Mother Tongue

Mother Tongue

(XXIII) 2021

In Memory of Merritt Ruhlen (1944–2021)

Edited by

Pierre J. Bancel

John D. Bengtson

GORGIAS PRESS

2022

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

www.gorgiaspress.com

2022

ISBN 978-1-4632-4441-5 **ISSN 1087-0326**

Library of Congress Cataloging-in-Publication Data

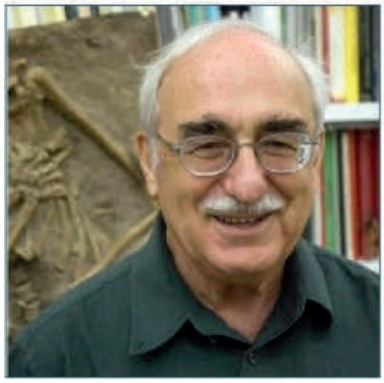
A Cataloging-in-Publication Record is available at the Library of Congress.

Printed in the United States of America

MEMORIALS

Ofer Bar-Yosef (1937–2020)

Ofer Bar-Yosef, Israeli archaeologist, was widely recognized as the world's leading authority on the transition from the Paleolithic into the early Neolithic Period. Beginning as Professor of Prehistoric Archaeology at Hebrew University in Jerusalem, in 1988 he accepted a similar post at Harvard University, and also served as Curator of Palaeolithic Archaeology at the Peabody Museum of Archaeology and Ethnology.



Ofer was a longtime member of ASLIP, deeply interested in the connections between archaeological artifacts and human language. At the honoree's request Ofer contributed an article, "Can Paleolithic stone artifacts serve as evidence for prehistoric language?," to ASLIP founder Harold Fleming's *Festschrift* (*In Hot Pursuit of Language in Prehistory*, 2008).

https://ioa.ucla.edu/content/ofer-bar-yosef-1937-2020-celebration-life

Tamaz (Thomas) Valerianis dze Gamkrelidze (1929–2021)

Thomas Gamkrelidze was a Georgian linguist, Orientalist, Hittitologist, and President (2005–2013) of the Georgian Academy of Sciences. He was best known in the West for his book (with V.V. Ivanov), *Indo-European and the Indo-Europeans: A reconstruction and historical analysis of a proto-language and a proto-culture* (1994), a translation of *Indoevropjskij jazyk i indoevropejcy: Rekonstrukcija i istoriko-tipologieskij analiz prajazyka i protokultury*.

https://titus.uni-frankfurt.de/curric/necrolog.htm

http://science.org.ge/old/cevrebi/Prof%20Thomas%20V%20Gamkrelidze.html

Wolfgang Schulze (1953–2020)

Prof. Dr. Wolfgang Schulze had been Professor of General Linguistics at Ludwig Maximilian University in Munich since 1992. Professor Schulze was probably best known as a specialist in East Caucasian (= Nakh-Daghestanian) languages. Apart from Avar, Lak, and Dargi, for which sparse documents are available from the 16th century onwards, the only East Caucasian language with older written attestation is Udi, whose immediate ancestor "Caucasian Albanian" is documented in two palimpsests recently found in the Katharine monastery on Mt. Sinai, at roughly 600 CE. Schulze was one of only a few linguists who extensively studied the development of Caucasian Albanian into Udi, now spoken in a few villages in Azerbaijan and Georgia. Schulze was always skeptical of the claim, by Trubetzkoy, Dumézil, Starostin, Nikolaev, Chirikba and others, that East Caucasian belonged with West Caucasian in a higher-order family, North Caucasian. These doubts were expressed in the article he wrote for *Mother Tongue* XXI (2016), "Comments on V.A. Chirikba's paper 'From North to North West'."

https://titus.uni-frankfurt.de/curric/necrolog.htm

THE ASSOCIATION FOR THE STUDY OF LANGUAGE IN PREHISTORY, INC.

Minutes of Meeting of the Board of Directors
December 4, 2021

The following members of the Board of Directors met on December 4, 2021 by means of a Zoom video conference: *Nicholas Davidson, John Robert Gardner, Michael Puett, Václav Blažek.*

They were joined by the following officers: *Peter Norquest, President; Michael Witzel, Vice President; John D. Bengtson, Vice President; Gregory Haynes, Secretary / Clerk;Alexander M. Kim, Treasurer.*

They were joined by the following additional individuals: *Pierre Bancel; John Saul; Roger Blench; Matthew R. Hillery.*

And the following took place:

(1) Quorum. A quorum of the Directors of The Association for the Study of Language in Prehistory, Inc. ("ASLIP") was recognized. Allan R. Bomhard resigned as a Director prior to the meeting, reducing the number of living Directors to seven. Three Directors, Murray Denofsky, Philip Lieberman and Ronald Christensen, were unable to attend.

(2) Call to Order: Peter Norquest, as President, called the meeting to order at 2:00 pm.

(3) Approval of Prior Minutes: On a motion by John Robert Gardner, seconded by Nicholas Davidson, the Board approved the minutes of the prior meeting with no one in opposition.

(4) Change to Number of Directors and Election of Directors: The Directors present discussed the appropriate number of Directors, which are set at twelve by the by-laws, but that number may be changed by the Board. Some individuals who previously had served as Director were deceased or had lost touch with ASLIP. After extensive discussion, on a motion by John Robert Gardner, seconded by John Bengtson and Nicholas Davidson, and without further discussion or objection, the Directors voted to decrease the number of Directors to six and to elect the following individuals to serve as Directors: *John Bengtson; Václav Blažek; Nicholas Davidson; John Robert Gardner; Michael Puett; Michael Witzel*.

(5) Election of Officers: On a motion by John Robert Gardner, seconded by Michael Witzel, the following individuals were elected as officers of ASLIP: *Peter Norquest, President; Michael Witzel, Vice President; John D. Bengtson, Vice President; Matthew R. Hillery, Secretary/Clerk; Alexander Mee-Woong Kim, Treasurer.*

(6) Council of Fellows: A discussion of the Council of Fellows then began. John Bengtson confirmed that membership on the Council is an honorary position with no specific duties and is meant as an appreciation for the holder. On a motion by John Bengtson, seconded by Nicholas Davidson, and with no opposition, the following persons were appointed to the Council of Fellows without further discussion: *Christopher Ehret; Irén Hegedűs; Roger Blench*.

(7) Financial Statements and Bank Balance: John Bengtson discussed the bank balance in his role as former Treasurer and keeper of ASLIP's bank account. The most recent balance was $28,433.95. This amount includes a recent grant from The Boston Foundation in the amount of $15,000. Nicholas Davidson stated that this amount should be viewed as a nest egg and held onto for a while. John Robert Gardner made a motion to accept the financial statement, which was seconded by Michael Witzel and then adopted without further discussion and with no opposition.

(8) Publishing Options for *Mother Tongue*: There next was extensive discussion of the manner in which the *Mother Tongue* journal is published. There was a discussion about whether or not to adopt Gorgias Press as the Journal's publisher that was led by Nicholas Davidson. The Press focuses on Syriac and Aramaic studies but has branched out, and publishes books that are available on Amazon. John Bengtson introduced a motion to adopt Gorgias Press as the publisher of *Mother Tongue*, which was then seconded by Michael Witzel and adopted by the Directors with no further discussion or opposition.

(9) Creation of Language in Prehistory Series with Gorgias Press: Nicholas Davidson then led a discussion of the creation of a language series to be published by Gorgias Press. Michael Witzel had negotiated with Gorgias for several years to see if there would be interest in such a series. The individuals in attendance discussed extensively whether such a series should be created. It was proposed that the first publication could be a book that John Bengtson previously had published semi-privately, that Václav Blažek may publish a work as part of the series, and that the series also may include a dictionary

created by Merritt Ruhlen. Eventually, on a motion by Nicholas Davidson, seconded by John Bengtson, with no discussion and no opposition, the Directors approved the creation of a Language in Prehistory Series with Gorgias Press.

(10) Management of Language in Prehistory Series with Gorgias Press: In connection with the foregoing vote, the Directors approved the following appointments with respect to the Language in Prehistory Series:

Michael Witzel, Founding Editor; Nicholas Davidson, Editor; John Bengtson, Senior Advisor.

(11) Distribution of *Mother Tongue* and Membership Structure: There was then an extensive discussion of several related topics. The first topic dealt with updating and consolidating multiple contradictory membership lists. The membership lists had been kept by the Secretary in the past but had become out of date. Alexander Kim was attempting to merge and consolidate several lists. The next topic dealt with the payment of dues. Dues paying members typically were entitled to receive a print copy of *Mother Tongue*. Some members paid yearly, some pre-paid for several years, and some had purchased lifetime memberships. The list of who-had-paid-what was lost. The general feeling was that existing members should be grandfathered in for the current year, an explanation should be given to everyone of the current situation, and then a request for dues should be made next year (subject to individuals who could credibly assert that they already had paid). Finally, there was a discussion of whether *Mother Tongue* should continue in print at all or whether it should be entirely digital. Nicholas Davidson suggested the Board should wait to hear what capabilities Gorgias Press may have and should defer to the officers on the issue. There was no formal board vote.

(12) Adjournment: The meeting adjourned at 4:15 pm. After the meeting adjourned, several members stayed to discuss forthcoming research and other issues of interest to ASLIP members on an informal basis.

Matthew R. Hillery
Secretary

NOTE: On 6 January, 2022, Nicholas Davidson submitted a letter of resignation from all ASLIP offices and duties. On 12 January, 2022, Matthew R. Hillery submitted a letter of resignation from the office of Secretary. As of this printing, ASLIP officers and Board have not elected or appointed successors to the vacant offices. **Ed.**

IN MEMORIAM

Merritt Ruhlen (1944–2021)

Frank Merritt Ruhlen was born on May 10, 1944. His father, Frank Merritt Ruhlen (1909-1997), was an administrative law judge for the Civil Aeronautics Board, and the family lived in Virginia, near the national capital. Merritt (junior) studied at Rice University, the University of Paris, the University of Illinois, and the University of Bucharest as a Fulbright fellow, and received his PhD in 1973 from Stanford University with a dissertation on the generative analysis of Romanian morphology. Subsequently, Merritt worked for several years as a research assistant on the Stanford Universals Project, directed by Joseph Greenberg and Charles Ferguson.

Beginning in 1994 Merritt was a lecturer in Anthropological Sciences and Human Biology at Stanford University. As such he was well positioned to combine historical linguistics with the other anthropological disciplines, especially human genetics (molecular biology). In 2001 Merritt Ruhlen, together with Murray Gell-Mann (1929–2019; Nobel Prize in Physics, 1969) of the Santa Fe Institute and Sergei Anatolyevich Starostin (1953-2005) of the Russian State University of the Humanities, co-founded the Evolution of Human Languages Project, based at the Santa Fe Institute and supported by a generous grant from the John D. & Catherine T. MacArthur Foundation, and by the Institute itself. Since 2005 Merritt had been on the advisory board of the Genographic Project and held an appointment as a visiting professor at the City University of Hong Kong. He had also been a Correspondant of the Museum National d'Histoire Naturelle in Paris, and served as a U.S. State Department interpreter in French and Romanian. Merritt knew and worked with Joseph Greenberg for three-and-a-half decades and became the principal advocate and defender of Greenberg's methods of language classification.

Merritt Ruhlen was the author of several books, including *A Guide to the Languages of the World* (1975), *A Guide to the World's Languages, Vol. 1: Classification* (1987), *The Origin of Language: Tracing the Evolution of the Mother Tongue* (1994: John Wiley & Sons), *On the Origin of Languages: Studies in Linguistic Taxonomy* (1994: Stanford U.P.), and numerous articles in journals, books, and encyclopedias. *The Origin of Language* (Wiley) was translated by Pierre Bancel and has been published in France in two editions (1997 and 2007) as *L'origine des langues*. The same book was published in Portuguese (1998), Italian (2001), and Turkish (2006) versions.

In 2011 an article, "The origin and evolution of word order," co-authored by Murray Gell-Mann and Merritt Ruhlen was published in *PNAS*. They concluded that the "distribution of word order types in the world's languages, interpreted in terms of the putative phylogenetic tree of human

languages, strongly supports the hypothesis that the original word order in the ancestral language was SOV."

Since studying at Stanford Merritt lived in Palo Alto, Calif., with his wife Anca. They had twin sons, Ricky and Johnny. Merritt was a hands-on father and cared for the boys at home while he worked on his books and articles, and Anca was at work at Stanford. Anca informs us that Merritt raised his sons to be polite and respectful, and their teachers praised them as being kind to their schoolmates. From their early years Merritt kept a detailed notebook of his sons' speech development. Merritt loved basketball and coached the sport at the YMCA, encouraging his sons to play. He coached for ten years and taped all the games with a video camera. Besides basketball Merritt played tennis for many years with the same guys from his student days, and they remained friends for life. Merritt also loved to play guitar and listen to classical guitar records, and amassed a large collection of LPs in that genre.

After a long illness Merritt succumbed on January 29, 2021. Since everything in recent times has been delayed by the Covid pandemic, a memorial gathering was finally held at his Palo Alto home on a lovely Sunday afternoon, 22 August 2021. It was attended by many of Merritt's friends, including some of his tennis buddies, and some who could not attend sent their tributes and memories.

Homages and Reminiscences from Colleagues and Friends

So very sad, indeed! We've been such good friends from 1994 on. Twenty-six years.

I admired his work so much, and admire it now still more as I am in the middle of the answer to Zamponi's paper and have been checking his proposed reflexes for **na* 'I' and **ma* ~ **mi* 'thou' in Amerind languages.[1] Most are good, of course. Alain and I planned to submit the paper to him and pray him to cosign it with us. Now it will definitely be dedicated to his memory.

He has changed my life forever, together with [John Bengtson], Hal [Fleming], and Michael [Witzel], and of course Joe [Greenberg] in the background. A fantastic bunch of trailblazers, you all.

Pierre Bancel

Association d'études linguistiques et anthropologiques préhistoriques (Paris, France)

Translator of Merritt's book: *L'origine des langues*: Éditions Belin 1997; Gallimard 2007

Merritt and I met in March 1986 at the Symposium on Genetic Classification of Languages (at Rice University, Houston, Texas), directed by Sydney Lamb, James Copeland, and Douglas Mitchell. Merritt was then putting the finishing touches on his trailblazing book, *A Guide to the World's Languages. Volume 1: Classification*, published the next year by Stanford University Press. At the conference we discovered that we had independently been working on global etymologies, and we resolved to co-author a paper listing some of the most promising sets of these worldwide lexemes.

[1] Zamponi, Raoul. 2017. First-person *n* and second-person *m* in Native America: A fresh look. *Italian Journal of Linguistics* 29.2: 189–230. See the response by Bancel and Bengtson in this volume.

At first this article was projected to appear in a book edited by a prominent Nostratic scholar, which never materialized. Next it was submitted to the journal *Language*, which rejected it. Finally it was published in Merritt's 1994 book, *On the Origin of Languages*. Throughout the 1980s and 1990s Merritt and I exchanged dozens if not hundreds of snailmail letters, and later numerous emails, about genetic classification of languages as well as global etymologies. As an unknown novice with no academic position, I found this relationship with Merritt to be highly supportive and encouraging. I remember those days as some of the most satisfying times of my life as an investigator in anthropological linguistics. Merritt was a bold visionary and I profoundly miss him as a scholar and a man.

John D. Bengtson
Association for the Study of Language in Prehistory
Evolution of Human Language Project

I met Merritt through Joe Greenberg, who was my PhD advisor at Stanford when I studied there in the early 1980s. Although my primary research interests were and are linguistic typology and universals, I spent many hours listening to Joe talk about ancient language families and Merritt's work with Joe on many of these families. I used Merritt's work from the beginning, in particular his great 1987 book surveying language classification—with an outstanding overview of the history and methodology of language classification as well as the first really comprehensive classification of all of the world's languages. I even used Merritt's 1975 book of inventories of sounds of languages to examine patterns in consonant systems.

Merritt's classification book came out the same year as Joe's book on the classification of Native languages of the Americas. This marked the crescendo of the firestorm of criticism within the field of linguistics towards Joe's and Merritt's work in genetic linguistics (though their work was welcomed by geneticists and anthropologists, not to mention myself and a few other linguists). I had graduated from Stanford and left the area, teaching first in Michigan and then in the UK. I would meet Joe whenever I returned to Stanford, and occasionally with Merritt. I gave Merritt feedback on some of his papers, and read everything Merritt (and Joe) published on genetic linguistics, including Merritt's 1994 collection of essays and his popular book on language classification—which I highly recommend to non-linguists wishing to understand what the real issues are. I used Merritt's examples from that book, and the maps he created of language families recent and ancient, in my Language Change class.

When Joe was diagnosed with cancer in 2000, I hopped on a plane the next day. I told Joe I would collect his essays on genetic linguistics and get them published. Merritt was a great help in tracking down some of the harder to find articles by Joe; he also played an important role in organizing a conference honoring Joe which, sadly, took place only after Joe had passed away. Merritt also told me excitedly at the time about his work on an ancient language family ranging from Africa to the Pacific. Of course, Merritt (and John [Bengtson]) had already established global etymologies demonstrating the genetic unity of all known contemporary human languages—"Proto-

World". But it is important for linguistics and for human prehistory to establish intermediate ancient families, which tell us more about human migrations. I also had the honor of reviewing the article that Merritt wrote with Murray Gell-Mann on the word order of Proto-World, which was published in the *Proceedings of the National Academy of Sciences* in 2011.

After Joe and then his wife Selma passed away, I rarely went back to Stanford. So I didn't see Merritt much after that, though we corresponded until around the onset of Merritt's illness. It was a shock for me to learn from Anca that Merritt had passed away.

Merritt was an intellectually courageous scholar, carrying forward the work on deep genetic classification even as the field turned its back on it. Merritt understood clearly the methodology; he responded thoroughly to the critics and kept advancing the field of genetic linguistics. Merritt's last major projects were the *Amerind Etymological Dictionary*, which greatly expands and reorganizes the lexical evidence supporting the Amerind family that Joe had published in 1987; and a much expanded version of his database of sound inventories. Merritt and I shared a tremendous respect for Joe's work on genetic linguistics, and Merritt was a true successor to Joe in that respect. Merritt's impact on linguistics is for the long term; but it will come.

William Croft
Professor of Linguistics, Emeritus
University of New Mexico

RIP, Merritt...We were friendly at Stanford, I audited his class and had a few memorable lunches with him. He was a highly intelligent and courteous man who was a pleasure to be around. A faithful follower of Joseph Greenberg, he had a natural-scientific view of language as a human faculty mixed with a romantic passion for the unbounded diversity of world languages. He had lofty goals and the unswerving boldness of thinking that he imparted in his students. He supported their ideas even if they diverged from his dreaming of a time when freedom of thinking and academic rigor would always go hand in hand.

German Dziebel
Resource Interactive

Merritt was like an eagle who soared high and taught us all to look at the global perspective of language and genes and culture. He was so unique and brilliant in his perspective on life – I learned so much from him and his marvelous conversations. I will miss you, buddy.

Roy J. King
Associate Professor of Psychiatry and Behavioral Sciences, Emeritus
Stanford University

Quelle tristesse, de perdre trop tôt un tel bonhomme, aussi pertinent sur le plan scientifique que chaleureux au niveau amical!

André Langaney, population geneticist

[Langaney was a great help in getting Merritt's book published in French, and whose preface he authored. He used to be lab head at the Musée de l'Homme, Paris, and was a teacher at Université de Genève in Switzerland.]

Today is a very sad day, as Merritt is not with us anymore. I have been lucky that I could have met this kind, elegant and brilliant man. This is how I came along into these linguistic journey with you folks. I also remember well the fascinating meetings and dinners with him in California in 2006.[2] I loved the way he was able to make his ideas accessible to us and keep his mind open to new avenues. We will miss him dearly and cherish his memory.

Sadness, sadness.

Alain Matthey de l'Etang

Association d'études linguistiques et anthropologiques préhistoriques (Paris, France)

Since I'm not able to be with you in person, I've written up some thoughts for someone else to share on my behalf.

I have fond memories of Merritt. I always found him so approachable. He was willing to help me understand any linguistic concept, and found such joy in any parallels between linguistic and genetic history.

My poster of the Language Families of the World has been hung near my desk at 23andMe whenever I've had some wall space. I was so proud to know Merritt as I showed it off. Whenever my colleagues have a question about language classification or history I refer them to Merritt's work.

Finally, I loved our foursome lunches at California Cafe – Merritt, Peter Underhill, Roy King and myself, once a year for a few years, usually in the spring. I miss those moments, and I miss Merritt.

I send my sympathy and loving thoughts to Anca and all of Merritt's family.

Joanna Mountain

Senior Director of Research at 23andMe

I was never as close to Merritt as my father, Sergei Starostin, to whom he always was a wonderful friend, helping him – and some of his Russian colleagues – both with his research and with helping

[2] Editor P.J. Bancel offered some minor corrections. Alain and other colleagues presented papers at the 105th annual meeting of the American Anthropological Association (2006) in San Jose at the Per Hage Memorial Workshop, in presence of Merritt who was very warmly greeted by the audience.

him adapt to the weird magical Western world after the fall of the Iron Curtain. But I do remember very well how, after Sergei's own untimely death in 2005, Merritt was the first to take me under his wing, helping me gather enough courage and find enough confidence to replace my father at a really huge interdisciplinary conference in Genoa. I know that, while Merritt could be very hard on those people whom he considered seriously flawed – often deservedly so – for most of my colleagues (Russian historical linguists) he always expressed respect and admiration, which actually felt like something to be proud about. Myself, I have always admired Merritt's integrity and determination in defending his points of view, never being afraid to go against the majority if he felt this was the right thing to do, even if it cost him rather dearly in terms of his own professional career. We may not have always seen eye to eye on every practical or methodological issue of our research (and neither did my father), had our share of disagreements and debates, but these were always conducted in the spirit of friendship and cooperation, as we were all striving to reach the same goals, occasionally taking different paths, and nothing in science can be healthier or more productive.

Sadly, with Merritt's passing, the third and last still living founder of the Evolution of Human Languages program (in 2001) is now gone; I am not at all a religious person, but it is still a very consoling thought to imagine that, perhaps, now Sergei Starostin, Murray Gell-Mann, and Merritt Ruhlen can all get together in Heaven and resume their friendly quarrelling over the best way to get to the origins of human language. All of these great people had their flaws and virtues, but ultimately, all of them turn out to be irreplaceable in a rapidly changing world where people seem less and less willing to be truly daring and risk-taking in any of the things they do – which is something that Sergei, Murray, and Merritt most definitely had in common. This is why I hope very much that Merritt's wonderful books will continue to be read for quite some time, and hopefully, every once in a while will inspire some young and daring researcher to boldly go where only Sergei and Merritt have gone before.

Merritt Ruhlen - may he never be forgotten.

George Starostin
Russian State University for the Humanities, Moscow
Evolution of Human Language Project, Santa Fe Institute, Santa Fe, NM

First and foremost, Merritt was a good human and a good friend. I will miss him.

Unlike most of us whose impact will quickly fade as do ripples from a sinking stone, Merritt’s innovative work will endure in high relief.

Peter Underhill
Senior Research Scientist (Ret.)
Department of Genetics, Stanford University School of Medicine

---

Merritt's *A Guide to the World's Languages* was a revelation to me when it came out, and I read it over and over again. I was even moved to write him a (truly idiotic) letter, to which his reply was sane and sensible. And it was he, in 1994, who put me in touch with Hal Fleming. But Merritt was always helpful and supportive, and I owe him a lot. It was a pleasure to finally meet up with him in Santa Fe.

Paul Whitehouse
Evolution of Human Language Project, Santa Fe, NM

Photographs

Merritt with evolutionary biologists Nicole Creanza and Sohini Ramachandran | September 2011

Merritt with his sons | Father's Day 2012

More Obituaries and Tributes on the Internet

Santa Fe Institute: https://www.santafe.edu/news-center/news/remembering-merritt-ruhlen

Thesaurus Indogermanischer Text- und Sprachmaterialien (TITUS) | University of Frankfurt: https://titus.uni-frankfurt.de/personal/galeria/ruhlen_scholarly_obit.pdf

Linguist List: https://linguistlist.org/issues/32.1860/

Stanford Magazine: https://stanfordmag.org/contents/obituaries-september-2021

Palo Alto Online: https://www.paloaltoonline.com/obituaries/memorials/frank-merritt-ruhlen?o=6809

Lectures by Merritt Ruhlen

HS Talks: Linguistic Evolution: https://hstalks.com/expert/282/dr-merritt-ruhlen/

Merritt Ruhlen's Legacy

Ruhlen was the author of several books, including *A Guide to the Languages of the World* (1975), *A Guide to the World's Languages,* Vol. 1: *Classification* (1987; second edition 1991), *The Origin of Language: Tracing the Evolution of the Mother Tongue* (1994a), *On the Origin of Languages: Studies in Linguistic Taxonomy* (1994b), and numerous articles in journals, books, and encyclopedias. *The Origin of Language* was translated into French by Pierre Bancel and has been published as *L'Origine des langues – Sur les traces de la langue mère* (1997a and 2007a); in Portuguese as *A Origem da Linguagem: Reconstituindo a Evolução da Lingua Mãe* (1998c); in Italian as *L'origine delle lingue* (2001c); and in Turkish as *Dilin Kökeni: Ana Dilin Evriminin İzinde* (2006).

Ruhlen's *Guide* (1987) came to be regarded as a classic, discussing in depth the genetic classification, and alternative taxonomies, of some 5,000 languages across the globe. It elucidated in detail the principles of genetic classification.

In many of his books and articles Ruhlen time after time revisited and rebutted what he considered misconceptions that had infected the field of historical linguistics. Among these were:

- A basic misunderstanding of linguistic taxonomy: that it was the result of identification of sound correspondences and the reconstruction of a proto-language. This is because most historical linguists work within the confines of a long-established language family,

or a subgroup thereof, and have never had the opportunity, or inclination, to execute an original language classification themselves. To do so would require the use of multilateral comparison (see next point).

- The idea that multilateral comparison (less aptly called "mass comparison") is a novel technique invented by Greenberg, and "totally different from the traditional methods." In fact MC is really the most traditional method of taxonomy, used consciously or unconsciously by Mahmud al-Kašghari, Jan Amos Komenský, Gaston Cœurdoux, Philip Johan von Strahlenberg, and others in their early identifications of linguistic taxa, long before any discussions of reconstruction and sound correspondences.
- The idea that there is a "temporal limit," variously delineated as between 6000 years (coincidentally the approximate age of Indo-European), or slightly earlier, 7000–8000 years, beyond which genetic relationships cannot be recovered. Ruhlen, Greenberg, and their colleagues such as Sergei Starostin and Harold Fleming argued that there is no factual basis for such a temporal limit. One strong argument is the existence of the Afroasiatic (macro) family, universally accepted, which has been glottochronologically dated by George Starostin at approximately 16.78 kya (Kassian 2010: 424, figure 8), which is consistent with Christopher Ehret's (2015: 379) view that "the first two divergences in the [Afroasiatic] family fell in the rough time span of 22,000–15,000 bce – coincident with the Last Glacial Maximum." This is clearly double the supposed maximum time depth of 8000 years.
- The idea that massive amounts of evidence are required to "prove" a classification, a concept unheard of before the twentieth century. "Instead of recognizing the simple basis of genetic classification, and thus linguistic relationships, twentieth century historical linguists put forth increasingly rigorous demands, generally involving reconstruction with regular sound correspondences, before genetic relationships will be acknowledged, demands in fact so rigorous that they can never be satisfied" (Ruhlen 2005).
- The idea that the existence of isolated errors, in the midst of large amounts of positive evidence, allows a critic to dismiss the entire proposal. "Historians and historical linguists – not to mention paleontologists working from handfuls of bashed fossils – use whatever material is available; they do not demand that the evidence be complete or immaculate" (1994f).

On the Origin of Languages (Stanford) was essentially a collection of Ruhlen's scholarly research essays, expressing in academic style the principles of genetic classification, once again, and defending or expanding on classifications that had been attacked or ignored. Possibly the best example of this is the chapter "Is Algonquian Amerind?," which starts with a discussion of Edward Sapir's (1913) discovery of a relationship between the Algonquian language family and the relatively isolated Wiyot and Yurok of Northern California, now universally recognized as the Algic language family. But in 1914 Sapir's hypothesis was attacked by the foremost Algonquianist of the time, Truman Michelson, on the basis of alleged "errors (real or imagined), an allusion to accidental resemblances and borrowings, a willful disregard of the most convincing positive

points, a dose of taxonomic nonsense, and ... an appeal to authority." These false criteria were characterized by Ruhlen as "Diffusionist." Ruhlen noted that "whatever talents Michelson may have had as an Algonquianist – and these apparently were considerable – he had little understanding of basic taxonomic principles, and the vast majority of his objections to Sapir's hypothesis were irrelevant." Ruhlen went on to demonstrate that the same "silly" methods used by Michelson were now being applied by specialists in Native American languages to criticize Joseph Greenberg's book *Language in the Americas* (1987) and its hypothesis of the Amerind language family encompassing most of the languages of North and South America. Notwithstanding these detractors, Ruhlen predicted that "in time *Language in the Americas* will be hailed as a monumental achievement." Since then we have witnessed a substantial confirmation of a population corresponding to the postulated speakers of Proto-Amerind by geneticists (including the very small size of the founding population: *e.g*. Kitchen, *et al*. 2008: "a founder group with an effective population size of ≈1,000–5,400 individuals"). David Reich entitled one section (pp. 173–76) of the American chapter of his 2018 book *Who We Are and How We Got Here* "The Genomic Rehabilitation of Joseph Greenberg," in which he states that "[t]he clusters of populations that [Greenberg] predicted to be most closely related based on language were in fact verified by the genetic patterns in populations for which data are available" (p.174).

Other chapters in *On the Origin of Languages* sought to buttress other more-or-less embattled taxonomic proposals: Sapir's Na-Dene, Starostin's Yeniseian and Sino-Caucasian, and Greenberg's Khoisan. Five of the chapters explore or expand on Greenberg's Amerind, including one that proposed that the closest relative of Amerind was his "Eurasiatic" (a large subgroup of the Nostratic megaphylum, largely accepted by the Russian Nostraticists and the American scholar Allan Bomhard). This concept, apparently developed by Greenberg and Ruhlen since the 1980s (Ruhlen 1989; Greenberg & Ruhlen 1992b; Ruhlen 1994b: 207–41) prefigures the discovery decades later by geneticists that there was an "Ancient North Eurasian" ghost population that eventually contributed to the Indo-Europeans and other West Eurasians, as well as to Native Americans, and "explains why Europeans are genetically closer to Native Americans than they are to East Asians" (Reich 2018: 82).

In the last chapter of *On the Origin*, "Global Etymologies" (co-authored with John Bengtson), Ruhlen dared to venture into even deeper taxonomic waters, with a list of twenty-seven etymologies meant to demonstrate that "the case for monogenesis of extant (and attested extinct) languages is quite strong." As a preface to the lexical sets the authors preemptively address the objections that would arise: borrowing, convergence, the alleged temporal limits of the comparative method, the absence of reconstruction and sound correspondences, unconvincing semantic changes, and whether errors in the data can override the hypothesis in general. Predictably, many critiques of this chapter employed the time-worn Diffusionist arguments, proving that the reviewers had not read, or at least not understood, the prefatory arguments. For example, one frequently encounters statements like "Bengtson and Ruhlen have tried to reconstruct the Proto-Human language," despite the authors' explicit denial that they were reconstructing anything.

Ruhlen's other 1994 book, *The Origin of Language*, covers many of the same themes as the Stanford book, but in a decidedly more popular style, aimed at educated laymen rather than linguistic specialists. It explains the methods of language classification in the simplest possible terms, showing, for example, how languages can be roughly classified on the basis of a dozen or so basic words. There is also stress on the implications of genetic classification for historical population movements: "out of Africa," the Dene-Caucasians, Native Americans, the Bantu, Austronesian and Indo-European expansions, and the Algonquian homeland. Of course, some of these issues are vigorously discussed to this day.

Ruhlen (1994a, chapter 6; 2000a) offered a significant proposal that kinship terms like papa 'father', mama 'mother' and kaka 'older brother' or 'uncle', etc., found in numerous "unrelated" languages throughout the world, did not evolve independently, and repeatedly, due to the babbling of children acquiring new consonants, but are actually inheritances from the original Proto-Sapiens language. Building on this the French scholars Pierre Bancel and Alain Matthey de l'Etang have conducted a thorough investigation of kinship terminology "that essentially confirms my hypothesis that it is indeed common origin—and not convergence—that is re[s]ponsible for the global distribution of these kinship terms" (Ruhlen, this volume: 11).

Some other valuable scholarly contributions by Ruhlen, apart from the ones already enumerated above, are his in-depth case studies, of Amerind T'A'NA 'sibling' and MALIQ'A 'throat' (in 1994a), and QETS' 'left (hand)' (in 1995c). These are fascinating examples of particular historical changes, similar to Greenberg's brief discussion of KWA 'eye, round object' (Greenberg 1987: 298-99, grammatical etymology #40), which can only be done in the context of language families with great time depth. Some Amerind etymologies (Greenberg & Ruhlen 2007b) feature a convincing structure of three sequential consonants, *e.g.* Amerind SUNIK ~ SINIK 'nose', attested in the Almosan, Penutian, Chibchan, Andean, and Equatorial subgroups, and C'E-PUQ' 'elbow, knee', in Keresiouan, Hokan, Central Amerind, Paezan, and Macro-Ge. These two etymologies are not found in Greenberg's 1987 book, and must have been assembled by Ruhlen thanks to meticulous searches of Greenberg's Amerind notebooks. That all of these triconsonantal resemblances (see also MALIQ'A 'throat'), attested throughout the Americas, are purely accidental seems highly unlikely.

In 1998 Ruhlen published an article, "The Origin of the Na-Dene," in which he called attention to a number of resemblances between the Yeniseian language family in Siberia and the Na-Dene family in North America. Other scholars before this had implied a relationship indirectly, *i.e.* between Starostin's Sino-Caucasian (including Yeniseian) and Na-Dene (Nikolayev 1991), but this was the first direct comparison of the two families, and it included a proposal of regular sound correspondences. Edward Vajda seized upon this idea and in the following years has greatly expanded on it.

