, eds. Routledge Language Family Series. London: Routledge, 162–192.
- Blench, Roger M. 2026. "Arguments for the coherence of Nilo-Saharan." *Nairobi Journal of African Languages* 12: 199–224.
- Botne, Robert & Adrien Pouille. 2016. *A Saafi-Saafi and English//French Dictionary*. Köln: Rüdiger Köppe. [B&P16]
- Boyd, Raymond. 1978. *A propos des ressemblances lexicales entre langues Niger-Congo et Nilo-Saharien*. SELAF, 65. Paris.
- Boyd, Raymond. 1994. *Historical perspectives on Chamba Daka*. Köln: Rudiger Köppe. [Bd94]
- Boyd, Raymond. 1999. "Congo-Saharan revisited." *Frankfurter Afrikanistische Blätter*, 8:15–48.
- Boyd, Raymond and Isa Sa'ad. 2010. *A Chamba-English Dictionary*. Lagos; Malthouse Press Limited. [BS10]
- Boyeldieu, Pascale. 1987. *Les Langues Fer (Kara) et Yulu du Nord Centrafricain*. Geuthner, Paris.
- Boyeldieu, Pascale. 1993. *Identité tonale et filiation des langues Sara-Bongo-Baguirmiennes (Afrique-Centrale)*. Thèse pour le doctorat d'état ès-lettres. Paris: Sorbonne.
- Boyeldieu, Pascale. ms., *Presentation du làà:l ou "Gori" (Moyen-Chari, Tchad)*. Paris: CNRS.

- Brisson, Robert. 2010. *Petit dictionnaire Baka-Français*. Paris: L'Harmattan. [Br10]
- Burrow, T. & Emeneau, Murray B. 1984. [2nd ed] *A Dravidian etymological dictionary*. Oxford: Clarendon Press.
- Canu, Gaston. 1976. *La langue Mò:re*. Paris: SELAF. [Ca76]
- Childs, T. 2000. *A dictionary of the Kisi language*. Köln: Rudiger Köppe. [Ch00]
- Christaller, J.G. 1933. *Dictionary of the Asante and Fante language called Twi*. Basel: Basel Evangelical Missionary Society. [Chr33]
- Cloarec-Heiss, France. 1995. "Emprunts ou substrat? Analyse des convergences entre le groupe banda et les langues du Soudan Central." In: *Proceedings of the Fifth Nilo-Saharan Linguistics Colloquium, Nice, 1992*. ed. R. Nicolai and F. Rottland. 321–355. Köln: Rudiger Köppe.
- Connell, Bruce A. 1991. *Phonetic aspects of the Lower Cross languages and their implications for sound change*. PhD. dissertation, University of Edinburgh. [Co91]
- Connell, Bruce (ined.) *Defaka dictionary*. ms. [BC]
- Crabb, D.W. 1969. *Ekoid Bantu languages of Ogoja, Pt. I*. Cambridge: Cambridge University Press. [Cr69]
- Creissels, Denis 1981. 162–192. "Songhay et Niger-Congo (mandé)." In: *Nilo-Saharan*. Schadeberg, Thilo C. and Marvin L. Bender eds. Holland: Foris Publications, 307–327.
- Creissels, Denis. (ined.), *Liste de mots Pre*. ms. [DC]
- Cretois, L. 1972–1977. *Dictionnaire Sereer-Français*. Dakar : CLAD. [Cr72]
- Cumberlidge, N. 1999. *The freshwater crabs of West Africa: family Potamonautidae*. (Vol. 36). IRD editions.
- Cyffer, Norbert. 1994. *English-Kanuri Dictionary*. Köln: Rudiger Köppe. [Cy94]
- Dalby, David. 1966. "Lexical analysis in Temne with an illustrative wordlist." *Journal of West African Languages*, 3(2): 5–26. [Da66]
- Dallet, J-M. 1982. *Dictionnaire Kabyle-Français*. SELAF, Paris. [Da82]
- Dasgupta, D. and S. R. Sharma. 1982. *A handbook of the Oge language*. Anthropological Survey of India, Linguistic Series, 5. Calcutta.
- Demolin, Didier (ined.). *Lexique comparatif des langues du groupe mangbetu*. ms. [DD]
- Dickens, Patrick. 1994. *English- Ju|'hoan, Ju|'hoan-English dictionary*. QKF8. Köln: Rüdiger Köppe.
- Diffloth, Gerald. 1994. "The lexical evidence for Austric, so far." *Oceanic Linguistics*, 33, 2:309–321.
- Dimmendaal, Gerrit J. 1988. "The lexical reconstruction of Proto-Nilotic: A first reconnaissance." *Afrikanische Arbeitspapiere*, 16:5–67. [Di88]
- Ducroz, Jean-Marie & Charles, Marie-Claire. 1978. *Lexique Sonney (Songay) Français*. Paris: Harmattan. [DM78]
- Edgar, John T. 1991. *Maba-group lexicon*. Berlin: Reimer. [Ed91]
- Ehret, C., Ehret, P. & Ten Raa, E. 2012. *A dictionary of Sandawe*. Köln: Rüdiger Köppe. [EPT12]
- Ehret, Christopher. 1987. "Proto-Cushitic reconstruction." *Sprache und Geschichte in Afrika*, 8:7–180. [E87]
- Elias, Philip, Jacqueline Leroy & Jan Voorhoeve. 1984. "Mbam-Nkam or Eastern Grassfields." *Afrika & Übersee* 48: 31–107. [ELV]
- Elugbe, B.O. 1989. *Comparative Edoid: phonology and lexicon*. (Delta Series No. 6.) University of Port Harcourt Press.
- Ferry, M-P. 1991. *Thesaurus Tenda*. Paris: Peeters.

- Frajzyngier, Zygmunt. 1992. *A dictionary of Mupun. (Sprache und Oralität in Afrika (SOA) 11)*. Berlin: Dietrich Reimer Verlag. [Fr92]
- Gardner, Ian 1980. *Abuan-English, English-Abuan dictionary*. Delta, 3. Jos: University of Port Harcourt and Nigeria Bible Translation Trust. [Ga80]
- Greenberg, Joseph H. 1963. *The Languages of Africa*. Published for Indiana University by Mouton, The Hague.
- Greenberg, Joseph H. 1971. "Nilo-Saharan and Meroitic." In Berry, J. & Greenberg, J.H. (eds.) *Current Trends in Linguistics Vol. 7, Sub-Saharan Africa*. Mouton, The Hague.
- Gregersen, Edgar A. 1972. "Kongo-Saharan." *Journal of African Languages*, 11,1:69–89.
- Guthrie, Malcolm. 1967-1971. *Comparative Bantu*. 4 vols. Farnham: Gregg International Publishers.
- Halaoui, Nazam, Tera, Kalilou and Trabi, Monique. 1983. *Atlas des langues mandé-sud de Côte d'Ivoire*. Abidjan: ACCT and ILA. [ALMCI]
- Hall, Beatrice L., Hall, R.M.R., Pam, M.D., Myers, A., Antell, S.A. and Cherono, G.K. 1974. "African vowel harmony systems from the vantage point of Kalenjin." *Afrika und Übersee*, LVII, 4:241–267.
- Hantgan, Abbie and Aminata Outtara. 2013. *Tiefo-English-French-Jula Dictionary*. ms. [HO13]
- Hayward, R.J. 1978. "Bayso Revisited: Some Preliminary Linguistic Observations--I." *Bulletin of the School of Oriental and African Studies, University of London*, 41(3): 539–570. [Ha78]
- Heath, Jeffrey. 1998. *Dictionnaire Songhay –Anglais- Français. Tome II, Djenné Chinni*. Paris: L'Harmattan. [He98]
- Hedinger, Robert. 1987. *The Manenguba languages (Bantu A.15, Mbo cluster) of Cameroon*. London: School of Oriental and African Studies. [He87]
- Heine, Bernd. 1999. *Ik Dictionary. (Nilo-Saharan - Linguistic Analyses and Documentation, 15)* Köln: Rüdiger Köppe. [He99]
- Hérault, G. 1983. *Atlas des Langues Kwa de Côte d'Ivoire, Tôme 2*. Abidjan: ACCT and ILA. [ALKwCI]
- Hoffmann, C. 1965. "A word list of Central Kambari." *Journal of West African Languages*, 2(1)7–31. [Ho65]
- Hudson, G. 1989. "Highland East Cushitic dictionary." *Cushitic Language Studies* Bd. 7. Buske, Hamburg. [Hu89]
- Hudson, R. n.d. *Beja dictionary*. ms.
- Jackson, Ellen W. 1988. *Lexique Français-Tikar*. ms. Yaoundé: SIL. [Ja88]
- Jakobi, Angelika. 1990. *A Fur grammar*. Hamburg: Helmut Buske Verlag. [Ja90]
- Jakobi, Angelika and Joachim Crass. 2004. *Grammaire du beria (langue saharienne)*. Cologne: Rudiger Koppe. [JC04]
- Jungraithmayr, H. and Dymitr Ibrizimow. 1995. *Chadic lexical roots*. [2 vols.] Berlin: Reimer. [JI]
- Kaliai, M.H.I. 1964. *Nembe-English dictionary*. Ibadan: Institute of African Studies. [Ka64]
- Keegan, John M. 1997. *Dictionary of Mbay*. München: Lincom Europa. [Ke97]
- Kießling, Roland & Maarten Mous. 2003. *The lexical reconstruction of West Rift Southern Cushitic*. Köln: Rüdiger Köppe. [KM03]
- Kleinewillinghöfer, Ulrich. 1991. *Die Sprache der Waja (nyan wiyáu)*. Frankfurt am Main: Peter Lang.
- Kra, K.A.E. 2016. *Dictionnaire Koulango-Français*. Paris: L'Harmattan. [Kr16]
- Kraft, C. 1981. *Chadic wordlists*. (3 vols) Berlin: Reimer. [Kr81]

- Kropp-Dakubu, Mary-Esther. 1973. *Ga-English Dictionary*. Legon: Institute of African Studies. [Kr73]
- LaCroix, P.F. 1962. "Notes sur la langue Galke." *Journal of African Languages*, 1(2): 94–121. [La62]
- Lamberti, M. and R. Sottile. 1997. *The Wolaytta language*. Köln: Rüdiger Köppe.
- Le Coeur, Charles. 1950. *Dictionnaire Ethnographique Teda*. Dakar: IFAN. [Le50]
- Lyth, R.E. 1971. *Murle Vocabulary*. (Linguistic Monograph Papers, 7.) Khartoum: Sudan Research Unit, University of Khartoum.
- Maddieson, Ian and Kay Williamson. 1975. "Jarawan Bantu." *African Languages*, 1:125–163. [MW75]
- Manessy, Gabriel. 1975. *Les langues Oti-Volta*. Paris: SELAF.
- Marchese, Lynell. 1983. *Atlas Linguistique Kru*. Abidjan: ACCT and ILA. [Ma83]
- Mbuagbaw, Tanyi Eyong. 1998. *Kenyang lexicon*. Yaoundé: CABTAL. [Mb98]
- Meeussen, A.E. 1980. *Bantu lexical reconstructions*. Tervuren: MRAC.
- Mensah, E.N.A. & Z. Tchagbale. 1983. *Atlas des langues gur de Côte d'Ivoire*. Abidjan: ACCT and ILA. [ALGCI]
- Micheli, Ilaria. 2007. *Profilo grammaticale e vocabolario della lingua Kulango (Côte D'ivoire)*. ms. [Mi07]
- Mikkola, Pertti. (ined.) *Nilo-Saharan revisited: some observations concerning the best etymologies*. ms. available from author.
- Miller, Kirk. (ined.) *Hadza dictionary*. ms. [KM]
- Moñino, Yves. 1995. *Le Proto-Gbaya*. Paris: Peeters. [Mo95]
- Moñino, Yves. ed. 1988. *Lexique comparatif des langues oubangiennes*. Paris: Geuthner. [Mo88]
- Mukarovsky, Hans G. 1976-1977. *A study of Western Nigritic*. 2 vols. Wien: Institut für Ägyptologie und Afrikanistik, Universität Wien.
- Munro, Pamela & Dieynaba Gaye. 1991. "Ay Baati Wolof: A Wolof Dictionary." *UCLA Occasional Papers in Linguistics*, 9. Los Angeles: UCLA. [MG91]
- Nicolaï, Robert. 1991. "Les apparentements génétiques du Songhay: éléments d'analyse." In : Bender, M.L. 1991a : 37–56
- Nlabephee, Kefas Othaniel. 2017. *Jen Cluster Comparative Wordlist*. ms. [N117]
- Nougayrol, Pierre. 1989. *La Langue des Aiki*. Paris: Geuthner.
- Noye, D. 1989. *Dictionnaire Foulfouldé-Français. Dialecte Peul du Diamaré, Nord-Cameroun*. Paris: Geuthner.
- Ohiri-Aniche, Chinyere. 1991. *A Comparative Phonology of Consonants in the Igboïd, Edoid and Yoruboid Languages of Nigeria*. University of Port Harcourt. (Doctoral dissertation). [CO91]
- Otero, M.A. 2019. *A historical reconstruction of the Koman language family*. Ph.D. University of Oregon. [Ot19]
- Perrin, M. J. and Mouh, M. 1995. *Lexique mambila-français*. Yaoundé: SIL, Cameroun. [PM95]
- Pohlig, J. n.d. *Mbe dictionary*. Electronic ms. [JP]
- Portman, M.V. 1887. *Manual of the Andamanese languages*. Government Printer, Port Blair.
- Prost, R.P.A. 1953. *Les langues Mandées-Sud du groupe Mana-Busa*. Dakar: IFAN.
- Prost, R.P.A. 1977. "Supplément au dictionnaire Sonay-Français (parler de Gao, Mali)." *Bulletin de l'Institut Fondamental d'Afrique Noire. Série B: Sciences Humaines* Dakar 34. 584–657. [Pr77]
- Reh, Mechthild. 1985. *Die Krongo-Sprache (niino mó-di)*. Berlin: Reimer. [Re85]
- Rongier, Jacques. 1995. *Dictionnaire français-éwé*. Paris: ACCT-Karthala. [JR95]

- Rottland, Franz. 1982. *Die südnilotischen Sprachen: Beschreibung, Vergleichung und Rekonstruktion*. Reimer, Berlin.
- Sachnine, Michka. 1982. *Le Lamé*. Paris: SELAF/ACCT. [Sa82]
- Santandrea, Stefano. 1950. "Indri-Togoyo-Ndogo-Feroqe-Mangaya-Mondu." *Comparative linguistics*. Verona : Museum Cobonianum. [Sa50]
- Sasse, Hans-Jurgen. 1982. *An etymological dictionary of Burji*. Hamburg: Buske. [Sa82]
- Schadeberg, Thilo C. 1981a. "The classification of the Kadugli language group." In: *Nilo-Saharan*. Schadeberg, Thilo C. and Marvin L. Bender, eds. 291–306. Holland: Foris Publications.
- Schadeberg, Thilo C. 1981b. *A Survey of Kordofanian. Volume 1: The Heiban Group*. Hamburg: Helmut Buske.
- Schadeberg, Thilo C. 1989. "Kordofanian." In: Bendor-Samuel, John (ed.) *The Niger-Congo languages*. 66–80. Lanham: University Press of America.
- Schadeberg, Thilo C. 1994. "Comparative Kadu wordlists." *Afrikanische Arbeitspapiere*, 40: 11–48. [Sch94]
- Schadeberg, Thilo C. and Elias, Philip. 1979. *A Description of the Orig Language*. MRAC, Ter-vuren. [SE79]
- Schadeberg, Thilo C. and Marvin L. Bender eds. 1981. *Nilo-Saharan*. Holland: Foris Publications.
- Schaefer, R.P. & F.O. Egbokhare. 2007. *A dictionary of Emai*. Köln: Rüdiger Köppe.
- Segerer, Guillaume. 2002. *La langue Bijogo de Bubaque*. Paris: Afrique et Language. [GS02]
- Shimizu, Kiyoshi. 1979. *A comparative study of Mumuye dialects*. Berlin: Reimer.
- Shimizu, Kiyoshi. 1980. *Comparative Jukunoid*. (3 vols) (Veröffentlichungen der Institut für Afrikanistik und Ägyptologie der Universität Wien. Vienna: Afro-Pub.
- Shimizu, Kiyoshi. 1983. *The Zing dialect of Mumuye: a descriptive grammar*. Hamburg: Helmut Buske. [Sh83]
- Smith, Kenneth D. 1972. *A phonological reconstruction of proto-North-Bahnaric*. Asian-Pacific Series, No. 2. SIL, Santa Ana, California.
- Snider, K.L. 1989. *North Guang comparative Wordlist*. Legon: Institute of African Studies. [Sn89]
- Stanford, Ron. n.d. *Bekwara dictionary*. Electronic ms. [RS]
- Sterk, Jan. (ined.) *Upper Cross Wordlists*. ms. [JS]
- Thelwall, Robin. 1981. "Subgrouping and Reconstruction of Daju." In Schadeberg, T.C. and Bender, M.L.: 167–184. [RT81]
- Tosco, M. 1991. *A Grammatical Sketch of Dahalo*. Hamburg: Buske. [To91]
- Traill, A. 1986. "Do the Khoi have a place in the San? New data on Khoisan linguistic relationships." *Sprache und Geschichte in Afrika*, 7 (1), 407–430. [Tr86]
- Traill, A. 1994. *A !Xóõ dictionary*. QKF9. Köln: Rüdiger Köppe. [Tr94]
- Vanderelst, John. 2016. *A grammar of Dagik*. Köln: Rüdiger Köppe. [Va16]
- Voßen, Rainer. 1982. "The Eastern Nilotes: Linguistic and Historical Reconstructions." *Kölner Beiträge zur Afrikanistik*, Volume 9. Berlin: Reimer.
- Voßen, Rainer. 1997. *Die Khoe-Sprachen: Ein Beitrag zur erforschung der Sprachgeschichte Afrikas*. Köln: Rüdiger Köppe. [Vo97]
- Werner, Roland. 1993. *Tidn-aal: A study of Midob (Darfur-Nubian)*, Berlin: Frankfurter Studien zur Afrikanistik 17. [We93]
- Westermann, Diedrich. 1911. *Die Sudansprachen*. Hamburg: Friederichsen.
- Westermann, Diedrich. 1927. *Die Westlichen Sudansprachen und ihre Beziehungen zum Bantu*. Berlin: De Gruyter.

- Williamson, Kay (ed). 1972. *Igbo-English dictionary: Based on the Onitsha dialect*. Benin: Ethiope Publishing Corporation. [Wi72]
- Williamson, Kay. 1973. *Benue-Congo comparative wordlist: Vol. 2*. Ibadan: West African Linguistic Society. [BCCW]
- Williamson, Kay and Shimizu, Kiyoshi. 1968. *Benue-Congo comparative wordlist: Vol. 1*. Ibadan: West African Linguistic Society.
- Williamson, Kay & Roger M. Blench. 2000. "Niger-Congo." In: *African languages: an introduction*, Heine, Bernd & Nurse, Derek (eds.). 11–42. Cambridge: Cambridge University Press.
- Williamson, Kay & Roger Blench. (ined.) *Comparative Ijo*. Electronic ms.
- Wilson, W.A.A. 2007. *Guinea languages of the Atlantic group: description and internal classification*. Frankfurt am Main: Peter Lang. [Wi07]
- Yepi, B.A. et al. 2009. *Dictionary Bakwé-French-English*. Méagui: CTAB. [Ye09]

REMARKS ON SPANISH *CONMIGO*, *CONTIGO*, *CONSIGO*

JAN HENRIK HOLST

ABSTRACT

The Spanish form conmigo ‘with me’ as well as its counterpart for the 2nd pers. sg. contigo and the reflexive consigo are more interesting than hitherto noticed. This holds in several respects. One requirement for linguistics is to arrive at synchronic analyses that do justice to the structure of the language; this leads to questions of grammar theory and typology. Another topic is provided by the diachronic background of these forms. Their history can be traced back to Latin and to Indo-European, and in a surprising way even beyond Indo-European.

CONTENTS

1. Introduction
2. A synchronic account
 - 2.1. Analysis I
 - 2.2. Analysis II
 - 2.3. Analysis III
3. A diachronic account
 - 3.1. A look at the more recent past
 - 3.2. A look at a rather distant past

1. INTRODUCTION

As is well known, the grammar connected with the Spanish preposition *con* ‘with’ differs from that of the remaining prepositions. What is usual for prepositions is that they are construed with various kinds of NPs (noun phrases), and if a personal pronoun is intended, this is put into its object form; for Spanish *yo* ‘I’, *tú* ‘you (sg.)’ and in the case of reflexivity this means that *mí*, *ti*, *sí* respectively are used, e.g. *a mí* ‘to me’, *para ti* ‘for you’. The preposition *con* is different in this respect, however, since in these three instances grammar requires the use of certain special forms; these are:

<i>conmigo</i>	‘with me’
<i>contigo</i>	‘with you (sg.)’
<i>consigo</i>	‘with himself / herself / itself / themselves’

As can be seen, the 1st and the 2nd persons of the singular as well as the reflexive pronoun are affected.

The origin of this peculiarity is known, see Lathrop (1980: 117), Penny (1991: 120f.), Ranson / Lubbers Quesada (2018: 254f.), Lara (2019: 51) and various other sources. In Latin, *cum* ‘with’ was mainly a preposition, as any other preposition, but with personal pronouns of the 1st and the 2nd persons (of both numbers) as well as with the reflexive pronoun it had been used as a postposition and had then become a suffix, which yielded:

<i>mēcum</i>	‘with me’
<i>tēcum</i>	‘with you (sg.)’
<i>sēcum</i>	‘with himself / herself / itself / themselves’
<i>nōbīscum</i>	‘with us’
<i>vōbīscum</i>	‘with you (pl.)’

In varieties of Early Romance, to these forms the preposition was added. Seen from a historical point of view, from then on ‘with’ existed twice in these forms – a fact that has repeatedly been noticed by researchers (e.g. by Lathrop 1980: 117, Penny 1991: 121). Stolz / Stroh / Urdze (2006: 501, n. 327) put it like this: “Comitativity is marked twice in the word-forms of Modern Spanish, etymologically speaking.”

Usual events of language history, such as sound change and morphological change, occurred as well. These led to the intervocalic “weakening” of *-c-* > *-g-* and established the vowel *i* (rather than regular *e*) in the singular forms due to analogy with *mí*, *tí*, *sí*. Old Spanish still had five forms: *comigo*, *contigo*, *consigo*, *con(n)usco*, *convusco*. Modern Spanish has only retained the first three (with the first one being spelled *conmigo* now). Also some other Romance languages and dialects show such forms: Portuguese still exhibits all five, and so does Galician; Sardinian (Nuorese variety of the Central Sardinian dialect) has *chin mecus* ‘with me’, *chin tecus* ‘with you (sg.)’ (Mensching 1992: 62).

At first glance it looks as if hardly anything new could be said about these facts. It turns out, however, that there do exist possibilities to say something new and interesting about the three Spanish special forms – this applies especially when grammatical, typological and diachronic considerations are brought into play. Section 2 will have a look at the synchronic structure of Spanish and will seek to do justice to it, while section 3 will make new points about the diachronic background of the phenomenon.

A brief note on terminology is perhaps in order. This paper will use the terms *preposition*, *postposition* and *adposition*, as is general practice in linguistics today. However, one should be aware of a few facts discussed by Holst (2001: 201f.). Prepositions and postpositions are one part of speech, not two. This can clearly be seen by comparison with adjectives, for instance, which are always called *adjectives* irrespective of whether they usually precede the noun (or NP), as in English, or follow it, as (mostly) in Spanish. The term *preposition* was coined with typical European languages in mind where these words precede the nouns (or NPs) they refer to. It was only modern linguistics that created the uniting term *adposition*, which is meanwhile standard in typological and other works, e.g. Croft (1990), Velupillai (2012), Herce (2023). (It is true that the term *adposition* is usually hardly required in Romance studies since in the Romance languages adpositions

are usually prepositions. A postposition can be seen, for instance, in Italian *due anni fa* ‘two years ago’, as opposed to Spanish *hace dos años*.)

2. A SYNCHRONIC ACCOUNT

Traditional grammars often express the facts in such a way that one does not combine *con* with *mí*, but uses *conmigo* instead (and correspondingly with the other persons). Some authors state that *con* and *mí* “become” *conmigo* together, though actually in a synchronic system nothing “happens”, hence nothing “becomes” anything; rather, there are nothing but facts and structures existing next to each other, side by side, and relationships between these facts and structures. It is just that explanations using the concept of “becoming” help the human mind. (For how grammars from the 16th and 17th centuries deal with *conmigo* etc. see Satorre Grau 2002: 376–379.)

Statements of the kind mentioned are unsatisfactory since they are not as explicit as one may wish, as well as for other reasons. One would be interested in an explicit description of what Spanish – or more specifically Spanish grammar – is like. What precisely is its morphology, and what precisely is its syntax? There are several options how to account for *conmigo*, *contigo*, *consigo*. (It needs to be added that there is nothing remarkable about the fact that there are several options. While in diachronic linguistics usually simply one solution is the one that matches language history, even though we linguists may be unable to identify it, in synchronic linguistics often several accounts are possible.)

In the following, three options how to analyze the facts of Spanish will be presented, and their pros and cons will be discussed.

2.1. Analysis I

The first option would be to insert *conmigo* into the paradigm of *yo* in some way, as well as the other two forms into the paradigms of *tú* ‘you’ and *él* ‘he’ etc. respectively. By this statement the following is meant. Typical modern western European languages, both of the Romance branch and of the Germanic branch, have reduced the Indo-European case system considerably. However, they still have case distinctions with personal pronouns. When in English, for instance, *I* and *me* (and *he* and *him*, etc.) are distinguished, this is best described as a difference in case. Spanish, too, is of course not a language without case; it is just that case is rather restricted in this language’s structure: to pronouns – while other nominals do not have case: nouns, adjectives, numerals; even many types of pronouns lack it.