In 2001 Nobel physicist Murray Gell-Mann (1929–2019) of the Santa Fe Institute received a generous multi-million dollar grant from the John D. & Catherine T. MacArthur Foundation, to be used for a project of his choosing. Gell-Mann, who had studied and cherished historical linguistics since his youth, decided to dedicate the grant to a project that would help to bring historical linguistics closer to its full potential of the genetic classification of the world's languages. Quite

naturally, Gell-Mann approached Merritt Ruhlen, together with Sergei Anatolyevich Starostin (1953–2005) of the Russian State University of the Humanities, a leader of the "Moscow School," and together they co-founded the Evolution of Human Languages Project, based at the Santa Fe Institute. For the first few years focused workshops were held at Santa Fe, with topics like "EHL Working Group Meeting" (February 2002), "The Khoisan Problem" (August 2002), "The Current Status of Sino-Caucasian" (October 2002) and "Linguistic Databases and Linguistic Taxonomy" (January 2003). EHL funding permitted transportation and lodging of scholars from around the world to the charming New Mexico city. Besides many eminent linguists, specialists in other anthropological fields, like archaeology and genetics, were brought into the discussions. Eventually the project sponsored linguistic databases, largely thanks to the technical skills of Starostin: see the Tower of Babel (https://starling.rinet.ru/main.html) and EHL (http://ehl.santafe.edu/) websites. Among other things, Ruhlen contributed **A global linguistic database** (phonetic and syntactic information on more than 5000 languages, similar to his 1975 book) and **Interactive maps** showing the distribution of dozens of language families and macrofamilies.

In 2011 Gell-Mann and Ruhlen collaborated on an article on "The origin and evolution of word order," which was published in *PNAS*. They concluded that the "distribution of word order types in the world's languages, interpreted in terms of the putative phylogenetic tree of human languages, strongly supports the hypothesis that the original word order in the ancestral language was SOV."

We have yet to see how Merritt Ruhlen's legacy plays out. But his career reminds us of what Dell Hymes (1971: 267) wrote about Morris Swadesh, of an earlier generation:

> Working at frontiers of knowledge, he could not always be sure of details, and sometimes went too far and fast for many of his colleagues to follow. He died before all could be woven together. Yet his explorations and the new dimensions he discovered have permanently extended our knowledge and conception of the contribution of linguistics to the understanding of the human past.

REQUIESCAT IN PACE

* * *

MERRITT RUHLEN: SELECT BIBLIOGRAPHY[3]

1975. *A Guide to the Languages of the World*. Stanford: Language Universals Project.
1976. The Geographical and Genetic Distribution of Linguistic Features. In: *Linguistic Studies Offered to Joseph Greenberg*, ed. by A. Juilland, vol. II, 137–60. Saratoga, Calif.: Anma Libri.
1979. On the Origin and Evolution of French Nasal Vowels. *Romance Philology* 32: 321–35.
1986. The Voice Warp: A Phonological Universal. In: *The Fergusonian Impact*, ed. by Joshua A. Fisherman, *et al.*, 17–19. Berlin – New York – Amsterdam: Mouton de Gruyter.

[3] Thanks especially to Pierre Bancel, William Croft and Anca Ruhlen for helping to assemble this bibliography.

1987a. *A Guide to the World's Languages, Volume 1: Classification.* Stanford: Stanford University Press.
1987b. Voices from the Past. *Natural History* 96.3: 6–10.
1989. Nostratic-Amerind Cognates. In: *Reconstructing Languages and Cultures,* ed. Vitaly Shevoroshkin, 75–83. Bochum: Brockmeyer. [Reprinted as "The Linguistic Origins of Native Americans" in Ruhlen 1994b, 207–241.]
1990. Phylogenetic Relations of Native American Languages. In *Prehistoric Mongoloid Dispersals*, No. 7 (Special Issue), University of Tokyo, Tokyo, 75–96.
1991a. The Amerind Phylum and the Prehistory of the New World. In *Sprung from Some Common Source*, ed. by Sydney M. Lamb and E. Douglas Mitchell, Stanford, Stanford University Press, 328–50.
1991b. Происхождение языка: ретроспектива и перспектива. [The origin of language: retrospective and perspective.] Вопросы языкознания [Voprosy jazykoznanija] 1: 5–19. [Translation printed in Ruhlen 1994b, 261–276.]
1991c. Postscript 1991 to reprint of Ruhlen 1987a, 379–407.
1992a. An Overview of Genetic Classification. In: *The Evolution of Human Languages*, ed. by John A. Hawkins and Murray Gell-Mann, 159–89. Redwood City, CA: Addison-Wesley Publishing Co. [Reprinted in Ruhlen 1994b, 9–38.]
1992b. (Joseph H. Greenberg and Merritt Ruhlen) Linguistic Origins of Native Americans. *Scientific American* 267(5): 94–99.
1994a. *The Origin of Language*. *Tracing the evolution of the mother tongue.* New York: John Wiley.
1994b. *On the Origin of Languages*. *Studies in linguistic taxonomy.* Stanford, Calif.: Stanford University Press.
1994c. Is Algonquian Amerind? In Ruhlen 1994b: 111–126.
1994d. Na-Dene Etymologies. [Based on J.H. Greenberg's notebook.] In Ruhlen 1994b: 93–110.
1994e. (John D. Bengtson and Merritt Ruhlen) Global Etymologies. In Ruhlen 1994b: 277–336.
1994f. First- and Second-Person Pronouns in The World's Languages. In Ruhlen 1994b: 252–260.
1994g. Review of *Linguistic Diversity in Space and Time*, by Johanna Nichols. *Anthropos* 89: 640–641.
1994h. On the Origin of the Amerind Pronominal Pattern. In: *In Honor of William S-Y. Wang: Interdisciplinary studies on language and language change*, ed. by Matthew Y. Chen and Ovid J.L. Tzeng, 405–07. Taipei, Taiwan, ROC: Pyramid.
1994i. Plus ça change, plus c'est la même chose. *Mother Tongue* (Newsletter) 23: 73.
1995a. Proto-Amerind Numerals. *Anthropological Science* (Tokyo) 103.3: 209–225.
1995b. Comments on R. L. Trask's Critique: Is Basque an Isolate? *Mother Tongue* 1: 149–156.
1995c. Proto-Amerind *QETS' 'left (hand)'. *Mother Tongue* (Newsletter) 24: 69–70.
1995d. (Jianteng Chen, Robert R. Sokal, and Merritt Ruhlen) Worldwide Analysis of Genetic and Linguistic Relationships of Human Populations. *Human Biology* 67.4: 595–612.
1995e. Linguistic Evidence for Human Prehistory. *Cambridge Archaeological Journal* 5.2: 265–68.
1995f. A Note on Amerind Pronouns. *Mother Tongue* (Newsletter) 25: 60–61.
1996a. Letter from Merritt Ruhlen to Larry Trask. *Mother Tongue* 2: 137–139. [In response to Trask, R.L., Response of Larry Trask to Merritt Ruhlen, *Mother Tongue* 2: 131–135]
1996b. Une nouvelle famille de langues: le déné-caucasien. *Pour la Science* (Dossier, Octobre): 68–73.

1997a. *L'origine des langues. Sur les traces de la langue mère*. Paris: Éditions Belin. [Translation of Ruhlen 1994a by Pierre Bancel.]
1997b. (John D. Bengtson and Merritt Ruhlen) In Defense of Multilateral Comparison. *California Linguistic Notes* 25.1: 3–4, 57. [A response to Joseph Salmons' "'Global Etymology' as Pre-Copernican Linguistics" in the same issue.]
1997c. Proto-Amerind *kapa 'Finger, Hand' and Its Origin in the Old World. In *Indo-European, Nostratic, and Beyond: Festschrift for Vitalij V. Shevoroshkin*. (JIES Monograph No. 22.) Eds. Irén Hegedűs, Peter A. Michalove, Alexis Manaster Ramer, pp. 320-25. Washington D.C.: Institute for the Study of Man.
1998a. The Origin of the Na-Dene. *Proceedings of the National Academy of Sciences* 95: 13994–13996.
1998b. Dene-Caucasian: A New Linguistic Family. In: *The Origins and Past of Modern Humans – Towards Reconciliation*, Keiichi Omoto & Phillip V. Tobias (eds.), 231–246. Singapore: World Scientific.
1998c. *A Origem da Linguagem: Reconstituindo a Evolução da Lingua Mãe*. Lisbon: Circulo de Leitores. [Translation of Ruhlen 1994a by Yolanda Salo.]
1999. Where Do Languages Come From? *Exploratorium* 23.1: 2–4.
2000a. Why *Kaka* and *Aya*? In: *Functional Approaches to Language, Culture and Cognition: Papers in honor of Sydney M. Lamb*, ed. David G. Lockwood, Peter H. Fries & James E. Copeland, 521–525. Amsterdam–Philadelphia: John Benjamins, 2000.
2000b. Some Unanswered Linguistic Questions. In: *America Past, America Present: Genes and Languages in the Americas and Beyond*, ed. by Colin Renfrew. Cambridge (U.K.), 163–175. The McDonald Institute for Archaeological Research.
2001a. Taxonomic Controversies in the Twentieth Century. In: *New Essays on the Origin of Language* (Trends in Linguistics. Studies and Monographs [TiLSM], 133), ed. by Jürgen Trabant and Sean Ward, 197–213. De Gruyter Mouton.
2001b. Il Dene-caucasico: una nuova famiglia linguistica. *Pluriverso* 2: 76–85.
2001c. *L'origine delle lingue*. (Biblioteca Scientifica, 31.) Milan: Adelphi Edizioni. [Translation of Ruhlen 1994a by Stefano Ravaioli.]
2003. (Alec Knight, Peter A. Underhill, Holly M. Mortensen, Lev A. Zhivotovsky, Alice A. Lin, Brenna M. Henn, Dorothy Louis, Merritt Ruhlen, and Joanna L. Mountain) African Y Chromosome and mtDNA Divergence Provides Insight into the History of Click Languages. *Current Biology* 13: 464–473.
2004a. (Paul Whitehouse, Timothy Usher, Merritt Ruhlen and William S-Y. Wang) Kusunda: An Indo-Pacific Language in Nepal. *Proceedings of the National Academy of Sciences* 101: 5692–5695. New York: Garland Science.
2004b. Opinion: Linguistic Evidence for Human Origins. In: M. A. Jobling, M. E. Hurles, and C. Tyler-Smith, eds., *Human Evolutionary Genetics: Origins, Peoples and Disease,* 5–6.
2004c. On the Amerind Origin of the Proto-Algonquian Numeral Suffix *-a:šyeka. In: *Traces of ancestry: studies in honour of Colin Renfrew*, ed. Martin Jones, 139–142. Cambridge: McDonald Institute for Archaeological Research.
2005. Taxonomy, Typology and Historical Linguistics. In: *Language Acquisition, Change and Emergence: Essays in evolutionary linguistics*, ed. by James W. Minett and William S-Y. Wang, 341–368. Hong Kong: City University of Hong Kong Press.
2006. *Dilin Kökeni: Ana Dilin Evriminin İzinde*. Ankara: Hece Yayınları. [Turkish translation of Ruhlen 1994a by İsmail Ulutaş.]

2007a. *L'Origine des langues. Sur les traces de la langue mère*. Paris: Éditions Gallimard. [Translation of Ruhlen 1994a by Pierre Bancel, with a new Postscript by the author.]
2007b. (Joseph H. Greenberg and Merritt Ruhlen) *An Amerind Etymological Dictionary*, version 12. Stanford: Department of Anthropological Science.
2009a. Migrations to the Americas. In: *Ancient Human Migrations: A Multidisciplinary Approach*, ed. by Peter Peregrine, Ilia Peiros & Marcus Feldman, 112–126. Salt Lake City: University of Utah Press.
2009b. Postscript to Ruhlen 1994a. Unpublished manuscript. [In this volume: *Mother Tongue* XXIII, 2021.]
2011. (Murray Gell-Mann and Merritt Ruhlen) The Origin and Evolution of Word Order. *Proceedings of the National Academy of Sciences* 108: 17290–17295.
2015. (Nicole Creanza, Merritt Ruhlen, Trevor J. Pemberton, Noah A. Rosenberg, Marcus W. Feldman, Sohini Ramachandran) A Comparison of Worldwide Phonemic and Genetic Variation in Human Populations. *Proceedings of the National Academy of Sciences* 112.5: 1265–1272.

Other References Cited

Ehret, Christopher. 2015. Africa from 48,000 to 9500 BCE. In *The Cambridge World History. Volume 1: Introducing World History, to 10,000 BCE*, D. Christian (ed.), 362–393. Cambridge University Press.
Greenberg, Joseph H. 1987. *Language in the Americas*. Stanford: Stanford University Press.
Hymes, Dell. 1971. Morris Swadesh: From the First Yale School to World Prehistory. In: Swadesh, M., *The Origin and Diversification of Language*, J. Sherzer (ed.), 228–270. Chicago / New York: Aldine · Atherton.
Kassian, Alexei. 2010. Hattic as a Sino-Caucasian language. *Ugarit-Forschungen* 41: 309-448.
Kitchen, Andrew, Michael M. Miyamoto, Connie J. Mulligan. 2008. A Three-Stage Colonization Model for the Peopling of the Americas. *PLoS ONE* 3(2): e1596.
Nikolayev, Sergei. 1991. "Sino-Caucasian Languages in America." In *Dene-Sino-Caucasian Languages*, ed. by Vitaly Shevoroshkin, 42–66. Bochum: Brockmeyer.
Reich, David. 2018. *Who We Are and How We Got Here*. New York: Pantheon Books.

POSTSCRIPT 2009

MERRITT RUHLEN

Fifteen years have passed since I wrote this book and obviously I would not write it today in exactly the same way. In this Postscript I would like to correct certain errors, reply to critics, and take into account what has happened in the past fifteen years that bears on the conclusions reached in this book.[1]

THE ORIGIN OF LANGUAGE OR LANGUAGES?

What I would most like to rewrite today is my inadequate discussion in Chapter 1 of the difference between the origin of Language—the language faculty, the ability to speak—and the origin of those languages which now exist. These are entirely different questions, though even today they are still constantly confused. So let me begin again.

Let us begin with a sketch of human evolution over the past 200,000 years. The earliest archaeological evidence for anatomically-modern humans has been found in Ethiopia and dated to 195,000 B.P. Anatomically modern humans are people who, on skeletal evidence, look the same as all people living today and do not look like Neanderthals, whose morphology was quite distinct. But while the Neanderthals' physical morphology might have been quite different from that of anatomically-modern humans, their cultures appear to have been almost identical. That they should have looked quite different is hardly surprising since the Neanderthals are descendants of the first Out-of-Africa migration, which took place between one and two million years ago, whereas anatomically-modern humans did not leave Africa until around 50,000 years ago, and the most recent common ancestor of the Neanderthals and modern humans is thought to have lived around 600,000 B.P.

What is most interesting, however, about the anatomically-modern humans who emerged in East Africa almost 200,000 years ago, is that for the next 150,000 years they continued to behave like Neanderthals and then suddenly, around 50,000 years ago, their behavior changed dramatically in numerous ways, though their physiology did not change at all. These new people are called behaviorally-modern humans and traits that distinguish them from the earlier anatomically-modern humans, as well as Neanderthals, include: (1) Artifacts are now made not just from stones, but

[1] "This book" refers to *The Origin of Language* (Ruhlen 1994a). So far a revised version has not appeared. This publication of the Postscript has been authorized by Merritt Ruhlen's executor, Anca Ruhlen. [Ed.]

from bones, shells, ivory and other materials, and there is greater artefactual diversity as well. (2) These artifacts begin to change rapidly in both time and space; before this time tools changed very little over hundreds of thousands of years. (3) Art appears for the first time. (4) Evidence for the spatial organization of camp floors is found. (5) Valuable raw materials are now transported over hundreds of kilometers. (6) There is evidence of ritual and elaborate burials. (7) Fishing appears for the first time. (8) Some of these behaviorally-modern people leave Africa, in the second Out-of-Africa migration, and eventually replace all the descendants of the first Out-of-Africa migration (Klein 1999). The Neanderthals, who had lived successfully in Europe for several hundred thousand years, disappear from the archaeological record around 30,000 years ago, only five or ten thousand years after the first behaviorally-modern humans arrived in Europe.

What could have been responsible for this sudden and profound transition in human evolution? Many people, including Richard Klein (Klein 1999), Luca Cavalli-Sforza (Cavalli-Sforza 2000), Jared Diamond (Diamond 1992), and Nicholas Wade (Wade 2006) have concluded that it must have been the sudden appearance of fully modern language. How this first fully modern language would have differed from earlier languages of the Neanderthals and others is a tantalizing but inscrutable question because all of these earlier languages went extinct along with the people who spoke them. Still, it seems likely that it *was* a new form of language that permitted these behaviorally-modern people to leave Africa and replace the earlier people in Eurasia by simply outcompeting them for the resources upon which they had lived quite successfully for over a million years.

Recent genetic studies indicate that this ancestral human population in East Africa, from which all modern humans descend, was quite small, perhaps as few as a thousand people, which would suggest that they spoke a single language from which all modern languages derive. Words such as *tik* 'finger, one' and *pal* 'two,' which we have seen are found in language families around the globe, would have been part of this original language. This does not mean that there were no other languages at this time, only that all the other languages went extinct. This is no different from the fact that all human Y-chromosomes found on the earth today derive from the Y-chromosome of one man who lived in Africa around 60,000 years ago, at a time when there were other men, and other Y-chromosomes, which have since disappeared without a trace.

If this scenario is correct—and it seems well-founded on archaeological and genetic evidence—then the question of the origin of those languages which now exist is resolved. The origin of these languages goes back to a single language spoken in East Africa 50,000 years ago. What then is the time and place of the origin of Language? The place must have also been Africa for the simple reason that, as Darwin saw, our closest relatives, gorillas and chimpanzees, have always lived exclusively in Africa and anatomically-modern humans only left Africa 50,000 years ago. Therefore the evolution of Language must have taken place in Africa.

But what about the date of the origin of Language? There are here two problems. First, since all intermediate varieties of Language between rudimentary chimpanzee communication and fully modern languages spoken by all humans today have been irretrievably lost, along with the people who spoke them, all we can say is that the origin of Language must have occurred some time in

the last 5 million years, after the separation of the chimpanzee and human lineages. Second, even if we knew exactly the nature of the languages of the australopithecines, *Homo habilis, Homo erectus,* Neanderthals, and other intermediate species, the origin of Language would be merely a matter of definition. On what grounds could one say that some particular stage of language evolution is the *origin* of Language? It would all depend on how one defined Language and it would thus be an arbitrary decision.

THE LIMITS OF THE COMPARATIVE METHOD

The chief criticism of this book has been that languages change so rapidly that after 6,000 years all traces of earlier relationships would have been erased by the shifting sands of time. Thus, even if languages such as Proto-Indo-European, Proto-Uralic, Proto-Austronesian, etc. did have relatives, evidence that might show this would have disappeared. This temporal limit of the Comparative Method became the mantra of twentieth-century historical linguists, even though there is no trace of it in the nineteenth century and it is clearly false. As pointed out in Chapter 3, Proto-Indo-European **nepo:t* 'nephew' has become *nepot* in modern Rumanian, essentially unchanged for 6,000 years, and this example is by no means unusual.

Austronesian is a family of almost 1,000 languages spoken on islands from Madagascar to Easter Island. Proto-Austronesian is thought to have been spoken on Taiwan some 6,000 years ago. Following is a list of Proto-Austronesian reconstructions followed by the modern forms in the Rukai language, also spoken on Taiwan: **dusa* 'two' > *dosa, *sepat* 'four' > *sepate, *'enem* 'six' > *eneme, *maca* 'eye' > *maca, *calinga* 'ear' > *calinga, *awang* 'canoe' > *avange, *kucuh* 'head louse' > *koco* (Bellwood 1991). As is obvious, the impression is not that everything has changed beyond recognition, but rather that almost nothing has changed at all. So much for the mythical 6,000 year limit on comparative linguistics.

Even if this 6,000 year limit is fiction, just how far back in time can the linguistic evidence take us? And how far back in time do we need to go to arrive at the origin of extant languages? It has been claimed that the origin of extant languages must go as far back as the origin of modern humans (McWhorter 2001), which we have seen is nearly 200,000 years ago. Such claims totally ignore the difference between anatomically-modern humans and behaviorally-modern humans, which I emphasized in the preceding section. While it is true that modern humans go back almost 200,000 years, the last common ancestor of those humans alive *today* only goes back around 50,000 years, to the advent of behaviorally-modern humans, which is, as we have seen, the apparent origin of extant languages.

Many linguists would claim that even 50,000 years would be long enough to have erased all evidence of genetic affinity among the world's languages, but these are the same people who claim that a mere 6,000 years is sufficient, so why should we believe them? Recently Mark Pagel, a phylogeneticist who usually studies biology, analyzed the rate of replacement of words in the Indo-European family and found that the rate of replacement varies greatly depending on the meaning. The half-life of a word is the amount of time required for a word to have a 50 per cent chance of being replaced by a non-cognate word. Most words have half-lives between 1,000 and 2,000 years, but some were found to have much longer half-lives:

> Fifteen words (7.5 per cent) have half-lives of greater than 13,000 years. One of these is 'one' with a half-life of 21,000 years. It has been suggested that the primitive reconstructed form *tok,* which variously represents the word 'one' or 'finger' or 'toe' among diverse languages, may be a candidate for a universal word, evidence of a 'Mother Tongue' of languages (Ruhlen 1994a). A word with a half-life of 21,000 years has a 22 per cent chance of not changing in 50,000 years (Pagel 2000: 205).

And Pagel notes that words such as 'I,' 'we,' 'who,' and 'two' have even longer half-lives.

HOW WERE THE TABLES MADE?

My main purpose in writing this book was to explain how Greenberg's method of classifying languages worked. Furthermore, I felt that merely explaining the method to the reader was unsatisfactory. To really understand how classification works the reader should do it himself, which is why I made the tables in this book.

How then were the tables actually made? Some critics have claimed that the tables are no more than a magic trick, where the words are arranged in advance so that the trick will work. I can't deny that I made all the tables in this book by myself. But how did I do it? First of all I chose 36 basic meanings that are known to be highly stable over time, such as 'I, you, two, who, not, name, tongue, eye, tooth, louse, water, dead.' Then, for each area, say Africa, I chose 12 languages, none of which I knew. I then went to the library at Stanford, found a dictionary for each language, looked up the 36 meanings and marked each word in a matrix of 12 languages by 36 meanings. I had no idea what sounds would represent which meanings in any given language and there were some surprises, as you may have noticed. In one language the word for 'snow' was *barf*. Next I classified the languages in the matrix, just as you have done for each table, except that for the tables in the book I only included the 12 words that gave the clearest evidence for the correct classification. These were not necessarily the same 12 words in each matrix.

Lest one wonder why I chose 12 words from the original table of 36, the answer is quite simple and is not, as has been alleged, because the other 24 words yielded a different classification. The answer was rather that the other 24 words gave exactly the *same* classification as the table based on 12. This should not really surprise you, if you did solve the tables in this book, since in each table the correct classification usually emerges after five or six words, so even the entire 12 is not really necessary.

The conclusion is that the tables are not magic tricks, but scientific experiments which can be replicated. Anyone can choose 12 other African languages, look up the words for 36 meanings, and then classify the table. I can guarantee that no such table will lead to a confusion of Khoisan, Chadic, Nilotic and Bantu languages with each other.

This method of classification contrasts sharply with the caricature of it that has been made by Larry Trask (1996), Lyle Campbell (Campbell and Poser 2008), and others. They claim that in looking for global etymologies—roots that are widely distributed around the world—John Bengtson and I first chose a meaning, say 'two,' and then chose a set of sounds, say *pal,* and then went through thousands of dictionaries looking for exactly this root, *pal* 'two.' We decided *in advance* what the sound-meaning correlation would look like. Nothing could be further from the truth. We decided nothing in advance and did not go through thousands of dictionaries looking for some specific sound-meaning correlation. What we compared was not thousands of languages, but

32 families that contain virtually all the world's languages, and what we compared was not words in several thousand languages, but the roots that had been identified as characteristic of each family by experts in that family. What we found was that a root like *pal* 'two' had been posited for 12 of these 32 families by experts who were usually unaware that a similar root existed in 11 other families. The root *tik* 'finger, one' was found in 21 of the 32 families. All we were doing was comparing language families with language families, which is no different from comparing languages with languages for, after all, a language family represents a single language that existed at some point in the past. Anyone who believes that you can pick out a sound-meaning correlation in advance and then go out and find it in hundreds of languages has almost certainly never done any taxonomic work and this would apply to the Indo-Europeanists and Americanists, who made no contribution to linguistic taxonomy throughout the entire twentieth century.

HOW MUCH EVIDENCE IS ENOUGH?

No matter how fast or slow language actually changes the bottom line is how much evidence is enough? How many cognates does it take to prove that two languages—or two language families—are related? Precisely this question was raised at an international conference of historical linguists at Stanford University in 1987 and it elicited an awkward silence, punctuated by coughs and clearings of the throat, because everyone knew that it was basically an unanswerable question. And then Ives Goddard stood up and gave a brilliant answer: "One, by definition."

The problem, of course, is that we can never have 100 percent certainty that two forms in two languages, or language families, are cognate. So how have linguists dealt with this problem over the past two centuries? How high does the bar have to be? I think it is instructive to consider this problem in historical perspective. The bar was, of course, first set by Sir William Jones in 1786 in his discovery of the Indo-European family. The evidence, as we saw in Chapter 1, was similarities "both in the roots of verbs and in the forms of grammar" which could only be reasonably explained by assuming them "to have sprung from some common source, which, perhaps, no longer exists." For Jones the bar was 30 cognates, involving five roots and six grammatical affixes, as seen in Table 1, and the validity of the Indo-European family has not been questioned since.

Table 1. An Indo-European Paradigm

	Sanskrit	Classical Greek	Latin	Old Irish	Gothic
I carry	bhár-āmi	phér-o	fer-ō	bir-u	bair-a
you carry	bhár-asi	phér-eis	fer-s	bir-i	bair-is
he carries	bhár-ati	phér-ei	fer-t	ber-id	bair-iθ
we carry	bhár-āmas	phér-omen	fer-ismus	ber-mi	bair-am
you carry	bhár-atha	phér-ete	fer-tis	ber-the	bair-iθ
they carry	bhár-anti	phér-ousi	fer-unt	ber-it	bair-and

Yet, as we saw in Chapter 6, for contemporary Indo-Europeanists what is required to prove the validity of a language family is a complete reconstruction of the Proto-language with regular sound

correspondences. Watkins assures us that "there is no other way," but the families he cites approvingly—Indo-European, Algonquian, and Austronesian—were accepted by everyone before anyone had reconstructed anything. They were accepted solely on the basis of similar words in different languages, exactly as Jones had proceeded. Furthermore, in classifying the languages in the tables in this book you were able to identify valid linguistic families throughout the world simply by comparing a dozen words of basic vocabulary, without attempting to reconstruct anything. The bar was raised so high by twentieth-century Indo-Europeanists that it became impossible to jump over it and even well-established families like Altaic were called into question (Campbell and Poser 2008). Were the Indo-Europeanists asked to devise the DNA fingerprinting test, they would certainly require a complete sequencing of both DNA samples, and anything less than that would be considered mere speculation. If, however, the Indo-Europeanists were told that the real DNA test, designed by geneticists, only looks at a few hundred nucleotides out of the 3.1 billion in the human genome, they would no doubt greet this news with shock and awe.

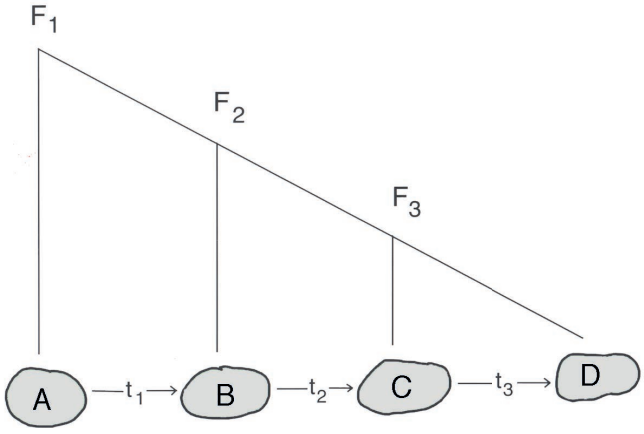
TIME AND INTERNAL STRUCTURE

One of the most constant criticisms of this book—and indeed of all taxonomic work that attempts to go beyond the obvious—is that as one goes back in time words are lost in all languages at a gradual rate so that things get dimmer and dimmer and finally turn to black. Whether things get black 6,000 years ago, or 100,000 years ago, was discussed above and is not relevant to the point I want to make here. While it seems obvious to anyone that as one goes back in time things do get dimmer, it's not true. Sometimes as one goes back in time things become more clear. The failure to understand this fundamental taxonomic principle is the source of much of the current confusion in historical linguistics.

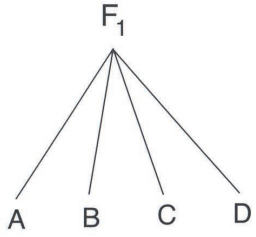
We have seen in the tables in this book that language families are defined by sharing certain words that are lacking in other families. This is why it was so easy to classify languages into language families. These shared words in a family are known in linguistics as "exclusively shared innovations" and they define any taxon at any level of a phylogenetic tree. What this means is that each node in a tree is defined by one or more innovations. However, whether such innovations develop, thus giving the internal subgrouping of a family, depends on the rate of migration, as illustrated in Figure 1.

Let us assume that there is a population on island A. At some time, part of that population moves to island B. Later part of that population moves to island C, and finally part of that population moves to island D. Let us further assume that there is no further contact between these populations. The correct phylogenetic tree will then be as in 1a and many historical linguists believe that F3 will be more obvious, and better supported, than F2, and F2 more obvious than F1. Whether this is true or not depends on the amount of time between the separation of the populations. Let us assume that some people on island A moved to B on Monday, then part of that population to C on Tuesday and part of that population to D on Wednesday, and then 500 years pass. Under these circumstances the phylogenetic tree, based on linguistic evidence, will be as in 1b since there will not have been any time for innovations to develop that would distinguish F2 and F3. The language

on each island would have been identical at the start and each would have then gone its own way, in a process similar to genetic drift.



1a. When the time of separation is large



1b. When the time of separation is small

Figure 1. Time of separation and phylogenetic structure.

Whether the correct phylogenetic tree can be ascertained on the basis of the linguistic evidence depends on the speed of migration. If the speed of migration is slow there will be time for innovations to develop, giving the correct tree in ***1a****. If, however, the speed of migration is fast there will be no time for innovations to develop, giving the incorrect tree in* ***1b****.*

If, however, the time of the separations was 500 years, not one day, then both F2 and F3 will be well defined by innovations that have accumulated during the 500 years. In both scenarios, however, F1 will be well defined by those words that have been preserved on islands A, B, C, and D. Thus, whether the intermediate nodes can be identified or not depends on the rate of migration. With a rapid migration it is only the *highest* level node that will be clear, the intermediate nodes often being unclear or even invisible. An example of this principle is the initial peopling of the Americas by the Amerind family. Archaeological evidence indicates that the initial peopling of the New World was a very rapid migration that filled two empty continents with people in 1,000 years. Under these circumstances it is the highest level node—Proto-Amerind—that will be the most obvious, as the *n/m* pronoun pattern and the *t'ina/t'ana/t'una* root clearly show. Similarly, the Austronesian family was recognized from the outset because Austronesian languages share certain words that are not found elsewhere, but the intermediate nodes in the Austronesian tree remain to a large extent unknown

because of the rapidity of the Austronesian expansion. Critics of Greenberg who claim that one cannot just jump to the top of a phylogenetic tree without first working out every intermediate node obviously do not understand this fundamental principle of taxonomy.

ORIGINS OF THE CONTROVERSY

How can it be that there is a unique pronoun pattern found throughout the Amerind family, as well as the complex *t'ina/t'ana/t'una* root—both exclusively shared innovations of Amerind—and yet Americanists maintain that there is no evidence of genetic affinity among any of the almost 200 language families in the Americas? There are four possible explanations for these patterns: (1) common origin, (2) borrowing, (3) accidental resemblance, and (4) sound symbolism. Only common origin can explain these two patterns. First- and second-person pronouns are almost never borrowed and words in general could not have been borrowed over entire continents—much less two—until very recently. The Amerind *n/m* pronoun pattern can hardly be due to accidental resemblance; accidental resemblances should happen randomly around the world. If they are all found essentially in North and South America, they can't be accidents. And that the distribution of the accidental *t'ina/t'ana/t'una* forms coincides with the distribution of the accidental *n/m* pronouns would truly be a remarkable coincidence.

What is the origin of the current vitriolic debate over all taxonomic hypotheses that dare to go beyond the obvious: Amerind, Eurasiatic, Dene-Caucasian, etc.? The answer, I believe, is that there was in the twentieth century a complete misunderstanding of what the Comparative Method really is. The Comparative Method in linguistics, as worked out in the nineteenth century, consists of two stages, taxonomy and historical linguistics, as shown in Figure 2.

Taxonomy, or classification, identifies language families by a comparison of basic vocabulary, exactly as Jones discovered Indo-European in 1786 and as you have solved the tables in this book. Historical linguistics investigates families that have thus been identified and attempts to (1) identify the sound correspondences connecting the family's languages, (2) reconstruct the words in the original proto-language, (3) locate the homeland of the family, (4) identify the time that the proto-language existed, etc. It should be obvious that the second stage must follow the first stage; it is after all impossible to investigate a language family until one has been found. Historically, too, the first stage had to precede the second. Jones identified the Indo-European family in 1786 by pointing out the similarities in the verbal paradigms of five languages. The second stage in the Comparative Method did not begin seriously until the 1870's, when the regularity of sound change was fully recognized by the Neogrammarians and its import for reconstruction was fully appreciated.

Greenberg made a strategic error in calling taxonomy first "mass comparison" and later "multilateral comparison," when what he was really doing was just taxonomy. Partially because of this the Indo-Europeanists and Americanists came to believe that Greenberg had invented a new approach to historical linguistics that was a substitute for historical reconstruction, when in fact he did nothing of the sort. As Greenberg opined more than once, "when I do historical linguistics I do it like anyone else." But what the twentieth century Indo-Europeanists, and the Americanists who fol-

lowed them, did not do was any taxonomy. Neither of these two groups made any taxonomic progress during the entire twentieth century; in fact, they actually regressed by breaking up valid families which had been accepted since the nineteenth century.

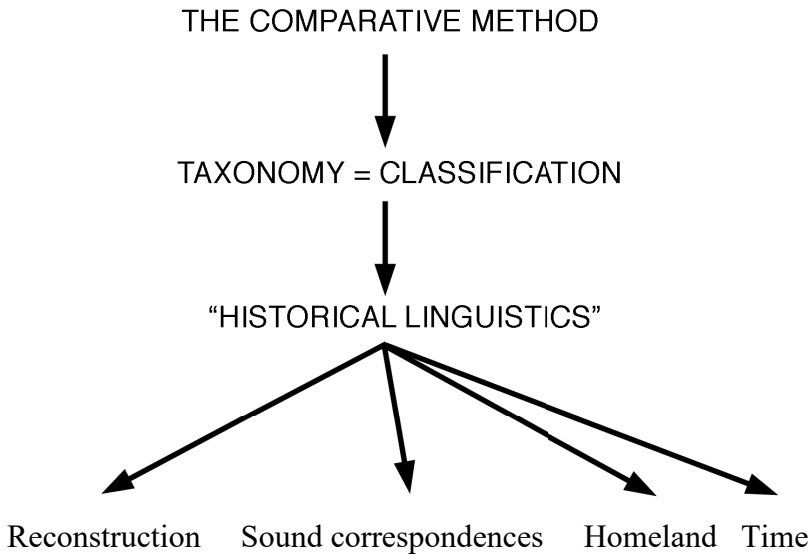


Figure 2. The comparative method.

The currently fashionable view is that Trombetti, Sapir, Greenberg, Illich-Svitych, Dolgopolsky, and Starostin are renegades who have deviated from the splendid foundation laid out by the nineteenth century Indo-Europeanists, while contemporary historical linguists are defending the faith. The reality is the exact opposite. It is rather Trombetti *et al.* who have followed rigorously the nineteenth century Indo-Europeanists, none of whom ever imagined that the reconstruction of Indo-European with regular sound correspondences had anything to do with "proving" Indo-European, whose validity was doubted by no one. It was only in the twentieth century, and only in linguistics, that reconstruction came to be viewed as necessary to prove any taxon.

THE PRONOUN CONTROVERSY

We saw in Chapters 3 and 4 that the Eurasiatic pronoun pattern *m/t* 'I/you' and the Amerind pattern *n/m* 'I/you' were both pointed out by Trombetti in 1905. The Eurasiatic pattern was also emphasized by the first Nostraticists and others, while the Amerind pattern was independently discovered by Sapir a decade after Trombetti's book.

But we also saw that the validity of these patterns as genetic traits, identifying specific families, was vehemently denied for myriad reasons. Or rather, as Trombetti noted acerbically, the *m/t* pattern was considered a valid genetic trait for both the Indo-European and Uralic families, but not for a family uniting the two. The evidence was the same, but the interpretation was different. Exactly the same thing happened with the Amerind family. While many Amerind subgroups had pronoun systems that had been reconstructed as *n/m* 'I/you' without complaint (for example, Proto-

Algic, Proto-Penutian, Proto-Hokan, Proto-Tanoan, Proto-Uto-Aztecan, Proto-Chibchan, Proto-Aruak, Proto-Quechuan, Proto-Guahiban), a family including all of these families—Amerind—was resolutely opposed. The evidence was the same, but the interpretation was different.

This specifically American pronoun pattern was also emphasized during the twentieth century in the work of Morris Swadesh, Sapir's student, as well as by Greenberg. In a survey of pronoun systems in the world's language families I showed that the *n/m* pattern is indeed essentially an American phenomenon; in fact it is an Amerind phenomenon since this pattern is not found in either Eskimo-Aleut or Na-Dene (Ruhlen 1994b). I also showed that the *m/t* pattern is found almost exclusively in northern Eurasia, characterizing the Eurasiatic family.

How then did Greenberg's critics explain the presence of the *n/m* pronoun pattern throughout Amerind languages on two continents and virtually nowhere else in the world? According to Lyle Campbell it simply wasn't true: "The *n/m* ['I/you'] pattern is not nearly as common in the Americas as Greenberg claimed . . . [and] his supposed *m/t* ['I/you'] pattern for his Eurasiatic languages is also found abundantly in the Americas (despite his and Ruhlen's assertions to the contrary)" (Campbell 1994a). Elsewhere he claimed that "several Amerind groups exhibit pronoun forms (*m/t* ['I/you']) that Greenberg attributes to Europe and Northern Asia" and "the *n* 'first person' / *m* 'second person' is by no means unique to, diagnostic of, or ubiquitous in American Indian languages" (Campbell 1994b). Johanna Nichols expressed similar views in 1992 (Nichols 1992). And then, at the end of the twentieth century, Campbell and Nichols had a debate in *Language* over which of them was the first to have noticed that the *n/m* pronominal pattern was found almost exclusively in the Amerind languages of North and South America, exactly as Trombetti had shown almost a century earlier, though Trombetti was not mentioned (Nichols and Peterson 1996, Campbell 1997).

But Nichols still denied that this pronoun pattern could have a genetic explanation since it must go back well beyond the 6,000 year limit of comparative linguistics and therefore these two pronouns could not be genetic traits. Their distribution in the Americas must be explained, according to Nichols, by a mysterious unknown force, not simple evolution from a common original source. I will leave it to the historians of science to sort this all out. In any event, a world atlas of language structures, published in 2005, contained maps of the distribution of the Eurasiatic *m/t* pronoun pattern and the Amerind *n/m* pattern (Nichols and Peterson 2005). Both maps confirm exactly what Trombetti had claimed in 1905. The great twentieth century pronoun controversy has officially come to an end.

MAMA, PAPA, AND KAKA REVISITED

In Chapter 5 I discussed the problem that the widespread kinship terms MAMA 'mother' and PAPA 'father' posed for those linguists who thought their broad distribution could not be explained by common origin since they were found in language families that were supposedly unrelated, as was the accepted view in the 1950's. Jakobson's explanation that these widespread words were an example of convergence—rather than common origin—was immediately accepted by almost all linguists because it solved a problem without recourse to an evolutionary hypothesis, which at the

time was unacceptable. According to Jakobson such words as MAMA and PAPA simply reflected the order of sounds and meanings in child language acquisition, with children perpetually reinventing the same sound/meaning correlations in languages around the world. The problem, as I pointed out, was that an almost equally widespread root, KAKA 'older brother,' was not susceptible to this same explanation since K, unlike M and P, is not acquired early in the acquisition of language and its subsequent association with the meaning 'older brother' is more than problematic. I concluded that all three of these roots go far back in human prehistory.

Subsequent to my proposal, Pierre Bancel and Alain Matthey de l'Etang (2002, 2004) have conducted a thorough investigation of kinship terminology that essentially confirms my hypothesis that it is indeed common origin—and not convergence—that is responsible for the global distribution of these kinship terms. They also pointed out that the immediate acceptance of Jakobson's hypothesis is bizarre in that there is no evidence in the literature on child language acquisition that such kin terms are continually reinvented each time a child learns a language. They are, rather, inherited just like all the rest of a language's vocabulary.

While I had only listed examples of KAKA from Eurasia, Oceania, and the Americas, they showed that it is equally well-attested in Africa, as are PAPA 'father' and TATA 'father,' all of which must be presumed to derive from the ancestral African language from which all extant languages derive. They also found, on the basis of an examination of over 1,000 languages, that the original meaning of KAKA was a male elder on the mother's side—grandfather, mother's brother, or older brother—with mother's brother being the main meaning. PAPA and TATA, they found, refer to a male elder on the father's side—father, father's brother, older brother, and grandfather. They are currently extending their investigation to MAMA, NANA, and YAYA, all of which are female kin terms. In their view, "if it is not the children who invent these words each time one of them learns an articulated human language . . . another childhood may have seen these words come to light: that of articulated language itself, or more exactly the stage corresponding to the invention of oral lexicon (that of syntax being later)" (Bancel and Matthey de l'Etang 2002: 221). "It seems that no other words might emerge as better candidates for the birth of articulated speech, that must have begun with a strange and unique invention: the discovery of consonants" (Bancel and Matthey de l'Etang 2004: 188). As has long been recognized, all kin terms conspire to avoid complexity, consisting of simple consonants and vowels, simple syllable structure, and reduplication. In unpublished work they have suggested that pronouns may have originally evolved from these kin terms and this suggestion is, in my opinion, worthy of further research.

YENISEIAN AND NA-DENE

The Dene-Caucasian family was amply discussed in this book. However, in 1998 I compared Yeniseian with Na-Dene—a comparison that had not previously been made by anyone—and found substantial additional evidence for the Dene-Caucasian family, as one would expect (Ruhlen 1998). The reason one would expect to find additional evidence involves the principle of transitivity in taxonomy, which says that if language (or family) A is related to language (or family) B, and

B is related to language (or family) C, then A and C are automatically related, even if they themselves show no obvious relationship. However, in virtually all cases there will be additional evidence for the higher-level family—Dene-Caucasian in this case—because there will be some cognates that are *only* found in A and C and thus considering only the form in A, or C, will fail to reveal its Dene-Caucasian status.

In this case the additional evidence shared exclusively by Na-Dene and Yeniseian suggests that Yeniseian may have been the branch of Dene-Caucasian from which Na-Dene emerged, though this is by no means certain. Table 2 gives some of these Yeniseian–Na-Dene comparisons.

The resemblance between the Yeniseian and Na-Dene words for 'birch bark' (not found in other branches of Dene-Caucasian) is particularly striking in that the words are virtually identical in sound and meaning and yet 'birch bark' is hardly part of the basic vocabulary, such as pronouns and body parts, which are often maintained for long periods of time. The reason this word has been retained for so long is that birch bark played a central role in the daily life of the Yeniseian people. Nearly all household items, including dishes and even the teepee, were made of birch bark, as well as the birch-bark boat. In addition, birch bark is impermeable to water so it is effectively always dry, a very useful quality in a cold and wet environment.

Table 2. Yeniseian-Na-Dene Cognates

	Yeniseian	**Na-Dene**
birch bark	qiʔy	*qʔəy
stone, rock	tɨʔś	tʔi:s (H)
bow, arrow	qiʔt	*qʔa:ʔ
squirrel	*saʔqa	-tsʔa:kw (H)
trough, dish	śiʔk	sʔixʔ (T)
river, ocean	*ses	si:skw (H)
boat	*qä(ʔ)p	*-qe:-
foot	*kiʔś	*-ke:ʔ
head	*tsiʔg-	*-tsiʔ
elbow, knee	*gid	*-god
skin	*sä:s	*-sətsʔ
nit, louse	*yok	*yaʔ
child	*gəʔt	gyi:tʔ (H)
hunger, hungry	qɔ:t	qʔut (H)
snow on ground	*ti:χ	tʔi:χʔ (T)
falling snow	*beʔč	wehs (E)

In the Yeniseian column Proto-Yeniseian reconstructions are preceded by an asterisk; forms without an asterisk are taken from Ket. In the Na-Dene column, forms preceded by an asterisk are Proto-Athabaskan reconstructions; forms without an asterisk are taken from Haida (H), Tlingit (T), or Eyak (E).

There is, however, an interesting difference between Yeniseian *qɨʔy* and Na-Dene *qʔəy*. The Na-Dene word begins with a glottalized consonant, *qʔ,* but in Yeniseian this glottalized consonant has broken up into its two components and the second part has switched positions with the following vowel, *qɨʔ*. This process is known as metathesis in linguistics and explains how Old English *waps* became modern English *wasp.* Usually metathesis affects single words, like *wasp,* but, as the words for 'birch bark,' 'stone,' 'bow/arrow,' 'squirrel,' and 'trough/dish' in Table 12 show, in this case glottalized consonants in Na-Dene, *CʔV,* regularly correspond in Yeniseian to *CVʔ,* where *C* represents any consonant and *V,* any vowel. There is thus a regular sound correspondence between Na-Dene and Yeniseian and most linguists consider such correspondences the ultimate proof of genetic affinity. Recently Edward Vajda has presented additional evidence for the genetic affinity of Na-Dene and Yeniseian (Vajda 2009), though he incorrectly excludes Haida from Na-Dene.

KUSUNDA: A MYSTERY LANGUAGE IN NEPAL

In 1848 an Englishman, Brian Hodgson, noted that there were two tribes living in the forests of central Nepal which appeared disconnected from other populations living in the area. He described them as follows: "Amid the dense forests of the central region of Népál, to the westward of the great valley, dwell, in scanty numbers and nearly in a state of nature, two broken tribes having no apparent affinity with the civilized races of that country, and seeming like the fragments of an earlier population" (Hodgson (1848). These two tribes were the Chepang and the Kusunda, both hunter-gathers living in forests.

When Hodgson obtained linguistic data on these two peoples in 1857 Chepang turned out to be simply a Tibeto-Burman language, closely related to the Lhopa language of Bhutan (Hodgson 1857). The Kusunda, however, whose name for themselves is *gilongdei mihaq* 'people of the forest,' showed no similarities, except for borrowings, with surrounding languages. Indeed, Kusunda showed no similarities with any other language and became known as a language isolate whose origin was a mystery.

This mystery, however, was finally solved in 2004 by a team of linguists working on the Evolution of Human Languages program at the Santa Fe Institute (Whitehouse *et al.* 2004), and the solution was surprising because Kusunda's closest relatives turned out to be the languages of the Indo-Pacific family, most of whose languages are found on New Guinea, with a few on surrounding islands. But how could a language in Nepal have its closest relatives on New Guinea, over 4,000 miles away? We will return to this question after we examine the linguistic evidence for this genetic affinity.

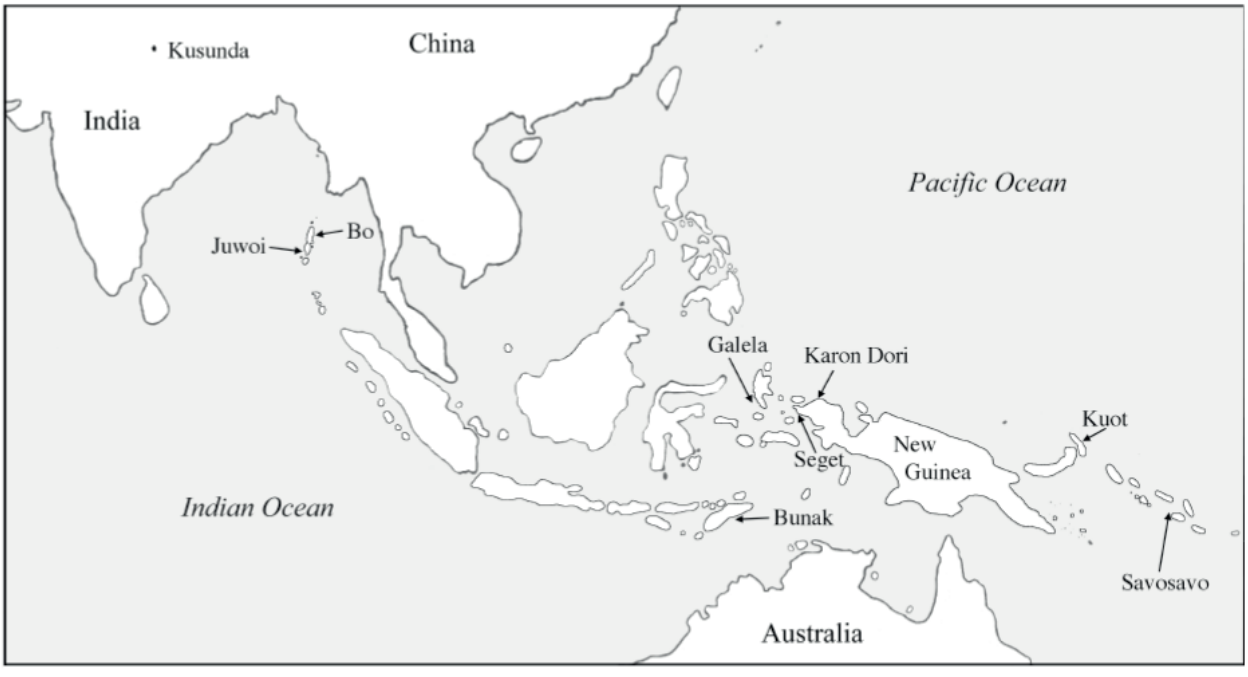
Table 3. An Indo-Pacific Pronominal Pattern

	Kusunda	Juwoi	Bo	Galela	Seget	Karon Dori	Kuot	Savosavo	Bunak
I	tsi	tui	tu-lʌ	to	tet	tuo	-tuo		ne
my	tsi-yi	tii-ye	ti-e	ǰi 'me'					n-ie
you	nu	ŋui	ŋu-lʌ	no	nen	nuo	-nuo	no	e-
your	ni-yi	ŋii-ye		ni 'thee'					Ø-ie
he / she	gida	kitɛ	kitɛ		gao			go	gi
his / hers	gida-yi								g-ie

In our paper we gave both grammatical and lexical evidence for this connection, but I will focus here on the grammatical evidence, specifically the pronominal similarities shared by Kusunda with other Indo-Pacific languages. The power of pronominal similarities in detecting deep genetic relationships has been emphasized throughout this book, with Eurasiatic *m/t* 'I/you' and Amerind *n/m* 'I/you' just two examples. The Kusunda pronouns, and those of eight other Indo-Pacific languages, are given in Table 3 and the location of these languages is shown in Map 1.

As can be seen in Table 3, the Kusunda pronominal system agrees with the other Indo-Pacific languages in five different parameters: (1) a first-person pronoun based on *t,* (2) a second-person pronoun based on *n* or *ŋ,* (3) a third-person pronoun based on *g* or *k,* (4) a vowel alternation in the first and second-person pronouns in which *u* occurs in subject forms and *i* in possessive forms, and (5) a possessive suffix *-yi* found on all three pronouns. The Kusunda first-person pronoun has, however, changed from the original system in two ways. First, *t* has changed to *ts* under the influence of the following *i,* just as Amerind *tina* 'son, brother' often became *tsina,* while *tuna* 'daughter, sister' and *tana* 'child, sibling' retain the original *t* because the vowels *u* and *a* do not cause the assibilation of *t* to *ts.* That the original consonant in the Kusunda first-person pronoun was *t* is shown by other forms of the first-person pronoun, *ton* 'me' and *to-ʔi* 'we,' in which *ʔi* is a plural suffix. In addition, the vowel alternation in the first-person pronoun has been eliminated by analogy, with *i* replacing *u* in the subject form, but the vowel alternation has been maintained in the second-person pronoun.

The Juwoi and Bo languages on the Andaman Islands show the greatest similarity with Kusunda, with other Indo-Pacific languages having retained less of the original system. The Bunak language on Timor has retained only third-person *g* of the three pronouns, but it shows the *-yi* suffix (in the form *ie*) on all three pronouns. The other languages show different parts of the original pronominal system.



Map 1. Kusunda and other Indo-Pacific languages.

But how can one explain the presence of an Indo-Pacific language in Nepal? It is hardly plausible that a small population could leave New Guinea and make its way to Nepal. The most likely explanation is that the Kusunda are a remnant of the Out-of-Africa migration that left Africa 50,000 years ago and eventually peopled the entire world, replacing the Neanderthals in Europe and various forms of *Homo erectus* in Asia. We will discuss the genetic and archaeological evidence for this migration below. The initial stage of this migration appears to have followed the south Asian coastline, moving through India and Southeast Asia and ultimately reaching the New Guinea-Australia region (which formed a single landmass at this time) by 45,000 years ago. The survivors of this migration were, however, later largely replaced by other migrations in South Asia of different peoples so that there remain today merely traces of the initial migration, one of which would be the Kusunda, who are clearly distinct linguistically from all other Eurasian populations, and perhaps also genetically, but we do not yet have any genetic evidence from the Kusunda. It would certainly be interesting to know whether the Kusunda share any genetic similarity with the people of the Andaman Islands, whose pronouns so closely resemble the Kusunda pronouns. It is believed that the Andaman Islanders are descendants of the Out-of-Africa migration discussed above.

Photo 1. Gyani Maiya Sen; Kusunda woman in Nepal.

There are also a few other small populations in South Asia that appear to be remnants of this migration since they appear physically different from most South Asians. The Vedda of Sri Lanka, who call themselves, like the Kusunda, "people of the forest," are probably one such example and are likely the original inhabitants of Sri Lanka who arrived on the island over 10,000 years ago, long before the Sinhalese, an Indo-European population, arrived around 500 B.C. The Vedda have,

however, adopted the Sinhalese language, thus losing their original language, for which there is no evidence.[2]

THE MYSTERIOUS ALGONQUIAN NUMERAL SUFFIX

In the mid 1990's I decided to investigate Amerind numeral systems in an attempt to figure out how the Amerind people counted when they first arrived in North America 13,000 years ago. To do this I collected as many numeral systems from Amerind languages in North and South America as I could find in the Stanford library. My analysis of these numeral systems led me conclude that the first three numerals in Proto-Amerind were most likely something like **ne-k'ʷe* '1,' **ne-pale* '2,' and **ne-qʷałas* '3' (Ruhlen 1995a).

Let us first consider Proto-Amerind **ne-k'ʷe* '1.' The numeral prefix **ne-* is found on many, though not all of the Amerind words for '1,' but it is synchronically identifiable as a prefix only in Proto-Algonquian, where it is found on the numbers 1–8, and in Algic, where it is found on the numerals 1, 3, and 4. Elsewhere within Amerind it has become fossilized on the root and is no longer recognized as a prefix, for example, Proto-Salish **nak'ʷ,* Kutenai *-nik?* 'one of,' and Kwakwala *n'axʷ* in the Almosan branch of Amerind. This fossilization is universal in all the other Amerind branches: Siuslaw *ni:k'a* and Proto-Atakapa-Chitimacha **nik'u* in Penutian, Washo *la:k'a-* (< *n'ak'a-*) in Hokan, Popoloco *nakua* and Amuzgo *nkwi* in Central Amerind, Millcayac *negui* in Paezan, Iquito *nóki* and Shimigae *nikíño* in Andean, Tucano *nike/niko/nika* '1 masculine/1 feminine/1 neutral' (with the Amerind gender pattern) in Macro-Tucanoan,Yabarana *enix-* in Macro-Carib, Atsahuaca *nikatsu* and Caduveo *nigot nuba:ř at* 'one hand' in Macro-Panoan, and Opaie *enex-* in Macro-Ge.

Not only did Proto-Algonquian preserve the *ne-* prefix intact, but it turned out that the Proto-Algonquian numeral 1, **ne-kwet-w-i,* resembles my hypothesized Proto-Amerind form, **ne-k'ʷe,* more than other Amerind languages, as we saw above. Why should this be the case? I believe the answer has to do with the Algonquian homeland, which I discussed in the Epilogue. I argued there that the Proto-Algonquian homeland was on the Montana-Canada border, which was also the Algic homeland and the Almosan homeland as well. Furthermore, it is also the location which the Amerind people would have arrived at when they left the ice-free corridor through Canada, and thus it would have been the Amerind homeland as well. Now it is well known that languages are more resistant to change on islands, where they are isolated from other languages, than when they spread through territory already occupied by other people. We saw above that words in the Rukai language on Taiwan are remarkably similar—sometimes identical—to the Proto-Austronesian words from which they derive, whereas in the Austronesian languages that left Taiwan and spread throughout the Pacific these same words become less similar the farther one gets from the Taiwan homeland. Now one might object that the Montana-Canada border is hardly an island, but from a demographic perspective it was, an island comprising two continents, with a small group of people

[2] Photo 1: Gyani Maiya Sen: Kusunda woman in Nepal, died on January 25, 2020 in Kulmor village of Dang district in western Nepal. See: https://www.imdb.com/title/tt12728670/ and https://globalvoices.org/2018/11/11/a-conversation-with-gyani-maiya-sen-one-of-the-last-speakers-of-a-dying-nepali-language/

on the extreme northern coast. One might also object that the Amerind people did not have to wade through other populations, as the Austronesians did, in their rapid occupation of two continents, but there would have been a great deal of contact between different Amerind tribes during this migration. They were, after all, just expanding into unoccupied territories, not going to South America. Apparently the Almosan people were able to stay above the fray better than other Amerind tribes.

On the morning I was looking at the Algonquian numeral systems (there are several), I came across a numeral whose origin was unknown: **pale·-nełkwi* '5.' While it was clearly composed of two morphemes, the meaning of each morpheme was unclear. The only purported explanation was that **-nełkwi* had derived from Proto-Algonquian **nełki* 'hand,' with the insertion of a *-w-* for unknown reasons. On semantic grounds one can hardly object because it is well known that around the world the word for 'hand' is often used for the numeral 5, just as 'finger' is used for 1, and 'body' is used for 20 (fingers+toes). The meaning of **pale·,* however, was unknown and it was suggested that perhaps it meant 'full,' though there is no evidence for this within Algonquian. This explanation is reasonable, but it leaves a number of loose ends and I realized that morning that is was not correct.

The reason this problem had not been resolved offers a cautionary lesson, for the reason was the Algonquianists' refusal to consider the extra-Algonquian context, Amerind. Despite the abundant evidence offered in this book that Algonquian is a run of the mill Amerind language family, Algonquianists persist to this day in maintaining that Algic has no known relatives. Had they dared consider the Amerind context they would have found numerous reflexes of Proto-Amerind **nepale* '2,' some of which you saw in Chapter 5. Many additional examples can be found in the *Amerind Etymological Dictionary* (Greenberg and Ruhlen 2007). However, **nepale* '2' had been replaced in Algonquian by a different root—except in **pale·-nełkwi.* The problem is now reduced to solving a mathematical equation. If $2 + x = 5$, what is x? Once the Algonqianists had solved this equation they would have noticed that the Proto-Algonquian numeral 3, **neʔłwi,* bore more than a passing resemblance to the second morpheme in **pale·-nełkwi.* The solution to this enigma was as simple as 2+3. What actually happened is illustrated in Table 4. The Pre-Proto-Algonquian numerals **nekłwi* '3' and **pale·-nekłwi* '5' are shown in the first row. First, metathesis affected 5, but not 3, changing *-kł-* into *-łk-.* Next, a regular sound change turned *k* into *ʔ* before *ł,* but since the *k* in 5 was now found after *ł*, not before, it was not affected. The final row shows the outcome of these two changes, the Proto-Algonquian numerals **neʔłwi* '3' and **pale·-nełkwi* '5.'

Table 4. The Origin of Proto-Algonquian **pale·-nełkwi*

Pre-Proto-Algonquian	**nekłwi* '3'	**pale·-nekłwi* '5'
(1) Metathesis	**nekłwi*	**pale·-nełkwi*
(2) k > ʔ	**neʔłwi*	**pale·-nełkwi*
Proto-Algonquian	**neʔłwi*	**pale·-nełkwi*

That same morning I bumped into yet another Algonquian problem that had never been resolved, the mysterious Proto-Algonquian numeral suffix *-w-a:šika* that is found on the numerals 6, 7, and

8, as can be seen in Table 5. There is an important process in linguistic evolution called *grammaticalization,* whereby ordinary lexical items become grammatical markers. We saw an example of grammaticalization in Chapter 5, in the discussion of the numeral 2, where the Proto-Australian numeral 2, **pula,* has been grammaticalized as a dual suffix on the second-person pronoun, **nyuN-palV* 'you-dual,' and as the third-person dual pronoun, **pula* 'they dual.' One important characteristic of grammaticalization is that the original meaning of the lexical item often changes in a way that is variously described as erosion, fading, or bleaching. Certainly this is what has happened to the numeral suffix *-w-a:šika,* which did not start off as a numeral suffix. At the beginning it had real meaning, but over time the original meaning faded away to the point where it wound up being nothing more than a numeral suffix and the original meaning was forgotten by the Algonquians and never discovered by linguists. This numeral suffix is also characterized by extraordinary length; usually suffixes are short, for example, *cat* and *cat-s*. I had already had a good morning, having found the correct etymology of **pale·-nełkwi,* and I didn't think it likely that I would also discover what *-w-a:šika* had originally meant. It was also lunchtime, and I was hungry, so I left Amerind numerals behind and went out to have lunch and, while having lunch, I discovered what the original meaning of *-w-a:šika* was.

Table 5. Proto-Algonquian Numerals 1–8

*ne-kwet-w-i '1'	*ne-kwet-w-a:šika '6'
*ni:š-w-i '2'	*niiš-w-a:šika '7'
*ne-ʔ ɬ -w-i '3'	*ne-ʔš-w-a:šika '8'
*nye:-w-i '4'	
*nya:łan-w-i '5'	

One thing that non-scientists are frequently not aware of is that there is often a happenstance, a serendipity, in scientific discovery and that day I had happenstance and serendipity for lunch. After buying a sandwich I picked up a copy of the *Stanford Daily* and went outside to have lunch. The first thing I noticed was a picture of one of my students on the front page of the *Daily*, showing her on a Zuni reservation in New Mexico. I was not surprised because she had told me she was going there and she looked to be having a good time. Now from a taxonomic perspective Zuni is an interesting language, having long been considered a language isolate, until Stanley Newman, a Zuni specialist, showed that it was in fact related to the California branch of Penutian (Newman 1964). The Indo-Europeanist Eric Hamp found "Newman's whole argument . . . persuasive" (Hamp 1975) and the evidence that Newman presented *was* overwhelming, except, of course, for Lyle Campbell, who still considers Zuni an "isolate" with no known relatives (Campbell 1997: 321). I have never studied Zuni, but in thinking about it that day I remembered the one word in Zuni that I knew and it was the right word: *wešik:ʔa* 'left hand.' The meaning of the mysterious Algonquian numeral suffix was solved in a second. Looking back at Table 5 one can see that the numeral 1 is clearly the first part of the numeral 6, just as 2 is the first part of 7, and 3 is the first

part of 8 (with one change). So the way the Algonquians counted was 1, 2, 3, 4, 5, 1 on the left hand [= 6], 2 on the left hand [= 7], and 3 on the left hand [= 8].[3] Zuni *wešik:ʔa* consists of three morphemes: *w-eši-k:ʔa* 'this-hand-left,' all three of which have an Amerind ancestry. Demonstrative *w-* is widespread in Amerind, though it has become a Stage III article in Algic and Penutian. [A Stage III article is the final result in the evolution of demonstrative pronouns. Demonstrative pronouns often become definite articles (Stage I), which in turn may become a general article, combining both definite and indefinite functions (Stage II), and finally the original meaning is totally lost and the article merely indicates that a noun follows (Stage III). Stage III articles are often lost, so that in closely related languages one language may have retained the Stage III article which was lost in the other. Stage III articles are frequently fossilized on the following noun.] Examples of Proto-Amerind **w-* 'this' include Proto-Algic **we* 'this (nonpersonal), **wa* 'this (personal), **wo* 'this (restricted),'Wiyot *w-,* Proto-Algonquian *w-,* Zuni *w-,* Chitimacha *w-,* Atakapa *w-,* Huave *ow-* 'this,' Proto-Hokan **wi,* Proto-Uto-Aztecan *w-,* Paya *w-,* and Fulnio *owa* 'this,' *awa* 'that' (Greenberg and Ruhlen 2007).

The Proto-Amerind word for 'hand' was something like **ʔasi* or **ʔiš,* and is sometimes preceded by demonstrative **w-*. Examples include Proto-Algic **-VʔS,* Wiyot *w-éʔs,* Wichita *ʔi·s,* Proto-Costanoan **ʔiṣ·u,* Zuni *ʔasi* 'hand,' *w-eši-kkʔa* 'left hand' (Proto-Algonquian *w-a:šika* and Zuni *wešik:ʔa* differ in the first vowel, but Zuni *ʔasi* 'hand' shows that *a* was the original vowel in *wešik:ʔa* and was raised to *e* under the influence of the preceding *w-,* and Wiyot *w-éʔs* shows the same development), Proto-Atakapa-Chitimacha **wɛši* 'hand, arm,' Atakapa *wiš ~ woš* 'hand, finger,' Chitimacha *waši*, Natchez *ʔi·š-,* Kashaya *ʔiša* 'arm,' Yuma *hisa-* 'hand,' Wichita *ʔi·s,* Proto-Uto-Aztecan **waci* 'finger,' Paya *wiš-ka* 'arm,' and Tehuelche *haš* 'arm.'

Greenberg (1987) did not posit an Amerind etymology for 'left hand' but he did posit an Almosan etymology for left hand (Nootka *qaʦ*) and a Macro-Ge etymology as well (Proto-Ge **-keʦ*). Since Almosan is the most northern branch of Amerind and Macro-Ge the most southern, with languages mostly in Brazil, I figured that there were probably other examples of *qaʦ ~ -keʦ* in between (Ruhlen 1995b). I then surveyed words for 'left hand' in Amerind languages and found that there were indeed many instances of this root in the other branches of Amerind: Biloxi *ḳaskani* (probably borrowed from Cherokee), Yuchi *kaš'ó,* Onondaga *sketsíuka,* Cherokee *agasgani,* Proto-Hokan **Kisár-iK,* Lake Miwok *k'eṣíili* 'to be lefthanded,' Zuni *wešikkʔa,* Proto-Atakapa-Chitimacha **keʔs,* Atakapa *keʦ ~ kuʦ,* Chitimacha *ki·s,* Tunica *ʔokešta,* Yuki *me-kač',* Creek *kaskaná,* Proto-Tzeltal-Tzotzil **k'ešam,* Manare *kuʦ(-maya),* Pehuelche *kesna,* Yabamasa *gaʦo-ʦatìa,* and Timote *kučumiya.*

I do not believe that Proto-Algonquian *-w-a:šika* and Zuni *wešikkʔa* are historically related. There is no evidence that the Algonquians and the Zuni have ever been in contact, so borrowing can be ruled out, and their isolation in just Algonquian and Zuni lead me to believe that both are independent inventions using the same three Amerind roots. The lesson from Proto-Algonquian *-w-a:šika* is the same lesson we learned from Proto-Algonquian **pale·-neɬkwi.* There are problems

[3] See also the entry 359 $HAND_3$ **(w-)ʔiš ~ *(w-)ʔasi* in the Amerind Etymological Dictionary (Greenberg & Ruhlen 2007) [Ed.].

in families that can *only* be resolved by taking into account the external context, Amerind. Until the Americanists take off their blinders such problems will never be resolved.

THE ORIGIN AND EVOLUTION OF WORD ORDER

There are two kinds of linguistic diversity, genetic and typological. Genetic traits are those which involve the arbitrary sound/meaning correlation, such as words and affixes, and it is these that have been the focus of this book. Typological traits, on the other hand, do not involve the arbitrary sound/meaning correlation, and they include such things as consonants (p,m,h), vowels (i,a,u), pronouns (first-person singular, third-person plural), and word order (subject-verb-object, subject-object-verb). A first-person singular pronoun, which all languages have, is a typological trait because it has a meaning, but no associated sounds. The English first-person pronoun 'I,' however, is a genetic trait because it has a meaning (first-person pronoun) associated with sounds ('I'). As we have seen in this book it is genetic traits that allow us to trace language evolution since genetic traits indicate forms that are historically connected. Typological traits, however, are not necessarily historically connected.

No linguist doubts that it is possible to trace the origin and evolution of genetic traits back in time, though we have seen that how *far* back in time has been—and remains—a controversial topic. But what about typological traits? The general consensus among linguists is that it is not possible to trace back typological traits, or only for a very short period of time. There are two reasons for this. The first is the ever present evil of rapid linguistic change, which, for genetic traits, is supposedly limited to only 6,000 years, as we have seen. The situation for typological traits is even worse because there are very few variants of typological traits, unlike the numerous possible forms for a word. For the order of the adjective (A) and noun (N) in a language there are only two possibilities, AN and NA, and for the order of the subject (S), verb (V), and object (O) in a language there are only six, SOV, SVO, VSO, VOS, OVS, and OSV. Since any possible state, whether two or six, can change into any other state in a short period of time, it is therefore impossible to trace typological traits back in time.