Hualde / Olarrea / Escobar (2001: 144) distinguish four cases. The table they provide shall be reproduced here for the singular forms:

	nominative	accusative	dative	prepositional
1st	<i>yo</i>	<i>me</i>	<i>me</i>	<i>mí</i> , <i>conmigo</i>
2nd	<i>tú</i>	<i>te</i>	<i>te</i>	<i>tí</i> , <i>contigo</i>
3rd m.	<i>él</i>	<i>lo</i>	<i>le</i>	<i>él</i>
f.	<i>ella</i>	<i>la</i>	<i>le</i>	<i>ella</i>
n.	<i>lo</i>	–	–	<i>ello</i>

The authors are correct in positing a case named “prepositional” by them (“preposicional” in their Spanish text), the fact that this case has no direct antecedent in Latin of course being irrelevant. However, what strikes the reader is that for the 1st and 2nd persons the fields of the prepositional are filled with two forms each. (Schpak-Dolt 2012: 45 proceeds in the same way with having one cell each.) Usually such grammatical tables should be designed in such a way that a field contains only one form (unless the goal is to register free variants, dialect differences, or the like). This raises the idea that perhaps the grammatical structure of the language has not yet been captured in an adequate way.

A possible solution is to claim that we require *another* case. This additional case could be called *comitative* as in Uralic studies and also in various other philological traditions. The job of a comitative case is to express what in other languages would be expressed with the adposition ‘with’ and an NP. In Hungarian, for instance, *hajó* ‘ship’ forms the comitative *hajóval* ‘with a ship’ (cf. Rounds 2009: 106); also Estonian possesses a comitative, while Finnish does not. Another option would be to call this case *sociative*, which is the term used in Basque studies, see e.g. Santazilia (2013: 269–271). Spanish *conmigo* corresponds in function to Basque *nirekin* ‘with me’ (*ni* ‘I’), and Spanish *contigo* to Basque *zurekin* ‘with you’ (*zu* ‘you’). (Note also the areal implication, since Spanish and Basque are neighbours.) Since *comitative* is the more frequent term (in relation to *sociative*), it will be used in the following. The comitative is clearly distinct from the prepositional; it is a fifth case.

For Hungarian, Basque and other languages researchers had no problem with positing a case (and along with this, inventing a case name) since these languages exhibit large case systems anyway. With Spanish, by contrast, scholars seem to have been guided less by the Spanish data themselves and more by the Latin model that they had in mind. A fact that should not hold researchers back from speaking of another case is that the formation of the comitative implies the use of a circumfix, *con-* *-go*: circumfixes have been reported from various languages (Harris 2002, Harris / Xu 2009: 287f., Marušič 2023). It would also be possible to speak of simultaneous prefixation and suffixation. With the circumfix mentioned, the comitative can be formed from the prepositional – not from the accusative or any other case; in this way it is ensured that the vowel *i* is arrived at correctly.

While the morphology has now been accounted for, what remains to be clarified is the syntax. If we formally define *conmigo* etc. as comitatives, there is a consequence that must not escape our attention. In order to make a statement containing the meaning ‘with’, for almost any NP the preposition *con* is used: *con el hombre* ‘with the man’, *con él* ‘with him’, etc. But for the 1st and 2nd persons singular as well as for the reflexive, no preposition *con* is used – the fact that the comitative form begins with *con-* must not distract from this. The comitative corresponds in function not to an NP, but to a PP (prepositional phrase). The syntax thus becomes slightly more complex here, and this is a disadvantage of this solution – which is otherwise rather straightforward in many respects.

2.2. Analysis II

The second option how to handle the forms *conmigo* etc. can now be explained by referring to the first option. It would be possible to say that actually we are dealing with two words each: *con* + *migo*, *con* + *tigo*, *con* + *sigo*. Of course, the Spanish orthography demands spelling these forms in one word each – but this is just a matter of the spelling. As is well known in linguistics, a spelling convention of course need not do justice to the structure of a language, and it must not prevent a linguist from positing any analysis that might suit the facts in a smooth way. (The relationship between *profesor* and *profesores* is the same as that between *cuestión* and *cuestiones*, even though the rules of placing the accent mark demand no accent mark in the first pair, and an interplay between the presence and the absence of an accent mark in the second pair.)

As to the elements *migo*, *tigo*, *sigo*, it would then be these that would be called the comitative forms (not the entire forms, as in analysis I), and it would be these that would be listed in morphological tables. These forms would not be formed with circumfixes, but only with suffixes – which is simpler. The forms have initial consonants that are known from other forms of the paradigms in question: *m*-, *t*-, *s*-.

This solution has the great advantage that it entails simpler rules in syntax. The complications arising with the first analysis, laid out above in 2.1., would be lacking; utterances containing the idea of ‘with’ would always contain the preposition *con*. While all other prepositions then govern the “prepositional” (to cite the name chosen by Hualde / Olarrea / Escobar 2001, see 2.1.), *con* governs the comitative. The fact that prepositions do not all agree as to which case they govern is of course well known from many Indo-European languages, e.g. German, Russian, Lithuanian, Albanian, etc.; German, for instance, has *mit mir* ‘with me’ with a dative form, but *ohne mich* ‘without me’ with an accusative form. Only a subset of (modern) Indo-European languages always exhibits the same case, after any preposition, and such languages are best known from the Germanic and Romance branches, e.g. English, Swedish, and Italian. English has *me*, for instance, for the 1st pers. sg. after any preposition: *to me*, *at me*, *for me*, *with me*. Spanish is usually perceived as a language in this typological category as well – but, considering what has been discussed, this can be questioned. When operating as proposed with two cases “prepositional” and “comitative” for Spanish, this would move this language from one like English into the group of German, etc. It should be added, though, that only a single preposition would govern the “comitative”, all others the “prepositional”, and in this way the complexity of German is not reached. Incidentally, even within Romance having more than one case after prepositions is not entirely unheard of: in Sardinian, prepositions usually go along with the forms *mene* ‘me’, *tene* ‘you (sg.)’ (Mensching 1992: 62), but *a* ‘to’ deviates by requiring *mime*, *tibe* (historically datives) (Mensching 1992: 62). Moreover, there is also a comitative *mecus*, *tecus*, see section 1.

Another analysis will be presented in the next subsection. All in all, however, possibly this second solution is the one with fewest disadvantages. (Herce 2023: 437 mentions this analysis as well, apparently taking it for granted.)

2.3. Analysis III

The third option that suggests itself should not go unmentioned. As underlined right from the start when discussing the previous option, the spelling is something that should not prevent the linguist from proposing an analysis. It turns out that another solution would be to segment into three words *con* + *mí* + *go*, and correspondingly for the 2nd pers. sg. and the reflexive. Again, this basic idea has consequences that require exploring.

This analysis has one surprising advantage. The elements in the middle of the phrases that this analysis would yield are simply *mí*, *tí*, *sí*, hence forms already known from the paradigms. (We ignore here the issue where accents marks are found in Spanish orthography, and where not.) These are the case forms called “prepositional” by some authors, cf. 2.1. Positing an additional case “comitative” is superfluous now, and it is even impossible. This constitutes an advantage over the two previous analyses.

In syntax, however, the situation is not as simple – which shows that whatever you do, there are always unsatisfactory effects somewhere, without there being an escape. The structure now contains not only one adposition, as in analysis II, but two: one comes before the pronoun and one after it. One is a preposition, *con*, and the other is a postposition, *go*. (The *g* in this postposition would always be not the plosive / stop, but the corresponding fricative or approximant, the “gamma”, but this can be attributed to sandhi.) Historically *con* and *go* are identical; synchronically we can choose whether we want to interpret them as the “same” element or as two different ones.

This raises the question whether it would be unnatural to have such an element twice. What can be clarified relatively quickly is that in general it is very well possible for grammatical elements to occur twice. This can be seen from an Icelandic example, for instance (data and analysis from Holst 2023: 106). The North Germanic languages mark definiteness in the noun, so that Danish *hestens* ‘of the horse’ consists of the stem *hest-*, the definiteness suffix *-en-* and the genitive suffix *-s*. In related Icelandic, the corresponding form is *hestsins* ‘of the horse’, consisting of the stem *hest-*, the genitive suffix *-s-*, the definiteness suffix *-in-* and again the genitive suffix *-s*; hence the genitive suffix occurs twice. Icelandic represents an older state; the form originated from a noun in the genitive followed by a definite article also in the genitive which were then univerted. Since each of the two former constituents was marked for genitive, the suffix for this case *-s* eventually appeared twice in the resulting form. Danish removed the first genitive marker. The example shows a suffix that appears twice, but it is also possible to have the same morpheme both as a prefix and as a suffix: in Berber, for instance, frequently the simultaneous occurrence of prefixed *t-* and suffixed *-t* marks feminines.

However, these examples are morphological, and seeking to back up the analysis *con* + *mí* + *go* typologically, the question of a double occurrence should rather refer to syntax. To be more specific, it narrows down to: is it possible that an adposition appears twice in a PP? This is indeed possible, as shown by Hungarian. In this language demonstrative pronouns added to an NP require the postposition to be doubled. An example, cited from Rounds (2009: 151), is *ez alatt a szék alatt* ‘under this chair’, literally “this under the chair under”. Step by step, the structure of this PP can

be understood as follows. The NP *a szék* is ‘the chair’, consisting of the definite article *a* ‘the’ and the noun *szék* ‘chair’. Demonstrative pronouns precede their heads and they require the definite article to be present (unlike in English and in fact most European languages with articles, where the demonstrative pronoun would replace the definite article), hence with *ez* ‘this’ one obtains *ez a szék* ‘this chair’. Moreover, Hungarian is a language with postpositions, so that *alatt* ‘under’ must follow the NP: *a szék alatt* is ‘under the chair’. If now an NP with both a demonstrative pronoun and a postposition is formed, hence semantically speaking *ez a szék* and *a szék alatt* are fused, it will not do to simply merge them but after the demonstrative pronoun the postposition must occur as well: *ez alatt a szék alatt* ‘under this chair’. It would be ungrammatical to omit the first *alatt*.

As can be seen, it is possible that an adposition occurs twice in a PP. However, the syntax of the analysis of the Spanish example remains slightly awkward in so far as both a preposition and a postposition are used under this analysis (one might speak of a “circumposition”). More importantly, however, a disadvantage appears again which already showed up in a similar way with analysis I. The 1st and 2nd persons singular as well as the reflexive do not behave in the same way as all other NPs: whereas the former require a preposition and a postposition, the latter require only a preposition.

3. A DIACHRONIC ACCOUNT

In the previous section three synchronic analyses were presented. It turned out that achieving an advantage in one issue goes along with a disadvantage in another. Therefore there is perhaps no optimal solution how to analyze Spanish *conmigo* etc. However, no decision needs to be made: it is a general fact of synchronic linguistics that often several analyses are possible. An important issue is that awareness has been raised about the facts of the Spanish language regarding the forms in question. As has been seen, it is possible to go beyond previous treatments of Spanish and be more explicit about how this language is structured.

But we will now turn to diachrony. This area is different in nature from synchrony since in diachrony there usually is one historical truth and the linguist’s task is to point this out or reconstruct it, or at least make reasonable proposals on how structures may have developed. In the following, first a look will be taken at the more recent past in 3.1., which means the connection to Latin. Subsequently a rather distant past will be explored in 3.2.

What will not be done is to address the possible future, this being hardly predictable. It is interesting to note, though, that occasionally forms such as *sinmigo* for ‘without me’ have appeared, for regular *sin mí*. Bruyne (2002) cites an attestation from almost half a millennium ago and calls it humorous. Ranson / Lubbers Quesada (2018: 255) report that “one sometimes hears” *sinmigo* and *sintigo* and cite as an example a song *Sinmigo* by Mexican artist José Madero Vizcaíno. It is obvious that the non-standard forms *sinmigo* and *sintigo* arise by analogy with *conmigo* etc., a fact that highlights the semantic connection between the concepts ‘with’ and ‘without’. In Finnish grammar the case indicating absence (‘without’) is called *abessive*; *raha* ‘money’ forms

the abessive *rahatta* ‘without money’. Theoretically, if one day *sinmigo* etc. should become standard usage, in analysis I a case *abessive* would have to be added, while in analysis II a form such as *migo* could not simply be called *comitative* but a compound *comitative-abessive* would be the appropriate designation (and further modifications of the three analyses would be required).

3.1. A Look at the More Recent Past

As laid out in section 1, the three modern Spanish forms under study go back to a slightly larger system in Latin, comprising the five forms *mēcum*, *tēcum*, *sēcum*, *nōbīscum*, *vōbīscum*. These forms are unusual within the system of Latin grammar since *cum* ‘with’ is otherwise a preposition. It must have been used as a postposition, however, after personal pronouns; this is the only explanation that can account for the five special forms – via univerbation.

Upon a closer look one discovers a link to the Silverstein hierarchy, or animacy hierarchy, as it is called by Croft (1990: 111–117); the former term has a slight disadvantage in that it was not only Silverstein who did research on it (with his seminal paper Silverstein 1976), while the latter term is probably still more misleading since it suggests a simple distinction between animate and inanimate – which is not what is intended. The Silverstein hierarchy is a hierarchy for nominals (or whole NPs) which helps describe and understand various phenomena in the languages of the world. From top to bottom, this hierarchy approximately consists of (it can vary from language to language): first and second person pronouns, third person pronouns, humans, other animates, inanimates, masses. If a language has two different behaviours of nominals in grammar, there usually is a point somewhere in this hierarchy, and all nominals above behave in one way, whereas all nominals below behave in a different way. See recently Holst (2019: 74f.) with further references. In language history this point can move downwards or upwards.

Apparently the positioning of the adposition Latin *cum* depended on the Silverstein hierarchy: the very upper end (personal pronouns) required postposing, while the remainder required preposing. On the way to Modern Spanish the point on the hierarchy shifted upwards, limiting the use of special forms to singular forms and the reflexive form, which is indifferent with regard to number; this presupposes that the singular is higher on the hierarchy than the plural. Recently a shift upwards on the Silverstein hierarchy has been shown to have occurred in the history of neighbouring Basque concerning the suffixes *-e* and *-en* for the genitive (Holst 2019: 193–195). Also the morphological and syntactic rules that the three synchronic analyses in section 2 imply can be formulated by referring to the Silverstein hierarchy.

3.2. A Look at a Rather Distant Past

Turning back to Latin, the question can be raised how the conspicuous syntactic system with prepositional and postpositional use of *cum* ‘with’ came into being, and whether the Latin facts perhaps tell something about a distant past. Note that learning about the past of Latin implies also learning something, indirectly, about the past of Spanish. In the following it will be argued that indeed something rather interesting can be gathered from the Latin facts concerning *cum*.

It will first be noted that Latin *cum* has cognates in other branches of Indo-European, e.g. Germanic and Slavic (Pokorny 1959: 612f.); these are simply prepositions, or the corresponding prefixes. As to the Italic branch of Indo-European (that Latin belongs to), we learn about Umbrian *kum* that it is attested both as a preposition and as a postposition (Untermann 2000: 409). According to Weiss (2009: 460, n. 45) Umbrian *kum* is postposed mostly with the meaning ‘at’, and Latin has some other, minor, instances where adpositions otherwise known as prepositions are postposed. However, while all of these facts are worth noting, they do not yet shed light on why Latin *cum* exhibited the syntactic behaviour it did.

In order to arrive at a really interesting insight, it is necessary to turn to Tocharian and Ossetic (from the Indo-European language family). The reason for doing this will be explained subsequently.

1. The two extinct Tocharian languages, Tocharian A and B, were once spoken in what is now the north-west of China. They clearly are languages with postpositions (Winter 1998: 156). However, two adpositions stick out in that they are prepositions: the semantically opposed A *śla* B *śale*, *śle* ‘with’ and A *sne* B *snai* ‘without’ (Krause 1971: 20). Tocharianists are aware of why this is so. The Indo-European languages originally had prepositions, as is still the case in Latin, Greek, Old Indian and many other key languages. Some branches or languages, however, switched to postpositions as the norm. To these belongs Tocharian, and ‘with’ and ‘without’ are remnants; the fact that they are used as prepositions is a syntactic relic. (There is also evidence that ‘in’ once was a preposition since it occurs as the first element in the Tocharian word for ‘today’, Krause 1971: 20.)

2. A similar situation is found in Ossetic, an Eastern Iranian language spoken in the Caucasus. This language, too, has postpositions as the norm. According to Abaev (1964), there are only two prepositions: *æd* ‘with’ and *ænæ* ‘without’ (data from the Iron dialect). Thus also in Ossetic these two prepositions withstood the syntactic change and remained in their position. Subsequent research made Thordarson (2009: 174) point out that there are some other prepositions, but nevertheless the role of ‘with’ and ‘without’ is striking and parallels that in Tocharian, as pointed out by Holst (2013: 26, n. 13).

To sum up, there are at least three languages (Tocharian A, Tocharian B, Ossetic) that indicate that there may be a special role of ‘with’ and ‘without’ among adpositions. In the cases just discussed, seen historically ‘with’ and ‘without’ indicate the former location of the adpositions; they are valuable relics permitting a view at earlier syntax.

It is now possible to try to apply this insight to Latin – and more broadly to Italic, since Umbrian shares properties of Latin. Here it becomes apparent why having discussed Tocharian and Ossetic is useful: if in these languages ‘with’ and ‘without’ can tell something about the former syntax with regard to the typological opposition of prepositions and postpositions, perhaps the Latin facts concerning ‘with’ can do so also. This holds even if in Latin ‘with’ is not paralleled by ‘without’ (which is the more marked preposition of the two and hence of lesser importance; in Latin this is *sine*), and even if postposed *cum* was restricted to the upper end of the Silverstein

hierarchy. It makes sense to interpret the position of ‘with’ as a relic again. In Tocharian and Ossetic, basic constituent order requires postpositions, the prepositions ‘with’ and ‘without’ are the exception and point to former prepositions. In Latin (and Italic) the situation is converse: prepositions are the norm, the role of ‘with’ as a postposition is the exception sticking out, and consequently one should infer a former state with postpositions. (There is an entirely legitimate method of arguing in syntactic reconstruction involved here.)

This may come as a surprise at first. The Indo-European languages typically use prepositions, as laid out, and this must have applied to Proto-Indo-European as well. This does not exclude the possibility, however, that relics of postpositions were part of the linguistic system, and these relics may have been handed down to daughter languages, even if in attenuated form. The Latin and Italic evidence concerning ‘with’ points to precisely that. Germanic and Slavic would have innovated in making ‘with’ exclusively a preposition (Czech *s* ‘with’, for instance, with *s* in this satem language corresponding to *k* in Umbrian *kum*, etc.). If the hypothesis laid out is correct, it would mean that Latin *-cum* in *mēcum* etc. is an archaism of great salience.

It is now time to draw attention to why this discussion is so important. A challenge which has been discussed for a very long time now (from the 19th century onwards) is the question whether the Indo-European languages have further relatives, or to put it in a different way, whether they are part of a larger “macrofamily” located in Eurasia. Evidence from syntax should not be forgotten in such discussions. It turns out that families such as Uralic, Turkic, Mongolic and others in Northern Asia fit the earliest Indo-European rather well with regard to constituent order. To be more specific:

1. Most of these families share(d) SOV (subject – object – verb) word order.
2. Most of these families share(d) GN (genitive – noun) word order.
3. Most of these families share(d) AN (adjective – noun) word order.
4. Most of these families share(d) NIN (numeral – noun) word order.

There had been a problem so far, however, with adpositions. Language families such as Uralic, Turkic, Mongolic and others in Northern Asia have postpositions, while Indo-European had prepositions. If now, based on the considerations in this paper, evidence has appeared from within Indo-European that postpositions occurred, this moves Indo-European towards the other language families in this respect.

It is interesting to become aware of what all this implies. It means, in fact, that Spanish still retains traces of an archaism which is many thousands of years old. In the Indo-European language family one would rather expect such an archaism in Lithuanian, famous for its often conservative traits, and one would not be surprised to find it in Modern Greek, given that this language has preserved many structures from Ancient Greek. However, one would not particularly expect it in Spanish, as a Romance language, and as a language rather remote from Proto-Indo-European, let alone from stages *before* Proto-Indo-European. This shows how surprising historical linguistics sometimes can be.

REFERENCES

- Abaev, Vasilij I. (1964): *A grammatical sketch of Ossetic*. Bloomington: Indiana University.
- Bruyne, Jacques de (²2002): *Spanische Grammatik*. Tübingen: Niemeyer.
- Croft, William (1990): *Typology and universals*. Cambridge: Cambridge University Press.
- Harris, Alice C. (2002): On the origins of circumfixes in Kartvelian. In: Bublitz, Wolfram / Roncador, Manfred von / Vater, Heinz (eds.): *Philologie, Typologie und Sprachstruktur: Festschrift für Winfried Boeder zum 65. Geburtstag*. Frankfurt am Main: Peter Lang. 305–322.
- Harris, Alice C. / Xu, Zheng (2009): Diachronic morphological typology. In: Brown, Keith / Ogilvie, Sarah (eds.): *Concise encyclopedia of languages of the world*. Oxford: Elsevier. 287–293.
- Herce, Borja (2023): Adpositions. In: Lier, Eva van (ed.): *The Oxford handbook of word classes*. Oxford: Oxford University Press. 420–442.
- Holst, Jan Henrik (2001): *Lettische Grammatik*. Hamburg: Buske.
- Holst, Jan Henrik (2013): Parallele Entwicklungen im Ossetischen und im Armenischen. In: *Georgica* 36. 9–30.
- Holst, Jan Henrik (2019): *Forschungsfragen zur baskischen Sprache*. Düren: Shaker.
- Holst, Jan Henrik (2023): *Das Rekonstruieren von Sprachen*. Düren: Shaker.
- Hualde, José Ignacio / Olarrea, Antxon / Escobar, Anna María (2001): *Introducción a la lingüística hispánica*. Cambridge: Cambridge University Press.
- Krause, Wolfgang (²1971): *Tocharisch. Handbuch der Orientalistik, 1. Abteilung, Band 4, Abschnitt 3*. Leiden: Brill.
- Lara, Victor (2019): The geolinguistics of the Ibero-Romance comitative. In: *Dialectologia et Geolinguistica* 27. 51–71.
- Lathrop, Thomas A. (1980): *The evolution of Spanish. An introductory historical grammar*. Newark, Delaware: University of Delaware.
- Marušič, Franc Lanko (2023): Circumfixation. In: Ackema, Peter et al. (eds.): *The Wiley companion to morphology. Vol. I*. Oxford: Wiley Blackwell. 419–444.
- Mensching, Guido (1992): *Einführung in die sardische Sprache*. Bonn: Romanistischer Verlag.
- Penny, Ralph (1991): *A history of the Spanish language*. Cambridge: Cambridge University Press.
- Pokorny, Julius (1959): *Indogermanisches etymologisches Wörterbuch. Vol. I*. Bern / München: Francke.
- Ranson, Diana L. / Lubbers Quesada, Margaret (2018): *The history of Spanish. A student's introduction*. Cambridge: Cambridge University Press.
- Rounds, Carol H. (²2009): *Hungarian. An essential grammar*. London / New York: Routledge.
- Santazilia, Ekaitz (2013): Noun morphology. In: Martínez-Areta, Mikel (ed.): *Basque and Proto-Basque. Language-internal and typological approaches to reconstruction*. Frankfurt am Main: Peter Lang. 223–281.
- Satorre Grau, Javier (2002): Revisión del sistema pronominal español. In: *Revista de Filología Española* 82 / 3–4. 345–380.
- Schpak-Dolt, Nikolaus (²2012): *Einführung in die Morphologie des Spanischen*. Berlin / Boston: De Gruyter.
- Silverstein, Michael (1976): Hierarchy of features and ergativity. In: Dixon, Robert M. W. (ed.): *Grammatical categories in Australian languages*. Canberra: Australian Institute of Aboriginal Studies. 112–171.

- Stolz, Thomas / Stroh, Cornelia / Urdze, Aina (2006): *On comitatives and related categories. A typological study with special focus on the languages of Europe*. Berlin / New York: Mouton de Gruyter.
- Thordarson, Fridrik (2009): *Ossetic grammatical studies*. Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- Untermann, Jürgen (2000): *Wörterbuch des Oskisch-Umbrischen*. Heidelberg: Winter.
- Velupillai, Viveka (2012): *An introduction to linguistic typology*. Amsterdam / Philadelphia: Benjamins.
- Weiss, Michael (2009): *Outline of the historical and comparative grammar of Latin*. Ann Arbor / New York: Beech Stave Press.
- Winter, Werner (1998): Tocharian. In: Giacalone Ramat, Anna / Ramat, Paolo (eds.): *The Indo-European languages*. London / New York: Routledge. 154–168.

LANGUAGE ORIGINS AND GROUP BEHAVIOR: PROTO-SPEECH, SYNTACTIC LANGUAGE, AND SPREAD OF LANGUAGE

PATRICK MANNING

ABSTRACT

The emergence of human language is explored here through a focus on group behavior of humans and other animals, leading up to the assertion that creolization was the final step in creating fully syntactic language, 70,000 years ago. Group behavior generally relies on animal perception of environmental information – audio, visual, and other – and on signals emitted and exchanged among organisms. Groups are widespread in the animal kingdom, with their communication enabling successful species to build new environmental niches and expanding domains. Applying this logic to hominin species has revealed social groups characteristic of primates, focusing on touch, grooming, gestures, and odor. Hominins later developed types of audio exchange, especially calls to assembly of groups for scavenging, which led with time to creation of words as vocalizations of specific concepts. Assembly of words by small groups yielded “proto-speech,” in which group members could exchange phrases of up to four words, but with only elementary knowledge of their links.

A process of creolization led from proto-speech to fully syntactic language. Young Homo sapiens had gained the genetically-based capacity to organize words into hierarchical order and sentences but could apply it only when they were collaborating in autonomous and persistent groups. This hypothesis, proposed in an abstract model, is supported by evidence on the invention of Hawaiian Creole English, Nicaraguan sign language, and debate on creolization, showing that native creole speakers rapidly created successive syntactic processes. The same studies give evidence of such non-linguistic aspects of creolization as social patterns of group size, brain-based individual language archives linked to group discourse, and social institutions to sustain the language. The model is then applied in a hypothetical narrative of the rise and expansion of fully syntactic language.

Keywords: group behavior, proto-speech, creolization, fully syntactic language, universal grammar, we-group, social evolution, institutions

INTRODUCTION¹

The historical study of human language continues to advance on many fronts. But for the origins and early development of spoken language, current study is complicated by several debates. In one debate, scholars ask: how far back in time can one explore the history of language and its groups of speakers? For those who see sharp limits, the practice of reconstructing ancestral languages can be pursued only as far back as mid-Holocene times, for lack of adequate lexical data. In contrast, contributors to *Mother Tongue* have explored aspects of language for earlier times, for instance by projecting certain words of the putative Proto-Human language.² A second debate centers on the pace of linguistic change.³ Did the complexity of syntactic speech require that it developed gradually over the millennia? Or could the creation of syntactic language be sudden? A third debate asks: is language a uniquely human capability?⁴ For those who defend human uniqueness, the dynamics of language can be explored only in human audio communication. But others see parallels in communication among species, arising from flows of information and group behavior among animals broadly.

Derek Bickerton has been prominent in the second and third of these debates. On the pace of change, Bickerton's theory of "language bioprogram" (1984) argued that creole languages were created rapidly by children who relied on universal grammar (UG) to create syntax within the span of a single generation.⁵ Bickerton then challenged the uniqueness of language, conducting cross-species analyses that combined linguistics and evolutionary biology. His expansive view of biology included not only natural selection but the newer analyses of epigenetics, niche construction, and expanding neural systems. But Bickerton gave little attention to patterns of group behavior, another rising area of biological study.⁶ Thus, a recent exploration of group dynamics offers the following summary of group behavior's significance in biological evolution:⁷

Collective behavior is common at the molecular, cellular, and organismal levels, from flocks of birds and insect task allocation to the coordinated movement of cell groups during embryonic development. . . . Individuals perceive their surroundings and the states of other individuals through visual, auditory, and chemical signals, making behavioral decisions accordingly. These local interactions can lead to global collective behavior.

This essay reaches across numerous species and long spans of time to identify links among group behavior, communication, and eventually language; it takes positions on all three debates about early language. For times from 400,000 years ago, it poses hypotheses tracing specific steps toward human spoken language. The essay acknowledges that human speech is unique among animals yet

¹ Acknowledgments: The author expresses appreciation to the late Christopher Ehret for decades of inspiration in study of language in African history; he offers thanks to scholars who have made specific advances to this study, especially R.M.I. Dunbar, S.O.Y. Keita, Gregory Haynes, Stephen Ambrose, and Joel D. Levine.

² Ehret 1998; van Driem 2024; Bengtson 2021.

³ Pinker and Bloom 1990; Tattersall 2009.

⁴ Traxler, Boudewyn, and Loudermilk 2012.

⁵ Universal grammar is the instinctive patterns of children in forming sentences, theorized in the generative linguistics of Noam Chomsky. Chomsky 1957, Bickerton 1981.

⁶ For instance, in the recognition of primate groups at intimate and community levels (Dunbar 1988; Gowlett, Gamble, and Dunbar 2012).

⁷ Zhang, Qiu, and Chen 2025.

reaches still further back to argue that all communication depends on group behavior, tracing animal groups and exploring parallels of human and animal groups that help explain communication through speech. Further, while many aspects of human and animal communication developed gradually, the breakthrough to syntactic language was arguably sudden and innovative.

In the absence of direct evidence on the origin of Pleistocene-era language, unraveling its puzzle can best rely on the modeling of modern creoles. Bickerton's 1981 formulation was that the rise of modern plantation creoles, in which children's experience with their parents' pidgin communication engaged their inherent syntactical skills, resulted in elegant, simplified, and remarkably similar creole languages. This, Bickerton reasoned, was a prototype for the Pleistocene-era creation of an initial creole out of the pidgins of parents. For his model, Bickerton chose Hawaiian Creole English as the best-known example of a modern creole. In addition, I have included the case of Nicaraguan sign language: it is well documented and adds important variance. This case shows that children, if denied audio communication because of deafness, turned their impulse to communicate toward the next sensory option, that of visual communication through gestures.

Human children, in initiating speech, articulated their genetic capability to organize concepts hierarchically but could not apply it either in audio or visual expression until they overcame the barrier to their group collaboration. Only once young children found themselves in groups of some permanence, autonomy, and sufficient size (about 15 children), were they able to exchange their syntactic formulations and create the beginning of a fully syntactic language. Modern children lived in a world filled with institutions facilitated by language, which gave order and structure to life—even if they were the oppressive institutions of plantations. Pleistocene children, in contrast, had no institutions until they created their own patterns governing language use.

Throughout this essay, an effort is made to maintain contact with issues raised in the opening section on early animals: sense perception, links among individuals, and the rise of group behavior. These processes have retained their roles in the developments of later steps in evolution, all the way to spoken language and beyond.

EARLY ANIMALS, FROM SENSE PERCEPTION TO GROUPS

Group behavior had its origin in early evolutionary times, with the interaction of animals and their environment.⁸ The global environment, even at the time of single-celled organisms over three billion years ago, encompassed the full extent of the Earth's waters, lands, and skies. Not just the masses and physical forces of the environment but also the waves of information about them echoed across the planet. Within this environment, multicellular animals first emerged some 800 million years ago. These organisms, with their nucleated and differentiated cells, evolved out of single-celled organisms.⁹ Their animal existence centered on material systems existing in a given environment, with intake of food, liquids and gases, accompanied by the internal circulation of that material and the expulsion of waste materials from the body. The successful early development

⁸ Vannier et al. 2019; Hsieh and Plotnick 2020.

⁹ Brunet and King 2017. Cyanobacteria, producing molecular oxygen through photosynthesis, contributed to the evolution of multicellular animals.

of these tiny yet differentiated animals is reflected in the wide range of their descendants, who have adapted in form and function to the immense variation in the Earth's environment.

But early animals, buffeted by all the forces of the environment, had to adapt to environmental pressures. In order to thrive, animals gathered information on the dynamics of their environment: information on the Earth's physico-chemical processes and also its biological organisms. Animals developed systems of sense perception, developed related patterns of neural exchange of information, and formed systems of bodily motion at individual and group levels.

The *input structure* of animal information systems began with sense organs. Within each type of early animal, specialized multicellular organs formed as receptors, able to identify specific elements of the environment, so that the organism could respond to them. Input categories included information on taste, motion, vision, touch (bodily touch and physical force), sound, odor, temperature, and balance. The *neural system*, transmitting information, linked neurons to receptor organs, translating chemical species and waves of light, sound, and pressure into electrical pulses transmitted through neurons, directing these electronic flows to ganglia and brains for interpretation and decisions on actions. The *output system* of motion and information received electronic messages and translated them into the flexing of selected muscles that moved body parts. Categories of output included bodily motion and touch, sound, and odor; in addition, information on bodily motion was carried by waves of ambient light as output. In sum, individual organisms gave *reflexive responses* to environmental information. For these reflexive motions, the reception of input data and response of movement completed the cycle of adjustment to the environment.¹⁰ For instance, the force of gravity is measured and then yields adjustments for bodily balance, while the touch of wind or water pressures may lead to body movement.

While some animals live mostly individual lives, most animals experience one or more types of group behavior, in which individual actions are compounded into group patterns. The first group pattern is the *reflexive movement* just described, in which a sense is received and an action results. The sense of taste tells an animal whether or not material is edible. For information on the touch of waves of water or sound waves, the organism may respond with motion. Where multiple organisms of the same species received a single environmental impulse, they responded in parallel, giving the impression of a group action even though the animals were responding to the impulse and not to each other.

The second group pattern is *collaborative interaction*, in which a sense is received from the environment and one organism responds by emitting a *signal* that reaches another. Signals included the emission of chemicals for smell, bodily action for touch, and sound for hearing. These types of signal output, translated from the neuron-based messages to the action of output, enter the environment and may be received by the receptors of other organisms, especially of the same species. Another signal is visual, emitted indirectly as bodily motion gives rise to a visual signal through reflection in ambient light.

¹⁰ Reflexive motions are also studied in human infants.

Beyond the emission of signals, a subsequent step allows for the creation of *group behavior*, especially groups of a single species. A first animal emits a signal that is received by a second; the second then selects a response and sends output into the environment. The first organism receives the output as input and makes its own response. In this way, *collaborative groups* arise, in which groups of organisms can generate coordinated responses to the impulses of the environment or the actions of each other.

The combination of the two types of group serves such functions as sexual reproduction, conflict (or peaceable relations) among organisms of the same species, defense against outside influences, and expansion of the species domain. The trans-species accumulation of individual, reflexive, and collaborative actions brought great diversity to the patterns of animal life. The diversity of animal life stretched across the scales of individual phenotype, sense perception, individual and group action, and species-level phenotype and domain.

The diversity itself arose through processes of natural selection.¹¹ Natural selection is now understood to take place through overlapping mechanisms, affecting organisms at levels from the molecular to that of whole species.¹² Darwin's 1859 theorization used the term "natural selection" to refer to the process of change in an organism's observable traits at both the individual scale and the species scale.¹³ After 1900, the field of genetics arose, focusing at the molecular level on DNA in nucleus and mitochondria for individuals and for sample populations. The slow pace of genetic change set the genetic instructions needed for an organism to develop and function.¹⁴ Epigenetic change, discovered in the 1960s, takes place through protein action, producing heritable changes in gene expression without altering the DNA sequence. It can lead to rapid change in phenotype and life-course development.¹⁵ Niche construction, unfolding at the phenotypical and species levels, reflects the actions of organisms in modifying the local environment, as in ways to get food.¹⁶ Niche construction feeds back with genetics and epigenetics in providing selective pressures on the genome and on life-course development. Further, neural evolution and expansion strengthened information transmission within organisms; while animal groups, both reflexive and collaborative, exploited varied environments and expanded species domains. The term "natural selection," still in use, is increasingly understood to apply to the ensemble of these mechanisms that contribute to evolutionary change.

Trilobites, ubiquitous early animals, exemplify this diversification through natural selection. These Cambrian-period arthropods, with external skeletons, three main body segments, and averaging 3-10 cm in length, were commonly preserved as fossils.¹⁷ Trilobites had thousands of species

¹¹ Schlosser (2018, 301) argues that, for sense organs among vertebrates, "... photoreceptors, mechanoreceptors, and chemoreceptors used in these sense organs have a long evolutionary history. and homologous cell types can be recognized in many other bilaterians or even cnidarians."

¹² Massari 2024; Dimijian 2012; Ashe, Colot, and Oldroyd 2021.

¹³ Darwin 1859.

¹⁴ The terms "phenotype" and "genotype," coined in 1911, distinguish the observable characteristics of an organism from the underlying genetic structure.

¹⁵ Moore, Le, and Fan 2013.

¹⁶ The term "niche construction" was coined by Odling-Smee, Laland, and Feldman (2020), based on earlier work.

¹⁷ Fossils revealing a sudden shift in trilobite forms gave the key to the punctuated history of evolution and to the mechanism of epigenetics. Eldredge and Gould (1972).

and varied lifestyles; many walked the ocean bottom, others swam. Their compound eyes had numerous units of transparent calcite. Their groups included single-file groups for migration and defensive clusters during the molting of their shells.¹⁸ Such sense receptions and signals are certain to have developed at the early, pre-Cambrian stage of animal evolution, since they are known throughout the Animal Kingdom. The same processes—transforming environments as well as species—have continued ever since.

REFLEXIVE AND COLLABORATIVE GROUPS: POST-CAMBRIAN

The early Cambrian Period, from 540 ma to 485 ma, brought a rise in Earth's temperature from a low of 12° to an average of 22° C, during which animals grew in size and formed almost all of the major animal subgroups – known today through fossilized remains of their shells and skeletons.¹⁹ The principal groups of species included *arthropods* (such as insects, crustaceans, and spiders), *vertebrates* (such as fish, amphibians, reptiles, and birds), and *cephalopods* (squids, octopi, and cuttlefish), along with smaller orders.²⁰ All had full sets of sense perceptions; all relied on their senses to form groups and enact group behavior. Mammals and marsupials formed later among vertebrates, about 210 million and 125 million years ago.

Group behavior of animals, arising from sense perception, falls into the two main categories of *reflexive* and *collaborative* group behavior. For *reflexive* behavior, individual motions, reflexive responses and reflexive groups yield the basic results of sense perception. These responses rely on the “environmental” senses: information on taste, gravity, magnetism, temperature, and the touch and pressure of wind, water, and solids. In present-day examples of reflexive responses, Atlantic herring and other fish form tightly linked “schools” of thousands of individuals for rapid movement, for instance to loci of spawning. The schools function through each organism's response to pressures and touch from the larger group.²¹ No fish sends a signal: instead, the physical waves generated by the swimming of each provokes adjustment to its course and resultant waves to be felt by others.²² In the atmosphere, somewhat parallel waves are created by the flight of birds, so that geese and some other large birds fly in V-formation. Each tailing bird benefits from the updraft formed by the birds ahead of it, making flight more efficient. On land, reflexive groups do not function as smoothly as in water and air, but antelope and other ungulate mammals run in formation, relying on vision to set their distance from each other. A quite different sort of environmental sense enables animals to choose direction. These animals, including some snails, contain fragments of iron oxide to sense the direction of the magnetic poles for orientation. In a spectacular such case, individual Arctic terns migrate annually between the Arctic and Antarctic, relying on

¹⁸ Trilobites survived from about 520 to 250 ma.

¹⁹ On debates about the Cambrian era, see Zhang and Shu 2021; on the variety of species after the Cambrian, see Droser and Finnegan 2003. Average Earth temperature today is 15° C.

²⁰ These are the principal categories of animals that survived waves of post-Cambrian extinction.

²¹ Niwa 1994.

²² On leadership within herring shoals, see Krause, et al. 2000.

this sense for navigation.²³ These few examples suggest how reflexive groups function. There appear to be great numbers of reflexive actions and groups at individual and group levels, and they overlap with collaborative groups. Over the course of evolution, growth in numbers of neurons in many animals is likely to have deepened the sensitivity of information and reflexivity.²⁴

Somewhat more complex are the *collaborative* groups, which arise from the smaller number of “collaborative” sense inputs, which may lead to emission of signals that ultimately link organisms into a group. The types of collaborative input and output information include: direct and bodily *touch*; *olfactory* information that diffuses the odor of small chemical species; and for *audio*, sound waves of various frequencies, emitted by types of bodily friction. For signals of *motion*, motions may be reflected in ambient light, both at the level of formulating motion and the level of propagating information on the motion.

Recent research has documented collaborative group behavior in a wide range of animals.²⁵ The brief examples included here give indications of the range in species known to act in groups and the distinctive group activities of each species. On the behavior of ants, naturalist John Lubbock showed in the 1880s that ants used odor trails in foraging.²⁶ An ant, after finding food, leaves a trail of pheromones or chemical traces on a trail back to the nest, which enables others to find the food.²⁷ Ants are now thought to make sounds as well, especially at the pupae stage. Bees rely especially on two types of visual links through “dance” in showing paths to flower blossoms. In the 1920s, Karl von Frisch described how bees returned to the hive and performed the “round dance” to locate nearby feeding places and the more complex “waggle dance” for distant blossoms, identifying the direction and elevation of the blossoms.²⁸ Bee observations depended on scent, color, magnetism, and using their compound eyes. Bees observing each dance translated and acted on the message. In each of these cases, collaborative groups exchange information and take action of value to the group.²⁹

Among primates, the vervet monkeys of East Africa developed distinct audio warnings against their main predators: leopards, eagles, and pythons.³⁰ The leopard call signals the need to climb a tree; the eagle call signals the need to look up and hide under a bush. Beavers in North America and Eurasia respond to the sound of running water by building dams; the resulting ponds provide homes for numerous plants and animals and include beaver lodges with underwater tunnels for entry. Beaver mates and their offspring maintain their groups with links of scent, vision, touch, and audio, working collaboratively in maintaining dams and lodges. Beavers eat wood and other

²³ Nordmann, et al. 2017.

²⁴ Frisch 1953.

²⁵ Wikipedia, “Animal Communication.”

²⁶ Lubbock 1882. Pheromone (a term coined in 1959) – an emitted chemical that serves the function of generating a reaction from other organisms.

²⁷ Hölldobler and Wilson 1990.

²⁸ Frisch 1953.

²⁹ George Williams (1966) convinced many readers that genetic evolution could only be effective at the level of individual organisms, not groups – thus stifling study of group behavior for decades. But as Bickerton (2009: 10) notes, group behavior is repeatedly shown to differ from individual behavior, even when the mechanism is unknown.

³⁰ Seyfarth et al. 1980.

vegetation; they are strong swimmers and can hold their breath for six minutes.³¹ They groom their own and others' hair for waterproofing; glands excrete castor oil for grooming and for marking territory against invading beavers. A slap of the tail warns against danger. Young beavers will travel as much as ten kilometers to find a space for building a dam.

These and other cases are studied within the field of Animal Communication Systems (ACS), a subfield of ethology or animal behavior.³² The forms of signals vary greatly by species, but their functions are much the same, conveying messages on survival, reproduction, and the social order.³³ Much early work in ACS was phenotypical description of animal groups. Animal signals include bodily movement or display, facial expression, gaze-following, color change, and bioluminescence. Audio signals are conveyed through *stridulation* (friction of body parts) and by vibration of swim bladders in bony fish. Chemical links are complex but are under study. In general, animal communication responds to external stimuli and focuses on the here and now, without details on time or distance. That is, Bickerton specified that ACS does not allow for *displacement* in time or space and can only convey whole messages.³⁴

The study of African fruit flies has opened possibilities for expanded understanding of animal communication. For over a century, fruit flies have been essential experimental animals for study of genetics, so they are well documented. From 2012, Joel D. Levine expanded the study of their group behavior.³⁵ Groups of from 6 to 24 flies were placed in containers and videotaped for 30 minutes, then reviewed in slow motion. Phenotypical description revealed links by scent, touch, sound, and vision, and showed that groups enable flies to balance their distance, facilitating feeding and egg-laying. Then the group interactions were coded for statistical analysis. As a framework for analysis, the researchers turned to Social Network Analysis (SNA), which had been applied for decades in study of human groups.³⁶ The advantage of SNA was the ability to calculate patterns of small and larger groups, though it is difficult to specify which links are by touch, sound, or vision. Levine's team found ways to quantify links among fruit flies, then went further and discovered a gene that influenced behavior in large vs. small groups, suggesting that flies could count their group size.³⁷ Up to this point the team had advanced the study of animal groups to trace the interaction of multiple scales: the genetic base, sense organs, data on sense reception and signals, links among organisms, functioning of a group under SNA analysis, and species-level interpretation of group behavior. Thus, they have further opened discussion on how organisms respond to environmental signals with sense assessment and group behavior.

³¹ Anon., 2024.

³² Wikipedia, "Animal Communication."

³³ Bickerton (2009, 16–19); Hauser 1996.

³⁴ Bickerton 2009 51–53.

³⁵ Schneider, Dickinson, and Levine 2012; Rooke et al. 2020.

³⁶ Freeman 2004.

³⁷ Rooke, et al., 2024.

SENSES AND GROUP BEHAVIOR IN HOMININ SPECIES

Hominin species came into existence some six million years ago among the primates of East Africa, as the region shifted from forest to savanna.³⁸ Some of the primates moved west into remaining forests, where their descendants became chimpanzees. Those who remained in East Africa evolved into hominins—surviving by gradually shifting their habitation from trees to the ground and their diet from mainly fruit to leaves, flowers, bark, insects, small animals, and a few fruits. Here we focus on hominin evolutionary processes for sense perception at the micro level and on niche construction at the level of species and groups.³⁹

By the time of the well-preserved Lucy skeleton of the *Australopithecine* genus, 3.2 ma, hominins had given up trees and fruit to become fully omnivorous and bipedal.⁴⁰ With the change in environment and phenotype, hominin sense capacities are certain to have changed: for instance, hearing became more precise and able to distinguish sounds at a distance.⁴¹ One may also expect that the sense of gravity adjusted to bipedal balance, the sense of taste expanded to assess the wider diet, and vision became sharper at distances. The sense of individual touch may have remained largely the same, relying on interpersonal grooming. The *Australopithecine* species of this era formed a *terrestrial omnivore* niche, working with tools of bone and wood for wide-ranging foraging.⁴²

Late in the era of *Australopithecine* species, from roughly 2.5 ma, a more productive lifestyle emerged: a *catchment scavenging niche*.⁴³ In it, members of *Homo habilis* used stones to break open bones left from carcasses eaten by other species, and then consumed the nutritious bone marrow.⁴⁴ Such bones were found near waterways or in stone outcrops. The pattern was confirmed since the cuts and scrapes on bones made by hominin were on top of the marks left by other predators who had exploited them first. This niche and the consumption of marrow helped to build the size of brains.

A new and adventurous hominin niche arguably emerged at about 1.8 ma. Derek Bickerton and biologist Eörs Szathmáry assembled the arguments of others to postulate *confrontational scavenging*, in which groups of hominin seized meat from newly deceased large animals.⁴⁵ A key element of this lifestyle was the development of a *recruitment call* by adults to gather a crowd and go to the carcass site. The recruitment call had some parallel to the alarm calls of vervets: each call was a complete message for the here and now, and thus an example of an Animal Communication System. Once at the site, males drove off competing scavengers while females cut off meat using

³⁸ For a concise but authoritative review, see Pontzer 2012.

³⁹ For other primate species, larger groups correlated with expanded auditory communication; galada monkeys, with unusually large groups, developed a sort of vocal choiring. Dunbar 2024.

⁴⁰ Hendry 2000.

⁴¹ Data available on early hominin sense capacities center on sense reception in the brain and on auditory capacity. See Kaas 2007 and Quam et al. 2015.

⁴² Bickerton 2009.

⁴³ Bickerton 2009.

⁴⁴ Bickerton 2009. The fossils known as *Homo habilis*, discovered and named by L. S. B. Leakey, are now widely thought to be closer to genus *Australopithecus* than to *Homo*.

⁴⁵ Bickerton and Szathmáry 2011. Bickerton (2009) initially used the term *territorial scavenging*. Bickerton and Szathmáry did not estimate the numbers of recruits assembled.

Acheulean hand axes, with edges able to cut through the skin and flesh of dead animals. The hominin scavengers then moved off with their meat, though their lifestyle also relied on omnivorous foraging. Meat consumption helped to expand brain size and neuron density, so that community groups grew to perhaps 70 members.⁴⁶ Members of *Homo ergaster* were the first such beings, tall and lanky by comparison with their predecessors, and better able to run. Paleontologist Ian Tattersall emphasizes that the new species arose rapidly, through a system-wide epigenetic process rather than through gradual genetic change.⁴⁷ In an important but still-undated transformation, at some time within the past two million years, hominin species switched gradually from multiple mating to pair-bonding, eventually leading to the formation of households of an average five members. In this change, battles among adult males declined and mortality of offspring declined.⁴⁸

Rapid creation of confrontational scavenging was followed by a long period of migratory expansion, from 1.8 to 1.2 ma, with little change in lifestyle or technology.⁴⁹ In this era, *H. ergaster* spread throughout African highland savannas and *H. erectus* spread to Asian highlands. For these occasional migrations, I assume that groups of about fifteen hominins hiked the distances from one highland region to another, managing to cross the lowland regions before settling and building communities of 70 at their destination.⁵⁰ Greater temperature tolerance and sensitivity presumably developed along with migration.

Hominin migrations were governed by at least two types of climate change, in response to variations in the Earth's orbit about the Sun. The first pattern, the East African monsoon, facilitated migration and biological interaction along Africa's Rift Valley, with alternate migrations to north and south.⁵¹ In addition to the yearly sequence of northerly and southerly winds, the monsoons brought relative humidity to Africa's northeast for 11,000 years, then to the south for the next 11,000 years. Second, major worldwide climate cycles were documented based on Marine Isotope Stages, with 100 stages measured within the past two million years.⁵² These alternating periods of warm and cool climate encouraged hominin migration in warm, humid times at the peak of odd-numbered MIS cycles. The cycles averaged 36,000 years between 1.9 and 1.0 ma (MIS 71–23) and then became longer, averaging 103,000 years from 1.0 ma to 100 ka (MIS 21–5).

Human remains in African archaeology are scarce between 1.2 ma and 400 ka, a period of most of a million years. Aside from periodic migration, it appears that the hominin lifestyle remained consistent during that time. Nevertheless, the physical type of the leading species underwent changes: of the African hominin remains dated between 800 and 300 ka, most have been classified as *H. heidelbergensis*.⁵³ For this species, one may propose three types of dates in its

⁴⁶ Gowlett, Gamble, and Dunbar 2012.

⁴⁷ Tattersall 2012.

⁴⁸ Chapais 2008; Gavrillets 2012; Wilson and Mesnick 1997; Manning 2023b. This complex transition might have been associated with the rise of a new species – of *H. ergaster*, *H. heidelbergensis*, or *H. sapiens*.

⁴⁹ Bickerton (2009, 211–215) described “the long stagnation.”

⁵⁰ Dunbar (2020: 8) illustrates the fractal structure of human social groups; each level expanded by a factor of roughly 3 (e.g. 1.5, 5, 15, 45, 140, 450, 1500 ...). See also Manning 2023a.

⁵¹ Governed by the Earth's precession cycles of 23,000 years.

⁵² Hirst 2019. Marine Isotope Stages, first published in 1955, are based on varying ratios of oxygen isotopes in fossil algae found in Swedish-mined cores from the ocean bottom, showing their correlation with glaciation.

⁵³ Buck and Stringer 2014.

speciation: the time at which its lineage separated from its last common ancestor (likely *H. ergaster*), as far back as 1.2 ma; the time of its full speciation, roughly 400 ka; and the time of its extinction, roughly 200 ka. Its estimated community size was 120 members.⁵⁴

The few known archaeological remains suggest only modest evolutionary change, such as greater precision in the shaping of hand axes.⁵⁵ Excavations also show occasional use of fire and some hunting with hafted spears, although these occasional advances did not reach general usage until the Middle Stone Age (MSA), after 350 ka. Human populations rose and fell, though genetic efforts to identify population bottlenecks have yet to yield definitive results.⁵⁶

During this long period of slow evolutionary change, continued reliance on confrontational scavenging eventually broke through the ACS limit on displacement. That is, vocalization gradually extended from the recruitment call to include a modest number of *words* and associated *concepts* that could specify time and space, which Bickerton categorized as the beginnings of “proto-language.”⁵⁷ Bickerton reasoned that early words arose from the high-end scavenging niche, from rare but recurring events of high impact such as flash floods, and from acts and words of negation to create the word “no.”⁵⁸ Like the initial recruitment call, each word and its concept was developed by adults and adolescents who shared the experience behind the word.⁵⁹ For many millennia there existed no syntax to link the words, yet each word was a step in complexity because it moved beyond the here and now, opening to distance in space and time: thus, a word might specify a “leopard” as a concept, designating a leopard at any time or place.⁶⁰ Bickerton, while assuming neural change, did not allow for structured relationships among words in early “protolanguage.”⁶¹ Chomsky and Fitch, however, each emphasized that internal concepts came before spoken words, so that concepts and communication were not the same.⁶² A further limiting factor on the development of proto-speech was the limits on size and intimacy of social groups: words and gestures might be shared among small and localized groups, but there was yet no process for sharing them over large populations.