Harris and Campbell (1995: 405) have claimed that "there is little or no evidence to support hypotheses that languages—or their syntax—are evolving in a single direction through non-renewable changes" and Fox (1995: 104) summarizes the current view as follows: "Syntactic reconstruction is a controversial area. . . . Indeed, there is a consensus among many scholars that it is difficult, if not impossible, to carry over into the field of syntax the methods—especially the Comparative Method itself—that have proved so successful in phonology."

Let us consider the problems involved in reconstructing word order. Today all six basic word orders are attested as the preferred word order in the world's languages, though at very different frequencies, as shown in Table 6.

Table 6: Distribution of the Six Possible Word Orders

SOV	50%	The man the bear killed. Japanese, Basque, Hindi, Navajo, Quechua
SVO	38%	The man killed the bear. English, Chinese, Finnish, Vietnamese, Zulu
VSO	8%	Killed the man the bear. Irish, Hawaiian, Maasai
VOS	2%	Killed the bear the man. Malagasy, Tzotzil, Chinook
OVS	1%	The bear killed the man. Hixkaryana, Ungarinjin
OSV	1%	The bear the man killed. Warao, Wik Ngathana

The main problem in reconstructing word order in a language family is well illustrated by the Indo-European family, in which different branches have different word orders, SOV in Indic, SVO in Albanian, and VSO in Celtic. The question of what the original word order was in Proto-Indo-European remains controversial even today. It is, however, possible to figure out the original Proto-

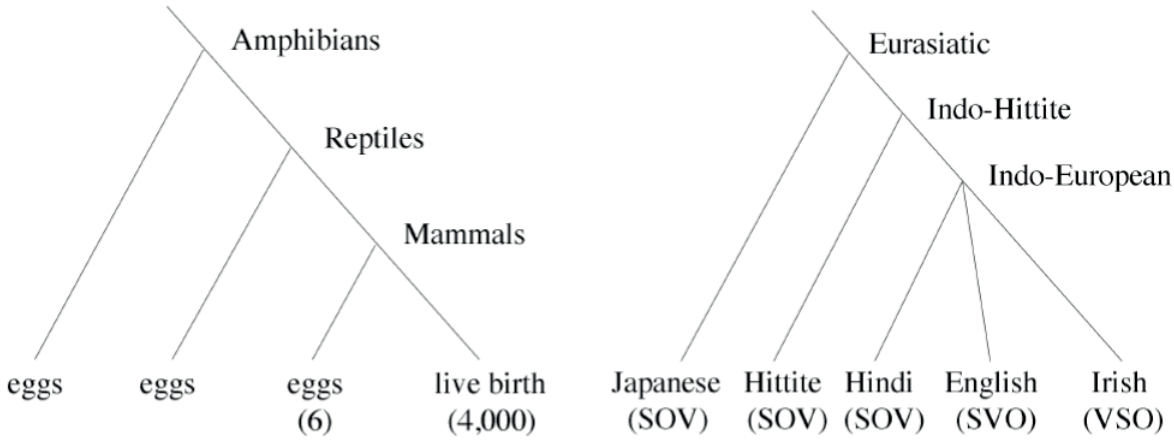


Figure 3. Outgroup comparison in biology and linguistics.

Indo-European word order if one uses outgroup comparison. We saw in this book that even though almost all mammals give live birth, and only a handful lay eggs, proto-mammal was an egg-layer because the nearest outgroups—reptiles and amphibians—lay eggs. In a similar fashion, as illustrated in Figure 3, Proto-Indo-European had SOV word order because the nearest outgroup, the

Anatolian family (which includes Hittite) was strictly SOV, as were more distant members of the Eurasiatic family.

Murray Gell-Mann and I have used such outgroup comparison to reconstruct the original word order for different language families (Gell-Mann and Ruhlen 2011), and we believe it is possible to reconstruct word order to very great time depths, indeed to the original word order in the language from which all extant languages derive. Our general conclusions are the following: (1) The word order in the original language was SOV. (2) Except for cases of diffusion, the direction of syntactic change has been for the most part SOV > SVO and, beyond that, SVO >VSO/VOS with a subsequent reversion to SVO occurring occasionally. (3) Diffusion, although important, is not the dominant process in the evolution of word order. (4) The two extremely rare word orders (OVS and OSV) derive directly from SOV.

In some families the reconstruction of the original word order poses no problem because all languages in the family have the same word order. The following families are strictly SOV: Basque, Caucasian, Burushaski, Yeniseian, Na-Dene, Kartvelian, Dravidian, Yukaghir, Altaic (with one exception due to diffusion), Ainu, Gilyak, Chukchi-Kamchatkan (with one exception), and Eskimo-Aleut. We see here that all six branches of Dene-Caucasian are strictly SOV, with the exception of Sino-Tibetan. However, of the 250 Sino-Tibetan languages all of them are SOV, with the exceptions of the Chinese dialects, Bai, and the small Karen family, which are SVO. It is generally assumed that Proto-Sino-Tibetan was SOV, a conclusion supported by the fact that all of its nearest outgroups are. All the evidence indicates that Proto-Dene-Caucasian was itself SOV.

What about Eurasiatic? We have seen that Proto-Indo-Hittite was SOV and Yukaghir, Altaic, Ainu, Gilyak, Chukchi-Kamchatkan, and Eskimo-Aleut were also SOV because all constituent languages still are. The other branch of Eurasiatic, Uralic, has languages that are SOV and SVO, but the phylogenetic tree given in Table 7 shows that SVO languages only appear at the lowest node in the family, and it is generally believed that these languages (Finnish, Estonian, Saami) have borrowed SVO word order from surrounding Indo-European languages. We can therefore conclude that Proto-Eurasiatic, like Proto-Dene-Caucasian, was SOV, as was the higher-level Nostratic family, which also includes Kartvelian and Dravidian, both uniformly SOV.

Table 7: Word Order in the Uralic Family

Uralic: SOV
- Samoyed: SOV
- Finno-Ugric: SOV
  - Ugric: SOV
  - Finnic: SOV/SVO

The Afro-Asiatic family, like Indo-Hittite, has three different word orders in its constituent branches, as seen in Table 8. This particular phylogenetic tree is taken from Ehret (1995) and is, as are all phylogenetic trees discussed here, based solely on genetic traits. We have simply added the word order after each branch. Once again one can see the progression SOV > SVO > VSO as

one moves from the highest node in the tree to lower nodes. As with Indo-Hittite, the three currently existing word orders—SOV, SVO, VSO—all derive from an original SOV.

Table 8: Word Order in the Afro-Asiatic Family

Afro-Asiatic: SOV
- Omotic: SOV
- Erythraic: SOV
  - Cushitic: SOV
  - Chado-Afro-Asiatic: SVO
    - Chadic: SVO
    - North Afro-Asiatic: VSO
      - Ancient Egyptian: VSO
      - Semito-Berber: VSO
        - Semitic: VSO

Let us turn now to the Amerind family in the Americas, where all six basic word orders are attested in different languages. However, SOV is the dominant word order in the Keresiouan, Hokan, Tanoan, Chibchan, Paezan, Andean, Macro-Tucanoan, Macro-Carib, Macro-Panoan, and Macro-Ge branches, and it has also been reconstructed for Proto-Uto-Aztecan, the languages of which exhibit SOV, SVO, VSO, and OVS word order. It seems very likely that the Proto-Amerind word order was SOV, which would agree with the word order in Amerind's closest relative, Eurasiatic.

In sub-Saharan Africa the situation is somewhat complex, but a case can be made that the original word order in Niger-Kordofanian, Nilo-Saharan, and Khoisan was SOV. Although SVO is by far the most frequent word order in Niger-Kordofanian, this merely reflects the fact that the hundreds of Bantu languages (a very low node in the tree as we saw in the text) are almost all SVO, but the higher levels of the tree tell a different story. Niger-Congo consists of two branches, Mande, which is exclusively SOV, and Niger-Congo proper, which is predominately SVO, but does have SOV in some of the higher nodes in the tree (Kru, Dogon, Gur), which, combined with the Mande outgroup, implies that Niger-Congo was in fact SOV. The other branch of Niger-Kordofanian—Kordofanian—has languages with SOV, SVO, and VSO word order, but the fact that its nearest outgroup, Niger-Congo, was most likely SOV suggests that Kordofanian too was originally SOV.

The situation is also complex in Nilo-Saharan, where SOV, SVO, and VSO are all found. However, Lionel Bender has investigated the distribution of the various word orders within Nilo-Saharan and he concluded that "the logical interpretation is that Nilo-Saharan was SOV and that an innovation to SVO spread through most of Satellite-Core. . . . VSO is also an innovation, found among neighboring parts of Surmic, Nilotic and also in Kuliak, which, as already noted, is surrounded today mostly by VSO languages" (Bender 2000: 59).

The distribution of word order in the Khoisan family is shown in Table 9. As can be seen, Khoisan consists of a South African Khoisan group, with three branches, and two outliers, Sandawe and Hadza. Since Sandawe, which has SOV word order, is the closest outgroup to South African Khoisan, one of whose branches is SOV, it seems likely that the word order in Sandawe-

South African Khoisan was SOV. The most divergent Khoisan language, Hadza, has SVO word order. Though we can not be sure that Proto-Khoisan had SOV word order, it would certainly not be incompatible with the data given in Table 9.

The Indo-Pacific family, found on New Guinea and surrounding islands, contains over 700 languages, almost all of which have SOV word order, except for some languages along the coast of New Guinea, which have borrowed SVO word order from neighboring Austronesian languages. It is not plausible that Indo-Pacific languages have been changing their word order over and over for a very long period of time and that today all of these languages have wound up with the same order, SOV. Indo-Pacific languages are all SOV today because they all preserve the initial state of word order unchanged. The Australian family contains languages with all six word orders and is known for its extraordinarily free word order. In some languages all six word orders yield grammatical sentences. Nevertheless, most specialists regard SOV as the most characteristic word order for the family as a whole.

Table 9: Word Order in the Khoisan Family

Khoisan:
- Hadza: SVO
- Sandawe–South African Khoisan: SOV
  - Sandawe: SOV
  - South African Khoisan: SOV
    - Central: SOV
    - Northern–Southern: SVO
      - Northern: SVO
      - Southern: SVO

The final family to be considered is the Austric family in Southeast Asia, which has four branches: Miao-Yao, Daic, Austroasiatic, and Austronesian. Miao-Yao and Daic are strictly SVO. Austroasiatic contains two branches, Mon-Khmer and Munda; Mon-Khmer is SVO and Munda is SOV. However, Munda languages are all spoken in India, where virtually all languages have SOV word order, and it is clear from internal evidence in Munda languages that they have borrowed SOV word order from other languages. It seems likely that the original word order in Austronesian was verb initial, VSO/VOS, though SVO is also found in certain branches. Proto-Austric was almost certainly SVO, with Austronesian having developed VSO/VOS word order. Austric is thus the only language family that cannot be traced back to an original SOV word order.

Our conclusions on the evolution of word order are shown in Figure 4. The two principal paths of evolution are first SOV > SVO and secondly SVO > VSO, as shown by the heavy arrows in Figure 4. The change SOV > SVO is unidirectional and SVO never reverts to SOV except by borrowing SOV from another language. Once a language arrives at SVO word order the most likely change, if one takes place, is to move to VSO word order. However, there are also cases of SVO evolving to VOS. VSO and VOS may evolve back and forth, and either may return to SVO. The two rare word orders, OSV and OVS, derive directly from SOV.

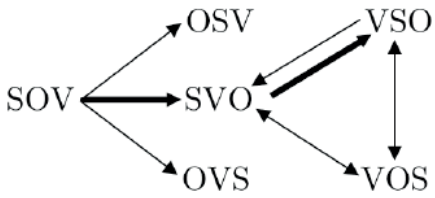


Figure 4. Evolution of word order

GENETICS AND LANGUAGE

The revolution in human genetics was really just beginning when I wrote this book in the early 1990's. The results of Cavalli-Sforza's work with classical genetic markers had recently appeared, as well as the first work using mtDNA, as I discussed in this book. The Y-chromosome had not yet appeared on the scene. The results of this work was a phylogenetic tree of human populations—a task that Cavalli-Sforza had begun in 1964—and an appreciation of the similarity between this tree, based on genes, and a similar tree, based on language. This language/genes correlation resulted in quite a controversy—even to this day—though I tried to show in this book that a language/genes correlation should be expected, since both language and genes provide independent windows on human prehistory and human prehistory only happened one way.

This does not mean, as Cavalli-Sforza stressed from the outset, that genes and languages always correlate. People can change their language quite quickly in a number of well-known ways. Hungarians speak a Uralic language, but their genes are basically European. The reason is that the Magyars were an Asian people who conquered Hungary around 1,000 years ago and imposed their language on the people already living there. The linguistic replacement was complete; the genetic contribution of the Magyars was, however, modest. The same is true of the Turks, who are linguistically Turkic, but genetically European. In both cases we have examples of what Renfrew called elite dominance and the linguistic consequences were the same.

On the other hand, genes may change in a population while the language stays the same. Basque is a classic example. Basque is the only remaining Paleolithic language in Europe, the other European Paleolithic languages having been replaced by Indo-European languages in the Neolithic expansion out of Anatolia, according to Renfrew, Cavalli-Sforza, and Dolgopolsky. The Basque genes, however, have over time come to resemble what are today general European genes, though the precise contribution of Paleolithic and Neolithic genes in this population is controversial. There are only a small number of genetic differences, such as the high incidence of the Rh-allele in the Rhesus blood system, that distinguish the Basques from (most of) the other European populations.

In this regard Basque is similar to the other members of Dene-Caucasian: Caucasian, Burushaski, Sino-Tibetan, Yeniseian, and Na-Dene. All of the branches of Dene-Caucasian appear to be genetically closer to surrounding populations than to each other. This is not surprising in that all six branches have been separated for a very long time and in contact with different populations. Just as the Basques are genetically closer to European populations, the Kets (the sole surviving

Yeniseian people) are closest genetically to the Evenki, a Tungus population, and the Na-Dene are closer to Amerind populations than to the other branches of Dene-Caucasian. Even within Sino-Tibetan the Han Chinese populations are genetically closer to surrounding non-Han Chinese populations than they are to each other.

The fundamental finding of this early work in human genetics was that all modern humans share a recent common origin in Africa. This had already been found using classical markers and mtDNA, but the Y-chromosome confirmed this in spades. There can no longer be any doubt that all modern humans alive today share a recent African origin. It should be noted, however, that this conclusion was first advocated by archaeologists.

It is also now clear that the greatest genetic diversity is found in Africa, which is not surprising since human evolution has been taking place in Africa for a much longer time than outside Africa, and greater time always implies greater genetic (and linguistic) diversity. The most divergent African population is the Khoisan, which appears to be a branch of the human family opposed to all other populations. In fact, there is greater genetic diversity *within* Khoisan than between Khoisan and other families (Knight *et al.* 2003). This suggests that clicks, which are essentially only found in Khoisan languages (ignoring borrowings into a few Bantu languages), were found in the original human population from which all modern people have evolved. Clicks appear to have been lost once and never to have been reinvented in languages deriving from this first non-click language. Clicks would thus represent an arrow of time that only goes in one direction: loss.

Genetics has also clarified the position of pygmies within the human population. As noted in this book pygmies today speak languages they have borrowed from neighbors (either Bantu or Nilo-Saharan), with their original languages having been lost, so there is no linguistic evidence for the position of pygmies within the human family. However, genetics indicates that pygmies, like the Khoisan, are quite divergent from other human populations.

The various paths that these people took upon leaving Africa 50,000 years ago have been the subject of intense research over the past fifteen years and a general outline, based on mtDNA and the Y-chromosome, has emerged. The maps showing the migration routes based on mtDNA and the Y-chromosome are not the same, which is hardly surprising since mtDNA traces female ancestry (everyone, male or female, inherits their mtDNA exclusively from their mother), while the Y-chromosome traces male ancestry exclusively since only males have a Y-chromosome, and the (pre-)history of men and women is not the same. An excellent summary of the current state of knowledge regarding mtDNA and the Y-chromosome may be found on the web at http://www.nationalgeographic.com/genographic. The state of linguistic evidence for human phylogeny can be found at http://ehl.santafe.edu/main and http://starling.rinet.ru/main.

The genetic evidence indicates that the first expansion out of Africa 50,000 years ago followed the South Asia coastline, reaching the New Guinea-Australia area (which was at the time one continent) around 45,000 B.P. Later there were migrations northward from southeast Asia as far as Japan, as well as a migration from south Asia into central Eurasia around 40,000 B.P. From central Eurasia there was a major bifurcation in the population, with part of it going west and entering Europe around 35,000 B.P., where they met the Neanderthals for the first time. The other branch

went eastward, through northern Eurasia, and eventually, around 13,500 B.P., entered the New World and colonized these two unoccupied continents in 1,000 years.

Finally, it should be noted that work with mtDNA and the Y-chromosome have not always confirmed the earlier results based on classical markers. An example regards the number of migrations to the New World, which I have argued in this book was three: Amerind, Na-Dene, and Eskimo-Aleut. This conclusion is strongly supported by the linguistic evidence, archaeological evidence, and classical genetic markers. However, work with mtDNA and the Y-chromosome has so far not confirmed this three migration hypothesis. In both cases different geneticists have favored one, two, or three migrations. Further research should resolve the current dilemma.

On the other hand, work on the Y-chromosome has been able to clarify the time of the first migration to the New World. Recently, Mark Seielstad *et al.* have found a mutation on the Y-chromosome (M242) that lies between two mutations that are known to have occurred in Asia (M45, M74) and a mutation that arose in the Americas (M3) (Seielstad *et al.* 2003). They have dated the M242 mutation to the range 15,000–18,000 B.P. and have argued that this date provides a fairly certain upper bound on the time of the first entry into the Americas.

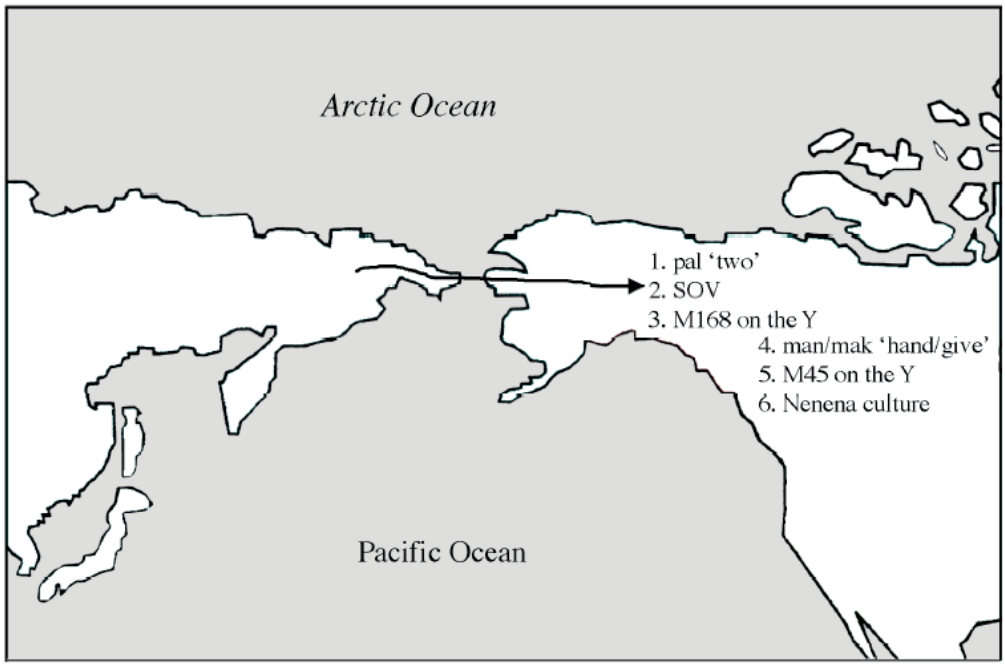
This date corresponds well with the archaeological dates, which support a first migration to the Americas at the time that the ice-free corridor opened up in Canada around 13,500 B.P. There is also other archaeological evidence that supports a late, rather than early, entry into the Americas. It is now believed that dogs were first domesticated in East Asia around 15,000 years ago (Wade 2006). Since the first Americans brought domesticated dogs with them they could not have left Asia before 15,000 B.P. or they would have had no dogs. It thus appears that the first entry into North America took place not long after the domestication of dogs. For unknown reasons all of these Asian dogs brought to America have gone extinct, replaced by European dogs who arrived much later.

With the explosion of genetic research over the past two decades, coupled with the advancement of our knowledge of linguistic taxonomy due primarily to the Nostraticists and Greenberg, and the much better evaluation of the archaeological evidence for human prehistory in the second half of the twentieth century, it is now possible to describe prehistoric human populations in terms of their specific linguistic, genetic, and cultural traits, a possibility that was already envisaged by Cavalli-Sforza in the 1950's. Let us look at three examples.

One of the most interesting prehistoric populations was the one that left Africa around 50,000 years ago and peopled the entire world, replacing the Neanderthals and various forms of *Homo erectus.* Map 2 shows a few of the traits that characterized this population. With regard to language they had the word *pal* 'two' and SOV word order. Men had the M168 mutation on the Y chromosome, as do all non-African men today. The archaeological record shows that they had an Upper Paleolithic culture.

The second example concerns the Asian population that first entered the Americas around 13,500 years ago and peopled the New World in 1,000 years. They brought with them *pal* 'two,' SOV word order and the M168 mutation on the Y chromosome, all of which had arisen in Africa, but this Asian population also brought with them traits that had arisen in Asia. These traits of Asian

origin include the morphologically related words *man*/*mak* 'hand/give,' the M45 mutation of the Y chromosome, and the Nenena culture, as we see on Table 4.



Map 2. The out-of-Asia migration, 13,500 BP

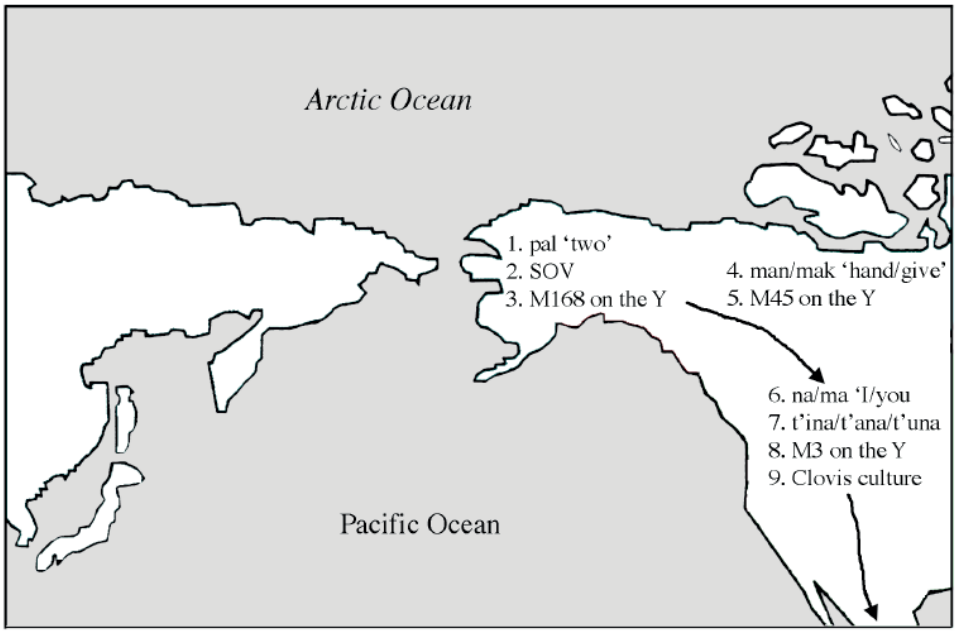
The final population is the Amerind population itself. This population retained the African traits *pal* 'two,' SOV word order, and the M168 mutation; it also retained the Asian traits *man*/*mak* 'hand/give' and the M45 mutation. The Amerind population, however, added specifically American innovations, among which were the *na/ma* 'I/you' pronoun system, the complex *t'ina/t'ana/t'una* 'son/child/daughter' paradigm, the M3 mutation on the Y-chromosome, and Clovis culture in North America, as shown on Map 3. We have thus arrived at a point in time where human prehistory can be investigated in a manner that would have been inconceivable a few decades ago and it is clear that this multidisciplinary project will be a major focus of scientific investigation in the coming years.

CONCLUSION

Although people who look morphologically like modern humans appeared in Ethiopia almost 200,000 years ago, they did not develop modern human behavior for another 150,000 years. Then, suddenly, around 50,000 years ago modern human behavior emerges in Africa for the first time and we have seen the fundamental changes that took place at this time: (1) tools made from ivory, shells, and bones, not just stone, (2) these tools change in style rapidly in both time and space, (3) art first appears, (4) spatial organization in houses, (5) religious burials, and (6) fishing begins.

The enigma is why all of these fundamental changes happened simultaneously. Do all of these disparate changes have something in common? It has been suggested that the underlying cause of

these changes was the appearance of symbolic thought in the form of fully modern human language.



Map 3. The Amerind migration through the Americas, 13,500–11,000 BP

Genetic evidence indicates that all people living today are descendants of a small East African population, perhaps as few as a thousand people, that lived around 50,000 years ago. Yet this small population succeeded in replacing all the people who had lived outside Africa for over a million years, as well as the other African populations that existed at that time.

The reason why this small African population succeeded in replacing everyone else was simply that they had developed the first fully modern language, which had an adaptive value so great that it allowed them to conquer the entire world in a short period of time, eliminating everyone else.

If there are clear genetic traces of this original small population, why aren't there linguistic traces as well? The answer, of course, is that there are, they have been recognized for over a century, and yet, as we have seen, many experts deny they exist or could exist. That's why I wrote this book.

The paths that this original population took in peopling the world, and the times of the bifurcations that led people in different directions, have been the subject of intense investigation by archaeologists, geneticists, and linguists over the past two decades. It now appears that the earliest path out of Africa led to New Guinea-Australia by 45,000 B.P. Later migrations seem to have followed the coast northward as far as Japan, while a different migration went north into central Eurasia, from which there were migrations both east and west. Another migration in this great human expansion brought the first people to the New World around 13,500 years ago. The final migration led to the occupation of Oceania by the Austronesian family, with New Zealand the final island occupied, around 800 years ago, in this vast migration.

There have of course been many later migrations that are now superimposed on the original migratory paths. One example would be the later Eurasiatic expansion overwhelming the earlier Dene-Caucasian expansion and leaving the six remaining Dene-Caucasian branches as isolated islands in a Eurasiatic sea. Later migrations are an obvious complicating factor in figuring out the relationships among different populations and the paths they took to arrive where they are today. While we now have a rough map of these migratory paths, it is clear that this map will continue to improve with further research in genetics, linguistics, and archaeology.

In Chapter 7 I gave a phylogenetic tree that was intended to represent the genealogical structure of the human population based on genetic and linguistic evidence. I would now revise that tree in a number of ways, as indicated in Figure 5. While I believe this tree represents an improvement over the original, it is by no means the final answer.

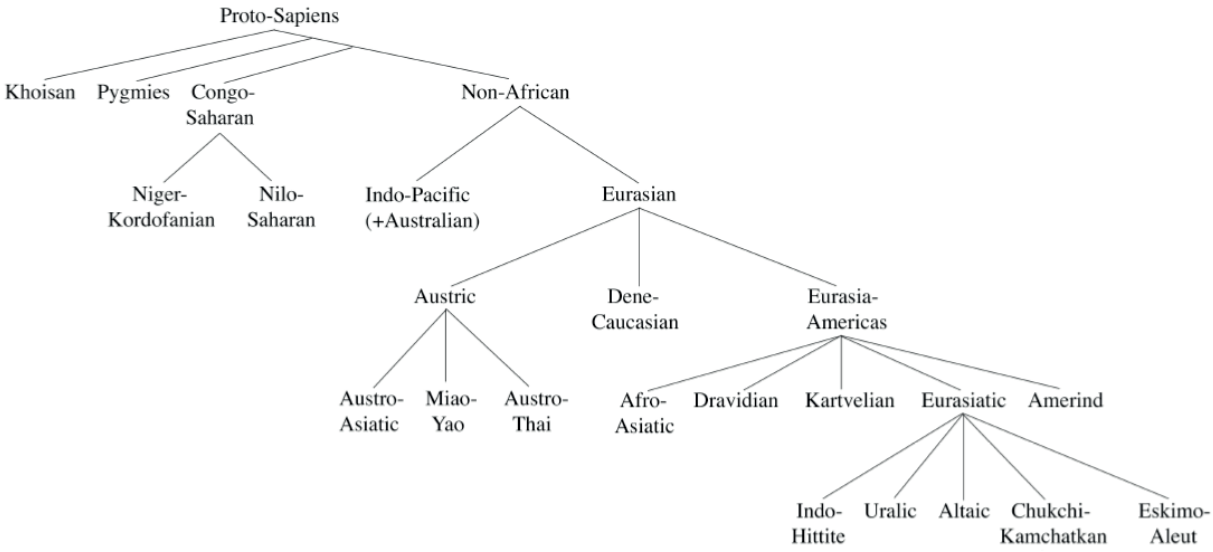


Figure 5. The genealogical structure of the human population.

REFERENCES

Bancel, Pierre J. and Alain Matthey de l'Etang. 2002. "Tracing the Ancestral Kinship System: The Global Etymon KAKA," *Mother Tongue* 7: 209–43.

———. 2004. "The Global Distribution of (P)APA and (T)ATA and their Original Meaning," *Mother Tongue* 9: 133–69.

Bellwood, Peter. 1991. "The Austronesian Dispersal and the Origin of Languages," *Scientific American* (July): 88–93.

Bender, Lionel M. 2000. "Nilo-Saharan," in *African Languages: An Introduction,* ed. by Bernd Heine and Derek Nurse, Cambridge, Cambridge University Press, 43–73.

Campbell, Lyle. 1994a. "Inside the American Indian Language Classification Debate," *Mother Tongue* [Newsletter] 23: 47.

———. 1994b. "Problems with the Pronouns in Proposals of Remote Relationships among Native American Languages," in *Survey of California and Other Indian Languages,* ed. by Margaret Langdon, Department of Linguistics, Univ. of California, Berkeley, pp. 3, 9.

———. 1997. "Amerindian Personal Pronouns: A Second Opinion," *Language* 73: 339–51.
Campbell, Lyle and William J. Poser. 2008. *Language Classification.* Cambridge, Eng.: Cambridge University Press.
Cavalli-Sforza, Luigi Luca. 2000. *Genes, Peoples, and Languages.* New York: North Point Press.
Diamond, Jared. 1992. *The Third Chimpanzee.* New York: Harper Collins.
Ehret, Christopher. 1995. *Reconstructing Proto-Afro-Asiatic (Proto-Afrasian).* Berkeley: University of California Press.
Fox, Anthony. 1995. *Linguistic Reconstruction.* Oxford: Oxford University Press, p. 105.
Gell-Mann, Murray and Merritt Ruhlen. 2011. "The Origin and Evolution of Word Order," *Proceedings of the National Academy of Sciences* 108: 17290–17295.
Greenberg, Joseph H. and Merritt Ruhlen. 2007. *An Amerind Etymological Dictionary.* Stanford: Department of Anthropological Science. Available on the web at https://starlingdb.org/Texts/Ruhlen_AED5.pdf
Hamp, Eric P. 1975. "On Zuni-Penutian Consonants," *International Journal of American Linguistics* 41: 310–12.
Harris, Alice C. and Lyle Campbell. 1995. *Historical Syntax in Cross-Linguistic Perspective.* Cambridge: Cambridge University Press, p. 343.
Hodgson, Brian H. 1848. ''On the Chepang and Kusunda Tribes of Nepal," *Journal of the Asiatic Society of Bengal* 17: 650–58.
———. 1857. "Comparative Vocabulary of the Languages of the Broken Tribes of Nepal," *Journal of the Asiatic Society of Bengal* 26: 317–32.
Klein, Richard. 1999. *The Human Career.* Chicago: University of Chicago Press.
Knight, Alec, *et al.* 2003. "African Y Chromosome and mtDNA Divergence Provides Insight into the History of Click Languages," *Current Biology* 13: 464–73.
McWhorter, John. 2001. *The Power of Babel.* New York: Times Books.
Newman, Stanley. 1964. "Comparison of Zuni and California Penutian," *International Journal of American Linguistics* 30: 1–13.
Nichols, Johanna. 1992. *Linguistic Diversity in Space and Time.* Chicago: University of Chicago Press.
Nichols, Johanna and David A. Peterson. 1996. "The Amerind Personal Pronouns," *Language* 72: 336–71.
———. 2005. "Personal Pronouns," in *The World Atlas of Language Structures,* ed. by Martin Haspelmath, Matthew S. Dryer, David Gill, and Bernard Comrie, Oxford, Oxford University Press, 546–553.
Pagel, Mark. 2000. "Maximum Likelihood Models for Glottochronology and for Reconstructing Linguistic Phylogenies," in *Time Depth in Historical Linguistics,* ed. by Colin Renfrew, April McMahon and Larry Trask, Cambridge, Eng., McDonald Institute for Archaeological Research, Vol. 1, 189–207.
Ruhlen, Merritt. 1994a. *The Origin of Language.* New York: John Wiley.
———. 1994b. "First- and Second-Person Pronouns in the World's Languages," in *On the Origin of Languages: Studies in Linguistic Taxonomy,* Stanford, Stanford University Press.
———. 1995a. "Proto-Amerind Numerals," *Anthropological Science* 103: 209–25.
———. 1995b. "Proto-Amerind *QETS' 'Left (Hand),' *Mother Tongue* [Newsletter] 24: 69–70.
———. 1998. "The Origin of the Na-Dene," *Proceedings of the National Academy of Sciences* 95: 13994–96.

Seielstad, Mark *et al.* 2003. "A Novel Y-Chromosome Variant Puts an Upper Limit on the Timing of First Entry into the Americas," *American Journal of Human Genetics* 73: 700–05.

Trask, R. L. 1996. *Historical Linguistics.* London: Arnold.

Vajda, Edward. 2009. "A Siberian Link with Na-Dene Languages," in *Archeological Papers of the University of Alaska,* Volume 6, ed. by Jim Kari and Ben Potter, Fairbanks.

Wade, Nicholas. 2006. *Before the Dawn.* New York: Penguin Press.

Whitehouse, Paul, Timothy Usher, Merritt Ruhlen and William S-Y. Wang. "Kusunda: An Indo-Pacific Language in Nepal," *Proceedings of the National Academy of Sciences* 101: 5692–5695.