⁵⁴ Gowlett, Gamble, and Dunbar 2012.

⁵⁵ I follow the convention of using *ergaster* for early *Homo* in Africa, and *erectus* for early *Homo* in Asia.

⁵⁶ On genetic bottleneck methods, see Bunnefeld, Frantz, and Lohse 2015. For a recent yet unverified assertion of a major bottleneck between 900 and 800 ka, see Hu et al. 2023.

⁵⁷ “But syntax, I began to realize, may have become possible only because two million years of protolanguage use brought about significant changes in its user’s brain” (Bickerton 2009, 50). This recruitment had some parallel with social insects, as bees and ants were able to develop signaling that enabled individuals to spread word about food; their brains may have been too small to develop words.

⁵⁸ Bickerton 2009, 218–222.

⁵⁹ Bickerton (2009), in contrast, explores infant learning of proto-speech.

⁶⁰ Bickerton 2009.

⁶¹ Bickerton 2009, 207.

⁶² Berwick and Chomsky 2016; Fitch 2019.

SOCIAL GROUPS, PROTO-SPEECH, AND THE MIDDLE STONE AGE

Recent decades have developed a picture of accelerating human evolution within the past 400,000 years. To begin with archaeological studies, recent digs have revealed five clear waves of migration from Africa into Arabia. The path of the migrants is not known, but archaeological finds in North Arabia show that the peaks in MIS humidity brought numerous settlers to the same region at 400 ka (MIS 11), 300 ka (MIS 9), 200 ka (MIS 7), 100 ka (MIS 5), and 55 ka (MIS 3). The settlements are documented by remnants of the stone tools in each period.⁶³ The timing and distance of these Arabian migrations provide a model for migration throughout



Assemblages from Khall Amsayah and Jubbah lake sites in north Arabia (Groucutt et al., 2021). From left to right: A (ca. 400 ka); B (ca. 300 ka); C (ca. 200 ka); D (ca. 100 ka); E (ca. 55 ka). Scale bar: 1 cm.

Africa, suggesting displacements of at least several hundred kilometers at times of the MIS peaks in humidity. The migrations at 400 ka and perhaps 300 ka were presumably of the *heidelbergensis* species; migrations of *Homo sapiens* likely included those beginning 300 ka and conceivably even 400 ka.

While archaeology and paleontology had long been the principal sources of information on human evolution, new understandings of biological evolution arose, from the 1960s, in adjoining disciplines. W. D. Hamilton's theoretical work in population genetics showed that kinship was not simply from parent to offspring but that siblings and cousins also carried forth a closely related genome.⁶⁴ The empirical notion of punctuated evolution, meaning occasional rapid change in phenotype (first documented for trilobites), led to theorizing the mechanism of epigenetics and to revived studies of life-course development.⁶⁵ The notion of niche construction gained increasing empirical confirmation, so that feedback with environmental change came to be treated as part of

⁶³ Groucutt et al. 2021.

⁶⁴ Hamilton 1964.

⁶⁵ Eldredge and Gould 1972.

natural selection overall.⁶⁶ Genomic studies of modern and ancient DNA have tended to suggest relatively early dates for separation among species ancestors, hence earlier dates for speciation.⁶⁷

For primates, after comparative studies revealed both intimate- and community-level groups in all species, comparison of fossil hominin crania showed community size to correlate with the size of the cerebral cortex, so that hominin communities grew from 70 for *H. ergaster* up to 150 for *H. sapiens*. All of these analyses have received some degree of empirical confirmation.⁶⁸ Other biological hypotheses formulated after 1960, known in general as *cultural evolution*, have sought to show expanded collaboration in hominin behavior.⁶⁹

How and when did *Homo sapiens* appear on the African scene? New data and analysis in paleontology have expanded knowledge and debate. The various dates of *sapiens* speciation may be estimated as 500-400 ka for separation of its lineage from that of its last common ancestor (separating *sapiens* from *heidelbergensis* and *neanderthalensis*); while its full speciation is estimated variously at from 350 to 200 ka. It may well be that such defining characteristics of *H. sapiens* as community size of 150 arose within the lineage before full speciation.⁷⁰ The paleontology of *H. sapiens* highlights the Moroccan Jebel Irhoud skull, recently dated at 300,000 years ago.⁷¹ But no uniform phenotype in *H. sapiens* appeared until after 100,000 ka. In recent years, two models of speciation have gained most attention. First, Tattersall's model of a distinctive population, arising through a broad epigenetic shift at about 200,000 years ago, led to rapid population growth and spread throughout the continent.⁷² Second, the model of Scerri et al. assumes an earlier and more gradual set of somewhat independent developments, with the nearly completed *sapiens* appearing widely by 300,000 ka.⁷³

The concept and experience of the Middle Stone Age (MSA) brought together the analyses and discoveries listed just above. The MSA took form as recently as 280 ka and as early as 350 ka. The path-breaking 2000 review of McBrearty and Brooks demonstrated that innovation in the material and cultural practices of African communities expanded substantially during the MSA rather than wait for a “revolution” a scant 40,000 years ago. MSA stone technology relied on prepared cores, from which major flakes were broken off. Blades and microliths were then hafted to spear handles with glue or binding, especially in East and Southern Africa.⁷⁴ Other expanded practices included hunting for large game, frequent fire for cooking; varying uses of ochre in decoration, and gift or exchange networks for obsidian. Meanwhile, Acheulean tools were abandoned by 130 ka.⁷⁵ Archaeological sites of the MSA are described as a “mosaic” in that they exhibited

⁶⁶ Odling-Smee, Laland, and Feldman 2020.

⁶⁷ Cann, Stoneking, and Wilson 1987; Stringer 2016.

⁶⁸ Dunbar 1988; Gowlett, Gamble, and Dunbar 2012

⁶⁹ These include social learning (Bandura 1971), dual heritage (Boyd and Richerson 1985), cumulative collaboration and intentionality (Denning 1987, Tomasello 2010), and domestication of humans (Thomas and Kirby 2018).

⁷⁰ It is not known when pair-bonded households became predominant among humans, whether the Merge capability spread among early *Homo sapiens*, or whether bottlenecks narrowed genetic diversity. Manning 2023b; Zaccarella and Friederici 2015; Hu 2023.

⁷¹ Hublin et al. 2017.

⁷² Schwartz and Tattersall 2010.

⁷³ Scerri et al. 2018

⁷⁴ Barham and Mitchell, 2008; Ambrose 2001; Greenfield 1991.

⁷⁵ MacDonald, et al. 2024; d'Errico, Henshilwood, Vanhaeren and van Niekerk 2005.

regional and temporal variation in place of the continental uniformity of earlier times. The term “mosaic” is used as well to describe variations in bodily remains, whether within *heidelbergensis*, *sapiens*, or other communities.⁷⁶

How were speech and language linked to the MSA? In exploring the development of hominin vocal communication, I follow a modified version of Bickerton’s “protolanguage”—a term that I replace with “proto-speech.”⁷⁷ Bickerton proposed a sequence of steps in communication, beginning ca. 1.8 mya with the call to assembly that launched the long-lasting system of confrontational scavenging, followed by the gradual emergence of individual words and, after more than a million years, the rise of proto-speech with phrases of up to four words ordered arbitrarily as beads on a string until they came to be combined through elementary syntax such as noun phrases. Bickerton gave general indications on the timing of these steps but without links to hominin biological change or group social organization.⁷⁸ More specific on both timing and topical connections was archaeologist Stanley Ambrose, who hypothesized two stages in the evolution of faculties for planning. For Stage #1, at the start of the MSA, he argued that manufacture of multicomponent artifacts, such as hafted spears, required an expansion of syntactic verbal communication, so that “composite-archive manufacture and grammatical language also coevolved.”⁷⁹ Meanwhile, such other practices as grooming, gestures, dance, singing, and teaching likely brought modest advances in communication.⁸⁰ Stage #2 of Ambrose’s argument, beginning about 70 kya, is discussed below.

Clearly, the MSA—including possible ties to proto-speech—amounted to a new hominin environmental niche. That is, hominin relations with the environment changed with the new uses of stone, fire, ochre, and more, linked by verbal and other communication. Altogether, these changes are likely to have generated feedback yielding selective pressures on the genome, bringing further phenotypic changes. While the changes in material culture are well documented, fuller analysis will be needed to clarify changes in group behavior and the scope of proto-speech. For instance, in what ways did hominin group organization change with the MSA? Few details are known. Overall community group size at 300 kya is estimated at near to 120, and it is argued that there were subgroups at various scales.⁸¹ But dates of emergence are not known for households based on pair-bonded couples, nor is it known what specific interactions among individuals formed the links in groups and subgroups (e.g. in voice, touch, gesture or other means).

For proto-speech, how widely and in what form did it spread during the MSA? Words developed in every region, yielding local versions of proto-speech rather than a widespread common lexicon. These versions of proto-speech, based on local invention, were spoken primarily by adults

⁷⁶ Scerri, et al. 2018.

⁷⁷ Bickerton (2009) used “protolanguage” for non-syntactic language, neglecting the long-standing use of the same term for ancestral language. I use “proto-speech” for non-syntactic language, in place of Bickerton’s “protolanguage,” with the difference that I refer simply to “words” until they begin to be assembled into phrases and then apply “proto-speech” for the era of MSA. Bickerton 2009, 18–43.

⁷⁸ Bickerton 2009. The era that I label proto-speech, he suggests, brought innovations in phonology, morphemes, and lexicon but not sentences.

⁷⁹ Ambrose 2010, S138–140; see also Greenfield (1991) and Schapiro and Ambrose 2015.

⁸⁰ Laland 2017.

⁸¹ Gowlett, Gamble, and Dunbar 2012, 696; Dunbar 2020.

but also by children; yet the versions were not genetically related to each other in the way that later, syntactic languages were related.⁸² Pronunciation and meaning were learned by each individual; they were not inherited. Individuals stored words in brain-based archives and shared discourse among local group members. In my view, such local groupings did not permit the development of pan-human words, nor perhaps even words shared by a full hominin community of 150.

This scattered nature of proto-speech may nevertheless have brought the beginnings of universal grammar. As toddlers sought to identify patterns in adult speech, they found rules for certain patterns in word order—common rules that could be applied within varying proto-speech versions.⁸³ Over time, the rules could become instinctual through gradual genetic change. Children thus gradually developed instincts in understanding some of the basic proto-speech of adults. But the MSA social order did not combine small children into persistent groups, so the children were unable to share their skills and expand their own syntax. This barrier remained in place for some time. Aside from these limits, however, proto-speech and its elementary syntax likely provided an important advance in communication during the MSA

COMPREHENSIVE SYNTAX, DISCOURSE, AND SOCIAL INSTITUTIONS

The Proto-Human language arguably arose not as the genetic offspring of an earlier language but as a creole invented *de novo*, in a fashion comparable to that for other creoles in later times. To illustrate this argument, this section begins by illustrating the dynamics of creolization through compact narratives of two modern cases of emerging pidgin and creole languages. I turn then to a review of creolization theory, especially from 1980 to 2000, showing how initially eclectic approaches came to be dominated by universal grammar in a generative approach, yet with continuing complexities. The section concludes with four groups of hypotheses on the rise of the Proto-Human language through creolization in the Pleistocene era.

Two outstanding cases of creolization. The story of pidgin and creole speech in Hawaii began in the 1820s, when a “stable and expanding pidgin” began to form in the small initial sugar industry, relying on immigrant adults, with a lexicon based on the dominant Hawaiian language.⁸⁴ This pidgin was destabilized in the late 1870s by a large-scale immigration of foreign sugar laborers, speaking over ten languages, continuing to 1930. Plantation workers developed a new pidgin: “Hawaiian Pidgin English” (HPE) had unstable structures from the many substrate languages and a lexicon from the English of plantation owners. It became the native language of the children of migrants from 1880 to 1900. The second generation of children, born between 1900 and 1920, drew on their capacity for universal grammar and shifted from HPE to a smoother and more sophisticated Hawaiian Creole English (HCE): this became the dominant language among young people of the plantations thereafter. Ironically, the creole language, HCE, is now popularly known as “Pidgin” by its speakers.⁸⁵

⁸² These assumptions are based on Dunbar 2020 and Manning 2023a.

⁸³ Bickerton 1984, 178–182.

⁸⁴ Bickerton 2014, 250.

⁸⁵ They now number an estimated 600,000 native speakers, with an additional 400,000 second-language speakers out of a Hawaiian population of 1.5 million. Yet Pidgin is not taught in schools and it is not an official language of the state. Academic organizations

In the Nicaraguan story, creolized sign language for the deaf arose through social revolution. The Sandinista Revolution in 1979 brought a program of schooling for all. From 1980, deaf children were brought from isolation to public schools – especially the Moreno primary school in Managua—with the intent of teaching them Spanish. The children did not learn Spanish but, on buses and at play, they developed a sign language of their own from about 1981. Their initially eclectic signs formed a “Lenguaje de Señas Nicaraguense” (LSN)—a signed pidgin or jargon for which the constituents had no domination by a superstrate language. Then as early as 1983, some of the younger children at the Moreno School had begun to share a creole, “Idioma de Señas Nicaraguense” (ISN). The children learning to sign at age 8 or more learned LSN, with its large gestures, while those learning to sign at age 7 or less learned ISN with its smaller gestures and more detailed sentences. Each language developed a growing lexicon and syntax. They interacted with each other and with a third language, an asymmetric pidgin bridging the gap between the deaf children and adults who were Spanish-speakers. MIT linguist Judy Kegl, arriving in 1985, worked for years with the program, researching student signing and supporting teaching in LSN and ISN.⁸⁶ With time, speakers of LSN and ISN pooled their resources to build a student association, first to support the existence of the group, then to perform other tasks in the society. By 1999, the signing community of Managua had reached some 500 members; at the present day it has reached roughly 3000 members.⁸⁷

Creolization theory. Scholarship in pidgins and creoles expanded rapidly from the 1970s, especially with empirical studies, until theoretical issues arose and brought two decades of intense debate, from 1980 to 2000. The results led to a substantial clarification and sharpening of theory, a great expansion in data, and wider recognition of the importance of creolization. Derek Bickerton did most to launch the debate with his two-part manifesto: a rambling but insightful 1981 book, *Language and Species*, and his more structured 1984 article, “Language Bioprogram Hypothesis,” which was accompanied by 25 commentaries and Bickerton’s response.⁸⁸ His publications supported generative analysis, inclusion of evolutionary biology, and a focus on plantation societies as the essential source of creole languages. As he saw the plantation model, immigrant laborers with varied first languages, in slave or indentured status, encountered each other and rapidly constructed a pidgin language without structure, drawing elements from many substrate languages but with a lexicon relying primarily on the superstrate—in Hawaii, the dominant English-speaking community. Within a few years, children of the plantation communities, speaking in what was then a native pidgin language, began to develop creole phrases relying on the instincts of universal grammar.⁸⁹

struggle to gain formal recognition of Pidgin, with modest success, while informal institutional structures seem to be successful in sustaining Pidgin in practice. Hawaiian language, with 2000 native speakers and 20,000 second-language speakers, is an official language.

⁸⁶ She had completed a 1985 MIT PhD on a topic in American Sign Language, directed by Noam Chomsky.

⁸⁷ Kegl, Senghas, Coffman 1999, 179.

⁸⁸ Bickerton 1981, and Bickerton 1984. He had settled in Hawaii in 1972 as a skilled creolist and immediately launched a three-year project, interviewing speakers of Hawaiian pidgin and creole aged 70 to 95.

⁸⁹ He accepted the generative assumption that the internal parameters could not be observed directly but could be estimated through their expression in spoken syntax..

Critiques of Bickerton came from all directions, some rejecting generative assumptions, some claiming that the original creole speakers were adults rather than children, and many other points. Claims that creoles developed only through European colonization after 1500 were gradually refuted by discoveries of creoles in wider regions and longer time periods. Bickerton, in response to critiques of the 1980s, expanded his Hawaiian research in the 1990s, including archival work on plantations and court documents by research assistant Sarah Roberts.⁹⁰

Nicaraguan sign language was introduced to scholars through a 1989 conference paper by Judy Kegl and Gayla Iwata. It confirmed the details of Nicaraguan sign languages, provided a broad review of the creolization literature, and advanced new theoretical arguments. The paper identified four overlapping systems of communication: (1) the minimal and unsystematic “home-sign” practices of isolated deaf children and their hearing parents that did not qualify as a language; (2) a jargon signing system developed rapidly by deaf adolescents in school-based groups yet without a superstrate; (3) a creole of more advanced syntax by young children that developed both from the pidgin and from home-signing; and (4) a pidgin intended to bridge the communication between the deaf children and the superstrate Spanish-speaking community, a language mixing signs and audio speech. The three languages interacted.⁹¹ Kegl and Iwata raised challenges to Bickerton on three issues: their argument for an additional asymmetric pidgin in Hawaii, the argument that pidgin was not necessary for the rise of creole, and their suggestion that certain of Bickerton’s proposed parameters did not arise from universal grammar.

Quite a different shift in linguistic thinking came with Noam Chomsky’s Minimalist Program, first circulated in a 1993 paper.⁹² In place of the complexity of earlier generativist arguments, it emphasized locating the simplest principles that could support a syntactical language. These changes in overall theory were too great to have an immediate impact on studies of creolization, but they did suggest that UG was to support the core of language rather than define every step, so that the expected number of parameters in UG declined; further, a bottom-up approach replaced the previous top-down structuralism.

In a comprehensive collaborative volume, MIT linguist Michel DeGraff gathered 20 creolists to contribute viewpoints not only on creolization but also on the related fields of language acquisition and language change.⁹³ In addition, DeGraff’s overview distinguished micro and macro scales of study for linguistic fields: language acquisition was micro-analysis of individuals, while creolization and language change were macro-analysis of group behavior.⁹⁴ DeGraff succeeded in his objective of defining continuities and contrasts in creolization, language acquisition, and language change, but he skipped over some of the emerging debates within creolization.⁹⁵

⁹⁰ Bickerton 2008.

⁹¹ Kegl and Iwata 1989.

⁹² Chomsky 1995; see also Bickerton 2014, 16-71.

⁹³ DeGraff 1997. DeGraff, a native speaker of Haitian Kreyol and French, is an MIT linguistics professor.

⁹⁴ “Micro” and “macro” here correspond to the I-language and E-language of Chomsky’s terminology. DeGraff’s volume remains valuable in providing perspectives and debates of a crucial moment in the field of study.

⁹⁵ DeGraff 1999, 473 – 543.

The DeGraff volume included a chapter by Bickerton on the acquisition of language and a chapter by Kegl, Senghas, and Coffman that extended the detail of the 1989 Kegl-Iwata paper.⁹⁶ The latter chapter restated the critique of Bickerton, arguing that he had likely missed the existence of a pidgin to bridge communication of superstrate-speakers and substrate-speakers, suggesting that court reporters had failed to record the part of the bridge pidgin that emphasized substrate perspectives. The 1999 chapter described institutional change and its centrality to the welfare of the community.⁹⁷ The chapter concluded with a forceful statement of the authors' discovery:⁹⁸

We have been able in Nicaragua to witness firsthand the source point of language genesis. We have discovered that the source of language is within us but that the conditions for its emergence depends crucially on community.

Academic knowledge of Hawaiian Creole English expanded with studies of Sarah Roberts, including a dataset of pidgin and creole speech based on archives of plantations and court records.⁹⁹ In her findings, locally-born children of age 5 were beginning to speak a new pidgin (HPE) by 1885, centered on an English lexicon as superstrate, while the second and later generations of children spoke a steadily developing creole, HCE, from 1905.¹⁰⁰ Bickerton, after retiring from University of Hawaii in 1996, published additional books beginning 2008 but returned to Hawaiian creole only in 2014.¹⁰¹ He praised Roberts for her research but rejected her interpretation of the results. She had organized results into 3 cohorts: a first speaking pidgin, a third speaking creole, and an intermediate group speaking a mix of pidgin and creole. Bickerton, instead, combined the second and third cohort as creole-speakers.¹⁰² Thus, Bickerton had reaffirmed his original pidgin-creole sequence and neglected the arguments that he had missed an asymmetric pidgin and that a creole could form without requiring a preceding pidgin.

Overall, the debates in creolization theory from 1980 to 2000 were productive and cordial.¹⁰³ Generative grammar became more widely accepted even as its logic simplified and its scope narrowed. Analysis extended to human biology and to social evolution; terminology became more consistent. Analysts sorted out the possible paths in language acquisition: abrupt creolization, nativization, and other ways to balance pidgin and creole.

A Pleistocene model of creolization. I turn now to a speculative model of the emergence of human creole. This section models the first hundred years of Proto-Human: its linguistic, biological, and social evolution. After setting the scene with hypothesized initial conditions, I present four

⁹⁶ Senghas 1995. Ann Senghas, an MIT linguistics doctoral student, joined the work in Nicaragua in 1990.

⁹⁷ Later studies, especially by Senghas (2003), traced details of development of additional rules and structures in creole syntax.

⁹⁸ Kegl, Senghas, and Coffman 1999, 223.

⁹⁹ Roberts 1998, 2002, 2004, and Roberts's undated database.

¹⁰⁰ Roberts 1995; 2000; Bickerton 2014, 243–255. Bickerton shows that Roberts confused two chronologies: the single chronology of overall creole development and the separate chronologies of immigrant cohorts arriving at different times. Kegl and Iwata (1989), in their study of Nicaraguan sign language, imply that Bickerton was correct in categorizing creole speakers but underestimated the number of pidgins in Hawaii.

¹⁰¹ Bickerton 2014, 218–256. His other studies were Bickerton 2008, 2009, 2012, and a 2016 reprint of his 1981 book.

¹⁰² He noted that various ethnic groups arrived with quite different timing. Bickerton 2014, 250–255.

¹⁰³ Nonetheless, other approaches to language development persisted: for an approach based on epigenetic development, see Tomasello 2005, 2008.

successive sets of dynamics for the creation of syntactic language: a newly defined children's social group and, within it, the creation of jargon and creole, archives and discourse, and institutions.¹⁰⁴ After successful implementation of these changes, by the end of its first century, the founding group had expanded to 3000 individuals in up to sixteen related language communities.¹⁰⁵

Setting the Scene—Initial Conditions. Here are the conditions assumed to have been in place before comprehensive, syntactic language could emerge some 70,000 years ago. In **time and place**, I chose savanna lands east of Lake Victoria, 70,000 years ago, based on four types of evidence: 1) this region's archaeological finds of innovative stone tool kits, now dated at 72 ka and labeled as Later Stone Age, which are consistent with technological innovation resulting from advances in communication;¹⁰⁶ 2) modern language-distribution data, which show that the projected homelands for Afroasiatic, Nilo-Saharan, and Niger-Kordofanian phyla all lie nearby in the upper Nile Valley;¹⁰⁷ 3) genetic analysis, which suggests migrations of Africans throughout Asia beginning about 60,000 years ago;¹⁰⁸ and 4) paleontological data in Southern and Central Africa from about 65,000 years ago, suggesting immigration to those regions from northeast Africa.¹⁰⁹ In **demography and society**, I hypothesize the experiences of a single *Homo sapiens* community which I call the "Home Community." It included 150 members, who occupied a territory of 150 square kilometers with a density of 1.0 persons per square kilometer.¹¹⁰ They lived in households and included working groups of up to 15 members for creating tools, shelters, and for hunting, relying on Middle Stone Age technology. I analyze population change in five-year groups, according to demographic conventions. In socio-environmental change, the Toba volcanic explosion of 74 kya was doubtless influential on human society.¹¹¹ In **verbal communication**, I assume that individuals in the Home Community had access to one or two levels of capability: "home-signs," eclectic steps for communication within isolated families; and "proto-speech," with words and concepts up to a lexicon of 200 words, in localized groups of up to fifteen members, with a capacity for noun phrases and verb phrases of up to four words, with the possibility that some words were used by many households.¹¹²

Dynamics of creating the Youth Group. To create a persistent group of children, a two-step process is assumed: formation of a "Youth Group" of 15 members, aged eight to fifteen, followed

¹⁰⁴ It is hoped that these hypothesized dynamics can be seen as testable. For suggested guidelines for the testing of such hypotheses on language evolution, see Manning 2023a, 17–18.

¹⁰⁵ Manning 2023a.

¹⁰⁶ Ambrose 1998; Rampino and Ambrose 2000.

¹⁰⁷ Manning 2023a. I argue that the homeland of Khoesan languages was in northern Tanzania.

¹⁰⁸ López, Van Dorp, and Hellenthal 2016.

¹⁰⁹ Ehret 2015a. In a contending thesis, Carina Schlebusch et al. (2020), argue from genetic data that ancestors of the Khoesan-speakers of southern Africa formed an early strain of *Homo sapiens*, with a population that became dense during the MSA and became the source of the worldwide African diaspora. Response has been critical.

¹¹⁰ Manning 2023a, 12. I selected areas and populations, intending them to fit with known demographic patterns. I assume that the Home Community lived within about 150 km² of terrain and included about 30 households of five or six members, headed by pair-bonded male-female couples.

¹¹¹ Rampino and Ambrose 2000.

¹¹² As an alternative hypothesis, one could assume that Home Community members spoke a more advanced proto-speech, spoken in groups larger than 15 but smaller than 150.

by gradual exposure of children aged seven and below to Youth Group activities.¹¹³ The self-selected members of the Youth Group met periodically by walking from their separated households to a common location, where they agreed to work and play together.¹¹⁴ They met every few days. Those 8 and older could walk for an hour to meet; those 7 and younger remained at the home site and could learn of the group and its activities only indirectly. Each year the Youth Group recruited additional children of age 8 and maintained membership of those over age 15.

Dynamics of creating jargon and creole language—the Language We-Group. The members of the Youth Group initially spoke their native versions of proto-speech or home-sign systems, as they varied by household.¹¹⁵ As they played together in the first three years, they developed communication through a common youth-camp jargon.¹¹⁶ In speaking this jargon, members of the Youth Group redefined their identity as the Language We-group, now confirming their purpose as one of collective intentionality.¹¹⁷ Jargon and eventually creole language spread beyond the Language We-group by two processes: members of the group introduced the jargon to their siblings and other children under 8; and members of the group gave birth to children and began teaching them to speak. Within five years or perhaps more, the Language We-group included a significant minority of members who had experienced the jargon before age 8. In this group setting and perhaps beyond it, young children engaged their UG and began speaking creole. From this time, the Language We-group supported the development of both the jargon and the creole, plus the interaction of the two languages.¹¹⁸ Pleistocene jargon survived for a generation or perhaps two, but creole was what grew. While there were innovators and leaders in speech, it seems that there was never a chief of language, so that both the persistence and the modification of both jargon and creole came about through an egalitarian consensus.¹¹⁹ Elder and jargon-speaking members of Language We-group acted as leaders; younger speakers of creole provided inspiration in expression.¹²⁰

Dynamics of archives and discourse—networked consensus. Individual archives of jargon gradually developed, as members of the Language We-group developed common usage. Later, the archive of creole-speakers, storing lexicon and syntax, became much larger than for the jargon.¹²¹ For both languages, individual speakers had to work steadily to build personal versions of the lexicon and syntax in order to participate in the speaking group by exchanging ideas as well as

¹¹³ Manning (2023a, 13–15) includes additional detail on the formation and growth of this Youth Group.

¹¹⁴ An alternative model would rely on an extended nursery assembling children under supervision of mothers or nannies—but the group needed to be large enough, persistent enough, and minimize adult supervision.

¹¹⁵ Kegl and Iwata 1989, 276, 289.

¹¹⁶ I follow the argument of Kegl and Iwata that the initial common tongue was a “jargon” rather than a “pidgin” in that the speakers were equals, without a dominant population or language. Kegl and Iwata 1989, 270–272; see also Kegl 2008.

¹¹⁷ On collective intentionality, see Tuomela (2013) and Manning (2020a). Tomasello (2010) and Dunbar each defined multiple levels of intentionality, but did not identify levels sufficient for syntactic language, as did Tuomela.

¹¹⁸ For comparison, examples of modern creole syntactical forms are proposed in Bickerton 1984, 178–182; Bickerton 1995, 219–229; Bickerton 2012, 456–457; and Kegl and Iwata 1989, 283–288.

¹¹⁹ It is conceivable that there was more than one original language in Africa, but difficult to demonstrate its existence.

¹²⁰ As for the Nicaraguan case, the jargon interacted with the creole and gradually stabilized: it remained part of the language scene, perhaps for generations.

¹²¹ The lexicon of a modern adult person is commonly above 20,000 words.

words and by accepting corrections from others.¹²² At the group level, members sought to ensure the close similarity of individual archives, to make exchange of precise ideas more likely. The exchange created discourse at two or more levels: individuals found contrasts between their own ideas and those learned from others; the group shared some ideas but also found that it shared differences and debates. This hypothesized formation of a new and networked social group is parallel to Ambrose's Step #2 in evolution of cognitive faculties for planning, which he identified as "regional social interaction networks" enhancing planning through information sharing 72 kya.¹²³ My treatment of networks in forming language adds both local and extended networks for creation and sharing of syntactic language.¹²⁴ The networked links among individual brains, through expressing and comprehending creole language, developed a group discourse of ideas, agreements and disagreements, including philosophical questions. Creole, with more precise sentences, led to the formulation of semi-contractual agreements among members, cementing loyalty to the group and agreement to serve its purposes: the character of the Language We-group changed to that of a *language community*.¹²⁵

Dynamics of creating institutions – Community and Ritual. The objectives of the tiny language community could only be met by expanding its size. This general task could be accomplished through sustaining the common association, recruitment of members, care for the youth camp, and creation of rituals and behavioral norms.¹²⁶ In Pleistocene society, no institutional structures were to be found, so that the children creating syntactic language also had to create its institutional support. Individual oaths of loyalty to the group and group rituals of common identity would have been extended to the jargon-speakers out of respect for their foundational work. Rituals and formal admission to the community could then take place. Every year, the community recruited additional children of age eight or above from within the Home Community, plus immigrant children from neighboring communities. At this rate, creole and jargon speakers grew by year 30 to about 75 members, half of the original Home Community.¹²⁷ From this point, the identity of the Home Community was transformed into that of the *Home Language Community*, a social institution and a language community—with a minority, mostly aging, speaking the antecedent proto-speech.

Once creole became widespread, instruction of first languages moved from youth camps to the household, where it remains. On the other hand, the institution of youth camps probably continued for a long time as a device for spreading creole language from one community to another. As the first institutions showed their value, other institutions—also relying on spoken language—formed for additional tasks, in what became a broader process of social evolution. One early institutional act of the Home Language Community was to carry out a partition, since its steady expan-

¹²² The common argument that children learn language "effortlessly" is refuted by these processes. Through the pain of error and disagreement, struggling to repeat enough times to learn, children's learning was a struggle by the laborer and not a gift from some superior. The reward for work was the benefits of communication.

¹²³ Ambrose 2010, S140–141.

¹²⁴ Manning 2025, 105–107.

¹²⁵ With this, the full conditions of Tuomela's we-group (2013) were met.

¹²⁶ Kegl, Senghas, Coffman (1999), 200–201.

¹²⁷ Manning 2023a.

sion had led it well beyond its historic maximum of 150 members. The Home Language Community thus divided its membership into Language Communities B and C—as early as year 35.¹²⁸ The two daughter communities established their institutions and, after continuing growth through incorporation of new creole speakers, they were each later to break into segments, each with its own institutional structure. Other speaking communities had been formed through migration of small numbers of speakers from the Home Language Community. By year 100, my projections suggest the existence of eight language communities formed by partition and perhaps another eight smaller language communities formed by migrant replication of the original process.¹²⁹

In sum, within the first hundred years of this hypothesized language community, a syntactical and lexically enriched creole had arisen, shared among some 3000 speakers even as it evolved. At what point should we identify the founding language community? One could apply the term “Proto-Human” to the language of the Home Language Community at a moment just before its partition in year 35. Or one could argue that it is best to reserve the term “Proto-Human” for the more fully developed language of year 100, treating the 3000 members of 20 language communities as speaking dialects of a single “Proto-Human” protolanguage. Either way, this protolanguage can be seen as the principal ancestor of all subsequent spoken languages.

LANGUAGE DISPERSION AND EXPANDING GROUP SIZE

The partition and consolidation of language communities brought steady replication, geographic spread, and gradual linguistic evolution for each speech community. This section begins with an interpretation of years 101 to 300 in northeast Africa, ending with a population of 9000 speakers, firmly established within a terrain of roughly 3000 km² or a square of 55 km on a side. Within this region, speaking populations dominated; beyond it, speaking communities lived and interacted with proto-speech communities. As speaking people moved into regions of differing environment, the immigrants sought to learn from local residents who knew the region but communicated only through the local versions of proto-speech. Relations between immigrants and local inhabitants may have been cordial or hostile.

The success of the institution of language community was such that institutional frameworks were applied to other tasks that required substantial technical or social skills, and which benefited from processes of ritual and governance. Workshops arose as small institutions, for instance to support visual art: a master artist led design and production but with artistic collaboration of some assistants and with material support from others, resulting in remarkable paintings on cave walls discovered and dated as far back as 50 ka.¹³⁰ Workshop participants formed we-groups confirming their obligations, conducted training, and adopted norms for the work—all relying on precise verbal communication.¹³¹ Other small institutions were marriage (linking the betrothed couple and the families of each in mutual obligations) and family structures (setting boundaries and norms for

¹²⁸ Manning 2025.

¹²⁹ Manning 2025, 107.

¹³⁰ Aubert et al. 2018. Workshops might have existed only for the duration of the master’s leadership.

¹³¹ Whitehouse 2022.

what became lineage or clan structures). In addition, such existing tasks as migration by groups of fifteen, stone tool manufacture or shelter construction could also be reformulated as workshops.

After year 300, one can assume that the creole language of Proto-Human was well established in its homeland and that two types of migration expanded: settling among communities of *Homo sapiens* in Africa and venturing across the Red Sea to Eurasia, where hominin populations were rare. Within Africa, syntactic language appears to have spread relatively rapidly in East, Southern, and Central Africa but to have spread more slowly to North and West Africa. In Southern Africa, archaeological records show an initial advance of peoples presumably from East Africa, then a recovery of the local population and a withdrawal of settlers, followed by a return of settlers and the spread of syntax.¹³² For Central Africa, settlers from East Africa entered the highlands west of the Nile to the Congo basin, where they developed Central African syntactical languages.¹³³ For Northeast and Northwest Africa in the time of the Last Glacial Maximum, Ehret has traced the Afro-Asiatic languages, using Later Stone Age culture, from a homeland in the Blue Nile Valley to the lower Nile, then to the Chad basin, the Sahara, and the Maghreb.¹³⁴ Nilo-Saharan languages spread west and south in humid post-LGM times.¹³⁵ Niger-Kordofanian languages spread from the Nile to the west, although archaeological work by Scerri shows that the MSA culture survived in much of West Africa, implying that syntactical speech did not come to those areas until 10 ka.¹³⁶

Migration to Eurasia was simpler than that across Africa, in that it rarely involved interaction with other hominin, though the ecology differed considerably. When did the migration to Asia begin? Since the distance from Lake Victoria to the Bab el-Mendebe was somewhat over a thousand kilometers, the discovery of Asia across the waters might have taken as little as a hundred years but perhaps several thousand years. Crossing the Bab el-Mendebe was not easy, yet the Yemeni shore would have been visible from Africa in those days of lower sea level and a narrower strait; humans almost surely had dugout canoes or at least rafts for the crossing. The apparent rapidity of migration to Australia suggests that movement along the coast speeded the movement east. On the other hand, the hominin archaeological remains in northern Arabia and the Levant indicate migration across the Sinai.¹³⁷ For those in Asia, populations grew, split, and languages diverged. Most areas of settlement were previously empty of hominin. But there was some encountering of Neanderthals and Denisovans—conceivably with proto-speech, conceivably with skills for learning syntactic speech.

The present-day distribution of language groups around the world ought to reveal much about the historical pattern of migration, fractioning, dispersal, and incorporation of language communities.¹³⁸ But efforts to classify the language phyla of the world and trace their spread from African

¹³² Ehret 2015a.

¹³³ Ehret 2015b. Those languages, however, were lost after later immigration from West Africa

¹³⁴ Ehret, 2023.

¹³⁵ Sutton 1974, 1977; Ehret 2023; Ehret 2015b.

¹³⁶ Scerri et al. 2021, Ehret 2015b.

¹³⁷ Groucutt et al. 2021.

¹³⁸ On the long-term history of syntactical language, I speculate that the rates of historical change varied by historical period for numbers of phonemes, vocabulary words, and syntactical rules. I define the periods as A (before 70 ka), B (70–20 ka), and C (since 20 ka). I estimate that the number of phonemes in use grew slowly in period A, grew rapidly in period B, and declined in period C.

origins have not led to consensus. Joseph H. Greenberg made a mighty effort in classification, with hints on migration, proposing four phyla in Africa, two phyla for the Americas, and two for Eurasia (for northern Eurasia and Oceania)—out of a total of perhaps fifteen phyla.¹³⁹ Outside of Africa, there has been a widespread rejection of Greenberg's classifications.¹⁴⁰ Such a rejection might logically entail further rejection of Greenberg's contributions to classification of African languages, but Africanist and other linguists continue to build on his modeling.¹⁴¹ Since Greenberg used similar methods and materials on his work throughout, all of his analyses thus retain some plausibility, so I have relied on Greenberg's work in my online classifications and maps.¹⁴² To clarify the patterns, I favor comparison of multiple hypotheses and models from various perspectives for formation and distribution of language communities across the world from 60,000 ya.

The story of language expansion goes beyond geographic spread. From about 20,000 years ago, language communities of the original 150 members began to consolidate. I have proposed a model of the process of consolidation in network models of community size, based on Dunbar's vision of a fractal structure of human social groups.¹⁴³ This model argues that the size of minimal social groupings, with lifestyles of hunting and gathering, expanded worldwide by a factor of ten from 150 to 1500 per community in the course of 10,000 years.¹⁴⁴ In much the same time frame, fishing and agricultural communities had increased to populations of 5000 and more—growth that has continued ever since. Explaining this rapid rate of community growth, after a long period of stability, seems to be a matter of importance.

CONCLUSION

The objective of this essay has been to argue that the debates on the origin and development of language can be clarified by focusing on an expansion of Bickerton's theory of creoles combined with analysis of group behavior in multiple species. I began with the basic and forceful dynamics of the Earth's environment in material impact and the information conveyed, especially the impact on animals. Then I turned to the capacity of early animals to sense information on environmental dynamics, in effort to prepare responses. These included reflexive responses to environmental impact and, with time, the emission of signals and the formulation of collaborative responses and animal group behavior. The diversity in animal species, including their numerous senses and group functions, shows that animals tinkered with signals—chemical, movement, sound, vision, touch—that occasionally created new environmental niches.¹⁴⁵ Ants found that selective use of odor could mark paths to food sources; bees found that specific motions could not only identify sources of

Total vocabulary rose very slowly in period A, rapidly in period B, and at a moderate rate in period C. Elements of syntax arose very slowly in period A, developed rapidly in period B, and continued change in period C as languages differentiated while also sustaining their universals. These initial speculations could easily be sharpened.

¹³⁹ Greenberg 1987, 2000, 2005. Ruhlen (1991) pursued and expanded the classificatory work of Greenberg.

¹⁴⁰ Campbell 1988, 2001.

¹⁴¹ Heine and Nurse, 2000; Ruhlen 2023.

¹⁴² Manning, 2022.

¹⁴³ Manning 2025; Dunbar 2020b. See also Ambrose (2001, 1752).

¹⁴⁴ Flannery and Marcus 2012.

¹⁴⁵ François 1977. This foundational article in epigenetics develops the concept of tinkering.

nectar but point to their direction and elevation. Beavers and fruit flies each relied on multiple senses to sustain their groups. In sum, it seems that animal groups did much to remake the Earth's environment.

Hominin species too are understood to have tinkered with signals in touch, motion, odor, vision, and sound in articulating their group behavior. Through touch, hominins fashioned stone tools, scavenging bone and then meat; with a vocal call, they assembled crowds to dominate the early seizure of meat from carcasses. With time, they codified audio signals into words and concepts, eventually grouping words into proto-speech. While communication took place along several lines of information, it did not reach a high level of precision.

A breakthrough at about 70 ka combined four related innovations in vocal communication to create a new group structure. Forming the *we-group* deepened the ties within existing groups of about fifteen youths by enforcing mutual loyalty, adopting a common task of communication, and requiring deep study by members. Members of this new group then assembled their individual versions of proto-speech into an expanded jargon, which included some basic links of noun phrases and verb phrases; then universal grammar kicked in and younger group members developed syntax in patterns that have been observed in creoles of the 19th and 20th centuries—inventing syntactical relations among words, developing norms for complete sentences, and encountering the inherent processes of linguistic change. Third, memorization of new words and syntactical norms required biological archiving; spoken conversations accessed the archives of each group member in formulating statements, hearing statements of others, and formulating ideas for expression. That is, the achievement was not just creation of syntactic language, it was also the social and intellectual processes that emerged. The encompassing institutional structure, an extension of *we-group* loyalty, coordinated the exchange of messages among group members and established rituals to sustain and propagate the language community.

The miracle of the transition was its apparent construction of the full model of the language community in a small group within as little as a single life-span. The language community, structured to grow demographically and to build vocabulary and syntax, expanded and differentiated without limit. The institution of language became the prototype for institutions beyond language itself, addressing such community tasks as governance, caring for health, and conceptualizing religion. On a smaller scale, institutions and workshops coordinated migration and cultural activities in visual art, music and dance. The exchange of verbal signals, modulated with phonology, morphology, and syntax that enabled infinite construction of sentences, opened the door to accelerated creation of ideas and social projects.

This was social evolution in human dynamics. Erected on the biological framework of hominin communities, the four overlapping components of language yielded a system of exchange among brains whose creativity largely escaped the dominion of natural selection. Purposeful and self-conscious groups, years of individual study of language, storage and access in an expanding cerebral lexicon, progressive creation and sharing of syntax, and institutional coordination of sharing language through ritual—these became the core of social evolution. The new environmental

niche for human society was to be the creation and exchange of knowledge and its application to successive tasks of material and conceptual change.

With time, visual signals were to grow in importance for human communication. Workshops in visual art, creating masterpieces of cave art, relied on vocal signals to plan the creation and interpretation of images and shapes. Visual creativity and communication became an important discipline, although art criticism was expressed through language. Later on, visual signals were used to create symbols for sounds and meanings, so that systems of writing became parallels to vocal language. Communication among human brains could now be achieved through the medium of the manuscript page, which had greater permanence than memorized archives. Human syntactic language—conveyed today through the audio medium of speech, the visual medium of text, and the electronic medium of files—is unique in the specifics of its communication but can be seen as parallel to various results of group behavior.

REFERENCES

- Ambrose, Stanley H. 1998. "Chronology of the Later Stone Age and food production in East Africa." *Journal of Archaeological Science* 25: 377–392.
- Ambrose, Stanley H. 2001. "Paleolithic Technology and Human Evolution." *Science* 291: 1748–1753.
- Ambrose, Stanley H. 2010. "Coevolution of Composite-Tool Technology, Constructive Memory, and Language: Implications for the Evolution of Modern Human Behavior." *Current Anthropology* 51, Supplement 1: S135–S147.
- Anon. 2024. "Beaver Behavior and Biology." Beaver Solutions, Southampton, MA. <https://www.beaversolutions.com/beaver-facts-education/beaver-behavior-and-biology/>
- Ashe, Alyson, Vincent Colot, and Benjamin P. Oldroyd. 2021. "How does epigenetics influence the course of evolution?" *Philosophical Transactions of the Royal Society B: Biological Sciences* 376 (1826): 20200111.
- Aubert, M., Setiawan, P., Oktaviana, A., et al. 2018. "Paleolithic cave art in Borneo." *Nature*, 564 (7735): 254–257.
- Bandura, Albert. 1971. *Social learning theory*. New York: General Learning Press.
- Barham, Lawrence, and Peter Mitchell. 2008. *The First Africans: African archaeology from the earliest tool makers to most recent foragers*. Cambridge: Cambridge University Press.
- Bengtson, John D. 2021. "Comments on 'Na-Dene and Beyond'; Sino-Dene (Updated), and the Position of Haida." *Mother Tongue* 22: 11–42.
- Berwick, Robert, and Noam Chomsky. 2016. *Why Only Us? Language and Evolution*. Cambridge: MIT Press.
- Bickerton, Derek. 1981. *Language and Species*. Ann Arbor, MI: Karoma Press; reprinted *Roots of Language* with a new preface in 2016 by Language Science Press, Berlin. First published as *Language and Species* (University of Chicago Press, 1981).
- Bickerton, Derek. 1984. "Language Bioprogram Hypothesis." *Behavioral and Brain Sciences* 7: 173–188.
- Bickerton, Derek. 1995. *Language and Human Behavior*. Seattle: University of Washington Press.
- Bickerton, Derek. 2008. *Bastard Tongues*. New York: Hill and Wang.
- Bickerton, Derek. 2009. *Adam's tongue: How humans made language, how language made humans*. New York: Hill and Wang.

- Bickerton, Derek. 2012. "The Origins of Syntactic Language." *The Oxford Handbook of Language Evolution*, eds. Maggie Tallerman and Kathleen R. Gibson, 456–468. Oxford: Oxford University Press.
- Bickerton, Derek. 2014. *More than Nature Needs: Language, Mind, and Evolution*. Cambridge, MA: Harvard University Press.
- Bickerton, Derek, and Eörs Szathmáry. 2011. "Confrontational scavenging as a possible source for language and cooperation." *BMC Evolutionary Biology* 11: article 11.
- Boyd, Robert, and Peter J. Richerson. 1985. *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.
- Brunet Thibault, and Nicole King. 2017. "The Origin of Animal Multicellularity and Cell Differentiation." *Developmental Cell* 43, 2: 124–140.
- Buck, Laura T., and Chris Stringer. 2014. "Homo heidelbergensis." 2014. *Current Biology* 24, 5: PR214–PR215.
- Bunnefield, Lynsey, Laurent A. F. Frantz, and Konrad Lohse. 2015. "Inferring Bottlenecks from Genome-Wide Samples of Short Sequence Blocks." *Genetics* 201, 3: 1157–1169.
- Campbell, Lyle. 1988. "Review of Language in the Americas by Joseph H. Greenberg." *Language* 64, 3: 591–615.
- Campbell, Lyle. 2001. *Historical Linguistics: An Introduction*. Cambridge, MA: MIT Press.
- Cann, Rebecca L., Mark Stoneking, and Wilson, Allan C. 1987. "Mitochondrial DNA and human evolution." *Nature*, 325 (6099): 31–36.
- Chapais, Bernard. 2008. *Primeval kinship: How pair-bonding gave birth to human society*. Cambridge, MA: Harvard University Press.
- Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- Darwin, Charles. 1859. *On the origin of species by means of natural selection*. London: John Murray.
- d'Errico, Francesco, Christopher Henshilwood, Marian Vanhaeren, and Karen van Niekerk. 2005. "Nassaurus kraussianus shell beads from Blombos Cave: evidence for symbolic behaviour in the Middle Stone Age." *Journal of Human Evolution* 48, 1: 3–24.
- Dimijian, Gregory G. 2012. "Darwinian Natural Selection: Its Enduring Explanatory Power." *Baylor University Medical Center Proceedings* 25, 2: 139–147.
- Droser, Mary L., and Seth Finnegan. 2003. "The Ordovician Radiation: A Follow-up to the Cambrian Explosion?" *Integrative and Comparative Biology* 43, 1: 178–184.
- Dunbar, Robin I. M. 1988. *Primate social systems*. London: Chapman and Hall.
- Dunbar, Robin I. M. 1993. "Coevolution of neocortical size, group size and language in humans." *Behavioral and Brain Sciences* 16, 4: 681–694.
- Dunbar, Robin. 2020. "Structure and function in human and primate social networks: implications for diffusion, network stability and health." *Proceedings of the Royal Society A* 476: 20200446.
- Dunbar, Robin I. M. 2024. "Structural and Cognitive Mechanisms of Group Cohesion in Primates." *Behavioral and Brain Sciences* 48: e162.
- Eldredge, Niles, and Stephen Jay Gould. 1972. "Punctuated equilibria: an alternative to phyletic gradualism." In Schopf, Thomas J. M., ed., *Models in Paleobiology*, 82–115 (San Francisco: Freeman, Cooper).

- Ehret, Christopher. 1998. "Nostratic — or Proto-Human?" In C. Renfrew and D. Nettle, eds., *Nostratic. Examining a Linguistic Macro-family*, 93–122. Cambridge: The McDonald Institute for Archaeological Research.
- Ehret, Christopher. 2015a. "Early Humans. Tools, Language, and Culture." In D. Christian, ed., *The Cambridge World History. Vol. 1: Introducing World History, to 10,000 BCE*, 339–361. Cambridge: Cambridge University Press.
- Ehret, Christopher. 2015b. "Africa from 48,000 to 9500 BCE." In D. Christian, ed., *The Cambridge world History. Vol. 1: Introducing world history, to 10,000 BCE*, 362–393. Cambridge: Cambridge University Press.
- Ehret, C. (2023). *Ancient Africa: A global history, to 300 CE*. Princeton: Princeton University Press.
- Fitch, W. Tecumseh. 2019. "Animal cognition and the evolution of human language: why we cannot focus solely on communication." *Philosophical Transactions of the Royal Society B: Biological Sciences*. 375, 1789: 20190046.
- Flannery, K., and J. Marcus. 2012. *The creation of inequality: How our prehistoric ancestors set the stage for monarchy, slavery, and empire*. Cambridge, MA: Harvard University Press.
- Freeman, Linton C. 2004. *The Development of Social Network Analysis*. Vancouver: Empirical Press.
- Frisch, Karl von. 1953. *The Dancing Bees*. New York: Harvest Books. (first published in German, 1927).
- Gavrillets, Sergey. 2012. "Human origins and the transition from promiscuity to pair-bonding." *PNAS* 109, 25: 9923–9928.
- Gowlett, John, Clive Gamble, and Robin Dunbar. 2012. "Human Evolution and the Archaeology of the Social Brain." *Current Anthropology* 53: 693–722.
- Greenberg, Joseph H. 1987. *Language in the Americas*. Stanford: Stanford University Press.
- Greenberg, Joseph H. 2000. *Indo-European and its Nearest Neighbors*, 2 vols. Stanford: Stanford University Press.
- Greenberg, Joseph H. 2005. "The Indo-Pacific Hypothesis (1971)." In Greenberg (ed. William Croft), *Genetic Linguistics: Essays on Theory and Method*, 193–276. Oxford: Oxford University Press.
- Greenfield, Patricia. 1991. "Language, tools and brain: The ontogeny and phylogeny of hierarchically organized sequential behavior." *Behavioral and Brain Sciences* 14: 531–551.
- Groucutt, Huw S., Tom S. White, Eleanor S. Scerri, et al. 2021. "Multiple hominin dispersals into Southwest Asia over the past 400,000 years." *Nature* 597, 7876: 376–380.
- Hamilton, William D. 1964. "The Genetical Evolution of Social Behavior." *Journal of Theoretical Biology* 7, 1: 1–16, 17–52.
- Hauser, Marc. 1996. *The Evolution of Communication*. Cambridge, MA: MIT Press.
- Heine, Bernd, and Derek Nurse. 2000. *African Languages: An Introduction*. Cambridge: Cambridge University Press.
- Hendry, Lisa. 2000. "Australopithecus afarensis, Lucy's species." London: Natural History Museum. <https://www.nhm.ac.uk/discover/australopithecus-afarensis-lucy-species.html>
- Hirst, K. Kris. 2019. "Marine Isotope Stages: Building a Paleoclimatic History of the World." *Thought & Co*. <https://www.thoughtco.com/marine-isotope-stages-climate-world-171568>.
- Hölldobler, Bert, and Edward O. Wilson. 1990. *The Ants*. Cambridge, MA: Harvard University Press.