ON THE PRONOUN ROOTS *N* '1SG' AND *M* '2SG' IN THE NATIVE LANGUAGES OF THE AMERICAS AND THEIR HISTORICAL MEANING[1]

PIERRE J. BANCEL[†] AND JOHN D. BENGTSON [†, ‡]

In memoriam **MERRITT RUHLEN**
(May 10, 1944 – January 29, 2021),
bold capital linguist

Zamponi (2017) gathered data on pronoun roots n *'1sg' and* m *'2sg' from an exhaustive sample of 172 Native language groups and isolates of the Americas. Then, discarding all of those taxa with only one member of the pair of roots and considering only those with the full pair, he improperly deems them essentially located in the western USA. His hazy conclusion is that these localized roots might derive either from a genetic link between a few groups, or from convergence due to contact. Here we argue that Zamponi's own data show that the two pronoun roots are very widely distributed in North and South America, while their geographic density is closely correlated with that of particular taxa. We also address the argument, asserted by Zamponi following various other linguists, that in most languages of the world pronoun roots use a restricted set of consonants, which has led some of them to explain this convergence by, well, "convergence," which is in itself not an explanation unless one finds the driving force behind it – a driving force that has never been discovered for the concerned set in more than a century. We conclude that the pronominal roots* n *'1sg' and* m *'2sg' found by Zamponi himself in 95 of his 172 taxa of the Americas, are a strong indicator of a genetic relationship between these 95 taxa and hence giving credence to the Amerind hypothesis, a concept otherwise fully supported by genetic studies.*

I. INTRODUCTION

The Italian linguist Raoul Zamponi has published in the *Italian Journal of Linguistics* (2017) a substantial paper in which he tackles, in his turn, the question of the pronominal root consonants *n* '1st person singular' (henceforth 1sg) and *m* '2nd person singular' (2sg), respectively, in the native languages of the Americas and how to interpret their distribution from a historical standpoint. These two pronoun roots are of particular significance, since their apparent wide distribution over the Americas has been considered a strong argument by all linguists supporting a common origin

[1] This paper is a follow-up to Raoul Zamponi's "First-person *n* and second-person *m* in Native America: A fresh look."

[†] *Association for the Study of Language in Prehistory;* [‡] *Evolution of Human Languages Project (Santa Fe Institute). Correspondence should be addressed to first author at pierrejbancel (at) hotmail (dot) com.*

of most languages of Native American peoples, with the exception of the Na-Dene and Eskimo-Aleut languages of northwestern North America.

This unity of origin has been envisioned and progressively refined for more than a century by several prominent linguists (Greenberg 1987; Greenberg *et al*. 1986; Greenberg & Ruhlen 2007; Ruhlen 1994a, 1994b; Sapir 1918; Swadesh 1954, 1960; Trombetti 1905). All these authors observed that the pronominal root consonants *n* '1st person' and *m* '2nd person' were widely distributed in Native American languages, a fact which seemed to them a powerful argument for a unique origin of the groups where they were found, given the extreme stability of personal pronouns over time in most languages of the world. For instance, Greenberg (1987: 48-49) claimed that "it would probably be easier to enumerate *[Native American languages other than Na-Dene and Eskimo-Aleut]* where *nV*- and *mV*- are not found than where they are."

Others, to the contrary (Campbell 1988, 1994a, b, 1997; Chafe 1987; Goddard 1987, 1990; Golla 1987; Matisoff 1990) have denied that it was possible to even hint at such a unity of origin – be it real or not – given the antiquity of the peopling of the continent and the pace of language evolution, allegedly making it impossible to retrieve any significant trace of an ancient language in its descendants after a few millennia. Campbell (1994b) went even as far as to write a paper entirely dedicated to "putting pronouns in proper perspective in proposals of remote relationships among Native American languages," concluding that pronouns do not have such a great historical significance after all.

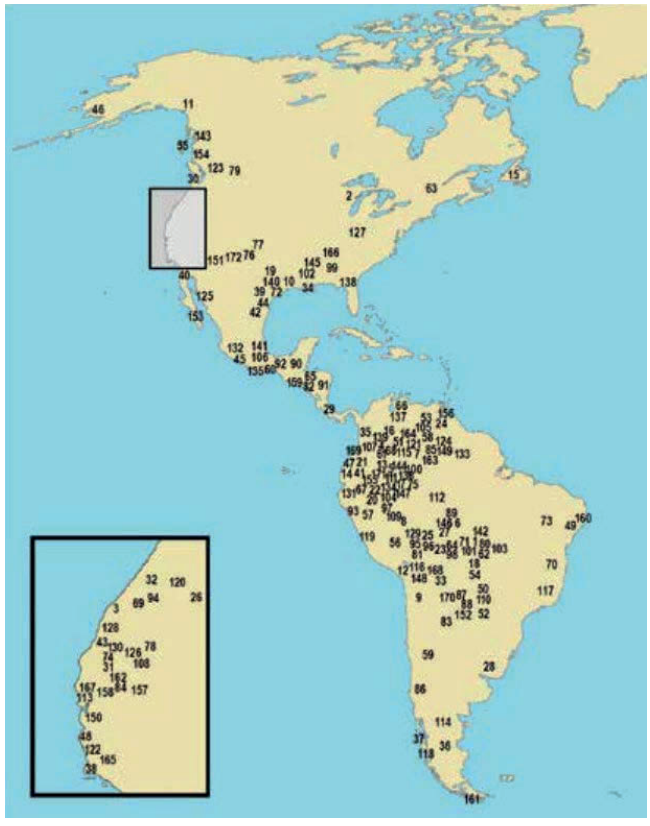


Figure 1. Zamponi's Map 1.

Location of the proto-languages of the linguistic families and of the language isolates of the Americas. Proto-languages and language isolates are identified by the same numbers in Table 2 in Zamponi's Appendix A.

Zamponi must be praised for having collected, like nobody had done before him, and presented in an orderly fashion an amount of data that seems to approach exhaustiveness in the present state of our knowledge of Native American languages. He performed the heavy job of looking for these pronominal roots in all of the 172 language families of North (including Central) and South America, as defined by the online site Glottolog 3.0 (Hammarström *et al.* 2017, as of writing available in version 4.3), amounting to 89 families comprising at least two languages and 83 so-called isolates, *i.e.* languages not

widely recognized as related to any other language and thus counted each as a separate taxon in and of itself. Zamponi opportunely locates each of these 172 taxa by numbers 1 to 172 on his Map 1 (Fig. 1 above), referring to Table 2 of his Appendix A for identification.

Zamponi's data constitute a powerful basis enabling us to assess hypotheses previously put forward by linguists, be it a common origin of the languages in which they are found, *à la* Greenberg and Ruhlen, or the impossibility of drawing any firm conclusion from them, Campbell- or Goddard-style, or the "Pacific Rim hypothesis" put forward by Johanna Nichols (1994) and hotly defended by her (Nichols & Peterson 1996, 1998), mainly against Campbell (1994b, 1997), or, as Zamponi himself does, restricting their distribution to Western North America and invoking a possible "convergence" phenomenon to explain their high frequency in this area, while ultimately finding the question "not answerable." Zamponi's conclusion is worth quoting and warrants careful pondering:

> … *[The* n *:* m*]* paradigm, observed through the proto-languages of all multi-member families and all isolates of the New World we know, appears essentially as a North American phenomenon that cannot be circumscribed precisely to a single area and is deeply-rooted only in the southern half of the Northwest Coast and in the neighboring California (Map 10). Not being possible to establish whether the conforming families of North America constitute a single areal cluster or not, it is impossible to draw a unique and sure conclusion about the spreading of the *n : m* paradigm in the continent. Limited to the southern half of the Northwest Coast and California, where the density of *n : m* is strikingly high, and, very likely, the close Southwest (to the presumed Uto-Aztecan Urheimat), we may suppose that the pronominal paradigm is a clear indicator of prehistoric contacts. But the question of whether these contacts have the form of a hypothetical proto-language or of a convergence between (perhaps few or very few) formerly distinct language branches is not answerable based on just the present-day knowledge of the history of the linguistic families of these areas. (p. 225)

In our present follow-up to Zamponi's fresh look at this *n* '1sg' / *m* '2sg' paradigm, we will argue that his interpretation of his data, and the unclear conclusions he draws from them, do not accurately account for the abundant linguistic facts he has gathered. While doing so, we will address the negative stance most consistently championed by Lyle Campbell. We will address Johanna Nichols' Pacific Rim hypothesis in another article. Welcoming again Zamponi's library work, we will highlight here the several ways in which his analysis strays from his own data.

II. DISCUSSION

1. Distorted frequency of occurrence of each of the two pronoun roots

Zamponi carefully documents occurrences of either *n* '1sg' and *m* '2sg' in the 172 language groups and isolates – *i.e.* 172 linguistic taxa – that he recognizes in the Americas, for the former under the condition that they be either reconstructed for the taxon or, absent an existing reconstruction, clearly attested among its constituent languages. He then indicates on his Map 4 (reproduced on Fig. 2 below), at the corresponding spots of Map 1, the taxa having either *n* '1sg,' or *m* '2sg,' or both, making it easy to assess their real importance among Native American languages. A short description of these pronouns in each concerned language, with full bibliographical references, is also given in Zamponi's Appendix B.

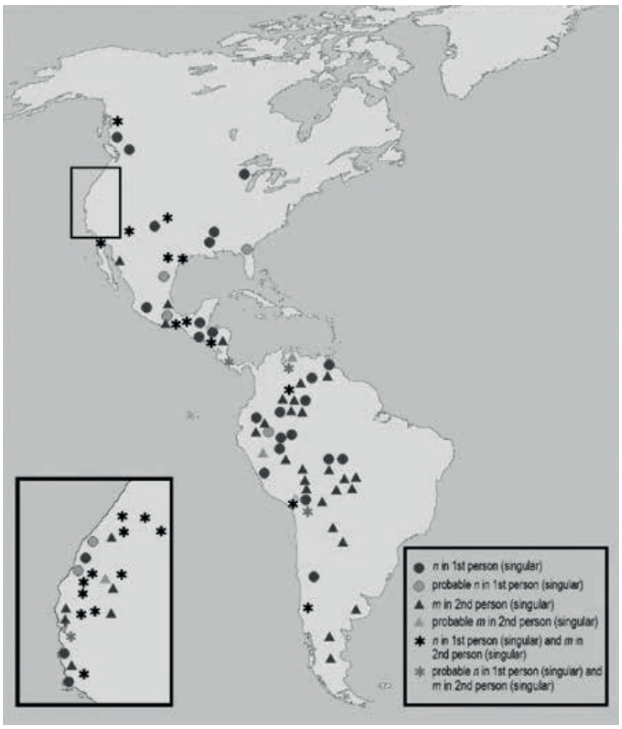


Figure 2. Zamponi's Map 4.

Location of either *n* '1sg" and/or *m* '2sg' in the proto-languages and language isolates of the Americas. See Zamponi's Map 1 above and his Appendices A for identification of taxa and B for data and their sources.

However, in the ensuing discussion of these facts, Zamponi quite unexpectedly takes into account only those languages exhibiting both members of the paradigm, without giving an explanation for this restriction. As far as historical linguistics is concerned, however, there is no reason why only languages with both the concerned pronoun roots should be given consideration. In the course of history, each member of the pair may have been replaced by another word, or have evolved into another sound, or even simply vanished, without the other one changing a bit.

A conspicuous example thereof is English, which almost completely lost in historical times its 2sg pronoun *thou*, which it had inherited from Proto-Indo-European (PIE) **tu* through Proto-Germanic **þū*, while it faithfully kept until today both *I*, inherited from PIE **eg^{h}om* '1sg, nominative,' and *me*, inherited from PIE **me* '1sg, accusative.'

The Middle English 2sg forms *thou*, *thee*, *thy* and *thine* were replaced by the 2pl forms *you*, *your* and *yours*, respectively, with the generalization of *vouvoiement* – introduced into English as a respectful form of address in imitation of French by the French-speaking nobility who were ruling over England after its conquest by William I the Bastard, Duke of Normandy, in 1066. However, the near complete disappearance[2] from English of PIE **t-* '2sg' pronouns did not in the least impair the descendants of **m-* '1sg', and, in spite of this loss, English nonetheless remains a *bona fide* Indo-European language.

In historical linguistics, each form of a paradigm is thus entitled to due consideration, on pain of losing sight of very real relationships between languages. Therefore, in leaving aside the many Native American languages with only *n* '1sg' or *m* '2sg,' Zamponi severely undercuts the significance of his data. Languages with these full pairs appear on his Map 4 as black (when Zamponi is fully assured of their presence) or grey asterisks (when he is less sure of them – although upon examining his report for these less certain taxa in his Appendix B, one may feel he could well have counted them as valid, as in every case his rationale is that he has only partial information, and no contrary information). He then observes that 23 (black asterisks) to 27 (adding 4 grey asterisks) of

[2] *Thou, thee, thy, thine* survived until recently in some local dialects of English. They also survive in Standard English in a few fixed phrases like *holier than thou*, as well as in religious usages, notably in the King James Bible, and linguists sometimes use them to gloss 2sg forms from other languages.

his 172 taxa of the Americas (or 13.4% to 15.7% of them), exhibit the full pair *n* '1sg' / *m* '2sg', and that this proportion is unevenly distributed over North and South America.

South America, with only 3 to 5 of its 102 taxa having the full set (2.9% to 4.9%), barely reaches the global occurrence level of 7%, found by Matthew Dryer (quoted in Campbell 1997b) in the 252 non-Native American languages of a worldwide sample of 333 languages.

In contrast, the full pair of pronoun roots is found in 20 to 22 (28.5% to 31.4%) of the 70 North American Native taxa. Looking at their respective locations on Map 4, Zamponi finds that a great part of these some 30% of North American taxa with the *n* '1sg' / *m* '2sg' pronoun pair are located in a restricted area of the continent. Hence, he concludes that,

> [l]imited to the southern half of the Northwest Coast and California, where the density of *n : m* is strikingly high, and, very likely, the close Southwest (to the presumed Uto-Aztecan Urheimat), we may suppose that the pronominal paradigm is a clear indicator of prehistoric contacts.

How does this conclusion fit with Zamponi's own data? Surprisingly, to say the least. As documented in detail in his Appendix B (pp. 32-40) and apparent from his Map 4 reproduced in Fig. 2 above, Zamponi found direct or near-direct apparent reflexes of *n* '1sg' in 50 taxa of North and South America without doubt (black circles on his Map 4, plus of course black asterisks), plus 9 others where he is less assured of its presence (gray circles and grey asterisks). These 50 to 59 taxa with *n* '1sg' represent 29.1% to 34.3% of his total 172 taxa.

As for *m* '2sg' (Appendix B: 40-49), Zamponi finds it in 58 secure taxa (black triangles and asterisks) plus 5 less certain ones (grey triangles and asterisks), or from 33.7% to 36.6% of his 172 taxa.

According to Zamponi's detailed linguistic data from his Appendix B, closely corresponding to Map 4, of his 70 North American taxa, 51 (72.9%) exhibit at least one of the two roots. More precisely, there are 39 taxa (55.7%) with *n* '1sg' and 32 (45.7%) with *m* '2sg,' including 22 taxa (31.4%) with both. These are certainly strikingly high figures for either root as well as for the full set of them. In South America, 44 of his 102 taxa (43.1%) exhibit at least one of the roots, while 19 taxa (18.6%) display *n* '1sg' and 29 (28.4%) have *m* '2sg,' including only 5 taxa (4.9%) with both.

Limiting the distribution of the full set *n* '1sg' / *m* '2sg' to a relatively small area of western North America is certainly bold, as nearly all taxa of Central America (which he counts with the North), or 12 out of 13 (92.3%), have at least one of the two roots. In turn, the lower density of the two roots in the eastern part of North America around, and in particular to the north of the Gulf of Mexico directly follows from the overall paucity of taxa having their homeland in this region, as appears evident from a look at his Map 1. While the full pair of pronoun roots appears in only 2 (11.8%) of these 17 eastern North American taxa, *n* '1sg' is found in 7 (41.8%) of them.

If we now turn to South America, in an apparently stark contrast with North America, out of the 102 taxa considered by Zamponi, only 3 to 5, or 2.9% to 4.9%, exhibit both pronoun roots. Nevertheless, due to a highly differentiated distribution of *n* '1sg' and *m* '2sg' over South American taxa, 44 of these 102 taxa (43.1%) exhibit at least one of the roots. This is not taken into account by Zamponi, who contents himself with discarding the unimpressive 4.9% of taxa with

both roots, and leaves unexplained the strikingly high occurrence levels of each individual root *n* '1sg' and *m* '2sg' separately. Actually, if one looks at his data in greater detail, it appears that in South America the root *n* '1sg,' found in 18 taxa out of 102 (17.6%), is only frequent, although *not extremely* frequent, while *m* '2sg,' found in 29 taxa (28.4%), approaches the levels found in North America.

Thus, to conclude, as Zamponi does, that *n* '1sg' and *m* '2sg' are essentially "limited to the southern half of the Northwest Coast and California" is a *non sequitur*. In North (and Central) America, these two pronoun roots are ubiquitous wherever there are taxa. And even if North America certainly wins the contest, South America also has plenty of them, as counting only those coming in pairs is not defensible. Zamponi's own data show that both pronoun roots *n* '1sg' and *m* '2sg' are frequent everywhere in both North and South America, and at least one of them appears in 95 (55.2%) of his 172 taxa.

2. Distorted spatial distribution

Another problem with Zamponi's analysis, closely linked to the preceding one, is that of the spatial distribution of *n* '1sg' / *m* '2sg' over the Americas. Taking into account only the taxa displaying both pronominal roots, he claims that they appear

> essentially as a North American phenomenon that cannot be circumscribed precisely to a single area and is deeply-rooted only in the southern half of the Northwest Coast and in the neighboring California.

It is just this restriction to considering only the full pairs of pronouns that allows him to make such a preposterous statement. The several dozens of *n* '1sg' or *m* '2sg' pronouns in eastern and southern North America, plus all of the 44 of his 102 South American taxa where he himself found one of these pronouns or both, are left in limbo.

A comparison of Zamponi's Map 4 with his Map 1 (reproduced in Figs. 1 and 2 above) shows that the spatial density of the two pronoun roots globally follows that of taxa throughout North and South America: high densities are found in North America along the Pacific Coast as well as around the Gulf of Mexico and in Central America, and in South America to the east of the Andes in two dense pockets north and south of the Amazon River.

Therefore, the mere geographic extent of the distribution of the two pronoun roots over the Americas excludes an explanation in terms of areal diffusion.

3. Languages do not "construct their pronominal systems" – they inherit them

A third problem arises from Zamponi's reckless use of the phrase "languages of the world construct their pronominal systems":

> [A]s indicated by Campbell (1994: 3-4, 1997a: 344-346), Goddard & Campbell (1994: 196-197), Willerman (1994: 38), Gordon (1995), Sasse (2015: 197), and other authors, most languages of the world construct their pronominal systems around a small set of basic consonants (including *n*, *m*, *t*, *k*, and *s*).

To his list of references, Zamponi could have added Bancel & Matthey de l'Etang (2010), a study of a worldwide survey of personal pronouns published by Ruhlen (1994c), in which we found that

the very same set of consonants (*n*, *m*, *t*, *k*, *s*) represented nearly three fifths (59.2%) of pronoun roots in the world's languages, as appears from Table 1 below.

Root consonant	1sg	1pl	2sg	2pl	Total	% of total
m	11	19	19	17	66	19.0
n	20	19	12	5	56	16.1
k	17	6	11	5	39	11.2
t	5	7	10	8	30	8.6
s	4	1	7	3	15	4.3
Subtotal	***57***	***52***	***59***	***38***	***206***	***59.2***
w	4	1	9	0	14	4.0
ŋ	8	1	4	0	13	3.7
j	7	3	0	1	11	3.2
ʔ	5	1	3	2	11	3.2
h	7	1	2	0	10	2.9
p alternating with *m*	2	2	4	1	9	2.6
i	6	0	1	0	7	2.0
u	4	0	2	0	6	1.7
b alternating with *m*	2	1	0	0	3	0.9
p	0	0	2	1	3	0.9
b	0	1	0	2	3	0.9
v, d, z, ð, r, tˡ, sw, ɬ, l, š, ž, č, ǰ, šʷ, šʲʷ, lž, ch, ñ, g, kh, kw, x, xw, ɢ, ħ, a	18	22	16	14	70	20.1
Total	**116**	**81**	**96**	**55**	**348**	**100.0**

Table 1. Number and percentage of occurrences of each root sound in Ruhlen's (1994c) worldwide lists of 1st and 2nd person singular and plural pronouns. (Adapted from Bancel & Matthey de l'Etang 2010.)

Root sounds are displayed in decreasing order of frequency. A few complex, isolated forms have been discarded. In CV, CVC, VC and VCV forms, C_1 is considered the root; in VV forms, V_1 is taken to be the root. Alternate forms with different C_1's have been counted under each of these, but alternate forms with the same C_1 have been counted only once. Symbols *j* and *y* most of the time transcribe a palatal glide and have been subsumed under *j* in the table. For both *b* and *p* root consonants, a separate count is given of forms alternating with *m* forms of the same pronoun in the same family. Each one of the 40 sounds appears on average 8.7 times in the 348 pronominal forms listed by Ruhlen, or 2.5%.

We surely would not deny that languages worldwide offer a very short list of prominent pronoun root consonants, exactly those claimed in his turn by Zamponi: *n, m, t, k*, and *s*, repeatedly documented by many linguists of all stripes. However, this list does not show that languages would "construct their pronominal systems around a small set of basic consonants." Let us have another look at this list. Do you notice anything? Anything that might be lacking, perhaps? Actually, there are absences that reveal more than presences: let us be structural just for a moment (Table 2).

	Labials	Alveodentals	Velars
Nasal stops	*m*	*n*	
Voiceless oral stops		*t*	*k*
Fricatives		*s*	

Table 2. Structural distribution of the most frequent pronominal root consonants.

The relatively less frequent fricative *s* stands apart. Meanwhile, among stops, the velar nasal *ŋ* and the unvoiced labial oral stop *p* are missing. How could these highly frequent consonants in the world's languages (11th most frequent for *ŋ* and above all 4th most frequent consonant for *p*) be nearly absent from the world stock of pronominal roots if these roots descended from many different, unrelated sources?

In the first two rows, both the velar nasal *ŋ* and the unvoiced labial oral stop *p* are missing. In Table 1, the velar nasal *ŋ*, which is the 11th most frequent consonant in the world's languages, as it is found in 63% of them (Moran & McCloy 2019), is present in only 13 pronouns (3.7%) of Ruhlen's global compilation. Above all, what about *p*? It is the fourth most frequent consonant in the world's languages, present in 86% of them, behind *m* (96% of languages), *k* (90%), and *j* (90%, also rather poorly represented in Ruhlen's pronominal compilation, at 3.2%). It is hardly conceivable that *p* might be such a rare pronominal root consonant, at 0.9% (or at most 3.5%, counting items alternating with *m* in the same language, possibly hinting at denasalization of a few of the so common *m-* pronouns), as compared with 19.0% for *m*, 16.1% for *n*, or even 11.2% for *k*.

Franz Boas, already in 1917, may have remarked these phonological gaps in the most frequent pronoun roots. Unable to find any phonetic (articulatory, acoustic or auditory) rationale – since there is none –, he was probably far off the mark when he ruminated

> how far the frequent occurrence of similar sounds for expressing related ideas (like the personal pronouns) may be due to *obscure psychological causes* rather than to genetic relationship. (p. 5, our italics)

A full century later, these psychological causes remain as obscure as they were in Boas' time, and no new hint has appeared that they should exist at all. Also, the reason why these causes would drive speakers of so many Eurasiatic (Bomhard 2012: 282-285; Greenberg 2000) or Niger-Congo (Babaev 2013; Ehret no date; Segerer 2002-2007) languages to associate the consonant *m* with '1sg' and descendants of the First Americans (not comprising the speakers of Proto-Na-Dene nor of Proto-Eskimo-Aleut, who arrived much later) to associate the same consonant *m* with '2sg' is indeed as obscure as it gets.

If languages really did "construct their pronominal systems around a small set of basic consonants," as Zamponi puts it, from a wide range of different, unrelated source words, or even from thin air, for the articulatory simplicity and the acoustic or even visual prominence of their sounds, a substantial fraction of them should use the highly frequent, highly distinctive plain labial voiceless stop *p*. The near absence of *p*'s from the global set of pronominal root consonants is not only surprising but indeed inexplicable in the framework of an imaginary independent "pronoun construction" by languages. The only plausible explanation, in the Americas and elsewhere, is that they have been inherited from common sources – in the American case, very likely from the language spoken by the small population of First Americans some 14,000 to 16,000 years ago.

Actually, Zamponi's idea to figuratively endow languages with a quasi-autonomous agency ("*languages construct* their pronominal systems") was not a good one. There are no such things as modern languages "constructing their pronominal systems."

A very long time ago, in the range of dozens or maybe hundreds of millennia, unknown *speakers* – not languages – began to construct the first pronominal system, most probably in an extremely long and hesitant process, because, obviously, they had no clue what they were doing since they had no example of a pronominal system at hand nor at ear – and no linguist in view to explain them what personal pronouns (nor, indeed, any person marker) were so terrifically good for. They were just speaking their archaic language, heretofore lacking personal pronouns, and modifying it in a convenient way – making it more quickly known to the hearer who was doing what to whom, and in a way which would reveal itself to be so convenient that no language having survived long enough to be known to us lacks 1st and 2nd person pronouns. Note that it would still be perfectly possible to use only the 3rd person, *i.e.* the non-person (Benveniste 1966, see section 4 below), and to refer in discourse to both the speaker and the addressee by their personal names.

Today, speakers *inherit* their pronominal system from their parents and *transmit* it to their children, and it must have been so from the very beginning. Phonetic change nibbles at pronouns bite by bite, morphosyntactic restructuring eats its share, a very few borrowings occur hither and thither, but alien words entering a pronominal system can probably be counted on the fingers of very few hands.

The short list of frequent pronoun root consonants we find today in the languages of the world, with its phonological gaps, may testify that many of these pronouns are the descendants of this process. Of course, some of them may (and some no doubt do) result from phonetic changes from a different source sound not belonging to this list. But many different, independent linguistic lineages having arrived at easy and highly distinctive sounds for their pronouns while excluding *p*'s is another statistical *non sequitur*.

4. Languages do not "converge," except with serious – not "obscure" – reasons

The fourth problem, closely linked to the preceding, arises from an inappropriate use of the concept of "convergence." Having taken into account only languages with a full pair of pronouns (the asterisks on his Map 4), Zamponi presents their spread as limited to western North America, and invokes "convergence" as a one of two possible sources of the abundance of the pair *n* '1sg' / *m* '2sg' in this area, the other being "a hypothetical proto-language" – an alternative he immediately dismisses as "not answerable":

> [T]he pronominal paradigm is a clear indicator of prehistoric contacts. But the question of whether these contacts *[in far western North America]* have the form of a hypothetical proto-language or of a convergence between (perhaps few or very few) formerly distinct language branches is not answerable based on just the present-day knowledge of the history of the linguistic families of these areas.

Of course, a proto-language is never "a form of prehistoric contacts." Rather, it is the *common ancestor language* of several daughter languages, parts of which may be discovered through systematic comparison of the latter's vocabularies and grammars, and assumed to have been spoken by a speech community.

In turn, the word "convergence," in the absence of any seriously supported material, physiological or psychological reason which could motivate it, as we have seen in the preceding section, and in view of the peculiar phonological restrictions bearing on the list of pronominal root consonants, may only be a fig leaf for another word, namely "borrowing." However, in spite of recent attempts at demonstrating the contrary, it is well known that borrowing of 1sg and 2sg pronouns is truly exceptional among languages.

An example of the rarity of pronoun borrowing, all the more telling as it was intended to establish the contrary, is offered by the study of Sarah Thomason and Daniel Everett (2010) on this very subject, expressly aimed at showing that pronouns "can be borrowed," so that one should not rely too much on them for language classification. In their worldwide quest for borrowed pronouns, Thomason & Everett came up with *barely two dozen examples*, one half of them being 3rd person pronouns, which are not really *personal* pronouns, as shown by the French linguist Émile Benveniste (1966) and which, though relatively resistant as compared to the overall vocabulary, are much more prone to be borrowed and replaced than those of 1st and 2nd person.[3] Consequently, Thomason & Everett submit hardly a dozen languages with a possibly borrowed 1st or 2nd person pronoun from the 6,500 languages of the world, or around 0.15% – one in some 650 of them.

Moreover, several of those numerically insignificant examples are indeed highly dubious. For instance, such is the case of the (unidentified) English pronouns that Thomason & Everett, in an early version circulated on the Internet, quoting a 1986 paper by William Foley, alleged are "borrowed" into the parlance of urban Thai youths, and that still have to demonstrate their ability to take root into Common Thai instead of vanishing after the generation of Anglophile students which found them fashionable will have gotten older and forgotten them, while the next generation might deem them obsolete and invent in turn their own ways of coolspeak. And it is precisely what happened. In the published version of their paper (2010: 307), they triumphantly add that "Thai has also other borrowed pronouns: Christopher Court says that Thai boys often use the Chinese pronouns for 'I' and 'you' when speaking to each other (1998)." In 1998, China had begun to mean a lot in Thailand – and the English-speaking world, somewhat less than before.

---

[3] A major difference between 1st and 2nd persons, on the one hand, and 3rd person, on the other, is that the first two always refer to *persons*, most of the time actual humans present at the scene of discourse – the speaker and hearer(s) – or *personified* beings like gods or animals, sometimes even things; in contrast, 3rd "person" indifferently refers to any living being, thing or abstraction without lending them any particularly human character.
Another major difference is that 1st and 2nd person pronouns are the only words, in any known human language, that automatically switch reference with the turn of speech: I am my own *I* and you are my *you*, while you are your own *I* and I am your *you*, making them really difficult to grasp for children, who learn to use them late in their speech development. For 3rd person, nothing happens when the speaker and the hearer switch turns: all *he*s, *she*s, *it*s and *they*s remain the same *he*s, *she*s, *it*s and *they*s whoever is speaking to whom.

Such is also the case of the 1st and 2nd person pronouns "interborrowed" between the Turkic, Mongolic and Tungusic language families, which could just as well have been inherited from the Proto-Altaic language hypothesized by many linguists to be ancestral to these three families – with the support of regular sound correspondences (Starostin *et al.* 2003). Thomason & Everett (2010: 303) are clearly conscious of this weakness of their example, as they take care to specify that they "have not carried out any kind of systematic analysis" of the case, and that their point "is therefore not to argue for an areal source for the semantic interchange in Altaic pronoun systems, but rather to argue against the too easy assumption that the partly shared Altaic pronominal systems must be inherited." A candid admission, indeed, but wilful ignorance may hardly be considered an argument.

Worse still, Malcolm Ross (2005: 55-57) also disconfirms two other examples of 1st and 2nd person pronoun borrowings in Papuan languages alleged by Thomason & Everett (2010: 303, quoted by Ross after the preliminary online version). The first one was between Iatmül, a language from the Ndu branch of the Sepik phylum, and Kambot, a Keram language from the Ramu branch of the Ramu-Lower Sepik phylum – two quite distinct phyla in spite of their similar names owing to their respective locations along the Sepik River. Ross shows with the relevant comparative data that both Iatmül and Kambot obviously inherited their pronouns from their respective ancestral languages.

The second example (Thomason & Everett 2010: 304) is the North Halmahera group of West Papuan, which would have borrowed pronouns from Austronesian, for which claim Ross finds that "comparative evidence does not strongly back up this view." As a result, next to nothing but dust remains of Thomason & Everett's study.

Confirming this exceptional nature of pronoun borrowing, in an unpublished study of 1st and 2nd person pronouns in the (disputed, which does not matter here) Eurasiatic macrophylum, comprising the Indo-European, Uralo-Yukaghir, Altaic, Japonic-Koreanic, Gilyak, Chukotko-Kamchatkan and Eskimo-Aleut phyla and totalling some 600 languages, we could find only one (one!) borrowed 2nd person pronoun, discovered and convincingly established by the German linguist Florian Siegl (2008) in Forest Nenets, an Uralic language from the Samoyedic branch, which borrowed *ū* '2sg' from the Yeniseian language Ket – and not a single case for 1st person.

Yet, an overwhelming majority of those Eurasiatic languages have been submitted, in the course of history, to extensive, intimate and lasting contacts with other languages, some of which had acquired a dominant status stemming from invasions, empires or trade routes. These dominant languages may have been Yeniseian, why not, or more often Sinitic, Tibetic, Altaic, Uralic, Indo-European, Semitic, or other, giving the numerous subservient languages ample occasion to borrow clearly foreign pronouns.

These subservient languages actually borrowed many dozens, often hundreds and sometimes thousands of words from their mightier, richer, or more prestigious neighbors, but – except Forest Nenets *ū* – did not borrow any 1st or 2nd person pronoun.

English itself, which borrowed some 10,000 words from French after the Bastard's conquest, and even borrowed the *vouvoiement* to the point of evicting from everyday use its *th-* '2sg' pronominal forms inherited from PIE **t-* '2sg,' did not borrow *vous* '2pl' from French for this use. Instead, it took over the French sociolinguistic use of '2pl' as '2sg', and applied it to its own *you* '2pl,' which it had inherited from PIE **jūs* '2pl, nominative' through Proto-Germanic **jūz* '2pl, nominative.'

On the same note, returning to the Americas, Campbell (1994b), in his effort directed against Nichols' "Pacific Rim hypothesis", quotes only two examples, namely Miskito, which "borrowed its personal pronouns from Northern Sumu," and Mednyj Aleut, which borrowed its "verb morphology, including pronominal endings, from Russian" (pp. 4-5). This latter example (which, by the way, does not concern an alleged Amerind language, as Aleut is a member of the Eskimo-Aleut phylum, itself a branch of the Eurasiatic macrophylum) was debunked by Nichols & Peterson (1998: 610), who remarked that Mednyj Aleut is a "mixed language," a category of languages close to pidgins and creoles, built on the fly by adult people with different maternal languages in order to communicate in a context of close and enduring contacts – here, the colonization by Russians of the Aleutian Islands. Such languages actually have not one but two ancestor languages, in this case Aleut and Russian.

In a later paper, Campbell (1997: 340) did not add any other American example than Miskito, and offered two dozen more languages from around the world which would have borrowed one or more personal pronouns, without quoting them precisely nor even specifying whether they be 1st, 2nd or 3rd person pronouns, which seem to him to be all the same as his only quoted data is "English *they, them, their* borrowed from Scandinavian" (*ibid.*). This may not be regarded as serious scholarship.

In the North American case, it is not 0.15% of taxa which would have "converged" pronouns, as Zamponi prefers to say, nor 1.5% or even 15%, but a stunning 72.9%, including 31.4% of them which would have both their 1sg and 2sg pronouns being "converged." This would be absolutely unheard of anywhere else on planet Earth – and is plain incredible.