- Hsieh, Shannon, and Roy E. Plotnick. “The representation of animal behaviour in the fossil record.” 2020. *Animal Behaviour* 169: 65–80.
- Hu, Wangjie, Ziqian Hao, Pengyuan Du, Fabio di Vincenzo, and Haipeng Li. 2023. “Genomic inference of a severe human bottleneck during the Early to Middle Pleistocene transition.” *Science* 380, 6661: 979–984.
- Hublin, J.-J., Ben-Ncer, A., Bailey, S. E., Freidline, S. E., Neubauer, S., Skinner, M. M., Bergmann, I., Le Cabec, A., Benazzi, S., Harvati, K., & Gunz, P. 2017. “New fossils from Jebel Irhoud, Morocco and the pan-African origin of *Homo sapiens*.” *Nature*, 546(7657): 289–292.
- Jacob, François. 1977. “Evolution and Tinkering.” *Science* 196, 4295: 1161–1166.
- Kaas, John H. 2007. “The evolution of the complex sensory and motor systems of the human brain.” *Brain Research Bulletin* 75: 384–390.
- Kegl, Judy, and Gayla Iwata. 1989. “Lenguaje de Signos Nicaragüense: A Pidgin Sheds Light on the “Creole?” ASL.” In Carlson, R., S. DeLancey, S. Gildea, D. Payne, and A. Saxena, eds., *Proceedings of the Fourth Meetings of the Pacific Linguistics Conference*, 266–294. Eugene, Oregon: Department of Linguistics, University of Oregon.
- Kegl, Judy, A. Senghas, and M. Coppola. 1999. “Creation through Contact: Sign Language Emergence and Sign Language Change in Nicaragua.” In Michel DeGraff, ed., *Language Contact and Language Change: The Intersection of Language Acquisition, Creole Genesis, and Diachronic Syntax*, 179–237. Cambridge, MA: MIT Press.
- Kegl, Judy. 2008. “The Case of Signed Languages in the Context of Pidgin and Creole Studies.” In Silvia Kouwenberg and John Victor Singer, eds., *The Handbook of Pidgin and Creole Studies*, 491–511. London: Blackwell.
- Krause, J., D. Hoare, S. Krause, C. K. Hemelrijk and D. I. Rubenstein. 2000. “Leadership in fish shoals.” *Fish and Fisheries* 1: 82–89.
- Laland, Kevin N. 2017. “The Origin of Language in Teaching.” *Psychonomic Bulletin & Review* 24: 225–231.
- López, Saioa, Lucy Van Dorp, and Garrett Hellenthal. 2016. “Human Dispersal Out of Africa: A Lasting Debate.” *Evolutionary Bioinformatics Online* 11: 57–68.
- Lubbock, John, Sir. 1882. *Ants, bees, and wasps: a record of observations on the habits of the social hymenoptera*. London: Kegan Paul, Trench & Co.
- MacDonald, Brandi L., Elizabeth Velliky, Bob Forrester, Gregor D. Bader, et al. 2024. “Ochre communities of practice in Stone Age Eswatini.” *Nature Communications* 15: 9201.
- Manning, Patrick. 2020a. *A History of Humanity: The Evolution of the Human System* Cambridge: Cambridge University Press..
- Manning, Patrick. 2020b. *Methods for Human History: Studying Social, Cultural, and Biological Evolution*. Cham: Palgrave Macmillan.
- Manning, Patrick. 2022. “Language Distribution.” https://patrickmanningworldhistorian.com/wp-content/uploads/2022/06/Language-Distribution_06-30-22.pdf . See also Manning, “Language Distribution Worldwide Dataverse,” <https://dataverse.harvard.edu/dataverse/language-distributionworldwide> .
- Manning, Patrick. 2023a. “The Origin of Social Evolution: Language and Institutional Evolution.” *Anthropos* 118: 7–21.
- Manning, Patrick. 2023b. “Households and Communities: Evolution in *Homo sapiens*.” *The History of the Family* 28, 4: 631–659.
- Manning, Patrick. 2025. “Networks in the World History of Human Evolution.” *Asian Review of World Histories* 13: 82–116.

- Massari, Paul. 2024. "A New Take on Natural Selection: How behavior shapes evolution." <https://gsas.harvard.edu/news/natural-selection>.
- Moore, Lisa D., Thuc Le, and Guoping Fan. 2013. "DNA Methylation and Its Basic Function." *Neuropsychopharmacology* 38: 23–38.
- Niwa, Hiro-Sato. 1994. "Self-organizing dynamic model of fish schooling." *Journal of Theoretical Biology* 171, 2: 123–136.
- Nordmann, Gregory C., Tobias Hochstoeger, and David A Keays. 2017. "Magnetoreception—A sense without a receptor." *PLoS Biology* 15, 10: e2003234.
- Odling-Smee, F. John, Kevin L. Laland, and Marcus W. Feldman. 2020. *Niche construction: The neglected process in evolution*. Princeton: Princeton University Press.
- Pinker, Steven, and Paul Bloom. 1990. "Natural language and natural selection." *Behavioral and Brain Sciences* 13, 4: 707–784.
- Pontzer, Herman. 2012. "Overview of Hominin Evolution." *Nature Education Knowledge* 3, 10: 8.
- Quam, Rolf, Ignacio Martínez, Manuel Rose, Juan Luis Arsuage, et al. 2015. "Early hominin auditory capacities." *Science Advances* 1, 8.
- Rampino, Michael R., and Stanley H. Ambrose. 2000. "Volcanic winter in the Garden of Eden: The Toba supereruption and the late Pleistocene human population crash." In Floyd W. McCoy and Grant Heiken, eds., *Volcanic Hazards and Disasters in Human Antiquity*, 71–82 (Boulder, Colorado: Geological Society of America Special).
- Roberts, Sarah J. 1995. "Pidgin Hawaiian: A sociohistorical study." *Journal of Pidgin and Creole Languages* 10: 1–56.
- Roberts, Sarah J. 1998. "The role of diffusion in the genesis of Hawaiian Creole." *Journal of Pidgin and Creole Languages* 10: 1–39.
- Roberts, Sarah J. 2000. "Nativization and the genesis of Hawaiian Creole." In *Language change and language contact in pidgins and creoles*, J. McWhorter, ed., 257–300 (Amsterdam: Benjamins).
- Roberts, Sarah J. 2004. "The emergence of Hawaii Creole English in the early twentieth century: The sociohistorical context of creole genesis." Ph.D. dissertation, Stanford University.
- Roberts, Sarah J. No date. "Language in Hawaii from 1789 to 1960. The documentary sources." Unpublished manuscript, University of Hawaii.
- Rooke, Rebecca, Amara Rasool, Jonathan Schneider, and Joel D. Levine. 2020. "Drosophila melanogaster behaviour changes in different social environments based on group size and density." *Communications Biology* 3, article 304.
- Rooke, Rebecca, Joshua J. Krupp, Amara Rasool, Mireille Golemic, Megan Stewart, Jonathan Schneider, and Joel D. Levine. 2024. "The gene "degrees of kevin bacon" (*dokb*) regulates a social network behaviour in *Drosophila melanogaster*." *Nature Communications* 15, article 3339.
- Ruhlen, Merritt. 1991. *A Guide to the World's Languages*, 3 vols. Stanford: Stanford University Press.
- Ruhlen, Merritt. 2023. *The Origin of Language: Tracing the Evolution of the Mother Tongue*. Piscataway, NJ: Gorgias Press.
- Scerri, Eleanor M. L., Thomas, Mark G., Andrea Manica, et al. 2018. "Did our species evolve in subdivided populations across Africa, and why does it matter?" *Trends in Ecology & Evolution* 33, 8: 582–594.

- Scerri, Eleanor M. L., Khady Niang, and Huw S. Groucutt, et al. 2021. "Continuity of the Middle Stone Age into the Holocene." *Scientific Reports* 11, article 70.
- Schapiro, Bob, and Stanley H. Ambrose. 2015. "On the Origin of Propaganda: Bio-Cultural and Evolutionary Perspectives on Social Cohesion." In *Neuroscience and Media. New understandings and representations*, Michel Grabowski, ed., 108–132. New York: Routledge.
- Schlebusch, Carina M., Per Sjödin, Gwenna Breton, et al. 2020. "Khoen-San Genomes Reveal Unique Variation and Confirm the Deepest Population Divergence in *Homo sapiens*." *Molecular Biology and Evolution*, 37, 10: 2944–2954.
- Schlosser, Gerhard. 2018. "A Short History of Nearly Every Sense—The Evolutionary History of Vertebrate Sensory Cell Types." *Integrative and Comparative Biology*, 58: 301–316.
- Schneider, Jonathan, Michael H. Dickinson, and Joel E. Levine. 2012. "Social structures depend on innate determinants and chemosensory processing in *Drosophila*." *PNAS* 109 (supplement), 17174–17179.
- Schwartz, Jeffrey H., and Ian Tattersall. 2010. "Fossil evidence for the origin of *Homo sapiens*." *Yearbook of Physical Anthropology*, 53, S51: 94–121.
- Senghas, Ann. 1995. "Children's Contribution to the Birth of Nicaragua's Sign Language." Ph.D. dissertation, MIT.
- Senghas, Ann. 2003. "Intergenerational influence and ontogenetic development in the emergence of spatial grammar in Nicaraguan Sign Language." *Cognitive Development* 18: 511–531.
- Seyfarth, Robert M., Dorothy L. Cheney, and Peter Marler. 1980. "Vervet monkey alarm calls: Semantic communication in a free-ranging primate." *Animal Behaviour* 28: 1070–1094.
- Stringer, Chris. 2016. "The origin and evolution of *Homo sapiens*." *Philosophical Transactions of the Royal Society B: Biological Sciences* 371, 1698: 20150237.
- Sutton, J. E. G. 1974. "The Aquatic Civilization of Middle Africa." *Journal of African History* 15: 527–546.
- Tattersall, Ian. 2009. "Language and the Origin of Symbolic Thought." In: Sophie A. De Beaune, Frederick L. Coolidge, and Thomas G. Wynn, eds. *Cognitive Archaeology and Human Evolution*, 109–116. New York: Cambridge University Press.
- Tattersall, Ian. 2012. *Masters of the Planet. The Search for Our Human Origins*. New York: Palgrave Macmillan.
- Thomas, James, and Simon Kirby. 2018. "Self domestication and the evolution of language." *Biological Philosophy* 33, 9: 1–30.
- Tomasello, Michael. 2005. *Constructing a Language: A Usage-Based Theory of Language Acquisition*, Cambridge, MA: Harvard University Press.
- Tomasello, Michael. 2008. *Origins of Human Communication*. Cambridge, MA: MIT Press.
- Tomasello, Michael, and Henrike Moll. 2010. "The Gap is Social: Human Shared Intentionality and Culture." In Peter M. Kappeler and Joan B. Silk, eds, *Mind the Gap*, 331–349. Heidelberg: Springer.
- Traxler, Matthew J., Megan Boudewyn, and Jessica Loudermilk. 2012. "What's special about human language? The contents of the 'narrow language faculty' revisited." *Lang Linguist Compass* 6, 10: 611–621.
- Tuomela, Raimo. 2013. *Social Ontology. Collective Intentionality and Group Agents*. Oxford: Oxford University Press.
- Van Driem, George. 2024. "The Father Tongues L, R, and P." *Mother Tongue* 24, 1: 45–56.
- Vannier, Jean, Muriel Vidal, Robin Marchant, et al. 2019. "Collective behaviour in 480-million-year-old trilobite arthropods from Morocco." *Scientific Reports* 9, article 14941.

- Whitehouse, Harvey. 2022. *The Ritual Animal: Imitation and Cohesion in the Evolution of Social Complexity*. Oxford: Oxford University Press.
- Williams, George C. 1966. *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought*. Princeton: Princeton University Press.
- Wilson, Margo, and Sarah L. Mesnick. 1997. "An Empirical Test of the Bodyguard Hypothesis." In Patricia Adair Gowaty, ed., *Feminism and Evolutionary Biology*, 505–511. Boston: Springer.
- Zaccarella, Emiliano, and Angela D. Friederici. 2015. "Merge in the Human Brain. A Sub-Region Based on Functional Investigation in the Left Pars Opercularis." *Frontiers in Psychology* 6: 1818.
- Zhang, Junqiao, Qiang Qu, and Xuebo Chen. 2025. "Understanding collective behavior in biological systems through potential field mechanisms." *Scientific Reports* 15:3709.
- Zhang Xingliang, and Degan Shu. 2021. "Current understanding on the Cambrian Explosion: questions and answers." *Paläontologische Zeitschrift* 95, 4: 641–660.

THE PRE-INDO-EUROPEAN SUBSTRATUM IN THE BRITISH ISLES AND THE PRE-URALIC SUBSTRATUM IN FENNOSCANDIA: TRACES OF COMMON ORIGIN

VÁCLAV BLAŽEK

ABSTRACT

Recent genetic studies of Mesolithic hunter-gatherers in Scandinavia reveal that the primary postglacial repopulation originated from Western Europe. This paper tests the hypothesis of a shared migration by identifying linguistic parallels between the Pre-Indo-European substratum in the British Isles and the Pre-Uralic substratum in Fennoscandia, providing evidence for a common origin of these ancient populations.

Shared lexical features between Insular Celtic and Saami are strikingly unexpected. Once Celtic influence on Saami and Saami influence on Insular Celtic—particularly Goidelic—are excluded, the most plausible explanation is a common substratum. This hypothesis finds robust extra-linguistic corroboration in genetic and archaeological evidence for a Mesolithic migration from Western Europe to Fennoscandia immediately following glacial retreat, dated to ca. 11,000 BP. This was swiftly succeeded by a second Mesolithic influx from Eastern Europe ca. 10,200 BP (Günther et al. 2018: 1–6; cf. Mallory 2013: 53). Ireland’s earliest inhabitants arrived ca. 10,000 BP, introducing Mesolithic industries, while Britain’s initial Mesolithic traces date to ca. 11,600 BP, with Late Palaeolithic hunter-gatherers attested as early as ca. 14,700 BP (Mallory 2013: 40, 48).

The following list of 12 lexical comparisons marks the first systematic effort to compile relevant data, though the items vary in semantic reliability. The strongest cases are ## 1, 6, and 12. Comparison #13 involves only Fennic forms and requires separate discussion.

1. Saami: South *aske* ‘moon’ (Lagercrantz 1939: 18, n. 105; Sammallahti 1998: 179) //

Goidelic: Old Irish *ésca(e)*, *éisce* n., later m. ‘moon.’ Matasović (2009: 118–19) connected it with Lithuanian *áiškus* ‘bright’ and Old Church Slavonic *iskra* ‘spark.’

2. PSaami **jōvę* ‘scree, field of jagged rocks’ (Lehtiranta 1989: n. 297) //

Goidelic: Old Irish *aife* ‘slope’ (*DIL* A-108; *LEIA* A-24); cf. Old Irish *aig* ‘ice’ < **jeg-* (Matasović 2009: 435).

3. PSaami **kājpē* ‘chin’ (Lehtiranta 1989: n. 341) //

Goidelic: Old Irish *cab* ‘mouth, muzzle’ (*DIL* C-1; *LEIA* C-1); less probable is the alternative comparison with Old Irish *clab* ‘hair (of the head), lock of hair’ (*DIL* C-174; *LEIA* C-93). In this case the semantic difference is comparable with Germanic **barda-* ‘beard’ vs. Slavic **borda* ‘chin’.

4. PSaami **kēδē* ‘temple’ (Lehtiranta 1989: n. 394) //

Goidelic: Old Irish *cod*, *cud* ‘head’ (*DIL* C-582; *LEIA* C-139).

5. PSaami **kēδkē* ‘stone’ (Lehtiranta 1989, n. 380) //

Insular Celtic: Old Irish *crec*, *creg* ‘rock’ (*DIL* C-513; *LEIA* C-225); Old Welsh (12th cent.) *creic*, Welsh *craig* ‘rock, boulder, stone’ (*GPC*).

6. PSaami **kumppā* ‘wolf’ (Lehtiranta 1989: n. 495) //

Goidelic: Old Irish *cuib*, var. *cuim* ‘hound, wolf’ (*DIL* C-583; *LEIA* C-267)

7. PSaami **lānte* ‘pond, pool’ (Lehtiranta 1989: n. 568); Finnish *lantto* ‘puddle’ was probably borrowed from Saami (Aikio 2009: 260). The origin of the Saami word in Finnish *lansi* ‘low-land, wetland’ or dial. *lanto* ‘low, damp place in terrain’ is not unambiguous—the opposite vector or a common heritage cannot be excluded. //

Brittonic **lindā* ‘lake, pool’ (Matasović 2009: 239–40).

8. Saami: North *liraš* ‘a bird belonging to the genus Totanidae’, Enontekiön *lirāš* ‘common sandpiper’ / Finnish *liro*, *lieru* ‘wood sandpiper / Totanus glareolus’ (*SKES* 298) //

Goidelic: Old Irish *rer* ‘blackbird’ (*DIL* R-48; *LEIA* R-21)

9. PSaami **sēplē* ‘slush’ (Lehtiranta 1989: n. 1142) //

Goidelic: Old Irish *slab* ‘mud, mire’ (*DIL* S-255; *LEIA* S-123).

10. PSaami **tāktē* ‘bone’ (Lehtiranta 1989: n. 1219); the comparison with Finnish *tähti*, gen. *tähden* ‘das Übriggebliebene, Aufbewahrte’ is problematic for semantic reasons and the comparison with Hungarian *tetem* ‘Leiche, Leichnam, die irdischen überreste, Gebein,’ Old Hungarian ‘Knochen, Bein; Rippe’ is, by reason of its isolation, not convincing enough to be generally accepted (cf. *UEW* 515–16) //

Goidelic: Old Irish *tec* & *tuc* ‘bone’ (*DIL* S-94, 352; *LEIA* T-38, 106).

11. PSaami **vēvlē* ‘hole in skin’ (Lehtiranta 1989: n. 1381; Lagercrantz 1939: n. 1001); cf. Koltta *veu’l* ‘loop, sling’ (Itkonen 2011: 738) //

Brittonic **ueḷā-* > Middle Welsh (14th cent.) *gwefl* f. ‘lip (of an animal)’, pl. ‘jaws’, Cornish *gwelv* ‘lip,’ Middle Breton *guefl* ‘mouth’ (Matasović 2009: 419 adds, in addition, Old Irish *bél* m. ‘lip’).

12. PSaami **vōjve* ‘liver’ (Sammallahti 1998: 179); the initial **v-* can be prothetic, cf. PSaami **vōlē* ‘down’ (Lehtiranta 1989: n. 1432) vs. Fennic **ala* ‘lower,’ etc. (UEW 6). It means, **vōjve* can be derived from Pre-Saami **ajvi* //

Insular Celtic: Old Irish *óa* f. gl. ‘iecur’, dat.pl. *óeib*; Middle Welsh *ahu*, *auu*, *afu* m. & f.; Old Cornish *au*; Middle Breton *auu*, Breton *avu* m. ‘liver’ (Matasović 2009: 49: **auV-*?).

13. Fennic **suti* ‘wolf’ > Finnish *susi*, gen. sg. *suden*, ill.sg. *suteen*, nom.pl. *sudet*, Estonian *susi*, Livonian *su’ž*, pl. *sudùD* (SKES 1129–1130) //

Goidelic: Old Irish *sod*, *sad* f. ‘bitch, she-wolf’ (*DIL* S-323; *LEIA* S-161).

CONCLUSION

This preliminary list demonstrates that the British Isles and Fennoscandia preserve traces of a distinctive Mesolithic substratum. The semantic range of the compared lexemes—spanning hunting, gathering, and basic environmental features—aligns closely with a pre-agricultural cultural horizon. The primary aim of this study is to stimulate further interdisciplinary research into these ancient linguistic layers.

REFERENCES

- Aikio, Ante. 2009. *The Saami Loanwords in Finnish and Karelian*. Ph.D. thesis: Faculty of Humanities of the University of Oulu.
- Aikio, Ante. 2012. “An essay on Saami ethnolinguistic prehistory.” In: *A Linguistic Map of Prehistoric Northern Europe*. Helsinki: Suomalais-Ugrilaisen Seuran Toimituksia / Mémoires de la Société Finno-Ougrienne 266, 63–117.
- DIL* = *Dictionary of the Irish Language. Based mainly on Old and Middle Irish materials*. ed. by E.G. Quin et alii. Dublin: Royal Irish Academy (Compact Edition) 1998.
- GPC* = *Geiriadur Prifysgol Cymru / A Dictionary of the Welsh Language* 1–4. R. J. Thomas et alii, eds. Cardiff: University of Wales Press 1967–2002.
- <https://welsh-dictionary.ac.uk/gpc/gpc.html>.
- Günther, Torsten, Malmström, Helena, Svensson, Emma M., Omrak, Ayça, Sánchez-Quinto, Frederico, Kılınç, Gülşah M., et alii. 2018. “Population genomics of Mesolithic Scandinavia: Investigating early postglacial migration routes and high-latitude adaptation.” *PLoS Biology* 16(1), 1–22.
- Itkonen, Terje I. 2011. *Wörterbuch des Kolta- und Kolalappischen*. Helsinki: Lexica Societatis Fenno-Ugricae XV.

- Lagercrantz, Eliel. 1939. *Lappischer Wortschatz*, I-II. Helsinki: Lexica Societatis Fenno-Ugricae VI.
- Lehtiranta, Juhani. 1989. *Yhteissaamelainen sanasto*. Helsinki: *Mémoires de la Société Finno-Ougrienne / Suomalais-ugrilaisen seuran toimituksia* 200.
- LEIA = Vendryes, Joseph. 1987. *Lexique étymologique de l'irlandais ancien*, Lettre C, par les soins de E. Bachellery et P.-Y. Lambert. Dublin: Institute for Advanced Studies – Paris: Centre National de la Recherche Scientifique.
- Mallory, James P. 2013. *The Origins of the Irish*. London: Thames & Hudson.
- Matasović, Ranko. 2009. *Etymological dictionary of Proto-Celtic*. Leiden – Boston: Brill.
- Sammallahti, Pekka. 1998. *The Saami Languages*. Karašjohka / Karasjok: Davvi Girji.
- SKES = *Suomen kielen etymologinen sanakirja*, I-VII. Y.H. Toivonen et al., eds. Helsinki: Lexica Societatis Fenno-Ugricae XII,1–7, 1955f.
- UEW = *Uralisches etymologisches Wörterbuch*. Károly Rédei et al., eds. Budapest: Akadémiai Kiadó 1986–88.

Václav Blažek
Department of Linguistics and Baltic Studies
Masaryk University – Brno
Czech Republic
blazek@phil.muni.cz

REVISITING GREEK ἄχθος

GREGORY HAYNES

ABSTRACT

*The etymology of Greek ἄχθος, ἄχθομαι ‘burden, to be loaded, to be vexed’ is uncertain. It has been compared to Hittite hatk-^i ‘to shut, to close,’ with derivatives appearing to connote ‘tight, pressing.’ Its analysis as a disyllabic root, $*h_2\acute{e}d^h\acute{g}^h\text{-e-}$, is also problematic since PIE roots were, in most cases, monosyllabic. The present paper will consider the possibility that Greek ἄχθος is a compound consisting of two separate PIE roots, one of which is $*d^h\acute{e}g^h\text{-}$ ‘earth’ as attested in Greek χθών.*

HOMER AND HESIOD’S USE OF ἄχθος ‘BURDEN, WEIGHT’

1. “Telemachus, no man is more unlucky in his guests than you, seeing that you keep such a filthy vagabond as this man here, always wanting bread and wine, and skilled neither in the works of peace nor those of war, but a mere burden on the earth.”¹

The Greek phrase here translated as “... burden on the earth” is ἄχθος ἀρούρης, where ἀρούρης is literally ‘arable land.’²

2. “They pelted us from the cliffs with rocks huge as a man could lift, and at once there rose throughout the ships a dreadful din...”³

The phrase here translated as “huge as a man could lift” is ἀνδραχθέσι. LSJ glosses this word as “loading a man, as much as a man can carry.”⁴

3. “And as when a shepherd easily carries the fleece of a ram, taking it in one hand, and but little does its weight burden him, so Hector lifted up the stone and carried it...”⁵

Here “little does its weight burden him” translates ὀλίγον τέ μιν ἄχθος ἐπείγει.

4. “He carried a mighty weight of dry wood...”⁶

¹ Homer, *Odyssey* 20.379, Murray translation, vol. 2, 308–309; see also *Iliad* 18.104 where the same phrase is repeated.

² Beekes, *Etymological Dictionary of Greek*, 137.

³ Homer, *Odyssey* 10.121, Murray translation, vol. 1, 366–367.

⁴ LSJ, s.v. ἀνδραχθήσ, 128.

⁵ Homer, *Iliad* 12.452, Murray translation, vol.1, 588–589.

⁶ Homer, *Odyssey* 9.233, Murray translation, vol. 1, 332–333.

This translates φέρε δ' ὄβριμον ἄχθος ὕλης ἀζαλέης...

5. "... it is bad if you put too great a load on your wagon and break the axle..."⁷

"put too great a load on your wagon" translates ἐπ' ἄμαξαν ὑπέρβιον ἄχθος αἰείρας.

Homer, Hesiod, and later authors also employed ἄχθος, ἄχθομαι and related forms to indicate the concepts: hated, vexed, pained, or grieved, but these uses are considered to be metaphorical derivations by Chantraine (2009: 143–144). With respect to both nominal and verbal forms of this Greek word, Frisk (1960: 200–201) considers them "nicht sicher erklärt."

MODERN LEXICOGRAPHERS

Risch (1964: 78) proposed an etymological connection between Greek ἄχθομαι and Hittite *hatk-i* 'to shut, to close.' This has been accepted by Puhvel, Beekes, Kloekhorst, and LIV (doubtfully).⁸ Beekes suggests that the semantic value 'shut' in Hittite must have arisen from the original concept, 'to squeeze.' The proposed root is **h₂ed^hg^h-*, which Kloekhorst calls an "awkward looking form." The vast majority (perhaps originally all) of PIE roots are monosyllabic and this, along with the semantic divergence, raises questions about the etymological correspondence.⁹

An Alternative Analysis based on PIE **d^héǵ^h-* 'earth'

Alternatively, ἄχθος could be considered an ancient compound ἄ-χθος. The second half of this compound would then be parallel to Grk. χθών 'earth, ground, land,' a reflex of PIE **d^héǵ^h-om-*, **d^hǵ^h-om-* 'earth, ground, man (as earth worker), slave' as seen in the following cognates: Hit *tēkan* 'earth, ground,' Ved *kṣám-* 'earth, ground,' Skt *kṣa* 'field, protector or cultivator of a field, peasant,' Lat *humus* 'earth,' *homo* 'human being,' OE *guma* 'man, (bride)groom,' Lith *žėmė* 'earth,' OCS *zemlja* 'earth, land,' Phrygian *zemel* 'slave,' TochA *tkam* 'earth, ground,' Alb *dhe* (from **d^hǵ^h-em-*) 'earth.'¹⁰

As is well-known, this root often appears in radical metathesis form (with inversion in the ordering of root consonants) and often is attested in zero grade (as in the Greek and Vedic).