5. Overlooked *n* '1sg' and *m* '2sg' forms as well as plural pronouns

Another point concerns the real scope of *n* '1sg' and *m* '2sg' in the Americas. We have until here assumed that Zamponi really had scoured the entire field in search of such possible forms. However, as we looked into the matter in more detail, we began to suspect that he might have overlooked some data, perhaps not so few of them. A rather cursory look into descriptive data, guided by Greenberg & Ruhlen's *Amerind Etymological Dictionary,* etymology 397: I_1 **na(ʔ)* (2007: 121-123) showed that several more taxa with *n* '1sg' reported by this source, often criticized for inaccuracies in its data, are indeed supported by other sources (Table 3).

Zamponi's taxon No. and name	*Proposed reflexes of Proto-Amerind* n *'1sg' in Greenberg & Ruhlen (2007) followed by our comments*
30 Chemakuan	Quileute *l* (< **n*) [no gloss is given, 'I' implied] Andrade (1933: 204) gives in a table a series of Quileute 1sg pronominal suffixes: indicative *-li*, interrogative *-la*, subjunctive *-al*, and conditional *ti?l*. In a later work, Andrade (1953:

Zamponi's taxon No. and name	*Proposed reflexes of Proto-Amerind* n *'1sg' in Greenberg & Ruhlen (2007) followed by our comments*
	215) gives the following equivalence between Chemakum and Quileute: Ch. *n* = Qu. *d*. We could not find another source for Ruhlen's alleged phonetic evolution **n* > Qu. *l*, although a further evolution **n* > *d* (? > *l*) is plausible. An uncertain case.
16 Yukian	Yuki *ʔin-* 'my (inalienable)' Verified in Balodis (2016: 106-107), which gives examples of inalienable (kinship) possessives *ʔin-* and *ʔitin-* 'my.' An unambiguous case.
14 Barbacoan	Guambiano *na*, Totoró *na*, Awa *na* [no gloss is given, 'I' implied] Zamponi takes at face value Curnow & Liddicoat (1998) who report the same forms for these three languages, but reconstruct Proto-Barbacoan **la* 'I' (p. 392, table 1, etym. 18), with a word-initial sound correspondence Pr.-Barb. **l* > Guamb., Tot., Aw. *n* (p. 397) based on only three words in their 41-item list in five languages, the other two languages being Cha'palaachi and Tsafiqui, of which only Ts. has *la* 'I', with data lacking for Ch. In the other two instances of word-initial **l* > *n* (etym. 15 and 40), data are lacking for both Guambiano and Totoró. This is really scant evidence for a regular sound correspondence. A possible case.
170 Zamucoan	Zamuco *ñu* [no gloss is given, 'I' implied] Bertinetto (2009: 19, Table 4 "Personal pronouns and affixes") gives alternating forms *yV- / ñV-* '1sg and 1pl personal prefixes' and *yV- // ñV* '1sg and 1pl possessive prefixes.' As these oral/nasal alternating forms are exactly parallel to forms *ba-/ma-* '2sg personal prefix' and *ba-/bV- // ma-/mV* '2sg possessive prefix,' which Zamponi counts as a valid case for Zamucoan *m* '2sg,' one is at a loss to understand why he did overlook Zamuco *ñV-* '1sg.' An unambiguous case.

Table 3. Zamponi's linguistic taxa with either certain or possible *n* '1sg' unreported by him, found in Greenberg & Ruhlen's (2007: 121-123) etym. 397: I_1 **na(ʔ)* and supported by independent sources.

Pursuing the quest looking for *m* '2sg' with the help of Greenberg & Ruhlen's etymology 778: $THOU_1$ **mi* ~ **ma* (2007: 229-230), we also wound up, as could be expected, with positive results (Table 4).

Zamponi's taxon No. and name	*Proposed reflexes of Proto-Amerind* m *'2sg' in Greenberg & Ruhlen (2007) followed by our comments*
145 Tunica	*ma* ([no gloss is given, 'thou' implied] Verified in Swanton (1921: 8). It is the masculine form of the 2sg independent pronoun (fem. 2sg *hă'ma*). Its *m* root is shared with 1sg *i'ma* (both genders), but it is nonetheless the root consonant of 2sg. An unambiguous case.
34 Chitimacha	*him?* [no gloss is given, 'thou' implied] Verified in Swanton's (1919: 21) table, giving *him* under column of "Independent forms (used as objective prefixes)." An unambiguous case.

Zamponi's taxon No. and name	*Proposed reflexes of Proto-Amerind* m *'2sg' in Greenberg & Ruhlen (2007) followed by our comments*
119 Quechuan	Proto-Quechuan **qam*, Quechua *kam* [no gloss is given, 'thou' implied] Verified in various grammars and dictionaries. Zamponi may not have missed this pronoun, but may have chosen to leave it aside according to his rule of considering only the initial consonant, here *q-*. However, in the neighboring Puelche, which he considers an isolate, he agreed to consider 2nd person pronouns *ki-ma-w* (sg., du.) and *ki-ma-n* (pl.) as instances of *m* '2sg' on account of the presence of an initial syllable *ki-* in 1st person pronouns as well, which he agrees to consider a prefix to the pronominal root. Given the geographic proximity of Puelche and Quechua, both of which Greenberg & Ruhlen group together in the Andean phylum, it may be suspected that the initial syllable *qa-* in Quechua *qam* '2sg' may be a former prefix as well. Zamponi also counts Guahiboan **xa-ni* 'I' and **xa-mi* 'thou' as prefixed forms – note the also velar C_1. See also Cahuapanan and Guaicuruan below. A highly likely case.
20 Cahuapanan	Cahuapana *kema* [no gloss is given, 'thou' implied] Like in Quechua, a second-syllable *-m-* in an undisputed form *kəma* (Rojas-Berscia 2019) or *kima* (Valenzuela Bismarck 2011) 'thou' in Cahuapana. Note again the initial syllable *kə-* / *ki-* in another Andean branch posited by Greenberg & Ruhlen. A highly likely case.
52 Guaicuruan	Proto-Waikuru **ham*, Toba *am* [no gloss is given, 'thou' implied] For Eastern Guaicuruan, Carpio (2014: 988, Fig. 4 "Independent pronouns") gives *ʔam* '2sg' in Western Toba and Grondona (1998) gives Mocoví *qami* '2sg,' while for Northern Guaicuruan Loukotka (1933) gives Kadiweu *akáme* '2sg' (cp. *okámi* '2pl'). Again a back (velar or uvular) C_1, not in Ruhlen's Andean but Macro-Panoan phylum. A highly likely case.

Table 4. Zamponi's linguistic taxa with either certain or possible *m* '2sg' unreported by him, found in Greenberg & Ruhlen's (2007: 229-230) etym. 778: THOU$_1$ **mi* ~ **ma* and supported by independent sources.

This relatively quick search comes up with 9 more instances of either *n* '1sg' or *m* '2sg' to the 95 already found by Zamponi among his 172 Native American taxa, making a total of 104 languages exhibiting at least one of the two roots, thus 60.5%.

Another weakness of Zamponi's study is to have considered only singular pronouns, while it is well known that there is some historical permeability between pronouns of the same person differing in number. For instance, Russian *мы* (мы) 'we, nominative' is the PIE 1sg non-subject root **m-* with a (Slavic) plural suffix *-ы*; another such case is Classical Armenian *mek'* (մեք) 'we.' Several more Native American taxa might have been added with either *n* '1pl' or *m* '2pl' pronoun roots.

Furthermore, the case of Barbacoan languages with *n* '1sg' overlooked by Zamponi on account of a weak reconstruction work based on a grand total of 41 word series raises the issue of *n* '1sg' and *m* '2sg' forms in taxa where they have not been reconstructed. For instance, Zamponi counts Caddoan as one of its 172 Native American linguistic units, but does not include it as a taxon with *n* '1sg.' True, it is not reconstructed for the whole Caddoan family, but Proto-North Caddoan **na(ˑ)t* 'I' is nevertheless posited by Taylor (1963: 122) – and Proto-North Caddoan represents all of the Caddoan languages save Caddo itself, which is the sole member of South

Caddoan. There are exactly equal odds for North Caddoan having preserved **n* '1sg' from Proto-Caddoan and Caddo having lost it as for North Caddoan having innovated its own **n* '1sg' (and Caddo having done whatever it pleased with its own 1sg roots). Actually, in a sound statistical assessment, Proto-North Caddoan **na()t* 'I' should be counted as half of one of the 172 taxa in addition to whole taxa displaying *n* '1sg.'

Moreover, not only personal pronouns in all numbers (singular, dual and plural) should be considered, but indeed all person markers, including verbal endings. Still in the familiar Indo-European phylum, corresponding to PIE pronoun root **m*- '1sg,' the reconstructed conjugation system contains both 1sg suffixes **-mi* and **-m* and 1pl suffixes **-mos* and **-me* – of the primary and secondary active paradigms of thematic verbs, respectively (Sihler 1995, Ringe 2006).

Still more interestingly, the PIE pronoun root **t*- '2sg' does not correspond to a PIE ***-t* 2sg verb suffix, but the 2pl suffix is **-te* in the primary, secondary and imperative active paradigms of thematic verbs. This clearly shows that a 2nd person root consonant may not only be found in pronouns but also in verbs, in singular as well as plural forms. Searching for related personal forms in languages whose relatedness is not *a priori* assumed, as does Zamponi, requires looking everywhere one may reasonably think they may hide themselves, instead of restricting the search to singular pronouns.

Finally, Zamponi may have performed a not so exhaustive quest for *n* '1sg' and *m* '2sg' forms. In no case does he report archaic forms such as English *thou*, almost out of use but nevertheless still found in phrases like *holier than thou*, and used in liturgic contexts which may spill into the common parlance under the form of proverbs like *Thou shalt not kill*. Finding remnants of this kind in Native American languages without *n* '1sg' or *m* '2sg' in their regular pronominal system would be a sure indication that they are relic forms from a previous stage of the language. With so many Native languages actually having the two pronoun roots, it is not unlikely that one might find some of those from which they are lacking which might have kept such telling traces of their former existence.

All in all, there may be a fair number of taxa and subtaxa left out by Zamponi for both *n* '1sg' or *m* '2sg' roots, which would have added still more etymological weight to the series.

Even with his overly constraining rules, Zamponi came up with 95 of his 172 language taxa, or 55.2% of them, having either *n* '1sg,' or *m* '2sg.' Unwillingly, as it seems, he reinforced Greenberg's remark already quoted in the Introduction above, that "it would probably be easier to enumerate *[languages]* where *nV*- and *mV*- are not found than where they are." At 60.5%, with the 9 additional taxa that we added to Zamponi's data (a number probably far from exhaustive, as shown *inter alia* by the case of Proto-North Caddoan**na(.)t* 'I'), Greenberg's remark begins to border on truism.[4]

[4] Another possible angle would be to study the loss and preservation rates of *n* '1sg' and *m* '2sg' roots in those of Zamponi's 95 taxa with more than one language where they are present, and to relate this rate to the estimated ages of their respective taxa. Given the preservation rates of pronouns in other parts of the world, we are rather confident that the calculated loss rates per millennium should be compatible with a common ancestor close to the date of the initial colonization of the Americas. We intend to perform this exercise in another article.

III. CONCLUSION

Zamponi's data show that 95 out of his 172 Native American linguistic taxa, or 55.2% of them, possess at least one of the two pronoun roots *n* '1sg' and/or *m* '2sg.' This huge proportion definitely excludes chance as a significant factor – as Zamponi himself agrees after Nichols and Peterson (1998). Moreover, as we have shown above, if all relevant forms were to be counted, one should add at least to this number one and possibly several dozens of taxa or subtaxa overlooked by him.

However, faced with this impressive load of pronominal forms from North and South America, Zamponi discards without a word of justification all the taxa displaying only one member of the pronominal paradigm, sweeping under the rug the greatest part of the data that he had himself carefully collected in his appendices and located on his maps. Moreover, he then downplays the real distribution of this drastically reduced set over the Hemisphere to a phenomenon essentially limited to western North America. Even then, he does not come up with a firm explanation and finally discards its surviving Californian data into the grey kingdom of the "not answerable."

Conversely, taking into account the whole set of the *n* '1sg' and *m* '2sg' pronominal forms – which obviously (and perhaps far) exceeds the 95 Native American taxa mentioned by Zamponi – necessarily entails conclusions that are radically different from his. While agreeing with him that chance cannot be a significant factor, we have shown that borrowing was very unlikely to have played an important role, either – nor any other kind of "convergence."

These considerations do not leave us with many explanations for these pronouns' distribution – excepted inheritance from a common ancestor language. As Sapir said in a letter written to Kroeber in 1918,

> [g]etting down to brass tacks, how in the Hell are you going to explain general American *n-* 'I' except genetically? It's disturbing to know but non-committal conservatism is only dodging after all, isn't it? (Sapir 1984)

In spite of himself, as it seems, Zamponi's fresh look on the two pronouns roots *n* '1sg' and *m* '2sg' in the Native languages of the Americas lends considerable support to Sapir's claim. Back to the idea that these languages – save the Na-Dene and Eskimo-Aleut languages – originated from a single ancestral language, for which Greenberg (1986) coined the name "Amerind." Two qualifications are in order, however:

(a) The two pronoun roots *n* '1sg' and *m* '2sg' say nothing about the taxa they are lacking from, and hence are by themselves not sufficient to pronounce the reality of the whole Amerind linguistic phylum. There still remain the 908 other etymologies proposed in Greenberg & Ruhlen's *Amerind Etymological Dictionary*, not all of which may eventually prove correct, and some of which may be entirely wrong, while neither can all of them collectively be due to chance, borrowing or convergence;

(b) A few of the numerous taxa having one or the other root may have acquired them through phonetic change, pronoun internal replacement, or, exceptionally, borrowing, although the documented frequency of this latter phenomenon does not warrant, from a statistical vantage point,

more than a handful of cases in all of the Americas' languages, and only an umpteenth fraction of percent in Zamponi's 172 taxa.

Finally, the Amerind ancestral language is not floating alone in the air. Linguists have argued endlessly as to whether one cannot reach beyond the close to 200 Native American language families recognized by Zamponi, or if there are enough data remaining from a potential common origin in the 1,000-odd languages having survived into historical times to assess this unity.

This squabble sharply contrasts with the unanimity of population geneticists. For the last three decades, they have concluded, often in multi-authored articles in high-profile publications such as *Nature*, *Science*, *PNAS USA* or *PLoS One*, that nearly all Native American populations derive their genetic ancestry from a single founding population (Cavalli-Sforza 1994; Greenberg *et al.* 1986; Kitchen *et al.* 2018; Moreno-Mayar *et al.* 2018a, 2018b; Reich 2018; Reich *et al.* 2012; Sikora *et al.* 2019; Skoglund & Reich 2016; etc.).

Called the "First Americans" by David Reich, they may have entered Alaska some 14,000 to 16,000 years ago, crossing from Siberia through Beringia – the landmass currently submerged and called the Bering Straits, at the time left to dry by the lowered ocean level from the previous Last Glacial Period, while a new warming period left an ice-free passage both along the Pacific coast and through an interior corridor between glaciers. Strikingly, two exceptions repeatedly found by geneticists to this unity of origin are the populations speaking Na-Dene and Eskimo-Aleut languages, whose genes are associated with more recent ingressions from Siberia, while their speakers also exhibit a strong genetic admixture with descendants of the First Americans (Reich *et al.* 2012; Reich 2018).

Furthermore, many geneticists (*e.g.* Fagundes *et al.* 2018; Tamm *et al.* 2007) assume that the founding group of First Americans was small, perhaps comprising no more than a few dozen individuals (Hey 2005: 193). Such a small group may not have brought with it several languages – much less 172 of them to give rise to the 172 "independent" taxa recognized by Zamponi. Most probably, this small population spoke only one language, from which all Native American languages known to us – save the Na-Dene and Eskimo-Aleut languages – descend.

Figure 3. The three Native American linguistic phyla.

The respective distribution of the three American phyla also makes sense in the perspective of the progressive warming event which allowed humans to colonize the Americas. The First Americans, coming from Siberia through Beringia, entered along the Pacific coast or through a corridor between glaciers, and spread over the whole Hemisphere south of the ice sheet. The speakers of Proto-Na-Dene, who arrived much later, occupied the new stretches of land opened by the melting ice cap, but their progression southwards was mostly blocked by the already established Amerind groups. Finally, thanks to their unique biological, technical and cultural adaptations to the extreme arctic conditions, Eskimo-Aleut speakers managed to spread along the northernmost fringe of the continent and to reach Greenland. (Map after Merritt Ruhlen.)

So much so that one section of David Reich's book *Who We Are and How We Got Here* (2018), in his chapter devoted to the genetic origins of Native Americans, is entitled "The genomic

rehabilitation of Joseph Greenberg." The Amerind hypothesis (Fig. 3), first conceived more than a century ago, way before the discovery of DNA, and further elaborated by Greenberg from the 1960s to the 1980s, then by Ruhlen down into the 2000s, thus receives strong support from the new science of population genetics.

While temporarily concluding our contribution to this discussion, we would like to thank again Raoul Zamponi for having collected the linguistic data necessary to re-open an in-depth discussion of the *n* '1sg' and *m* '2sg' pronoun roots in Native American languages. In line with Sapir's assumption, we believe that explaining their pan-American distribution by inheritance from a single proto-language is, for the time being, the most economic hypothesis, answering the criteria of simplicity, internal consistency and falsifiability.

This hypothesis is also strengthened by its conspicuous external consistency with the findings of geneticists positing that a single population entered the Americas perhaps as soon as 16,000 years ago and quickly spread over the whole Hemisphere, reaching and settling in South America no later than 14,000 years ago. This would imply that the two pronoun roots *n* '1sg' and *m* '2sg' would have been part of the pronominal system of the First Americans' language.

REFERENCES

Andrade Manuel J. 1933. Quileute. In Franz Boas (ed.), *Handbook of American Indian languages*, vol. 3: 151-292.

Babaev 2013 = Бабаев Кирилл В. 2013. *Нигеро-конголезский праязык: личные местоимения*. [Niger-Congo protolanguage: personal pronouns]. Moscow: Languages of Slavic Culture.

Balodis Uldis. 2016. *Yuki grammar. With sketches of Huchnom and Coast Yuki*. Oakland: Univ. of California Pr.

Bancel Pierre J. & Alain Matthey de l'Etang. 2010. Where do personal pronouns come from? *Вопросы языкового родства* **3**: 127-152.

Benveniste Émile. 1966. La nature des pronoms. In *Problèmes de linguistique générale*, vol. 1. Paris: Gallimard: 251-257.

Bertinetto Pier Marco. 2009. Ayoreo (Zamuco). A grammatical sketch. *Quaderni del Laboratorio di Linguistica della Scuola Normale Superiore di Pisa* **8**.

Boas Franz. 1917. Introduction. *International Journal of American Linguistics* **1**-1: 1-8.

Bomhard Allan R. 2012. *Reconstructing Proto-Nostratic: Comparative phonology, morphology, and vocabulary*. Charleston, SC.

Campbell Lyle. 1988. Review of *Language in the Americas* (by Joseph H. Greenberg). *Language* **64**-3: 591-615.

———. 1994a. The American Indian classification controversy: An insider's view. *Mother Tongue* **23**: 41-55.

———. 1994b. Putting pronouns in proper perspective in proposals of remote relationships among Native American languages. *Survey Reports* **8**, Survey of California and Other Indian Languages. Berkeley: Univ. of California.

———. 1997. Amerind personal pronouns: a second opinion. *Language* **73**-2: 339-351.

Carpio María Belén. 2014. 'Restricted group' and 'group' within the pronominal system of Western Toba (Guaicuruan, Argentina). *Studies in Language* **38**-4: 982–994. doi 10.1075/sl.38.4.09car.

Cavalli-Sforza Luigi Luca, Paolo Menozzi & Alberto Piazza. 1994. *The history and geography of human genes*. Princeton: Princeton Univ. Pr.

Chafe Wallace. 1987. Review of *Language in the Americas* by Joseph H. Greenberg. *Current Anthropology* **28**-5: 652-653.

Curnow Timothy J. & Anthony J. Liddicoat. 1998. The Barbacoan languages of Colombia and Ecuador. *Anthropological Linguistics* **40**-3: 384-408.

Ehret Christopher. No date. *Reconstructed pronouns in Niger-Congo*. Unpubl. ms.

Fagundes Nelson J.R., Alice Tagliani-Ribeiro, Rohina Rubicz *et al.* 2018. How strong was the bottleneck associated to the peopling of the Americas? New insights from multilocus sequence data. *Genetics and Molecular Biology* **41**-1, Suppl. 1: 206-214. doi.org/10.1590/1678-4685-GMB-2017-0087.

Goddard Ives. 1987. Review of *Language in the Americas* by Joseph H. Greenberg. *Current Anthropology* **28**-5: 656-657.

———. 1990. Review of *Language in the Americas* by Joseph H. Greenberg. *Linguistics* **28**: 556-558.

Golla Victor. 1987. Review of *Language in the Americas* by Joseph H. Greenberg. *Current Anthropology* **28**-5: 657-659.

Greenberg Joseph H. 1987. *Language in the Americas*. Stanford: Stanford Univ. Pr.

———. 2000. *Indo-European and its closest relatives: The Eurasiatic language family*, vol. 1: *Grammar*. Stanford: Stanford Univ. Pr.

Greenberg Joseph H. & Merritt Ruhlen. 2007. *An Amerind etymological dictionary*. Stanford: Stanford Univ. Online version 12, https://newstar.rinet.ru/Texts/Ruhlen_AED5.pdf, accessed on March 20, 2021.

Greenberg Joseph H., Christy G. Turner II & Stephen L. Zegura. 1986. The settlement of the Americas: A comparison of the linguistic, dental, and genetic evidence. *Current Anthropology* **27**: 477-497.

Grondona Verónica María. 1998. *A grammar of Mocoví*. Ph.D. dissert. Pittsburgh, PA: Univ. of Pittsburgh.

Hammarström Harald, Robert Forkel & Martin Haspelmath. 2017. *Glottolog 3.0.* Jena: Max Planck Institute for the Science of Human History. http://glottolog.org.

Hey Jody. 2005. On the number of New World founders: A population genetic portrait of the peopling of the Americas, *PLoS Biology* **3**-6: e193.

Kitchen Andrew, Michael M. Miyamoto & Connie J. Mulligan. 2008. A three-stage colonization model for the peopling of the Americas. *PLoS One*, **3**-2: e1596, doi: 10.1371/journal.pone.0001596.

Loukotka Čestmír. 1933. Nouvelle contribution à l'étude de la vie et du langage des Kaduveo. *Journal de la Société des américanistes* **25**-2: 251-277.

Matisoff James. 1990. On megalo-comparison: A discussion note. *Language* **66**: 106-120.

Moran Steven & Daniel McCloy (eds.). 2019. *PHOIBLE 2.0*. Jena: Max Planck Institute for the Science of Human History. phoible.org, accessed on 2020-12-23.

Moreno-Mayar J. Víctor, Lasse Vinner, Peter de Barros Damgaard *et al*. 2018. Early human dispersals within the Americas. *Science* **362**-6419: eaav2621, doi: 10.1126/science.aav2621.

Moreno-Mayar J. Víctor, Ben A. Potter, Lasse Vinner *et al.* 2018. Terminal Pleistocene Alaskan genome reveals first founding population of Native Americans. *Nature* **553**: 203–207.

Nichols Johanna. 1994. The spread of language around the Pacific Rim. *Evolutionary Anthropology* **3**-6: 206-215. doi.org/10.1002/evan.1360030607.

Nichols Johanna & David A. Peterson. 1996. The Amerind personal pronouns. *Language* **72**-2: 336-371.

———. 1998. Amerind personal pronouns: A reply to Campbell. *Language* **74**-3: 605-614. doi:10.1353/lan.1998.0126.

Reich David. 2018. *Who we are and how we got here: Ancient DNA and the new science of the human past*. London: Oxford Univ. Pr.

Reich David, Nick Patterson, Desmond Campbell *et al.* 2012. Reconstructing Native American population history. *Nature* **488**-7411: 370-374. doi:10.1038/nature11258.

Ringe Donald. 2006. *From Proto-Indo-European to Proto-Germanic*. London: Oxford Univ. Press.

Rojas-Berscia Luis Miguel. 2019. *From Kawapanan to Shawi. Topics in language variation and change*. Ph.D. dissert. Nijmegen: Radboud Univ..

Ross Malcolm. 2005. Pronouns as a preliminary diagnostic for grouping Papuan languages. In Andrew Pawley, Robert Attenborough, Jack Golson & Robin Hide (eds.), *Papuan pasts: cultural, linguistic and biological histories of Papuan-speaking peoples*. Canberra: Pacific Linguistics: 15-65.

Ruhlen Merritt. 1994a. *The origin of language: Tracing the evolution of the mother tongue*. New York: John Wiley.

———. 1994b. The linguistic origins of Native Americans. In Merritt Ruhlen (ed.), *On the origin of languages: Studies in linguistic taxonomy*. Stanford: Stanford Univ. Pr.

———. 1994c. First- and second-person pronouns in the world's languages. In Merritt Ruhlen (ed.), *On the origin of languages: Studies in linguistic taxonomy*. Stanford: Stanford Univ. Pr.

Sapir Edward. 1984. Letter to A. L. Kroeber (1918). In Victor Golla (ed.), *The Sapir-Kroeber correspondence: letters between Edward Sapir and Alfred L. Kroeber, 1905-1925*, Berkeley: University of California at Berkeley.

Segerer Guillaume. 2002-2007. *Les marques personnelles dans les langues africaines*. Base de données en ligne du LLACAN, http ://sumale.vjf.cnrs.fr/pronoms.

Siegl Florian. 2008. A note on personal pronouns in Enets and northern Samoyedic. *Linguistica Uralica* **XLIV**. doi:10.3176/lu.2008.2.04.

Sihler Andrew L. 1995. *New comparative grammar of Greek and Latin*. London: Oxford Univ. Press.

Sikora Martin, Vladimir V. Pitulko, Eske Willerslev *et al*. 2019. The population history of northeastern Siberia since the Pleistocene. *Nature* **570**: 182-188.

Skoglund Pontus & David Reich. 2016. A genomic view of the peopling of the Americas. *Current Opinion in Genetics & Development* **41**: 7-35.

Starostin Sergey A., Anna V. Dybo & Oleg A. Mudrak. 2003. *Etymological dictionary of the Altaic languages.* Leiden: Brill.

Swadesh Morris. 1954. Perspectives and problems of Amerindian comparative linguistics. *Word* **10**: 306-332.

———. 1960. *Estudios sobre lengua y cultura*. Mexico City: Escuela nacional de antropología e historia.

Swanton John R. 1919. *A structural and lexical comparison of the Tunica, Chitimacha, and Atakapa languages*. Washington: Govt. Printing Office.

———. 1921. The Tunica language. *International Journal of American Linguistics* **2**-1/2: 1-39. doi:10.1086/463732.

Tamm Erika, Toomas Kivisild, Maere Reidla *et al.* 2007. Beringian standstill and spread of Native American founders. *PLoS One* **2**-9: e829. doi.org/10.1371/journal.pone.0000829.

Taylor Allan R. 1963. Comparative Caddoan. *International Journal of American Linguistics* **29**-2: 113-131.

Thomason Sarah G. & Daniel L. Everett. 2010. Pronoun borrowing. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* **27**-1: 301-315. doi: 10.3765/bls.v27i1.1107.

Trombetti Alfredo 1905. *L'unità d'origine del linguaggio*. Bologna: Libreria Treves di Luigi Beltrami.

Valenzuela Bismarck Pilar. 2011. Contribuciones para la reconstrucción del proto-cahuapana: Comparación léxica y gramatical de las lenguas jebero y chayahuita. In Willem F. H. Adelaar, Pilar Valenzuela Bismarck & Roberto Zariquiey Biondi (eds.), *Aru, Simi, Taqu, Lengua: Estudios en homenaje a Rodolfo Cerrón-Palomino*. Lima: Fondo Editorial, Pontificia Universidad Católica del Perú: 271-304.

Zamponi Raoul. 2017. First-person *n* and second-person *m* in Native America: A fresh look. *Italian Journal of Linguistics* **29**-2: 189-230. Appendices pp. 1-68 online only at http://www.italian-journal-linguistics.com/wp-content/uploads/Volume_29_Issue2_Appendice_Zamponi.pdf.

REVIEW OF LYLE CAMPBELL AND MARIO MIXCO'S *A GLOSSARY OF HISTORICAL LINGUISTICS*

ROBERT LINDSAY

The following quotes are from the book under review, a textbook or general text on historical linguistics by Lyle Campbell and Mario Mixco, *A Glossary of Historical Linguistics* (2007). The purpose of this paper will be misrepresented as critics will say that I am a long-ranger who is criticizing conservatives for their opposition to long-range language families.

There is some of that here, but what I am trying to do here more than anything else is to show how Campbell and Mixco have been untruthful about linguistic specialist consensus regarding these families. In most cases, they are openly misrepresenting the state of consensus in the field.

As will be shown, Campbell and Mixco repeatedly seriously distort the state of consensus regarding many language families, particularly long-range ones. They usually favor a more negative and conservative view, saying that a family has little support when it has significant support and saying it is controversial when the consensus in the field is that the family is real. Campbell and Mixco engage in serious distortions of fact all through the book**.**

First of all, I would like to pose a few interesting questions and then elaborate on the methodology of this paper. Who is a specialist? All you have to do to be a specialist is to specialize in a particular language family. In my opinion, it means to have written at least one article about the language family.

In this paper, I will show that in many cases, most linguists reject the family, but most of them are non-specialists. However, once we look at actual specialists, the view changes completely and in many cases, consensus among them either supports or is optimistic about the family. I don't see why the opinions of non-specialists are so important. Go study the languages and get back to us! Clearly with a view towards consensus on support or rejection of a language family, the only people's opinions that should count would be specialists.

Now to the math.

I want consensus to be at least 70% of specialists. That's a high bar for consensus and simple majority rule won't cut it. At 60%, I say that the consensus is either optimistic about the family or pessimistic about it, depending on which side has the 60-69% view, supporters or opponents. However, from 60 to 69% support for a view, I refuse to say that consensus on this family is either

supportive or rejecting because that needs a 70% metric to achieve that status. The field is actually divided on the question and can't answer whether the family is valid or not. Instead, experts are either optimistic or pessimistic about the proposal, but they haven't concluded one way or the other yet.

Between 50 and 59%, I refer to consensus as split or divided on the question. It neither supports nor rejects the proposal. It's simply uncertain and experts may well be fighting it out as regards this proposal, which is up in the air.

I tried to be as fair as possible. As a long-ranger, clearly I am biased towards long-range proposals and against conservative views. However, you will see below a number of cases where I acknowledge that consensus either rejects or even strongly rejects the proposal. Sometimes I will try to show how the proposal might be valid nonetheless, but this paper isn't really about my opinions on the validity of language families, which I consider pretty irrelevant except in a few cases where I actually undertook an investigation into the question.

Another question is what is the actual view of conservatives? Do they really think that language was invented via polygenesis by 160 different groups in the Americas alone, since that is how many language families and isolates they find in the Americas? Are they really opposing monogenesis of language and arguing for this extremely high level of polygenesis? If that's true, then their proposal doesn't even pass the laugh test. You can't even take it seriously.

Or are they arguing something else? If they are arguing that there was probably monogenesis of language and that ultimately all human languages are probably related, then I would agree with them. Beyond that, they would say, sure, they are all related at some very deep level, but we want to know which ones cluster together to form a language family? Now that is a theory that even I can get behind.

So what exactly are they arguing?

Campbell and Mixco, p. 5:

> **Afroasiatic:** Enjoys wide support among linguists but it is not uncontroversial, especially with regard to which of the groups assumed to be genetically related to one another are to be considered true members of the phylum.
>
> There is disagreement concerning Cushitic, and Omotic (formerly called Sidama or West Cushitic) is disputed; the great linguistic diversity within Omotic makes it a questionable entity for some. Chadic is held to be uncertain by others. Typological and areal problems contribute to these doubts. For example, some treat Cushitic and Omotic together as a linguistic area (Sprachbund) of seven families within Afroasiatic.

Afroasiatic is part of Greenberg's four-family analysis of the African languages.[1] One of his papers focuses specifically on Afroasiatic (Greenberg 1950).

Campbell and Mixco's judgment is simply wrong, though it is popular with some conservatives. Afroasiatic is not controversial at all (Sands 1998). There is widespread consensus that the family exists and that all of the subfamilies are correct.

[1] Greenberg 1948, 1955, 1970.

The "we can't reconstruct the numerals" argument is much in evidence here. See the Altaic debate below for more on this. One argument against Altaic is "We can't reconstruct the numerals." However, Afroasiatic is a recognized family, and not only has reconstruction itself proved difficult, but the numerals in particular are a gigantic mess (Kaye and Daniels 1992). It seems that one does not need to have a fully reconstructed numeral set after all to have a proven language family.

In fact, primitive hunter-gatherers may have few numerals at all. Baka, a Pygmy language in equatorial Africa, only has numerals up to three. Informants could not agree on whether there was a numeral for four or not. The entire numeral set was borrowed from Fang, a Bantu language spoken in Gabon. As they borrowed their entire numeral set, it's possible that their original set may have been sparse indeed (Bancel 2021).

Numerals may also not be as well-preserved as people think. In the hypothesized Eurasiatic language, Indo-European ***duwo* "2" has been preserved in all IE descendants but has only been preserved in derivatives in Altaic, Uralic, and Kartvelian and has been lost altogether in Eskimo-Aleut and Chukchi-Kamchatkan (Bancel 2021).

There is consensus that Cushitic is a valid entity (Mous 2012). Chadic is also uncontroversial by consensus (Newman 1977; Blench 2006). Granted, there has been some question about Omotic (Newman 1990; Diakonoff 1998; Theil 2006), but in the last ten to fifteen years, consensus has settled on an agreement that Omotic is part of Afroasiatic (Lecarme 2003; Dimmendaal 2008; Ehret 2010). The great diversity of Omotic is no surprise. Omotic is probably 13,000 years old! Where do we get the idea that a language family cannot possibly be highly diverse?

Afroasiatic itself must be still more ancient. Expert views put the age of the language as old as 17,000 years, making it the oldest recognized language family. However, lexicostatistics tests on Greenberg's four African stocks show 1% match, which means it has been at least 26,000 years since they were part of each other. So Afroasiatic may be 26,000 years old.

I am not aware of any serious proposals to see Cushitic and Omotic as an Altaic-like Sprachbund of mass borrowings. Campbell and Mixco's comments above are simply not correct.

Adding up consensus, it is clear that Greenberg's Afroasiatic has been accepted. The only issue is Omotic, which a number of conservatives have grown pessimistic about in the last thirty years. Nevertheless, a good theory is that expert consensus now believes that Omotic is Afroasiatic, albeit a highly divergent branch.

Campbell and Mixco:

> **Berber** is usually believed to be one of the branches of Afroasiatic.

This is far too pessimistic. Berber is recognized by consensus as being one of the branches of Afroasiatic (Hayward 2000).