If such a compound form was originally a Greek construction, the initial element of ἄ-χθος would either be the well-known alpha-copulative signifying 'with, together with'¹¹, or else an assimilated form of Grk. ἄγω 'drive, lead, carry.'¹²

⁷ Hesiod, *Works and Days*, 692.

⁸ See Beekes, *Etymological Dictionary of Greek*, 183; Kloekhorst, *Etym. Dict. Of Hittite Inherited Lexicon*, 134–135; LIV 255; Puhvel, HED 3, 266f.

⁹ In Haynes (2023:59), I argue that PIE root structure can be generalized as **(s)-C [+/- R (R)] -C-*, where (s) is the s-mobile, C is any consonant, and R is any resonant or laryngeal (or a zero-grade of the same), and that apparently disyllabic roots, including any additional element that follows the final consonant of the monosyllabic root, is a root-extension, a derivational ending, a suffix, or the remnant of some ancient compound that will not have been a part of the original root.

¹⁰ Haynes 2023, Table 17; IEW 414–416; EIEC 174; NIL 86–88; Mallory & Adams 120; Watkins 20; DELG 143; Ringe 19; EDHIL 858–862; Bomhard 145; EIEC 247–248; Illič-Svityč no. 69; Ruhlen and Bengtson 323–326; Fortson 461 (*zemel*); MW 324; Blažek, *Tocharian 'man' vs. 'god' in Perspective of Semantic Oppositions*, 105–145

¹¹ Smyth, 250; Beekes 1.

¹² Beekes 18–19; LSJ 17–18.

If the first, this would then signify ‘with earth,’ and would originally have been applied to commoners or slaves tasked with carrying loads of earth, clay, mud, or stones for construction projects, indicating that they were loaded, weighed down, burdened. From here it is not difficult to imagine the semantic shift to vexed, pained, or grieved, since those emotions would be a natural outcome of being burdened with heavy loads and worked to one’s extreme limits.

If the second case is the correct one, then the compound would be ἄγ-χθ-ος, where the gamma was assimilated to the following chi, signifying: *carrying earth* → *loaded, burdened*.

On the other hand, if ἄχθος reflects an ancient PIE rather than a later Greek compound, then the two possibilities would be the following:

- 1) PIE **sṃ-* or *som-* as attested in Lat *sem-* ‘at the same time, together,’ OHG *samn* ‘together,’ Lith *sam-* ‘with,’ OCS *so-* ‘with,’ Av *ha(m)-* ‘together,’ Skt *sam-* ‘with,’ Grk ἄ ‘with, together with.’¹³

Plus: PIE **dʰéǵʰ-* **dʰǵʰ-* ‘earth, ground, land’ as attested in the previously listed examples.
yielding:

**sṃ-dʰǵʰ-* or (in metathesis form) **sṃ-ǵʰdʰ-* → Grk ἄχθ-ος ‘with earth, weighted, loaded.’

Or, if the first element of the PIE compound was an assimilated form of **h₂eǵ-* ‘lead, carry, drive’ (instead of **sṃ-* as above), then the components would be:

- 2) PIE **h₂eǵ-* ‘lead, carry, drive’ as seen in the following PIE cognates: Lat *agō*; Grk ἄγω ‘lead, carry, fetch, bring; Ved *ájati* ‘to drive’; Arm *acem* ‘leads,’ OIr *-aig, -agat* ‘drive, lead,’ *tāin* (from **to-ag-no*) ‘raid,’ ON *aka* ‘go, travel, drive,’ Mlr *āg* ‘fight, warrior’s ardor,’ TochAB *āk-* ‘lead, guide, drive,’ *āsām* ‘lead.’¹⁴

Plus: PIE **dʰéǵʰ-* **dʰǵʰ-* ‘earth, ground, land’ as described above
yielding (with metathesis):

**h₂eǵ-ǵʰdʰ-* (then with assimilation) → **h₂e-ǵʰdʰ-* → Grk ἄχθ-ος ‘carrying earth, weighted, loaded.’

A third (more remote) possibility is that the initial ἄ in ἄχθ-ομαι is a prothetic vowel¹⁵ parallel to Grk. χθές ‘yesterday’ and ἐχθές ‘yesterday,’ in which case the meaning of ἄχθομαι would be simply ‘to be *earthed* → to be weighed down with earth.’

¹³ Beeks 1; M&A 291; de Vann 553; IEW 902–905; Blažek, *The Prefix *Sṃ-/Som- in the Indo-European Tree- and Plant-Names*, 5–11.

¹⁴ Haynes 2024, Table 7; LIV 255–256; IEW 4–6; LSJ 8, 14, 17–18; OLD 85, 87; Monier-Williams 9; DELG 9, 16; Bosworth and Toller 5 (see LIV 256, n. 3); Mallory and Adams 280, 403; Buck 191; EIEC 201, 284, 348; Frisk I-18, II-348; EWAia 50–52; Beekes 18–19; de Vries 3; Adams 36; Anttila 1ff and Anttila, “Aggression and Sustenance, 121; NIL 267–270; Watkins 1; Bomhard 706–707; Dolgopolsky no. 17.

¹⁵ Smyth §41, p. 17; LSJ 748, 1991.

COGNATE WITH HITTITE?

If either of the above cases is correct, then it could not be possible for Hittite *h^hatk-i* ‘shut, close’ to be cognate with Grk ἄ-χθ-ος for the following reasons:

- No reflex of PIE **h₂eg-* is found in Hittite.
- The PIE prefix **sm* ‘with, together with’ is attested in Hittite in the form =(š)šan, not as the *ha* in *h^hatk-i*.

WHAT ABOUT THE GODDESS Ἀχθεία ?

Hesychius defines Ἀχθεία as ἡ Δημήτηρ, μυστικῶς ‘Demeter, according to the mysteries.’ Since Demeter is typically taken as “Mother Earth,” this would confirm a link between the word ἄχθος and the concept *Earth*. The second element of Δημήτηρ is obviously ‘mother.’ The first is thought by some as a pre-Greek word for ‘earth,’ and by others as Illyrian (attested by Alb. *dhe* ‘earth’) and therefore a cognate of Grk χθών, arising from PIE **d^hǵh-* ‘earth, ground, land.’¹⁶

ADDITIONAL EVIDENCE FROM THE SANSKRIT

As mentioned above, PIE **d^hǵh-em* became Ved *kṣám-*, Skt *kṣā́* ‘earth, ground.’ The initial and final consonants of this zero-grade root (via an intermediate Proto-Indo-Iranian *džh-*) fused to become Skt *kṣ* as in *kṣám-* and in a number of other derivatives, not all of which are typically viewed as being based on the root **d^hǵh*. The following are a few such instances (with comments):

1. *kṣmā́* ‘the earth,’ *kṣmā́tala* ‘the surface of the earth,’ *kṣā́ma* ‘ground, earth.’¹⁷

These are straightforward attestations of PIE **d^hǵh-em*.

2. *kṣa* ‘field, protector or cultivator of a field, peasant’¹⁸
3. *kṣoná* ‘a multitude of men, people (as opposed to the chief),’ *kṣonī* ‘the earth,’ *kṣonimaya* ‘containing or representing the earth,’ *kṣauṇī* ‘the earth,’ *kṣauṇītala* ‘the surface of the earth.’¹⁹

In an earlier article (Haynes 2023: 106–109) I have argued that attestations of PIE **d^hǵh-em* with the semantic value ‘man’ (Lat *homo* ‘human being,’ Goth *guma* ‘man,’ Phrygian *zemel* ‘slave’), as well as another root of the same consonantal form, **d^h—ǵh-* (i.e., **d^heuǵh-*, with the Greek attestation τυγχάνω carrying as one of its meanings ‘common, or vulgar men’), refer not to men in general, but rather to men of a lower class, i.e., earth workers, cultivators of the soil, or builders employing earth, stones, mud, or clay. The preceding two Sanskrit forms: *kṣa* ‘cultivator of a field, peasant’ and *kṣoná* ‘a multitude of men, people (as opposed to the chief)’ clearly fit into the same

¹⁶ Beekes, 324; see also EIEC, s.v. Earth Goddess, 174; and OCD, s.v. Demeter, 263.

¹⁷ MW 333; KEWA 288; 295; EWAia 423–425. (Note: MW employs the notation /sh/ to indicate what in modern usage is written /ṣ/. In the list to follow, these will be normalized to /ṣ/.)

¹⁸ MW 324.

¹⁹ MW 333.

rubric. These terms do not refer to the rulers or chiefs, but rather to the commoners and slaves. According to the ancient logic: they are low, they are dirty, they are dirt.

4. *kṣétra* ‘landed property, land, soil, field, place region, country, enclosed plot of ground,’ *kṣétrakara* ‘cultivating a field,’ *kṣétrakarṣaka* ‘soil-plougher, husbandman,’ *kṣétraja* ‘produced in a field,’ *kṣétrajeśá* ‘contest for landed property, acquisition of land,’ *kṣétratā* ‘seat, place of residence,’ *kṣétrapati* ‘owner of a field, land owner, landlord, farmer,’ *kṣétrasāti* ‘acquisition of fields or land,’ *kṣatrá* ‘dominion, supremacy, power, might,’ *kṣatríya*, ‘governing, endowed with sovereignty, dominion, supremacy, power, might, the second or military caste.’²⁰

KEWA expresses uncertainty as to the etymology of these terms, as does EWAia. I would argue that the derivation from **d^hǵ^h-* is very likely, considering the many references to ‘earth, land, soil, field, plot of ground, cultivation, plowing, acquisition of land, possession of land, lordship over lands and countries.

5. *kṣubh*, *kṣubhita*, *kṣubdha*, *kṣobha*, *kṣóbhaṇa*, *kṣobhya* ‘shake, tremble, be agitated or disturbed, be unsteady, stumble, cause to shake, stir up, excite, frightened, alarmed, angry, enraged, causing emotion’²¹
6. *kṣip* ‘to utter abusive words, insult, revile, abuse.’²²

The semantic values of these two Sanskrit terms are directly parallel to those derivations of Grk ἄχθομαι and ἄχθος that have been glossed as: ‘mental oppression, to be hated, vexed, pained, or grieved; burden, trouble; burdensome, oppressive; annoyance, suffering; vexed at being conquered, etc.’²³ I would argue that these are compound roots based on PIE **d^hǵ^h-* as the first element with some root containing a strong labial consonantal component as the second.

7. *kṣi* ‘possess, have power over, rule, govern, be master of,’ *kṣi* ‘abide, stay, dwell, reside, inhabit,’ *kṣití* ‘dwelling, habitation, house, the earth, soil of the earth, settlements, colonies, nations,’ *kṣitíja* ‘earth-born, produced of or in the earth, son of the earth, daughter of the earth,’ *kṣititrāṇa* ‘protection of the earth’ *kṣitidhārin* ‘carrying soil or earth,’ *kṣitípati* ‘lord of the earth,’²⁴

These Skt words are typically ascribed to the PIE roots **tk-ej-* and **te-tk-*. In that previously mentioned earlier article (Haynes 2023: 107–108) I have argued that these two PIE roots:

1. **tk-ej-* ‘cultivate soil, settle a land, dwell in a place’

²⁰ MW 332; KEWA 285, 293; EWAia 422, 432.

²¹ MW 331; KEWA 292; EWAia 440 s.v. *kṣobh*.

²² MW 328.

²³ Beekes 183; MW 296.

²⁴ MW 324–328; KEWA 289; EWAia 427, 431

Ved *kṣéti* ‘dwells, lingers,’ Myc *ki-ti-je-si* = /*kʰiensi*/ ‘to build on, cultivate, or work land,’ Lat *pōnō* ‘put, place, sit down,’ Grk *κτίσις* ‘settlement,’ *κτίζω* ‘people a country and build houses and cities in it,’ Av *šiti* ‘settlement,’ Arm *šēn* ‘dwell, build on, farm, town.’²⁵

and

2. **tek-s*, **te-tk-* ‘establish, produce, hew, cut, fabricate, fashion, axe’

Lith *tašyti* ‘hew, trim,’ OCS *tesati* ‘hew,’ Skt *tákṣati* ‘fashions, creates, carpenters, cuts,’ Grk *τέκτων* ‘architect,’ *τέχνη* ‘art, craft, skill, technique,’ Skt *tákṣan* ‘carpenter,’ Hit *taksanzi* ‘undertake, prepare, cause, joint,’ OHG *dehsa* ‘axe.’²⁶

are de-voiced and de-aspirated variants of PIE **dʰǵʰ-*, showing phonetic development similar to that which occurred in Tocharian. The first shows a clear connection to the earth: ‘settling, building, and dwelling on land; cultivating and farming the ground.’ The second refers fundamentally to the practice of building with earth as the basic material of walls and houses constructed of mud bricks. Later, builders began using wood and other materials, but the term for building (originally working with stones, mud, and clay) was later applied to other materials and techniques.

Examples of this phonetic reduction (loss of voicing and aspiration) were given in that earlier article. The part of the introduction of that article that describes this process is reproduced below by way of example, but many more instances are given there.

PHONETIC REDUCTION

Another type of root mutation could be called *reduction*. This concept is also familiar, since something very close to it is seen in Tocharian (and to some extent in Hittite) where the rich PIE obstruent inventory has been reduced to include only the simple, unvoiced, unaspirated (lenis) forms.²⁷ It has been suggested that this change may have been due to the influence of a substrate language with a similarly limited range of obstruents.²⁸ This same dynamic can be seen in distinct synonymous PIE roots. Obvious examples are often remarked upon in the standard handbooks. The following are some examples of this transformation:

- **gol(H)uos* ‘bare, bald’ **k_lHuos* ‘bald’²⁹
- **-dʰro-*, **dʰlo-* = instr. suffix **-tro-*, **-tlo-* = instr. suffix³⁰

²⁵ Beekes 788–789; LIV **tkei-* 643; IEW 626; Watkins 95; Mallory and Adams 223; EIEC 622. Compare possible metathesis form: TochB *²keta* ‘parcel of land, estate, field,’ Adams, *Dictionary of Tocharian B*, 191; and Adams, *History and Significance of Some Tocharian B Agricultural Terms*, 373.

²⁶ LIV **tek-* 638; IEW **tekp-* 1058–1059; Watkins 92; Mallory and Adams 220, 243, 283; Bomhard 206; EIEC 139; Beekes 1460; EDHIL 813–814.

²⁷ See EIEC 14, 28, 592. See also Kloekhorst, “Chapter 5: Anatolian,” in Thomas Olander, ed., *The Indo-European Language Family*, 2022, “...the merger of PIE mediae and aspiratae into a single series that is called lenis (PIE *d, *dʰ > PANat. */t/)...” See Hodge, “Indo-European Consonant Ablaut,” 143–162, for an early attempt to systematize some of these features along with a good survey of the prior literature on the subject.

²⁸ Peyrot, “The deviant typological profile of the Tocharian branch of Indo-European may be due to Uralic substrate influence,” 72–121.

²⁹ EIEC 45; IEW 554.

³⁰ EIEC 52; IEW 692; Mallory and Adams 57.

- | | |
|--|--|
| • <i>*h₂eng-</i> ‘bend’ | <i>*h₂enk-</i> ‘bend’ ³¹ |
| • <i>*peh₂ǵ-</i> ‘fasten securely’ | <i>*peh₂k-</i> ‘fasten securely’ ³² |
| • <i>*peǵ-</i> ‘draw, color’ | <i>*peǵk-</i> ‘draw, color’ ³³ |
| • <i>*sred^h-</i> ‘boil, be agitated, move’ | <i>*sret-</i> ‘boil, be agitated, move noisily’ ³⁴ |
| • <i>*b^hend^h-ǵros</i> ‘relation’ | <i>*pent-h₂ǵros</i> ‘father-in-law’ ³⁵ |
| • <i>*h₂eug-</i> ‘increase, become strong’ | <i>*h₂euk-s-</i> ‘grow, become large’ ³⁶ |
| • <i>*greh₂b^h-</i> ‘hornbeam’ | <i>*karp-</i> ‘hornbeam’ ³⁷ |
| • <i>*g^hab^h-</i> ‘take, seize’ | <i>*kap-</i> ‘have, hold, seize’ ³⁸ |
| • <i>*pleh₂g-</i> ‘strike, beat’ | <i>*pleh₂k-</i> ‘strike, beat’ ³⁹ |
| • <i>*k₂uoidis</i> ‘white’ | <i>*k₂uoi₂ós</i> ‘white’ ⁴⁰ |
| • <i>*sab-</i> ‘sap’ | <i>*sap-</i> ‘sap’ ⁴¹ |
| • <i>*steǵb-</i> ‘make stiff’ | <i>*steǵp-</i> ‘make stiff’ ⁴² |
| • <i>*deǵ-</i> ‘teach, show, indicate’ | <i>*deǵk-</i> ‘preach, say, index’ ⁴³ |
| • <i>*ueǵb^h-</i> ‘vibrate, be agitated’ | <i>*ueǵp-</i> ‘move back and forth, vibrate’ ⁴⁴ |
| • <i>*g^heb^hól</i> ‘head’ | <i>*kapolo-</i> ‘head’ ⁴⁵ |
| • <i>*g^hrd-</i> ‘heart’ | <i>*kér₂d-</i> ‘heart’ ⁴⁶ |

Many more examples of this dynamic can be observed once the variation of medial resonants in PIE roots is allowed for. The evidence suggests that an ancient dialectical subset of PIE speakers experienced a phonetic influence similar to that which occurred in Tocharian, and then, during a later period of reunification with a group that had not experienced this linguistic change, the dialects became merged. The result is that, after this merger, synonymous pairs (doublets) coexisted within the basic vocabulary of PIE and these have persisted down into the various daughter languages. These synonyms are now considered separate roots, but they should, it will be argued, be seen as variants of an ancient original.

In their most strict formulation, these phonetic reductions can be summarized as follows:

- d, d^h became t
- b, b^h became p

³¹ EIEC 61; IEW 45–46.

³² EIEC 64; IEW 787–788.

³³ EIEC 64; IEW 794–795; LIV 464.

³⁴ EIEC 76; IEW 1001–1002.

³⁵ EIEC 196; IEW 127; Beekes 1171.

³⁶ LIV 274–275, 288–289; EIEC 248; IEW 84–85.

³⁷ EIEC 273; de Vaan 94; Mallory and Adams 161.

³⁸ EIEC 563; IEW 407–409, 527–528; Watkins, s.v. “*kap-*” 38.

³⁹ LIV 484–485, see 485, n. 1 regarding the original identity of these roots.

⁴⁰ Mallory and Adams 332; Watkins 46; IEW 628–629; see below, Table 19.

⁴¹ Mallory and Adams 158; IEW 880.

⁴² LIV 592, 594.

⁴³ Watkins 15; IEW 188.

⁴⁴ de Vaan 674; IEW 1131; LIV 671.

⁴⁵ See Haynes 2023, Table 18.

⁴⁶ IEW 580; EIEC 262–263; Mallory and Adams 187; Michael Witzel, “Comparison and Reconstruction,” 48.

- $\hat{g}$, $\hat{g}^h$ became $\acute{k}$,
- g , g^h became k
- g^u , g^{uh} became k or k^u

The semantic parallels of **tek-s*, **te-tk-* to the voiced and aspirated form of this root can best be seen by way of comparison to the following, also quoted from that earlier article:

**d^heġġ^h-*, **d^hiġġ^hs-* ‘form, build, mold mud or clay, knead, smear, plaster; wall of mud’ Skt *dēhmi* ‘spread, fill,’ *dēhī* ‘wall, rampart, dam,’ Goth *digan* ‘form, fashion, knead, make pottery,’ ON *deig* ‘dough,’ *digr* ‘thick,’ NE *dough*, Lith *žiedžiù* ‘form from mud,’ TochB *tsikale* ‘to form,’ Lat *fingō*, *finxī* ‘form, shape,’ *figūra* ‘form, shape, figure,’ *fictilis* ‘fashion out of clay, made of earth or clay,’ *figulus* ‘potter,’ Av *pairi-daēza-* ‘enclosure’ (> NE *paradise*); Grk *τειχος*, *τοιχος* ‘wall, embankment,’ possibly Grk *θιγγάνω* ‘touch with the hand,’ OIr *digen* ‘build, firm, solid, hard, strong, fixed.’⁴⁷

Mallory and Adams (223-224, 371) write, “The underlying semantics of **dheġġh* indicate that it was specifically associated with the working of clay (e.g. Lat *fingō* ‘fashion,’ Skt *dēhmi* ‘smear, anoint,’ TochAB *tsik-* ‘fashion [pots, etc.],’ hence the English cognate *dough*; in Greek and Indo-Iranian it is also associated with building walls, e.g. Av *pairi-daēza* ‘build a wall around’ ... but there are also cognates of more general meaning, e.g. OIr *con-utainc* ‘builds,’ Lith *diežti* ‘whip, beat,’ Arm *dizanem* ‘heap up’.” And in EIEC (629) they write: “The substance from which the walls were made, [earth] came to be applied both to the finished product, e.g., Grk *τειχος* ‘wall,’ Av *uz-daēza-* ‘wall,’ and clay-like substances, e.g. Germanic *dough*.”

Another probable example of a reduced form of *ἄχθος* ‘load, burden, mental oppression; vexed, pained, or grieved; burden, trouble,’ can be seen in a Greek word denoting the same semantic field: *ἀγανακτέω* ‘be indignant or irritated.’ This is a word of uncertain etymology, with Beekes even doubting its source in Greek at all.⁴⁸ This word could be analyzed as *ἀγαν-α-κτέω*, with the first element being the familiar Greek adverb *ἄγαν* ‘much, too much,’ and the remaining compound *-α-κτέω* being a phonetically reduced version of *ἄχθος*, *ἄχθομαι*. The original meaning of the compound would therefore be ‘overloaded, overburdened, literally: too much with earth, or carrying too much earth.’

CONCLUSION

The preponderance of the evidence suggests that Greek *ἄχθος*, *ἄχθομαι* is a compound based on the PIE root **d^heg^h-* ‘earth, ground, field, earth worker.’

⁴⁷ LIV 140; IEW 244; NIL 118; de Vries 194; Mallory & Adams 223–224, 228; Watkins 18; EIEC 283, 649; Bomhard 166.

⁴⁸ Beekes 8.

ABBREVIATIONS OF LITERATURE

Adams	Douglas Q. Adams, <i>Dictionary of Tocharian B</i>
AHD	<i>American Heritage Dictionary</i>
ALEW	Hock, Wolfgang, <i>Altltitaisches etymologisches Wörterbuch</i>
CLL	Melchert, Craig, <i>Cuneiform Luvian Lexicon</i>
DELG	Chantraine, Pierre, <i>Dictionnaire étymologique de la langue grecque</i>
de Vaan	de Vaan, Michiel, <i>Etymological Dictionary of Latin & other Italic Languages</i>
EDHIL	Kloekhorst, Alwin, <i>Etymological Dictionary of the Hittite Inherited Lexicon</i>
EIEC	Mallory, James P., and Douglas Q. Adams, <i>Encycl. of Indo-European Culture</i>
EWAia	Mayrhofer, Manfred, <i>Etymologisches Wörterbuch des Altindischen</i>
IEW	Pokorny, Julius, <i>Indogermanisches etymologisches Wörterbuch</i>
KEWA	Mayrhofer, Manfred, <i>Kurzgefaßtes etymologisches Wörterb. des Altindischen</i>
LIV	Rix, Helmut, <i>Lexicon der indogermanischen Verben</i> . 2 nd edition
LSJ	Liddell, Scott, and Jones, <i>A Greek–English Lexicon</i>
Mallory and Adams	Mallory and Adams, <i>The Oxford Introduction to Proto-Indo-European and the Proto-Indo-European World</i>
NIL	Wodtko, Irslinger, and Schneider, <i>Nomina im Indogermanischen Lexikon</i>
OCD	<i>Oxford Classical Dictionary</i>
OLD	<i>Oxford Latin Dictionary</i>

REFERENCES

- Adams, Douglas Quentin. “On the History and Significance of Some Tocharian B Agricultural Terms.” *The Bronze Age and Early Iron Age Peoples of Eastern Central Asia*. Victor H. Mair, ed. Washington D.C.: Institute for the Study of Man, 1998, 372-378.
- Adams, Douglas Quentin. *A Dictionary of Tocharian B*. Amsterdam: Rodopi, 1999.
- Anttila, Raimo. “Aggression and Sustenance: Driving (*ag-) and Beating (*g^when-) Symbiosis in (Proto-) Indo-European.” *UCLA Indo-European Studies* 1, edited by Vyacheslav V. Ivanov and Brent Vine. University of California at Los Angeles, Program in Indo-European Studies, 1999.
- Anttila, Raimo. *Greek and Indo-European Etymology in Action: Proto-Indo-European *aĝ-*. Amsterdam: John Benjamins, 2000.
- Beekes, Robert. *Etymological Dictionary of Greek*. Leiden and Boston: Brill, 2009.
- Blažek, Václav. “The Prefix *Sṃ-/Som- in the Indo-European Tree- and Plant-Names.” *Studia Minora Facultatis Philosophicae Universitatis Brunensis*. A 51, 2003—Linguistica Brunensis.
- Blažek, Václav. “Tocharian ‘man’ vs. ‘god’ in Perspective of Semantic Oppositions.” *The Journal of Indo-European Studies*, Volume 51, Number 1 & 2, Spring/Summer 2023, 105-145.
- Bomhard, Allan R. *A Comprehensive Introduction to Nostratic Comparative Linguistics: With Special Reference to Indo-European*. 4th ed. Charleston, South Carolina, 2021.
- Bosworth and Toller. *Anglo-Saxon Dictionary*. Oxford: Oxford University Press, 1921.
- Buck, Carl Darling. *A Dictionary of Selected Synonyms in the Principal Indo-European Languages*. Chicago: The University of Chicago Press, 1949.
- Chantraine, Pierre. *Dictionnaire étymologique de la langue grecque: Histoire des mots*. new edition. Klincksieck, 2009.
- de Vaan, Michiel. *Etymological Dictionary of Latin and the other Italic Languages*. Leiden: Brill, 2008.