Campbell and Mixco, p. 139:

> **Niger-Kordofanian** (now often just called **Niger-Congo**): A hypothesis of distant genetic relationship proposed by Joseph H. Greenberg (1963) in his classification of African languages. Estimated counts of Niger-Kordofanian languages vary from around 900 to 1,500 languages. Greenberg grouped 'West Sudanic' and Bantu into a

> single large family which he called Niger-Congo after the two major rivers, the Niger and the Congo, 'in whose basins these languages predominate' (Greenberg 1963: 7).
>
> This included the subfamilies already recognized earlier: (1) West Atlantic (to which Greenberg joined Fulani, in a Serer-Wolof-Fulani [Fulfulde] group), (2) Mande (Mandingo) (thirty-five to forty languages), (3) Gur (or Voltaic), (4) Kwa (with Togo Remnant) and (5) Benue-Congo (Benue-Cross), with the addition of (6) Adamawa-Eastern, which had not previously been classified with these languages and whose classification remains controversial.
>
> For Greenberg, Bantu was but a subgroup of Benue-Congo, not a separate subfamily on its own. In 1963 he joined Niger-Congo and the 'Kordofanian' languages into a larger postulated phylum, which he called Niger-Kordofanian.
>
> Niger-Kordofanian has numerous supporters but is not well established; the classification of several of the language groups Greenberg assigned to Niger-Kordofanian is rejected or revised, though most scholars accept some form of Niger-Congo as a valid grouping.
>
> As Nurse (1997: 368) points out, it is on the basis of general similarities and the noun-class system that most scholars have accepted Niger-Congo, but 'the fact remains that no one has yet attempted a rigorous demonstration of the genetic unity of Niger-Congo by means of the Comparative Method.'

As Campbell notes here, there is consensus that at least some form of Greenberg's Niger-Congo is a valid grouping. Some conservatives say that Mande is not proven and that smashes the credibility of Niger-Congo, but that does not seem to be correct because Niger-Congo could still be said to be ratified were Mande not proven; it would just have one fewer family as a constituent. Since Mande only equals 60-70 languages out of 1,540 Niger-Congo tongues, even if we pull Mande, we still retain 95-96% of Greenberg's Niger-Congo. How pulling 4-5% of the languages out wrecks the family eludes me.

As the argument suggests, Mande is indeed a highly divergent branch of Niger-Congo, and this family may have been the first to branch off in Niger-Congo. Some put it in a sister position with all of the rest of Niger-Congo, in this sense similar to the best theory about the place of Anatolian in Indo-European.

Ijo is also highly diverse and some experts say it is not Niger-Congo. Ijo is not a separate branch like Mande. Instead, it is part of Benue-Congo. Like Mande, Ijo is also often put on a separate branch. It is also placed as one of the earliest branches to branch off, typically the second after Mande.

Both Ijo and Mande lack the large amount of morphological evidence that links Kordofanian to Niger-Congo. Most of this takes the form of verbal extensions. There's either no trace or barely a trace of this in Mande or Ijo, which is disturbing.

It is for this reason that some experts feel that Mande and Ijo are not Niger-Congo. The noun-class system that characterizes Niger-Congo as a whole is barely extant in Ijo, the only possible match being a nominal *a-* prefix which may be related to the plural noun-class prefix *a-* in the rest of Niger-Congo, but that's not a lot to go on. It is not extant at all in Mande. Mande does retain nasal vowels typical of the family as a whole, but that could just as easily be a typological feature which are widely borrowed as opposed to a genetic link.

As noted above, indeed, some conservatives now toss Mande out of Niger-Congo and list it as a separate family along with Ijo.

To this they add some parts of Gur, particularly Dogon. It is not even certain that Dogon is properly placed in Gur. Others have it as an isolate family within another major branch, Volta-Congo. Dogon is indeed very ancient and there are suggestions that the pre-Niger-Congo substrate is heavily present here. At the moment, I am not sure why some scholars exclude Dogon from Niger-Congo. Also, Dogon, like Mande and nearly like Ijo, completely lacks the characteristic Niger-Congo noun-class system.

Also for this reason, a number of Niger-Congo trees have Dogon splitting off early in addition to Mande and Ijo. Roger Blench, who is neither a conservative nor a long-ranger but instead seems to be simply a non-dogmatic pure scientist, accepts all of Niger-Congo but Dogon. He doesn't exclude Dogon but he says it may not even be a part of Niger-Congo. I think it is safe to say the expert consensus at the moment is that Dogon is not Niger-Congo. It may in fact be an isolate family in Africa.

However, Dogon's 1st and 2nd person singular pronouns line up well with the basic Niger-Congo system, although the 1st and 2nd person plural pronouns do not line up (Babaev 2009). However, Dogon's pronouns, like those of Ijo, are highly eroded, often leaving a mere V or V(C) segment left (Babaev 2009). It gets hard to line up such eroded forms with the rest of Niger-Congo in which the pronouns are much more intact. Ijo's pronouns are as eroded as Dogon's, and they line up worse with the rest of a family, with a match only in the 1st person singular (Babaev 2009).

In general, personal pronouns in Niger-Congo are pretty scattered and don't make up nearly such a coherent system as we see among other recognized language families. In fact, Mande's pronouns are highly divergent and barely match the main pattern, and Kordofanian's pronouns don't line up at all, except for possibly in the 2nd person plural (Babaev 2009).

If Niger-Congo is indeed over 26,000 years old, this ancient heritage may be the reason for the eroding of a fair amount of its pronoun system. Certainly, Niger-Congo's personal pronouns are in much worse shape than even those of Eurasiatic or even Amerind (Babaev 2009). As Amerind may be 13,000 years old and Eurasiatic 15,000 years old, to have a more dissolved pronoun system than those two is quite a feat. But if the ancient pedigree of Niger-Congo is a fact, even this odd fact starts to make sense.

The theory is that the Dogon people spoke an ancient pre-Niger-Congo language and then switched to Niger-Congo. Similarly, Shabo and Mimi-D in Nilo-Saharan may be similarly former hunter-gatherers who spoke a pre-Nilo-Saharan language who switched over to Nilo-Saharan.

There are many words in those languages that have no cognates anywhere in Africa. Even some non-conservative linguists think that some Dogon languages are not Niger-Congo – that is, they are isolates. They may well be the residue of the older pre-Niger-Congo language spoken earlier. Based on lexicostatistics tests of Greenberg's four African stocks, Niger-Congo may be at least 26,000 years old. See the Afroasiatic section for the methodology used to arrive at that figure.

To this in some cases, they add most of Ubangian, which is really Eastern Adamawa, to the list of families not included in the stock. The eastern group of Ubangian, the Mba group, has retained the noun-class system typical of Niger-Congo. However, in the rest of Ubangian, there remains no trace of it, hence some experts have removed these languages from Niger-Congo. The

personal pronouns for the entire Ubangian family match in the 1st person singular and possibly in the 2nd person plural, but for the rest, there is no match (Babaev 2009). Nevertheless, a highly eroded or divergent pronoun paradigm is typical of the Niger-Congo family as a whole.

Some conservatives exclude Kordofanian, but there is extensive morphological evidence in the form of verbal extensions to tie Kordofanian to the Niger-Congo family. Kordofanian also has the ancient aspectual morphology in Niger-Congo. This takes the form of suffixation to denote aspect, typically perfective aspect. Since Kordofanian retains ancient Niger-Congo verbal extensions and aspectual suffixes, the argument that it is not Niger-Congo is without merit.

Nevertheless, Kordofanian is still highly divergent. Its pronouns barely match typical Niger-Congo pronouns at all (Babaev 2009).

A few specialist views of Niger-Congo seem to dominate. The extreme position, evidenced by Campbell and Mixco above and the quote from Nurse (who actually supports Niger-Congo), is that Niger-Congo is not a valid entity. This is considered an extreme conservative view even by other conservatives, some of whom take the view that Niger-Congo is Greenberg's most-validated family, while holding that Nilo-Saharan, Khoisan, and even Afroasiatic are not valid entities (Hombert and Philippson 2009).

However, most specialists feel that Niger-Congo is indeed a valid entity, while there are arguments that some of the families included in it have not been proven to be part of the family. One view regards Niger-Congo as a valid entity, however, they take a rather conservative view that the Mande, Dogon and Ijo families as well as most of the Ubangian languages are not part of it (Storch and Dimmendaal 2016). Blench characterizes this as an extreme position (Blench undated a), but it's actually fairly common.

Others take the position that Mande and Ijoid are not Niger-Congo but that Dogon and Ubangian are (Güldemann 2018). Blench is uncertain if Dogon is actually Niger-Congo and suggests some Dogon languages are actually isolates (Blench undated b). However, Blench does accept Mande, Ijoid and Ubangian as part of Niger-Congo, with the exception of the highly divergent Bangi me and the Ijoid language Defaka, both of which may be isolates (Blench undated a).

A good moderate or judicious position seems to be to ratify Niger-Congo as a valid entity and to include Dogon and the Ubangian languages, while rejecting Mande and Ijoid languages as members (Babaev 2013).

A recent study looked at all branches of Niger-Congo on the basis of personal markers. A number of Nilo-Saharan languages were set up as controls to control for chance resemblances. All of the Nilo-Saharan languages had radically different person marking systems.

On this basis, the paper validated all of Niger-Congo including Ubangian and Dogon, which both had the characteristic person markers, with the exception of Mande and Ijoid (Babaev 2013). The eastern half of Ubangian lacks the classic noun-class system (Storch and Dimmendaal 2016). Dogon also lacks this system and also lacks the extensive system of verbal extensions (Storch and Dimmendaal 2016).

Mande and Ijoid lack the noun-class system, the verbal extensions (Storch and Dimmendaal 2016) and the personal markers (Babaev 2013), which makes their membership in Niger-Congo quite dubious. Blench is the only specialist who includes Ijoid and Mande in Niger-Congo, but I am not sure on what basis he includes them (Blench undated a).

Therefore, consensus holds that while Niger-Congo is a valid family, the Mande and Ijoid families are not parts of the stock (Babaev 2013), and Defaka and Bangi me may be isolates (Blench undated a).

Campbell and Mixco, p. 140:

> **Nilo-Saharan**: One of Greenberg's (1963) four large phyla in his classification of African languages. In dismantling the inaccurate and racially biased 'Hamitic' of which Nilo-Hamitic was held to be part, Greenberg demonstrated the inadequacy of those former classifications and argued for the connection between Nilotic and Eastern Sudanic.
>
> He noted that 'the Nilotic languages seem to be predominantly isolating, tend to monosyllabism, and employ tonal distinctions' (Greenberg 1963: 92). To the extent that this classification is based on commonplace shared typology and perhaps areally diffused traits, it does not have a firm foundation. Nilo-Saharan is disputed and many are not convinced of the proposed genetic relationships. It is generally seen as Greenberg's wastebasket phylum into which he placed all the otherwise unaffiliated languages of Africa.

First of all, Nilo-Saharan is not classified based on its language typology, which was perhaps areally diffused. There is also a great deal of the more typical lexical and morphological evidence in favor of this language family.

After a few preliminary groupings (Westermann 1912), Greenberg published his revolutionary four-family division of the African languages (Greenberg 1948, 1955, 1970), with one paper focusing on Nilo-Saharan (1950b). It quickly became settled consensus, though some opposition has risen in later decades. However, most specialists have treated Nilo-Saharan as a valid entity since Greenberg's book (Blench and Lameen undated).

Since the publication of *The Languages of Africa* (henceforth *TLA*), a flurry of specialist papers have issued forth, all supporting Nilo-Saharan.[2] What is not certain is exactly which languages are included in it.

There is still substantial opposition to Nilo-Saharan as an entity at all other than Campbell and Mixco.[3] Both sides, pro and anti, have substantial momentum, but the critical mass seems to be forming around the pro side who have picked up supporters and papers in the last decade.

Adding up supporters and leaving out Westermann, we get 12 supporters of at least 85% of the languages that Greenberg put in the stock. There are seven opponents, five if we leave out Campbell and Mixco, neither of whom can be considered Africanist specialists, Mixco in particular. This leaves us 12-5 or 71% in favor of a large chunk of Nilo-Saharan. That should be considered consensus, but it is not a large consensus. Therefore, Campbell and Mixco's statement below is incorrect:

---

[2] Gregersen 1972; Blažek undated; Bender 1991, 1996, 2000; Dimmendaal 1992, 2008, 2011, 2016; Mikkola 1999; Ehret 2001; Blench and Ahland 2010; Blench 2015, 2017, 2019; Starostin 2016, 2017.
[3] Heine 1992; Matthews 2007; Hombert and Philipson 2009; Güldemann 2018.

Nilo-Saharan is disputed and many are not convinced of the proposed genetic relationships. It is generally seen as Greenberg's wastebasket phylum into which he placed all the otherwise unaffiliated languages of Africa.

However, it must be noted that even among supporters, there continues to be dispute about which members of Greenberg's phylum are valid parts of the family. Nevertheless, all supporters, even the most conservative, agree that 85% of the languages Greenberg put into Nilo-Saharan are valid.

One otherwise supporter of Nilo-Saharan feels that Saharan and Kuliak are not Nilo-Saharan (Dimmendaal 2016; Starostin 2017), which gives 65% for inclusion of Saharan and Kuliak as members of Nilo-Saharan. Specialists are not convinced but are optimistic about the inclusion of Saharan and Kuliak as members of the family.

There is even less support for Komuz (= Gumuz) and Koman, Kadu, and Songhay as Nilo-Saharan, as two erstwhile supporters are opposed to the inclusion of these four languages in the stock (Dimmendaal 2016; Starostin 2017). This leaves us with 58% support among specialists for Gumuz, Koman, Kadu, and Songhay as members of the family. There is no consensus that these languages are part of the stock; instead, experts are divided on the question.

With six opposed to Shabo in Nilo-Saharan[4] and four in opposition,[5] consensus (60%) is split but pessimistic about whether Shabo is part of the family. Furthermore, the opponents have the momentum, have published more recent papers, and continue to refine their arguments.

The field is more united that Mimi-D or the Mimi of Decourse, known only from a wordlist (Gaudefroy-Demombynes 1907), is not Nilo-Saharan, with four in opposition[6] and one in support (Greenberg 1970). The discoverer of the language was included among opponents as he also felt it was an isolate. That leaves us with 3 to 1 or 75% solid consensus against Mimi-D belonging to Nilo-Saharan.

It is possible that Mimi-D and Shabo speakers are the remains of hunter-gatherers who spoke a non-Nilo-Saharan language who then went over to Nilo-Saharan, leaving a large layer of substrate behind. The reason for this theory is that there are many words in these languages that seem to have no cognates anywhere in Africa. Nilo-Saharan may be at least 26,000 years old. See the Afroasiatic section for the methodology used to arrive at that figure.

Another form of "Mimi" has also been questionable. This is known as Mimi-Nachtigal or Mimi-N, named after the man who collected a wordlist of it a century ago. Specialists are unanimous that Mimi-N is Nilo-Saharan, in particular a member of the Maban family, albeit being quite distant from the rest of the Maban languages.

This leaves us with a hard core of the Eastern Sudanic (including Berta), Central Sudanic, Maban, Kunama-Ilit, Fur-Amdang families accepted via specialist consensus as Nilo-Saharan. These languages represent 80% of Greenberg's original stock.

Therefore, any conception of Nilo-Saharan based on expert consensus must leave out Saharan, Kuliak, Gumuz, Koman, Kadu, Songhay, Shabo and Mimi-D, which add up to 43 languages out

[4] Bender 2000, Schnoebelen 2009; Taye 2015; Dimmendaal 2016; Starostin 2017; Blench 2019.
[5] Anbessa and Unseth 1989; Fleming 1991; Blench and Ahland 2010 – Ahland is counted as a supporter, while Blench is not as he changed his mind later.
[6] Gaudefroy-Demombynes 1907; Blench and Ahland 2010; Starostin 2011.

of 204. Expert views are optimistic about the inclusion of Saharan and Kuliak – 6% of the languages – in Nilo-Saharan, though recognizing that neither has been shown to be part of the family.

Consensus is still quite divided on 15% of the languages – Gumuz, Koman, Kadu, and Songhay – although there is a trend towards slight optimism. Consensus is pessimistic about the inclusion of Shabo in the family. Furthermore, consensus is that Mimi-D is simply not Nilo-Saharan. The last two together make up 1% of the languages.

Campbell and Mixco, pp. 89-90:

> **Khoisan**: A proposed distant genetic relationship associated with Greenberg's (1963) classification of African languages, which holds some thirty non-Bantu click languages of Southern and Eastern Africa to be genetically related to one another. Greenberg originally called his Khoisan grouping 'the Click Languages' but later changed this to a name based on a created compound of the Hottentots' name for themselves, Khoi, and their name for the Bushmen, San.
>
> Khoisan is the least accepted of Greenberg's four African phyla. Several scholars agree in using the term 'Khoisan' not to reflect a genetic relationship among the languages but rather as a cover term for all the non-Bantu and non-Cushitic click languages.

A look around implies that the claims that Khoisan has been abandoned – that "Khoisan" is now three separate language families and two isolates – are not a complete picture of consensus judgements about the family (see the end of this section for more). Starostin (1998) says Khoisan is a proven entity, citing Bonny Sands (Sands 1998a, b, c) as providing the proof earlier that year.

Starostin (2003) claims that the hard negative position that there is no such thing as Khoisan (Westphal 1962, 1963, 1965, 1971) is no longer popular, although he admitted that Khoisan had not yet been proven. By a decade later, he had assembled an actual Khoisan family (Starostin 2013), including 17 languages: Hadza, Sandawe, Kwadi, the 10 Khoe languages, !Kwi, Taa, Western Hoan, and Ju. He included Sandawe and Hadza, which remains controversial:

> A final point, probably obvious, but one that I still feel is worth mentioning, is that this article, unlike Starostin 2003, is not primarily dedicated to proving the fact of genetic relationship between the various Khoisan subbranches. Rather, it assumes such a relationship as a given and proceeds from there.
>
> This may sound like a bold statement, considering that a general consensus on the issue has not been reached, but, when taken in relation to the goals of the article, it should be viewed as a methodological convenience rather than a categorical statement. The logic is as follows: There exists significant linguistic evidence for Khoisan and Macro-Khoisan, accumulated through lexicostatistical calculations, typological analogies, Greenberg's 'mass comparison', and B. Sands' various methods of testing... (Starostin 2013)

In addition to Greenberg (1948, 1950c, 1955, 1970) and Starostin (2013), Christopher Ehret (1996, 2003) and Bonny Sands (1998a, b, c) also support Khoisan.

Anthony Traill (1986) originally supported it but then changed his mind later on (Traill 1998), saying it was not possible to prove it by our current methods. Khoisanist Henry Honken (1988, 1998) spoke in favor of the family, although he also later changed his position (Heine and Honken 2010). Güldemann (2018) also rejects the family, but he's a noted conservative who also rejects Nilo-Saharan, which is supported now by consensus. Dimmendaal (2008), who is more open-minded, also rejects the family.

There has always been a lot of controversy about the position of Sandawe and Hadza, both of which are far away linguistically and geographically from the other languages, as they are located in East Africa, whereas the rest of the family is in South Africa, Namibia, and Botswana. They also have some of the oldest genes among mankind, dating back 53,000 years. Khoisan itself may be ancient. Lexicostatistics tests show that it may be at least 26,000 years old. The methodology used to calculate that figure is described in the Afroasiatic section.

We count nine specialists in favor of Sandawe as an isolate,[7] four in favor of it as Khoisan,[8] and two who feel that it is related to Khoe, though they do not believe in Khoisan as a whole.[9]

Seven out of thirteen or 54% say it is an isolate, four or 31% say it is Khoisan, two or 15% say it is related to Khoe but that neither is part of Khoisan. At this point, critical consensus on Sandawe is completely up in the air, with specialists unable to decide if it is an isolate, a Khoisan language, or a Khoe language outside of Khoisan.

At this point, Starostin and Sands both say that Hadza is not Khoisan, leaving only Greenberg (1955) and possibly Ehret (Ehret 1996, 2003) to argue for its inclusion in the family. On the other hand, 11 specialists[10] say that Hadza is not Khoisan. 11 against 2 means that 83% – an overwhelming majority – of specialists say that Hadza is not Khoisan. The vast majority of them see it as an isolate.

In recent years, Starostin and Militarev have postulated that there is just as much evidence that Hadza is an Afroasiatic language as there is that it is a Khoisan language. They propose a theory where Hadza is the remains of an old branch of Afroasiatic whose speakers originally spoke a Khoisan language but abandoned it in favor of a new Afroasiatic language. The clicks and click words were retained as substratum. A similar thing has been documented in the Cushitic language Dahalo (Bengtson 2021).

Adding up the views among specialists and leaving out Campbell and Mixco who are not specialists, we find nine opposing the existence of a Khoisan family,[11] with another four arguing that it is a valid entity.[12]

This leads us to nine (69%) against the proposal versus four (31%) in favor of it. This means that critical consensus, while not explicitly rejecting the possibility of a Khoisan family, is on the verge of doing so, and is very pessimistic about whether it is a valid entity. This is not quite the same as Campbell and Mixco's excessively pessimistic view.

Starostin notes that at least among Khoisan professionals, the debate seems to have calmed down from the extreme view of vehement denial of Khoisan's existence (Westphal 1965, 1980) over time to a more nuanced view today.

---

[7] Westphal 1962, 1963, 1965, 1971; Traill 1998; Dimmendaal 2008; Hombert and Philippson 2009; Heine and Honken 2010.
[8] Greenberg 1955; Ehret 1996, 2003; Sands 1998a, 1998b, 1998c; Starostin 2013.
[9] Güldemann and Elderkin 2010.
[10] Westphal 1962, 1963, 1965, 1971; Sands 1998a, 1998b, 1998c; Traill 1998; Dimmendaal 2008; Hombert and Philippson 2009; Güldemann and Elderkin 2010; Heine and Honken 2010; Starostin 2013; Güldemann 2018.
[11] Westphal 1962, 1963, 1965, 1971; Traill 1998; Dimmendaal 2008; Hombert and Philippson 2009; Heine and Honken 2010; Güldemann and Elderkin 2010; Güldemann 2018.
[12] Greenberg 1955; Ehret 1996, 2003; Sands 1998a, 1998b, 1998c; Starostin 2013.

The extreme point of view on this problem, propagated chiefly by the late Ernst Westphal in works such as Westphal (1965, 1980), is not very popular today, for obvious reasons. While these and other works rightly emphasize the current lack of substantial evidence proving the existence of a genetic relationship between the several established groups of 'Bushmen' languages, this by no means gives Westphal a right to claim that such a relationship definitely does not exist.

Furthermore, such an approach can hardly be called constructive when it comes to actually explaining what evidence there is. This is well understood by modern day scholars; therefore, an approach of "moderate skepticism" rather than "decisive denial" is much more popular in Khoisanology today.

The difference between the two approaches is summarized well in an article in which Anthony Traill (1986) both presents the reader with a good selection of comparative material and explains the problems related to its interpretation.

Most of the proposals listed in this article are more deserving of moderate skepticism than decisive denial if one is inclined to reject them. If more conservatives talked like this, the debate would become less political and perhaps we could get back to science again.

You can see above that Campbell and Mixco take the decisive denial road instead of the one of moderate skepticism. Here again they distort consensus in the field, this time by a fine-grained look at how deep the pessimistic views of some of these proposals actually is.

Furthermore, a number of people used to support Khoisan but went over to pessimism later on (Traill 1998; Heine and Honken 2010), which should at least be noted. In addition, not all opponents feel that Khoisan is a Sprachbund. For instance Traill (1998) says it may be a family, but we are not able to make that determination with current methods, which is not the same thing as saying it is a Sprachbund. So while Campbell and Mixco have the basic tune and the words of the song near to correct, they're still singing in the wrong key about Khoisan.

Campbell and Mixco, p. 143-144:

> **Nostratic** (< Latin *nostra* 'our'): A proposed distant genetic relationship that, as formulated in the 1960s by Illich-Svitych would group Indo-European, Uralic, Altaic, Kartvelian, Dravidian and Hamito-Semitic (later Afroasiatic), though other versions of the hypothesis would include various other languages (see Kaiser and Shevoroshkin 1988). Nostratic has a number of supporters, mostly associated with the Moscow school of Nostratic, though a majority of historical linguists do not accept the claims.
>
> There are many problems with the evidence presented on behalf of the Nostratic hypothesis. In several instances the proposed reconstructions do not comply with typological expectations; numerous proposed cognates are lax in semantic associations, involve onomatopoeia, are forms too short to deny chance, include nursery forms, and do not follow the sound correspondences formulated by supporters of Nostratic.
>
> A large number of the putative cognate sets are considered problematic or doubtful even by its adherents. More than one-third of the sets are represented in only two of the putative Nostratic branches, though by its founder's criteria, acceptable cases need to appear in at least three of the Nostratic language families. Numerous sets appear to involve borrowing (see Campbell 1998, 1999.) It is for reasons of this sort that most historical linguists reject Nostratic.

It is correct that consensus among specialists is to reject Nostratic, but serious papers taking apart the proposal seem to be lacking. However, there have been a fair number of critiques.[13] Sadly, it is beginning to enter into the same emotionally-charged terrain as Altaic and Amerind, particularly the former, and belief in it is becoming a thing of ridicule as it is for Altaic. Nevertheless, there have been a few excellent linguists doing work on this very long-range family for decades.[14]

Also, since Nostraticists are using the comparative method of reconstruction of proto-families and finding sound correspondences, they win the approval of even critics like Campbell, who at least appreciate that derided methods like mass comparison are not being used (Campbell 1998).

I am not going to bother to count up pro and con experts to look for consensus. Recall that by our criteria to be an expert, all you have to do is to have published a paper on the language family as a whole or even on a single group of the family.

So for Altaic, the views of all experts in Turkic, Tungusic, Mongolic, and in the relation between Koreanic and Japonic and core Altaic families. I limited it like this because so many Japanese linguists reject Japonic-Koreanic, possibly on nationalist or ethnocentric grounds, that this would seriously tip the scale against Altaic. Besides, simply saying "Japonic is not related to Koreanic" says nothing about the hypothesis involving the three core families (Turkic, Tungusic, Mongolic) or the relation of any of these core families with either Japonic or Koreanic.

For Nostratic, I would probably count an expert in Indo-European, Uralic, Altaic, Yukaghir, Kartvelian, Dravidian, Nivkhi, Chukchi-Kamchatkan, Eskimo-Aleut, and Afroasiatic who had commented on the subject, which to me implies that they have looked into it. You don't have to publish a paper on Nostratic as a whole, and few papers opposing the proposal have been published anyway. I should note that some consider Afroasiatic and Dravidian to be sister languages to Nostratic as a whole instead of constituent families of Nostratic itself.

So once you cast the net that wide, a survey of experts in any of the constituent Nostratic families who has an opinion on the matter would probably tilt consensus towards strong rejection. Some linguists are agnostic on the matter, which probably isn't a lot different: "It doesn't exist" versus "It hasn't been proven yet" seems to be a distinction without a difference. Since we will be looking at experts in any of the constituent families of Eurasiatic too (a subset of Nostratic), this will also encompass many folks, most of whom no doubt will reject the proposal.

Nevertheless, just looking at the pronouns, the Eurasiatic part of Nostratic (plus Kartvelian) must be true, with repetition of the *m*- '1st person' and *t*- '2nd person' pronoun paradigm strikingly obvious across all of those families. You have to see the families lined up in a chart of the pronouns of the world's language families to see how shocking the agreement is.

There's nothing like this sort of striking agreement across multiple unrelated families anywhere else on Earth except the Americas, where we are again shocked by a striking lineup of *n*- '1st person' and *m*- '2nd person' pronouns, a paradigm which does not exist in any large family on Earth. Everywhere else, the conservatives are somewhat correct that "pronouns go everywhere."

[13] Trask 1996; Campbell 1998, 2004; Campbell and Poser 2008; Georg 2013; Kallio and Koivulehto 2017.
[14] E.g. Illich-Svitych 1965; Dolgopolsky 1998; Bomhard 2008.

At any rate, the lineup between unrelated families is poor elsewhere, which makes the Eurasiatic and Amerind lineups so striking in the first place.

I can't think of any other reason for the matches among these pronoun paradigms than genetics. Any other theory is so bad it's almost laughable. However, that's never stopped conservatives, who posit absurd and comical theories all the time to explain away vast morphological and lexical resemblances in long-range proposals. If theories must be judged against their competitors, then many long-range proposals would win simply because they are so much better than the ridiculous conservative theories thrown up against them.

That said, it's not completely apparent to me that Dravidian and Afroasiatic match well with the rest of the somewhat expanded Eurasiatic. I will side with conservatives with that theory and say it's not yet proven.

Campbell and Mixco, p. 58:

> **Eurasiatic:** Greenberg's hypothesis of a distant genetic relationship that would group Indo-European, Uralic-Yukaghir, Altaic, Korean-Japanese-Ainu, Nivkh, Chukotian and Eskimo-Aleut as members of a very large 'linguistic stock'. While there is considerable overlap in the putative members of Eurasiatic and Nostratic, there are also significant differences. Eurasiatic has been sharply criticized and is largely rejected by specialists (Georg and Vovin 2003, 2005).

Greenberg's (2000, 2002) Eurasiatic phylum is almost certainly not supported by the consensus of specialists. See the Nostratic theory above for my definition of a specialist, which in the cases of these macrofamilies is quite broad. Nevertheless, critiques of the proposal have been uncommon, though there have been a few (Georg and Vovin 2003; Daniels 2004).

However, an interesting computerized study of ultraconserved words in the highly prestigious journal *Proceedings of the National Academy of Sciences*, or *PNAS* (Pagel *et al.* 2013), claimed to have validated the Eurasiatic proposal, or at least the part of it consisting of Indo-European, Uralic, Altaic, Eskimo-Aleut, Chukchi-Kamchatkan, and Kartvelian. Dravidian, not part of the Eurasiatic proposal but instead part of the Nostratic theory, was also validated. I assume that Macro-Altaic with Japonic and Koreanic was the Altaic tested. Yukaghir, Nivkh, Ainu, Tyrsenian, Etruscan, and Raetic were either not validated or possibly not even tested for.

A critique of the study showed up later that year in the same journal (Heggarty 2013). Sarah Thomason also rejected it in a post on *Language Log* (Thomason, May 8, 2013). However, it should be noted that Thomason is as extreme of a conservative as Campbell, and the *Language Log* website is a hornet's nest full of belligerent conservatives or frankly just nasty folks all around where the snark rises to radioactive levels. It suffers from "Internet anonymity syndrome" where you can act as awful and mean as you want as long as you are hiding under a pseudonym and there's no accountability for what you say. That doesn't make her wrong, but I think the connection should be noted.

I haven't read either the study or the rejoinder, and I'm not quite clear how whatever they were doing proves or disproves a linguistic relationship.

On the other hand, I don't really understand the methodology used in either paper, as it's all rather technical to me. Therefore, I can't comment on the study itself or the rebuttals which do not

appear to have good arguments. All of the rebuttals were successfully answered by the study's authors. It does strike me though that along with the pronouns, both studies seem to be onto something.

A computational phylogenetic comparison was published two years later in the same journal (Jäger 2015), which validated much though not all of Eurasiatic, this time finding a clade consisting of: Core Altaic (Mongolic, Turkic, and Tungusic), Yukaghir, Nivkh, Uralic, Chukotko-Kamchatkan, and Indo-European. Notably, Dravidian and Japanese were part of other clades. These ASJP studies seem similar to mass comparison in that they give a good, broad overview and are often correct in a broad sense – that is, most of the families linked up are probably either related or linked together as a result of massive borrowing (Sino-Tibetan and Hmong-Mien).

In a few places, its results are simply erroneous. For instance, in this study (Jäger 2015), Dravidian forming a clade with Yeniseian and Nakh-Daghestanian is clearly false. Nakh-Daghestanian may be linked in a deep long-range proposal, but it is widely rejected and will not be discussed here. But Dravidian should not be there at all.

Other links validated include Austronesian and Thai as Austro-Tai, and, interestingly enough, Austroasiatic, Japonic, and Ainu. Austro-Tai is already validated by consensus.

The latter theory is interesting but it ties in well with what linguists have been noticing for a long time about Japanese, which is that there are two layers – an Altaic layer and a so-called "Austric" layer of which Austroasiatic and Ainu would be a part. The Austric layer may be substrate and was probably brought to the islands by the Ainu, who themselves link back to Thailand 25,000 YBP on cranial studies which also link with Aborigines and other Australoid peoples. The Altaic layer may have come from the Yayoi immigrants from Southwestern Korea who invaded Japan 2,300 years ago and slowly subjugated the Ainu over centuries.

It's conceivable that the two languages mixed somehow. Whether Japanese is an Altaic language with an Austric substrate or a true mixed language consisting of Altaic and Austric is not known, but the truth is probably somewhere around theories linking up those huge macro-families in the formation of the Japonic languages.

Greenberg's Eurasiatic proposal includes the following macro-families, existing families, and isolates: Indo-European; Uralic; Yukaghir; Nivkh; Eskimo–Aleut; Chukotko-Kamchatkan (Chukotian); Macro-Altaic, consisting of Micro-Altaic (Turkic, Mongolic, and Tungusic) along with Koreanic and Japonic; Ainu; Tyrsenian; Etruscan; and Raetic. Eurasiatic is one of the better-supported long-range proposals out there.

Based on the data from the *m-/t-* 1st and 2nd person pronoun paradigm alone (see also above under Nostratic), a good argument could be made that most but not all of Eurasiatic is a valid entity, namely including Indo-European, Uralic-Yukaghir, Altaic including Japanese and Korean, Chukotian, and Eskimo-Aleut, leaving out Nivkhi for the time being and certainly leaving out Ainu. Nivkhi does seem to be a Eurasiatic language, but it may not be a separate node. Instead, it may be a part of the Chukotian family (Fortescue 2011). Or even better yet it seems to be part of a family connected to the New World via the Algonquian-Wakashan or Almosan family in the Americas (Fortescue 1998; Nikolaev 2015a, 2015b).

Ainu is very hard to place and most specialists continue to regard it as an isolate. It may be closer to Austric languages than to Eurasiatic. I have no idea if the inclusion of Tyrsenian, Etruscan, and Raetic is valid.

Eurasiatic may be a more solid entity than Nostratic. Correlatively, the parts of Nostratic that overlap with Eurasiatic may be the best supported.

Campbell and Mixco, p. 82-83:

> **Indo-Anatolian:** The hypothesis associated with Edgar Sturtevant that Hittite (or better said, the Anatolian languages, of which Hittite is the best-known member) was the earliest Indo-European language to split off from the others. That is, this hypothesis would have Anatolian and Indo-European as sisters, two branches of a Proto-Indo-Hittite.
>
> The more accepted view is that Anatolian is just one subgroup of Indo-European, albeit perhaps the first to have branched off, hence not 'Indo-Hittite' but just 'Indo-European' with Anatolian as one of its branches. In fact, the two views differ very little in substance, since, in either case, Anatolian ends up being a subfamily distinct from the other branches and in the view of many the first to branch off the family.