- de Vries, Jan. *Altnordisches etymologisches Wörterbuch*. 2nd ed. Leiden: Brill, 1977.
- Dolgopolsky, Aharon. *Nostratic Dictionary*. Cambridge, UK: McDonald Institute for Archaeological Research, 2008.
- Fortson Benjamin W. *Indo-European Language and Culture: An Introduction*. 2nd ed. Chichester, U.K.: Wiley-Blackwell, 2010.
- Fraenkel, Ernst. *Litauisches etymologisches Wörterbuch*. 2 vols. Heidelberg: Carl Winter, 1962-1965.
- Frisk, Hjalmar. *Griechisches etymologisches Wörterbuch*. Heidelberg: Carl Winter Universitätsverlag, 1960.
- Greenberg, Joseph H. *Indo-European and Its Closest Relatives: The Eurasiatic Language Family*. vol. 2 *Lexicon*. Stanford: Stanford University Press, 2002.
- Haynes, Gregory. *Tree of Life, Mythical Archetype: Revelations from the Symbols of Ancient Troy*. San Francisco: Symbolon Press, 2009.
- Haynes, Gregory. "Resonant Variation in Proto-Indo-European." *Mother Tongue* 22 (2020): 151-222.
- Haynes, Gregory. "Resonant Variations on Immortality" *Mother Tongue Journal* 23 (2021): 151-162.
- Haynes, Gregory. "Root Transformations in Proto-Indo-European." *Mother Tongue Journal* 24 (2023): 57-132.
- Haynes, Gregory and Michael Witzel. "Of Dice and Divination." *Journal of Comparative Mythology* 2 (2016): 1-26, https://www.academia.edu/44802729/Of_Dice_and_Divination.
- Hesiod. *Works and Days*. Evelyn-White, trans. revised edition. Cambridge, Massachusetts: Harvard University Press, 1936.
- Hesiod. *The Homeric Hymns and Homerica*. Hugh G. Evelyn-White, trans. *Works and Days*. Cambridge, MA: Harvard University Press, 1914.
- Hock, Wolfgang, et al. *Altltauisches etymologisches Wörterbuch*. Berlin: Humboldt-Universität zu Berlin, 2019.
- Hodge, Carleton. "Indo-European Consonant Ablaut." *Diachronica* 3, no. 2 (1986): 143-162.
- Homer. *Odyssey*. trans. A. T. Murray, rev. George E. Dimock, 2nd ed. Cambridge, Massachusetts: Harvard University Press, 1995 (first published 1919).
- Homer. *Iliad*. Translated by A. T. Murry. Cambridge, Massachusetts: Harvard University Press, 1999.
- Illič-Svityč, Vladislav Markovič. *Materials for a Comparative Dictionary of the Nostratic Languages (Indo-European, Altaic, Uralic, Dravidian, Kartvelian, Hamito-Semitic)*. Этимология 'Ėtimologija' (1965): 321-373.
- Illič-Svityč, Vladislav Markovič. *Опыт сравнения ностратических языков*. 'Comparative Dictionary of Nostratic Languages.' 3 vols. Moscow: Nauka, 1971, 1976, 1984.
- Kloekhorst, Alwin. *Etymological Dictionary of the Hittite Inherited Lexicon*. Leiden and Boston: Brill, 2008.
- Kloekhorst, Alwin. "Chapter 5: Anatolian." *The Indo-European Language Family*, edited by Thomas Olander. Cambridge: Cambridge University Press (2022): 63-82.
- Kluge, Friedrich. *Etymologisches Wörterbuch der deutschen Sprache*. 19th edition. Berlin: De Gruyter, 1963.
- Mallory, James Patrick, and Douglas Q. Adams. *Encyclopedia of Indo-European Culture*. London and Chicago: Fitzroy Dearborn Publishers, 1996.

- Mallory, James Patrick, and Douglas Q. Adams. *The Oxford Introduction to Proto-Indo-European and the Proto-Indo-European World*. Oxford: Oxford University Press, 2006.
- Liddell, Henry George, Robert Scott, and Henry Stuart Jones. *A Greek–English Lexicon*. Oxford: Clarendon Press, 1968.
- Mayrhofer, Manfred. *Etymologisches Wörterbuch des Altindiarischen*. 2 vols. Heidelberg: Carl Winter Universitätsverlag, 1992.
- Mayrhofer, Manfred. *Kurzgefaßtes etymologisches Wörterbuch des Altindischen*. 4 vols. Heidelberg: Carl Winter Universitätsverlag, 1956-1980.
- Melchert, Craig. *Cuneiform Luvian Lexicon*. Chapel Hill, North Carolina, 1993.
- Monier-Williams, Sir Monier. *A Sanskrit–English Dictionary*. new ed. Delhi: Motilal Banarsidass, 1981. Originally published Oxford: Oxford University Press, 1899.
- Oxford Classical Dictionary*, edited by Max Cary, Herbert J. Rose, H. Paul Harvey, and Alexander Souter. Oxford: Clarendon Press, 1964.
- Oxford Latin Dictionary*, edited by Peter G. W. Glare. Oxford: Clarendon Press, 1982.
- Peyrot, Michaël. “The deviant typological profile of the Tocharian branch of Indo-European may be due to Uralic substrate influence.” *Indo-European Linguistics* 7, no. 1. Leiden: Brill (2019): 72–121.
- Pokorny, Julius. *Indogermanisches etymologisches Wörterbuch*. Bern: Francke, 1959.
- Puhvel, Jaan. *Hittite Etymological Dictionary. Volume 3, Words beginning with H*. Berlin, Boston: De Gruyter Mouton, 2011.
- Ringe, Don. *From Proto-Indo-European to Proto-Germanic*. Oxford: Oxford University Press, 2006.
- Rix, Helmut, et al. *Lexicon der Indogermanischen Verben*. 2nd ed. Wiesbaden: Dr. Ludwig Reichert Verlag, 2001.
- Ruhlen, Merritt and John Bengtson. “Global Etymologies.” *On the Origin of Languages: Studies in Linguistic Taxonomy*. Merritt Ruhlen. Stanford, California: Stanford University Press, 1994. 277-336.
- Smyth, Herbert Weir. *Greek Grammar*. Cambridge, Massachusetts: Harvard University Press, 1956.
- Thesaurus Linguae Graecae*. Univ. of California at Irvine. Maria Pantelia, Director. <https://stephanus.tlg.uci.edu/>.
- Vigfusson, Gudbrand. *Icelandic–English Dictionary*. Oxford: Clarendon Press, 1874.
- Watkins, Calvert. *American Heritage Dictionary of Indo-European Roots*. 3rd ed. Boston: Houghton Mifflin Harcourt, 2011.
- Witzel, Michael. “Comparison and Reconstruction: Language and Mythology.” *Mother Tongue Journal* 6 (2000/2001): 45-62.
- Wodtko, Dagmar S., Britta Irslinger, and Carolin Schneider. *Nomina im Indogermanischen Lexikon*. Heidelberg: Universitätsverlag Winter, 2008.

ABBREVIATIONS OF NAMES OF INDO-EUROPEAN LANGUAGES

Alb	Albanian	Corn	Cornish
Arm	Armenian	Cymr	Cymric
Av	Avestan	Gall	Gallo-Roman
Bret	Breton	Gaul	Gaulish
Bulg	Bulgarian	Goth	Gothic
CLuv	Cuneiform Luvian	Grk	Greek

HLuv	Hieroglyphic Luvian	OLat	Old Latin
Hit	Hittite	OLith	Old Lithuanian
Illyr	Illyrian	ON	Old Norse
Khot	Khotanese	OPers	Old Persian
Lat	Latin	OPrus	Old Prussian
Latv	Latvian	ORus	Old Russian
Lith	Lithuanian	OSax	Old Saxon
Luv	Luvian	Osc	Oscan
Lyc	Lycian	Oss	Ossetic
Lyd	Lydian	OSwed	Old Swedish
Mcymr	Middle Cymric	OWels	Old Welsh
ME	Middle English	Phryg	Phrygian
MHG	Middle High German	PIE	Proto-Indo-European
Mlr	Middle Irish	Pol	Polish
MPers	Middle Persian	Rus	Russian
MWels	Middle Welsh	SC	Serbo-Croatian
Myc	Mycenaean Greek	Skt	Sanskrit
NE	New English	Slav	Slavic
Norw	Norwegian	Sogd	Sogdian
NPers	New Persian	Swed	Swedish
NWels	New Welsh	TochA	Tocharian A
OAv	Old Avestan	TochB	Tocharian B
OCS	Old Church Slavonic	Ukr	Ukrainian
OE	Old English	Umb	Umbrian
OFris	Old Frisian	Ved	Vedic
OHG	Old High German	YAv	Young Avestan
OIr	Old Irish		

REVIEW OF JAN HENRIK HOLST, *PROPOSALS FOR THE DIACHRONIC INTERPRETATION OF TOCHARIAN*, DÜREN (SHAKER VERLAG) 2025, ISBN 978-3-8191-0155-7, 282 PAGES, 32.80 €.

CARSTEN PEUST, KONSTANZ

Among the twelve branches of the Indo-European language family, Tocharian (which is a shorthand term for two related languages Tocharian A and B) is certainly the least known and studied. This is not so much because Tocharian would be particularly irrelevant to Indo-European studies, but rather due to extralinguistic accidents: The language came to scholars' attention only late; nobody comes with Tocharian experience from home or from schooling; the texts, often translations of Buddhist content, are hard to digest for those familiar with the more classical Indo-European languages. This implies that the Tocharian evidence for reconstructing Proto-Indo-European (henceforth: PIE) is most likely underrated. In turn, this suggests that research on Tocharian should be particularly promising for providing new insights into the reconstruction of PIE. We should therefore welcome any new and accessible piece of literature—whether of scientific or didactic nature—that encourages working with this language. Such a specimen of an up-to-date, accessible linguistic contribution about Tocharian is found in the book reviewed here. This reviewer must confess that he does not read it as a Tocharian specialist, but only as a linguist interested in historical linguistics and Indo-European studies.

After an introductory chapter, Holst addresses—in this order—a number of issues of historical phonology (covering both shallow and deeper diachronic levels down to PIE), topics of morphology and syntax, and finishes with a section on the genetic and areal position of Tocharian. However, it is not a grammar book; rather, it singles out a number of more or less independent issues about which the author has something interesting to say.

Going through Holst's issues of historical phonology, I skip over his treatment of a number of consonantal correspondences between the two Tocharian varieties A and B, moving on directly to the so-called “dental problem” discussed in section 3.1, which, as we will see, has a considerable relevance for Indo-European studies in general. PIE is usually reconstructed with three series of stops, traditionally described as voiceless, voiced and voiced-aspirate, e.g. */k/, */g/, */gh/. It is also known that these three series merge into one in Tocharian at least for the velar and labial places of articulation. The only potential exception is */d/. While both */t/ and */dh/ clearly merge into Toch. /t/, the outcome of */d/ has been subject to dispute. Holst singles out and evaluates all

the different and contradictory proposals that have been made on the issue and begins with a newly compiled dataset of 87 etymologies that is large and as reliable as possible at the same time. The clear conclusion emerges that a development PIE */d/ > Toch. Ø (zero), initially proposed by W. Winter in 1962 (though without a satisfying explanation), is the one to be regarded as correct. One example out of many is Toch. A *ākār* B *akrūna* (pl.) = Greek δάκρυ “tear(s)”.

Competing with the traditional reconstruction of the PIE stop system (*/t/, */d/, */dh/), there is the “glottalic theory”, which reconstructs the same stops as */t/, */t'/ and */d/ respectively, mainly for reasons of typological naturalness. Holst sympathizes with this theory, which happens to align well with the zero representation of the aforementioned dental. According to the glottalic interpretation, the lost sound is */t'/ rather than */d/, which amounts to a process known as debuccalization. This interpretation also allows for grouping the surviving dentals (now conceptualized as */t/ and */d/) into a natural class “non-glottalic”. As acknowledged by Holst, the linking of the change */t'/ > Ø with the glottalic theory was briefly hinted at already by M. Peyrot in 2022. However, it is only in Holst’s book that the argument is exposed in full.

A second issue surrounding the “dental problem” is the frequent shift of the PIE dentals to /ts/ or /c/. Holst confirms a previous view according to which there were two stages of palatalization of the dental stops before front vowels, which led to *ts* (first palatalization) and *c* (second palatalization), respectively. For example, in Toch. A *pats* B *pets* = Greek πόσις (< **pótis*) “spouse”, the first palatalization *t* > *ts* occurred before the front vowel *-i-, which was later lost in Toch. In A *pācar* B *pācer* = Greek πατήρ “father”, by contrast, the first palatalization was suppressed by analogy with inflected forms of the type **patr*-, in which no front vowel was present. The dental stop therefore survived as such for a while before undergoing the second palatalization, *t* > *c*, at a later stage. This chronology should open an opportunity to also integrate the relative dating of various vowel shifts into this network, which still largely remains a task for the future.

Holst now finds that PIE */t'/ (= traditionally /d/), though regularly yielding zero, could be affected by the first palatalization (as in AB *tsār*- = Greek δέρω “to separate”), but not by the second: */t'/ never yields *c*. He explains this observation with the following chronological scenario: (1) In an early stage of Toch., when PIE */t'/ was still a stop, it could undergo the first palatalization. (2) Subsequently, the reflex of PIE */t'/ was lost. (3) The second palatalization took place after that, which is why */t'/ could no longer be affected. Careful handling of relative chronology and clear separation of successive diachronic stages of development and logically deriving one from another is one of the methodological strengths found here and elsewhere in this book.

As Holst says, “the glottalic theory fits the interpretation of the data in an excellent way. However, of course the decision whether the glottalic theory should be adopted for PIE should not be based solely on Tocharian and its dental problem” (p. 71). So he still remains cautious about the extent to which this evidence should be used to draw far-reaching conclusions. I personally would be a little braver and claim that Toch. makes a very strong case for the glottalic theory being preferable to the traditional reconstruction of PIE. It is almost impossible to imagine how a language could shift *d* to Ø, but *dh* to *t*, in all positions including word-initially, which is what the traditional reconstruction of PIE presupposes. Only in light of the glottalic theory does the loss of

this specific dental become a plausible sound change. When Winter discovered the development PIE */d/ (= /t'/) > Toch. Ø in 1962, the glottalic theory had not yet been proposed, which is why his idea did not gain widespread acceptance.

Tocharian offers yet another piece of evidence in support of the glottalic theory that was not seen or not claimed by Holst. He says (p. 69): “Why among the ejectives precisely the dental was affected in Tocharian in such a way [i.e. was lost], and not a different plosive, is difficult to tell.” An answer to this question will be proposed here. Departing from the glottalic interpretation of PIE, there was a system of ejectives including *t' and *k' (leaving aside the problem of the number of velars), but no *p', a position that, as is well known, seems to have been a gap within the PIE consonant system. In Toch., the first palatalization added a *ts', but while *t' was lost, *ts' was not. It probably coincided with the non-glottalized ts only much later. Thus, we have to reconstruct a stage at which Toch. possessed ts' and (one or more) k', but no *p' or *t'. This is consistent with a well-known preference of ejectives for back places of articulation. Cross-linguistically, k' is more common than t', and t' is more common than p'. (Note that voiced stops, on the contrary, have a preference for front places of articulation, which is why some languages, such as Arabic, have b and d, but no g.) Just as Toch. in the early stage just mentioned, Hausa is a language that has k' and ts', but no *t' or *p'. I believe this tendency is the reason why specifically t' (and not ts' or k') was debuccalized. Conversely, the debuccalization specifically of a dental stop provides even more support for the glottalic interpretation of early Toch. and, consequently, PIE. With a voiced series, a *g would have been more likely to be lost or to be shifted to something else. We can see that Toch., somewhat surprisingly, helps bring one of the core problems of PIE phonology close to a definite conclusion.

While several individual components surrounding the “dental problem” had been proposed by one or the other scholar before, we have to credit Holst for weighing all the evidence based on a large set of etymologies coupled with typological plausibility, sorting out the numerous misleading views, and, in this reviewer's opinion, having finally reached solid ground concerning the various issues that have plagued this problem in the past. And as I have tried to show, the conclusions have implications not only for the narrow field of Tocharian, but far beyond.

In the subsequent section 3.2, Holst examines a number of issues related to the history of Tocharian vocalism, a field which is still poorly understood in several respects. The two varieties A and B often diverge in their vowels, which is a pain and a blessing at the same time. Both inherited apophony and younger umlaut phenomena complicate the picture. Holst focuses on cases in which previous scholars have diagnosed different vocalic outputs from a single PIE source. The common denominator is that he links the different outputs to the presence or absence of PIE word accent. This is again accomplished by assembling a number of relevant cognate sets.

Holst adopts a minority view brought up by N. van Brock in 1977, who proposed that PIE *ū yields u under stress but /ə/ elsewhere in Tocharian (where /ə/ is represented as ä in A, ä or a in B, or may become zero in certain environments). It is clarified that the stress of PIE, rather than historical Tocharian, is meant; the latter has shifted in a number of cases, as far as we can tell. (Toch. stress is not marked directly in writing but can be inferred to some extent from the vowel

stressed *ǐ:	A <i>trit</i> B <i>trite</i> = Greek τρίτος “third”
unstressed *ĩ:	A <i>wās</i> B <i>wase</i> (i.e. Proto-Toch. * <i>wəsæ</i>) = Sanskrit <i>viṣá-</i> “poison”
stressed *ǔ:	A <i>-us</i> B <i>-usa</i> = Greek -υῖα (< * <i>-úsja</i>) (participle suffix)
unstressed *ũ:	A <i>rtär</i> B <i>ratre</i> = Greek ἐρυθρός “red” (i.e. Proto-Toch. * <i>ratræ</i> with <i>a</i> , reflected as <i>a</i> in B and metathesized in A.)

A pracar B *procer* = Sanskrit *bhrātā* = Greek φράτηρ “brother”
A mācar B *mācer* = Sanskrit *mātā* = Greek μήτηρ “mother”

These suggestions all seem plausible to me, though not always strictly provable, and are underpinned with a number of typological or areal parallels in each case. I would like to highlight

two subchapters here, beginning with section 4.3, which deals with “The gender distinction in the 1st pers. sg.” Tocharian A is one of the few languages worldwide that distinguish gender in the 1st sg. personal pronoun. The forms are masc. *nās*, fem. *ñuk*. Tocharian B has a common form *ñās* for both genders. The etymology of the feminine form *ñuk* has remained disputed. Building upon a suggestion made by J. Schindler in 1967, Holst derives this pronoun from a pronominal base **mene-*, otherwise only found in Slavonic and Indo-Iranian (a fact that this reviewer would like to consider as a hint as to the genetic position of Tocharian), combined with a relative pronoun **k^wis* (masc.) ~ **k^wā* (fem.). He explains this as having evolved from a former cleft sentence “(it is) me who ...”, which evolved into an emphatic (focalizing) personal pronoun and hence into the regular personal pronoun. As a support of this theory, I would like to mention that a very close parallel to what Holst reconstructs for Tocharian has been suggested for Laal, a language spoken in Chad. Like Tocharian, Laal is among the few languages that possess a gendered 1st sg. personal pronoun: *já* masc., *jí* fem. This pronoun is homonymous with the relative marker in Laal. R. Voigt concluded from this fact in *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 136 (1986), p. 185 that the personal pronoun arose from a periphrasis *v-já* / *v-jí* “(it is) me who ...”, where the original pronominal stem was a vowel (*v-*) that was lost through phonetic erosion, so that all that remained from the cleft construction was the former relative marker.

As a further typological parallel, I would like to draw attention to the Proto-Semitic personal pronouns **anāku* “I” and **anta* masc., **anti* fem. “you”, which evidently suggest a segmentation into *an-āku*, *an-ta*, *an-ti*. The element *an-* must originally have been a focalizing morpheme, as revealed by Ancient Egyptian, where *ʔn-* is still the productive focalizer before any noun phrase and is also contained in the personal pronoun *ʔnk* “I” (presumably $\approx$ *anāku*, vowels are not written in this language). In fact, there are many more languages in which some or all of the personal pronouns share a common element, which at least in some of the cases may indicate a history of grammaticalization from focalizers or cleft periphrases. I would like to briefly suggest some potential candidates with speculative segmentation, admitting that I have not conducted detailed research in each case. All of the following examples cite the pronouns “I” and “you (thou)” respectively: Sanskrit *ah-am*, *tv-am*; Welsh *f-i*, *t-i*; Finnish *mi-nä*, *si-nä*; Turkish *b-en*, *s-en*; Chechen *s-o*, *h-o*; Kannada (Dravidic) *nā-nu*, *nī-nu*; Somali *an-iga*, *ad-iga*; Margi (Chadic) *nà-yà*, *nà-gà*; Bemba (Bantu) *i-né*, *i-wé*; Zulu (Bantu) *mì-ná*, *wè-ná*.

The last grammaticalization topic to be discussed here is located on the borderline between morphology and syntax. It is Holst’s explanation for the genesis of the infinitive (section 4.6), whose marker is AB *-tsi* (= Balto-Slavic *-ti* with the first Tocharian palatalization applied, see above). This suffix attaches to the present stem of the verb in A, but to the subjunctive stem in B. The latter formation is typologically unexpected and has thus far defied explanation. Holst suggests that at an earlier stage of Tocharian, speakers employed a periphrasis built upon a finite subjunctive verb as in Albanian *ai do të shkojë* “he wants to go”, lit. “he wants that he go (subjunctive)”. Holst describes how such a construction could have given rise to a verbal form interpretable as an infinitive but retaining traces of subjunctive morphology. This can be the background of the Toch. B peculiarity, whereas Toch. A has restored its infinitive formation towards a

more common typological state. This scenario, though by necessity somewhat speculative, seems convincing to the reviewer. Holst also reminds us (p. 144) that PIE may not have possessed an infinitive at all, noting that no common infinitive morphology can be reconstructed for the proto-language. The idea that PIE may have employed a finite construction similarly to modern Balkan languages is a remarkable assumption about PIE syntax that probably no one has ever considered. It is striking how a neglected language such as Tocharian can help to establish far-reaching hypotheses about PIE syntax.

The final chapters of Holst's book address the areal and genetic position of Tocharian. Holst argues that contact with an early Turkic language probably shaped some of the language's typological traits. I will not go into the details here except for one point made by Holst that I consider perhaps the weakest. This point concerns the presence of only one plosive type, which Holst relates to the (allegedly) limited distinction between voiced and voiceless plosives in many Turkic languages (section 6.2.2). But we saw above that earlier Tocharian had two plosive types, plain and ejective. This fact does not seem to have been coordinated with his statement (p. 193) that Tocharian speakers may have arrived with a contrast of the *t-d* type when they encountered a language with only one plosive type from the Turkic stock. I would like to suggest a slightly different scenario along the following lines: The Tocharians arrived in their later homeland with a contrast of plain and ejective plosives. The areal contact languages—whether Turkic or other—did not use ejectives but other types of plosive distinctions (not necessarily an absence of plosive types altogether). This could have sufficed to break down the inherited binary contrast. I am reminded of the German dialect area, where a northern zone with a voiceless–voiced distinction of plosives and a southern zone with a fortis–lenis distinction are separated by a central contact zone in which the distinction of plosive types has been abandoned altogether.

The final question regarding the genetic position is, naturally, where Tocharian fits into the Indo-European family tree. It is an enigma that, despite two centuries of intense study of the Indo-European stock, its internal subgrouping has not yet been established. Much of the earlier research was misled by a dichotomy between *kentum* and *satem* languages, which has proven to be fruitless. More recently, various highly technical approaches using sophisticated phylogenetic methods have emerged, but so far without reaching a consensus. Note also that despite their high level of sophistication, they inherently do not distinguish between common innovations and common retentions. Holst discusses the genetic position of Tocharian within the family tree in section 7.2. While remaining cautious on this point, he leans towards grouping Tocharian with Anatolian. On this occasion, I would like to advocate for a simple classification method that assembles lexical isoglosses of the basic vocabulary and turns them into fragments of an unrooted family tree. Consider the following representations of the verb “to lie down” (incidentally, this Tocharian verb is also the subject of a specific discussion in section 5.3.1 of Holst's book):

“to lie down”: Tocharian AB *lyäk-* = Gothic *ligan* ≠ Hittite *ki-* = Sanskrit *śay-*

We find one lexical representation in Tocharian + Gothic and another one in Hittite + Sanskrit. This must not be taken as straightforward evidence of two genetic subgroups { Tocharian +

Gothic } and { Hittite + Sanskrit }, because one of the two terms is most probably a shared retention, which is not probative in terms of genetic subgrouping. However, (at least) one term should represent a shared innovation, if we assume (1) that borrowing is improbable for such a basic vocabulary item and (2) that two independent lexical replacements are unlikely to produce the same outcome. This isogloss does not allow us to decide on which side to localize the innovation or the retention, but we can conclude that at least this single lexeme is a piece of counterevidence against grouping Tocharian and Hittite together (or if we do, we must argue for either borrowing or independent but converging lexical replacements). I only use this single lexeme to exemplify the method. Many more examples would, of course, be needed to build a robust argument. I took this approach in a paper titled “The subgrouping of East Chadic” in *Folia Orientalia* 55 (2018), to which any interested readers are referred.

As we have seen, Holst’s book brings up a number of solutions of long-standing riddles concerning Tocharian, based either on arguments forwarded by him for the first time, or dug out from footnotes or from highly specialized literature where they are likely to go unnoticed by general linguists. This accessible and well-written book, which is full of original ideas, makes the reader want to learn more about the underrated language that is Tocharian. The reader is made aware of how many basic questions about this language have still remained unsettled. The argument repeatedly extends into the realm of other languages, offering remarkable typological insights and suggestions of general interest. Holst’s insightful piece of work can be recommended not only to every Tocharian specialist, but also to anyone who wants to broaden their knowledge of Indo-European in general, or even as a source of inspiration for linguists working on reconstruction or typology in quite different parts of the world.