As noted, the original Indo-Hittite argument was laid out in a series of articles by Edgar Sturtevant (1931, 1932, 1933, 1940, 1942). It was later supported by Greenberg (1987). Starting in 1938, a number of linguists attacked the original Indo-Hittite argument,[15] and after a while, it fell out of favor.

Papers defending Sturtevant's strong Indo-Hittite argument continue to be published (Fournet 2021a). However, the author does not use the term "Indo-Anatolian" as he considers it better to see Indo-European as a single language with two branches, Anatolian and non-Anatolian (Fournet 2021b). Nevertheless, while the author indeed has Anatolian as a single node and divides Indo-European into two major nodes – Anatolian and non-Anatolian –, on the other hand, he derives Indo-European itself, of which Anatolian is a major node, from something called "Proto-Anatolic" language (Fournet 2021a). This is more extreme than Sturtevant's position, but it may not be completely unwarranted.

The debate is confusing and gets tied up in semantics a lot. A recent paper noted the contradictions and difficulties in dealing with Anatolian without a specific term for the non-Anatolian branch (Forston 2010: 11).

> 1.16. The discovery of Anatolian and Tocharian in the Twentieth Century has further fueled the debate over the internal structure of the PIE family tree. Anatolian in particular is significantly different from the picture of PIE that scholars had developed by the close of the nineteenth century on the basis of the other branches, even though Anatolian is the oldest attested branch.
>
> In the 1930s, the American linguist Edgar Sturtevant proposed that PIE was not the ancestor of Anatolian, but a sister of it, and that both PIE and Anatolian were descended from a language he called Indo-Hittite. Sturtevant's theory was not widely followed, and the term "Indo-Hittite" has largely been abandoned. But the hypothesis that Anatolian, and then Tocharian, split off from the family first, and that the remainder of PIE underwent further common development before the other branches emerged has found increasing support in recent years and is in many ways little different from Sturtevant's original claim.

[15] Pedersen 1938; Jasanoff 2003; Holm 2008; Holm 2008, 2010; Melchert 2012.

A new paper (Melchert 2012) claims to represent the new consensus view on the Indo-Hittite question. According to the author, the new view rules out both the strong Indo-Hittite hypothesis and the view of Campbell and Mixco above that Anatolian is just another branch of Indo-European, albeit one that split off first in favor of a third, different position.

A third response to the new evidence of Hittite was exemplified by the work of Jerzy Kuryłowicz, in both phonology and morphology (see respectively 1927 and 1964 as representative). Kuryłowicz and others rejected both the "Schwund Hypothese" and "Indo-Hittite", contending that proper integration of the Hittite evidence demanded a radical and far-reaching revision of reconstructed PIE – meaning PIE as the source of not only Hittite, but also the non-Anatolian languages including Sanskrit, Greek and the rest.[16]

It is uncertain how much this argument differs from Forston's or whether both positions are saying the same thing in different ways.

Nevertheless, everyone seems clear that both the hard Indo-Hittite hypothesis and Campbell and Mixco's argument above have been largely abandoned. In their place a new argument, something along the lines of what Forston and Melchert seem to be saying, is the new consensus.

Therefore, Campbell and Mixco seriously distort consensus on this question by referring to their argument as consensus when it has been largely abandoned and then by saying that their position is actually little different from Sturtevant's hard Indo-Hittite.

Campbell and Mixco, p. 84:

> **Indo-Uralic:** The hypothesis that the Indo-European and Uralic language families are genetically related to one another. While there is some suggestive evidence for the hypothesis, it has not yet been possible to confirm the proposed relationship.

This summary seems too negative. Indo-Uralic is probably one of the most promising long-range proposals out there. There is persuasive evidence that the two are related at some level via the Eurasiatic hypothesis, which seems proven by the pronouns alone. So they're clearly related. But was there a special Indo-Uralic node? Frederik Kortlandt has done a lot of good work on this idea (Kortlandt 1989, 2002). Kortlandt is fairly conservative and a number of other fairly conservative linguists have worked on Indo-Uralic, which implies there may be something there.

Also, perhaps Campbell is more open to Indo-Uralic than he lets on above:

> I should mention that, in person, Campbell is not quite as conservative as he appears to be in print. He, himself, has even delved into a comparison of Indo-European and Uralic tree names. (Bomhard 2021)

A recent paper undertook a very strict mathematical comparison of IE and Uralic and came up with 14% cognates using very conservative criteria to qualify as a cognate (Kassian *et al.* 2015). The paper generated a flood of responses from the usual suspects such as Donald Ringe (2015),

[16] Other representatives of this viewpoint include Watkins (1969), Meid (1963, 1975), Neu (1976, 1985), and Adrados (1963, 1982, 2007). Unsurprisingly, these scholars disagree, sometimes markedly, on just what radical revisions should be made.

but it seemed very well done. In particular, the mathematics was sound. Chance resemblance between the two languages would have resulted in 7% cognates (Kassian *et al.* 2015), hence the paper appears to be proof that there is a connection between Indo-European and Uralic.

Starostin used the same precise mathematical method in the Khoisan and Nilo-Saharan families.

Furthermore, the alternative hypothesis pushed by Campbell (1998) , himself a noted Uralicist, that almost all of the differences are due to borrowings, seems impossible. He continually references Rédei (1990, 1999) who wrote the original etymologies for Proto-Uralic, but Rédei automatically assumed all similarities between Indo-European and Uralic must be due to borrowing.

This is because his theory was this:

How can we explain these similarities? They are either indicative of genetics and Indo-Uralic is a thing, or they must be borrowings. Since Indo-Uralic clearly doesn't exist, they must be borrowings. This is a typical conservative feint that has been recently used by Mithun in the Americas. Conservatives are going from theory to facts instead of the other way around as in proper science.

In the last twenty years, Uralicists have gone back over the original Proto-Uralic etymologies and gotten rid of fully half of them (from 2,000 down to 1,000) for a variety of poor reasons, mostly irregular sound correspondences. It appears that while there were some obvious bad etymologies in there, most of the ones that were thrown out were perfectly good.

This is the work of the Janhunen and Aikio school of Uralistics, an excessively conservative school of thought. Everything in an etymology must be phonologically explained, including very complicated vowel correspondences, or else the etymology is thrown out. This is excessively cautious conservatism. Some 50% of Indo-European etymologies have irregular sound correspondences. This way of thinking is dangerous because some of these people have suggested to me that we throw out much of Pokorny's Indo-European dictionary on this basis. It doesn't seem reasonable at all to throw out half of Pokorny because the correspondences aren't perfect.

Ante Aikio's important paper on Mari sound correspondences (Aikio 2014b) meant that this school of Uralistics ended up putting back in a lot of the etymologies that they had previously taken out. This boils down to dismantling the wheel and then putting it back together again in order to get the cart rolling. It seems like a lot of unnecessary work.

What the Uralicists have done is vandalism. This is not just conservatism. It is out-and-out reaction. It is akin to a city council declaring that all of the old beautiful buildings in the city are going to be torn down because they were not constructed properly. Will they be rebuilt? Well, of course not. Most of the top Uralicists are involved in this silly and destructive project.

Campbell and Mixco, pp. 225-226:

> **Yukaghir:** A small language family of Siberia, composed of Tundra (Northern) Yukaghir and Kolyma (or Southern) Yukaghir. It is often thought possibly to be related to Uralic, though the evidence has not yet been sufficient to confirm this proposal.

The idea that the Yukaghir languages at the Kolyma River in far northeastern Siberia may be distantly related to the Uralic family has a long pedigree.

The closeness between the two language families was first observed a hundred years ago by Heikki Paasonen (1907) and Ernst Lewy (1928), although neither felt that a genetic relation was

provable at the time. Holger Pedersen (1931, 1933) included the family in his original Nostratic family. The first serious attempts to link the two languages came a decade later (Bouda 1940, 1952; Collinder 1940).

Collinder later elaborated on this theory in subsequent papers (Collinder 1957, 1965a, 1965b). Further efforts to elaborate the theory continued until recently.[17] And although he has not published anything on the connection yet, noted Indo-Europeanist and Altaicist Václav Blažek also supports the theory (Piispanen 2021).

Long-rangers building on Collinder's original evidence include Uralo-Yukaghir in their Eurasiatic and Nostratic proposals.[18] A computerized phylogenetic ASJP study found close links among the two languages, with Uralic and Yukaghir sharing a clade with Nivkhi (Jäger 2015).

The links between the two language families in core vocabulary are vast and include personal pronouns, demonstratives, numerals, kinship terms and verbs. Verbs are important because they are much less likely to be borrowed than nouns, yet there are far more verb matches in the two families than noun matches, as is expected in a typical borrowing scenario. In addition, the 1st and 2nd person pronoun paradigms between the two groups are so similar it is almost eerie.

There are indeed opponents of Uralic-Yukaghir, to which we will include Campbell, as he is a notable Uralicist. All of them are Uralicists.[19] The attempts at opposing the theory date back thirty years. The noted Uralicist Rédei (1990, 1999), as he did in opposing Indo-Uralic, explained all of the resemblances as borrowings. Later, more objections were written by Häkkinen (2012) and Aikio (2014), whose theories are now received wisdom for Uralistics (Piispanen 2021).

Häkkinen (2012) made an attempt to explain resemblances between the families by moving Proto-Uralic far to the east of its homeland in the West Siberian Plain and Proto-Yukaghir far to the west of its homeland at the Kolyma River. He also had a time problem as Proto-Uralic dates back 6,000 YBP in the Siberian Plain and Proto-Yukaghir dates back a mere 2,500 years at the Kolyma River.

Hence speakers of the languages would have had to time-travel 3,500 years to meet up with each other to trade words. Opponents of the Uralo-Yukaghir theory are also forced to deal with a large amount of core vocabulary dating back a long time, which they must explain by unlikely borrowing scenarios. Häkkinen (2012) tried to solve this problem by pushing a heavy borrowing of core vocabulary all the way back to not just Proto-Uralic and Proto-Yukaghir but Pre-Proto-Uralic and Pre-Proto-Yukaghir.

Heavy borrowing of core vocabulary at the proto-language level is another way of saying genetics. "Pre-Proto-Uralic" at 8,000 years means nothing less than Uralo-Yukaghir. What else could it mean? Further, as Yukaghir only goes back 2,500 years, how is it that Pre-Proto-Yukaghir goes back 8,000 years? Pre-Proto-Yukaghir sat in place for 5,500 years before transforming into mere Proto-Yukaghir? That could not have happened.

---

[17] Harms 1967; 1977; Nikolaeva 1998a, 1998b, 2006; Piispanen 2013; Krejnovich 1978; van Driem 2009; Babaev 2009; De Smit 2019.
[18] Ruhlen 1987; Greenberg 2000, 2002; Bomhard 2008.
[19] Rédei 1990, 1999; Häkkinen 2012; Aikio 2014; Zhivlov 2020.

Of course, mass borrowing at the pre-proto-language level was originally used by Doerfer in his anti-Altaic arguments.

This was the first attempt at alternate facts that Uralicists have come up with to explain their unlikely borrowing scenarios between the two languages.

Aikio (2014) did the same to explain his mass borrowings between Samoyedic and Yukaghir. Proto-Yukaghir goes back to the Kolyma River 2,500 YBP (Piispanen 2021). Proto-Samoyedic dates from 5,300 YBP in the Sayans. The proto-languages would have had to time-travel over a period of 2,500 years because at the time of Proto-Samoyedic, there was no Proto-Yukaghir to borrow into and at the time of Proto-Yukaghir, Samoyedic was long since broken up into small pieces (Aikio 2014).

The two proto-languages were never adjacent in either time or space, so Aikio had to make up "alternative facts" positing new homelands for both language families in places where they never were. Aikio (2014) was forced to shove Samoyedic 1,000 miles to the north of the Sayans and Yukaghir 1,000 miles south of the Kolyma River to make them meet so they could borrow from each other. Yet there is no evidence that either group was in any of these made-up homelands.

But they must have been in these places because conservatives are once again going from theory to facts, and where the facts don't match the theory, alternative facts need to be created because the theory is obviously true, I suppose in a Biblical sense. They're doing it backwards. You go from facts to theory in science and when the two don't match up, it's back to the drawing board to get a new theory to explain the facts better. You don't get to change the facts into "alternative facts." Theories are malleable; facts are not.

Such unverifiable alternative facts that must be true because the theory behind them is infallible is a case of the theory being *unfalsifiable*. Therefore, many conservative arguments are not only not provably true, but, so much worse, they are not provably false either. These arguments are literally *not even wrong*!

Resorting to "alternative facts" shows that you are no longer doing science but instead you are doing what I call "politics," as I see the two in polar opposition. You're either doing one or the other. There's no in-between.

Aside from consensus, Uralo-Yukaghir supporters feel that they have compiled an impressive degree of evidence of such quantity and quality that they are certain that the language families are related. One notable supporter of the Uralo-Yukaghir hypothesis put it like this:

> Vast geographic distance (borrowing of basic vocabulary over such extreme distances is supposed to have happened whereas no such thing would ever have been suggested between any other two languages in the world), severe chronological problems (borrowings from several different geographically disparate locals from several layers of different protolanguages are supposed to explain lexical similarities), and "borrowing" of basic vocabulary (which might happen but is not common on such a scale; only Altaic is a parallel and that too is a valid genetic language relationship on some level) are huge problems for the anti-Uralic-Yukaghir camp, but such matters are not discussed at all.
>
> On the other hand, the sharing between the families of a near-perfect phonetic system, several identical suffixes, typological parallels, similar pronominal and numeral systems, similar basic and non-basic vocabulary (of Paleolithic age) and, actually, several regular sound laws, along with results from population genetic studies from

> outside the field, are very strong pro-Uralic-Yukaghir camp arguments, along other matters not even yet properly discussed.
>
> The Uralic and Yukaghir languages are simply genetically related languages and are two branches of an ancient Uralo-Yukaghir language (which might actually be additionally related to the Eskimo languages), where the Yukaghir languages are more closely related to the Eastern Uralic languages than to the Western group (Piispanen 2021).

If we leave out the primitive groupings which were not conclusive (Paasonen 1907; Lewy 1928), we have considerable specialist support for the family.[20] If we include Blažek, we get 11 supporters of Uralo-Yukaghir. The proposal dates back a long time, and efforts continue to this day and have not slowed down. Instead, as with Hokan and Penutian, evidence in favor of a relationship continues to pile up.

Opponents got off to a late start thirty years ago, but they have continued to publish powerful arguments up until the last decade. They are all coming out of the school of Uralistics. If we count Campbell, we end up with five opponents of Uralic-Yukaghir (see footnote 19 above). We are leaving out Mixco's judgement because he's not a specialist. Granted, opponents are rather strident, though avoiding the out-of-control rhetoric of opponents of Amerind and Altaic, which is to their credit. All are superb scholars. On the other hand, they just don't have the numbers on their side and are badly outnumbered by supporters of the theory.

We end up with 11 against 5, or 69% in favor of Uralic-Yukaghir. That would be right on the edge of it not quite all the way to critical consensus, which continues to be very optimistic about the relationship if not fully supportive of it.

Campbell and Mixco, pp. 7-8:

> **Altaic**: While 'Altaic' is repeated in encyclopedias and handbooks, most specialists in these languages no longer believe that the three traditional supposed Altaic groups, Turkic, Mongolian and Tungusic, are related. In spite of this, Altaic does have a few dedicated followers.
>
> The most serious problems for the Altaic proposal are the extensive lexical borrowing across Inner Asia and among the 'Altaic' languages, lack of significant numbers of convincing cognates, extensive areal diffusion, and typologically commonplace traits presented as evidence of relationship.
>
> The shared 'Altaic' traits typically cited include vowel harmony, relatively simple phonemic inventories, agglutination, their exclusively suffixing nature, (S)OV ([Subject]-Object-Verb) word order, and the fact that their non-main clauses are mostly non-finite (participial) constructions.
>
> These shared features are not only commonplace typological traits that occur with frequency in unrelated languages of the world and therefore could easily have developed independently, but they are also areal traits shared by a number of languages in surrounding regions, the structural properties of which were not well-known when the hypothesis was first framed.

It is true that Altaic is still up in the air, and anti-Altaicists are not rare, even currently (Nichols 2012; Vajda 2020), but Campbell and Mixco are not correct when they say that the idea has been abandoned. Most linguists in the West regard it as a laughingstock, and if you say you believe in it, you will experience intense bullying and taunting from them. Oddly enough, in Russia, Altaic

[20] Pedersen 1931, 1933; Bouda 1940, 1952; Collinder 1940, 1957, 1965a, 1965b; Harms 1967, 1977; Krejnovich 1978; Nikolaeva 1998a, 1998b, 2006; Babaev 2009; van Driem 2009; Piispanen 2013; De Smit 2019.

is regarded as obviously valid. The anti-Altaicists are frankly belligerent, hysterical, and typically violate the rules of academic decorum (see the Amerind debate for a similar situation). A recent overview (Blažek 2019) looked at the history of pro- and anti-Altaic views over time and showed that neo-pro-Altaicists are quite common.

In fact, many papers have been published both recently and historically that make an excellent case for Altaic.[21] In addition, a computerized ASJP phylogenetic study (Jäger 2015) found Turkic, Mongolic, and Tungusic represented a clade within a larger more or less "Eurasiatic" pattern. However, Japanese was not included in this clade and instead matched more with Austric languages.

The problem is that most of the linguists who will laugh in your face and call you an idiot if you believe in Altaic *are not specialists in these languages*. I did a quick study of Altaic specialists, current and historic (Lindsay 2021), and 79% of them believe in some form of Altaic, if only some reduced form, with at least two of the five "Transeurasian" language families (in Martine Robbeets' sense) being related. Examples would be Japonic-Koreanic, Tungusic-Mongolic, Tungusic-Koreanic, Turkic-Mongolic, etc. The figure also included supporters of the traditional tripartite system (Turkic-Mongolic-Tungusic) and the expanded five-family proposal (same three, plus Koreanic and Japonic). Only 21% wholly reject the Altaic proposal.

A more elaborate and better referenced chronology of the Altaic debate listing both supporters and opponents both historically and currently is available in Blažek (2019). Though it doesn't count numbers of specialists, a quick overview shows it arrives at much the same conclusion as I did. At the moment, consensus among specialists is that at least two of the five Transeurasian families are related to each other. So consensus is that Transeurasian exists, if only in a truncated or reduced form.

Robbeets (2004) tested Japanese against the four other Transeurasian families on the Swadesh's 100-word list. The results showed a 41% match between Japanese and at least one of the other four languages. Further, in one-on-one matches between the four other families, scores ranged from 17 to 24%. So it's clear that Japanese is related to all four of the other families on some level.

The anti-Altaicists are pushing a massive lie – that critical consensus has completely abandoned Altaic and regards it as a laughingstock – but their project is more propaganda than science. In particular, it's a fad. So Altaic is in the preposterous position where almost all of the people who know nothing about it will laugh in your face and call you an idiot if you believe in it, and the overwhelming majority of specialists will say it's real, at least in a two-family "micro" form.

Altaic must be the only nonexistent family that has an incredibly elaborate 1,000-page etymological dictionary, full reconstructions of the proto-languages, etymologies of over 2,000 Altaic terms, and elaborate sound correspondences running through it (Starostin *et al.* 2003). The anti-Altaicists use the "we can't reconstruct the numerals so it's not real" line here. See the entry on

[21] Ramstedt 1924; Hashimoto 1986; Bisang 1998; Whaley and Fengxiang 2000; Nevskaya 2010; Robbeets 2005, 2010; Blažek 2019; Robbeets *et al.* 2021.

the universally accepted Afroasiatic for how an obvious language family can have few if any numerals to reconstruct.

Altaic is obviously true based on 1st and 2nd person pronoun paradigms at an absolute minimum (Babaev 2009). Because the pronoun paradigm argument is so powerful in favor of Altaic, conservatives have begun to attack the argument in papers (Nichols 2012).

However, a quick look at the pronouns in the world's major language families shows the power of the pronoun argument. Pronouns are actually quite varied around the world, and even in obvious language families do not always line up well (Babaev 2009).

However, the *m-/t-* pronoun paradigm in Altaic and in Northern Eurasia in general stands out as a shocking agreement compared to other families. The agreement extends across a number of language families and is stunning in its accuracy (Babaev 2009). The only other place on the globe where we can find a similar lineup on pronoun paradigms is in the Americas, with the *m-/n-* paradigm.

Once again, the resemblance is so blatant compared to those between other families that it almost knocks you out of your chair (Babaev 2009). The *m-/t-* and *m-/n-* paradigms in Eurasiatic and Amerind, respectively, stick out like a sore thumb and demand an explanation, but only a genetic explanation is available. There's simply no other argument.

The anti-Altaic arguments are often ridiculous. As noted, they dismiss a 1,000-page etymological dictionary with 2,300 reconstructed etymologies (Starostin *et al.* 2003) as a hallucinated work. Where else on Earth can we find an equally documented language family that is said not to exist?

There are vast parallels in all three families at all levels, in particular in the Mongolic-Tungusic family, which gets a 100% score with computer phylogeny programs. The go-to argument here has always been that these changes are all due to borrowings, but for this to have occurred, borrowing would have had to occur between large, far-removed language families on a scale the likes of which has never been seen anywhere else on Earth.

Furthermore, the borrowings must have occurred at an early stage. Gerhard Doerfer, the famous Turkologist opponent of Altaic, had to keep pushing his massive borrowings of core vocabulary further and further back until he eventually had the scenario taking place at the Proto-Turkic, Proto-Tungusic, and Proto-Mongolic levels.

In my opinion, massive borrowing of core vocabulary at the proto-language level is simply another word for genetics. This argument – mass borrowing of core vocabulary at the proto-language level – has also been used by conservatives in the cases of Uralo-Yukaghir (Häkkinen 2012) and Quechumaran (Adelaar 1992) to explain away genetic explanations for those families.

The argument that the entire 1st and 2nd person pronoun paradigms in Altaic have been borrowed is particularly absurd because 1st and 2nd person pronouns are almost never borrowed anyway (Ruhlen 1986, Bancel and Bengtson 2021), and there has never been a single case on Earth of the borrowing of a 1st and 2nd person pronoun paradigm (Babaev 2009), much less twice in a row between three different proto-languages. So the anti-Altaicists are arguing that something that has never happened anywhere on Earth not only happened but happened more than once among

five different proto-languages. The anti-Altaic argument is that something that could not possibly have happened actually occurred.

Campbell and Mixco, pp. 90-91:

> **Korean (Japonic-Koreanic)**: A language isolate, spoken by over 70 million in the two Koreas and adjacent areas. Korean is sometimes cited as an example of an agglutinative language, and a language with SOV (Subject-Object-Verb) word order. Some scholars classify Korean in a single family with Japanese; however, this is a controversial hypothesis.
>
> Korean is often said to belong with the Altaic hypothesis, often also with Japanese, though this is not widely supported.

Japonic-Koreanic has a fair amount (~35%) of support among Altaicists. Campbell and Mixco are excessively negative about the level of support for Japonic-Koreanic (which has long had passionate supporters) and an expanded Altaic or Transeurasian. The notion of Japanese as an Altaic language is an old one (Ramstedt 1924), and evidence has only accumulated over a century (Bisang 1998; Robbeets 2005, 2010; Robbeets *et al.* 2021).

Nevertheless, to be fair, it should be noted that among Japanologists in Japan itself, the vast majority reject Japonic-Koreanic, but this may have more to do with sociology, particularly the noted hyper-ethnocentrism of Japanese society, than it does with science. Notably ethnocentric ethnic groups are more likely to think that their family has no relatives and is therefore special and not influenced by outsiders.

But if we include Japanese linguists – and we should (ethnocentric or not) for questions of only a relationship between the two families — then current consensus definitely opposes Japonic-Koreanic, perhaps strongly so. It should also be noted that a computerized phylogenetic ASJP study that included Core Altaic and Japanese found that Japanese did not match with the rest of Altaic; instead, it lined up more with Austric languages (Jäger 2015).

Robbeets (2004) compares Japanese with the other four families in Altaic since a lot of people say that expanded Altaic is not supported. Japanese was compared to all four other families on the Swadesh 100-word list. Sifting criteria were very strict, possibly excessively so. At any rate, this is not a permissive comparison at all. It's quite a conservative one. Japanese scores 41% against each of the other families, with at least one match necessary for a positive score.

Based on this study, Japanese is absolutely related to at least one of these other languages. There's no other possible conclusion. Most of the 41% match with more than one family, which is particularly impressive. Chance would be 7%. Japanese is surely related to at least one of these other families. There's no interpretation other than genetics for a 41% score with the Swadesh list.

Campbell and Mixco. p. 225:

> **Yeniseian, Yenisseian**: Small language family of southern Siberia of which Ket (Khet) is the only surviving member. Yeniseian has no known broader relatives, though some have been hypothesized (see the Dené-Caucasian hypothesis).

Campbell and Mixco state a serious untruth here, including some weasel words. By discussing Dene-Caucasian in the same breath as relatives of Yeniseian, they are able to deflect away from the more widely accepted proposal of a link between Yeniseian in the Old World and Na-Dene in

the New World. This is Edward Vajda's Dene-Yeniseian proposal (Comrie 2008; Vajda 2010, 2012).

This long-range proposal has the support of many people.

Of the 20 experts who have weighed in on Dene-Yeniseian, 19 of them had a positive view.[22]

Campbell and Mixco are the only two who are negative, but Mixco is not an expert on the family. Almost all specialists in either or both families support the proposal. When 19 out of 20 is not enough, one wonders at what point the field reaches a consensus.

Campbell and Mixco, p. 141:

> **Nivkh** (also called **Gilyak**): A language isolate spoken in the northern part of Sakhalin Island and along the Amur River of Manchuria, in China. There have been various unsuccessful attempts to link Nivkh genetically with various other language groupings, including Eurasiatic and Nostratic.

A recent paper by Sergei Nikolaev made an interesting case that Nivkhi is related to Algonquian-Wakashan, a family of languages in the Americas (Nikolaev 2015a, 2015b). One of these languages is Wakashan, and there has been talk of links between Wakashan and the Old World for some time.

Michael Fortescue places Nivkhi in Chukotko-Kamchatkan (Fortescue 2011). But he also says that it is related to Mosan (Fortescue 1998). Greenberg places it in Eurasiatic as a separate node (Greenberg 2000). But as Chukotko-Kamchatkan is part of Eurasiatic as a primary node, they are both saying the same thing in a way. A computerized phylogenetic ASJP study (Jäger 2015) linked Nivkhi with Uralic and Yukaghir in a sort of "Siberian" clade. Chukotko-Kamchatkan was also part of this clade together with Indo-European, but at a somewhat removed distance. This study validated both Greenberg and Fortescue's theories.

Another theory would be that Nivkhi is Eurasiatic, possibly related to Chukotko-Kamchatkan, and like Yeniseian, is also connected to languages in North America, as some linguistic relatives of the Nivkhi's ancestors probably migrated to North America and became part of the American Indians. In this way, we can reconcile both hypotheses.

It is true that Campbell and Mixco oppose the proposal, but as they are not specialists, their opposition will not be recorded.

There are three specialist views on Nivkhi. One says it is Eurasiatic (Greenberg 2000), the other that it is Chukotian (Fortescue 2011), and the third that it is part of the Algonquian-Wakashan or Almosan family in the New World (Fortescue 1998; Nikolaev 2015a, 2015b). Consensus among these specialists is that Nivkhi is related to one of two entities – other languages in Northeastern Asia or a New World Amerindian family. So expert consensus seems to have moved away from the view of Nivkhi as an isolate. However, there are still experts who feel that this language is an isolate.

---

[22] Bernard Comrie 2008, 2010; Peter Daniels 2016; Michael Dunn 2012; Willem de Reuse 2010; Michael Fortescue 1998, 2010; Victor Golla; Eric Hamp 2004-2010, 2010; Gary Holton (Sicoli and Holton 2014); John Ives (2010, Ives *et al.* 2010); James Kari (2010, Kari and Potter 2010); Paul Kiparsky 2015; Michael Krauss; Jeff Leer 2010; Johanna Nichols in Daniels 2016; Bill Poser 2008; Ben Potter (Kari and Potter 2010); Keren Rice 2011; Sally Rice (in Ives *et al.* 2010); Merritt Ruhlen 1998; Mark Sicoli (Sicoli and Holton 2014); Edward Vajda 2010a, 2010b, 2012; and Heinrich Werner.

Campbell and Mixco, p. 148:

> **Paleosiberian** languages (also sometimes called **Paleoasiatic, Hyperborean** languages): A geographical (not genetic) designation for several otherwise unaffiliated languages (isolates) and small language families of Siberia.
>
> Perhaps the main thing that unites these languages is that they are not Turkic, Russian or Tungusic, the better-known languages of Siberia. Languages often listed as Paleosiberian are: Chukchi, Koryak, Kamchadal (Itelmen), Yukaghir, Yeniseian (Ket) and Nivkh (Gilyak). These have no known genetic relationship to one another.

Taken as a broad statement, of course this is true. However, Chukchi, Koryak, and Kamchadal or Itelmen are part of a family called Chukotko-Kamchatkan. This family has even been reconstructed. Campbell and Mixco's statement that these languages have no known genetic relationship with each other is false.

Campbell and Mixco, p. 18:

> **Austroasiatic:** A proposed genetic relationship between Mon-Khmer and Munda, accepted as valid by many scholars but not by all.

The fact is that Austroasiatic is not a "proposed genetic relationship." Instead, it is now accepted by consensus. That there may be a few outliers who don't believe in it is not important. I'm not aware of any linguists who doubt Austroasiatic other than Campbell and Mixco, and neither is a specialist. Austroasiatic is also part of the expanded version of the Austric hypothesis (Benedict 1942). A computerized phylogenetic ASJP study (Jäger 2015) linked Austroasiatic with Ainu and Japonic. Japonic is not part of any theoretical link with the family, but Ainu is linked to Austroasiatic as part of a "Macro-Austric" proposal.

Campbell and Mixco, p. 125:

> **Mon-Khmer:** A language family of southeast Asia, thought by many to be a branch of the proposed larger Austroasiatic family, together with the Munda languages of India. The Mon-Khmer languages include Cambodian (or Khmer), Mon (or Talaing), Vietnamese, Nicobarese, and numerous others.

This judgment is simply absurd. Mon-Khmer is not thought by *many* to be part of the Austroasiatic family. Instead, it has been an absolutely uncontroversial member of the family for a long time now. However, we now get into semantics. Austroasiatic has traditionally been seen as two families – Mon-Khmer and Munda (Diffloth 1974). However, a recent reclassification by Diffloth of his original view cut it into three families – Nuclear Mon-Khmer, Munda, and Khasi-Khmuic (Diffloth 2005). The newest view by Paul Sidwell, one of the deans of Austroasiatic studies, abandoned the Mon-Khmer branch altogether and turned the word into a synonym for Austroasiatic as a whole.

Campbell and Mixco, p. 132:

> **Munda:** A family of some sixteen languages, mostly spoken in India. Most scholars classify Munda together with Mon-Khmer as members of the broader postulated Austroasiatic family. Some of the Munda languages are Santhali, Mundari, Sora, Korku, etc.

More absurdity along the lines of the previous entry. It is not true that *most* scholars classify Munda as Austroasiatic, nor is it true for that matter that Austroasiatic is a postulated family. Austroasiatic

is simply a valid family by consensus of all serious scholars. I'm not aware of a single scholar other than the authors who doubts the validity of Austroasiatic. And Munda is an uncontroversial family within Austroasiatic (Hoch and Bashir 2016; Anderson 2017).

Campbell and Mixco, p. 123:

> **Miao-Yao** (also called **Hmong-Mien**): A language family spoken by the Miao and Yao peoples of southern China and Southeast Asia. Some proposals would classify Miao-Yao with Sino-Tibetan, others with Tai or Austronesian; none of these has much support.

This seems to be more weasel wording on the part of the authors. By listing Tai or Austronesian and Sino-Tibetan as possible relatives of Miao-Yao and then dismissing these theories, they leave out a better proposal linking Hmong-Mien to Austroasiatic (Ratliff undated, 2010; Haudricourt 1951). This shows some promise, but the relationship is hard to see amidst all of the Chinese borrowings (Genographic Consortium *et al.* 2011). A computerized phylogenetic ASJP study (Jäger 2015) found that Hmong-Mien and Sino-Tibetan formed a clade of their own on a chart of Eurasian families. However, the close relationship is thought to be due to mass Sino-Tibetan borrowing into Hmong-Mien.

This shows that while ASJP does tend to show accurate relationships most of the time (but not always), the relationship may be one of genetics, borrowing, or even substrate. Usually, it finds a genetic match. Sometimes it finds a substrate or nearly a mixed language as in the link between Japonic and Ainu and Austroasiatic, other times it finds a major borrowing scenario as with Sino-Tibetan and Hmong-Mien above, and a few times it finds a completely false and spurious link (Dravidian in a clade with Nakh-Daghestanian and Yeniseian (Jäger 2015).

As noted, the relationship between Hmong-Mien and Sino-Tibetan is one of borrowing (Handel 2008). Hmong-Mien's relationship with Tai or Austronesian is part of Paul Benedict's (1942) original Austric proposal. He later turned against this proposal and supported a more watered-down Austric with Austronesian and Tai-Kadai (Benedict 1975). Tai-Kadai now has consensus support.

Campbell and Mixco, p. 18:

> **Austric:** A proposed distant genetic relationship that would group Austroasiatic and Austronesian. The hypothesis is mostly doubted, but has some supporters. It was first proposed by Wilhelm Schmidt (1906).

The paragraph is in error. Benedict's original proposal linked Austronesian, Tai-Kadai, Austroasiatic, and Hmong-Mien (Benedict 1942). Later formulations added Nihali and Ainu (Bengtson 2006). See discussion under Austro-Tai below.

Campbell and Mixco, p. 20:

> **Austro-Tai:** A mostly discounted hypothesis of distant genetic relationship proposed by Paul Benedict that would group together the Austronesian, Tai-Kadai, and Miao-Yao.

This does not appear to be correct. Austro-Tai only groups Austronesian and Tai-Kadai.

Benedict's (1942) original Austric proposal, which included Austroasiatic, was abandoned even by Benedict (1975) himself. However, a more watered-down grouping that he later supported, consisting of Austronesian and Tai-Kadai and called Austro-Tai, has much more support (Benedict

1975). Campbell and Mixco get around discussing the watered-down Austro-Tai with good support by confounding it with Benedict's Austric theory which even he rejected later in life. In this sense, they misrepresent the debate.

In fact, evidence is building towards acceptance of Austro-Tai after papers by Weera Ostapirat (2005, 2013) and Laurent Sagart (2005, 2019) proved the case using the comparative method.

Roger Blench (2008, 2018) also supports the concept. In addition to Benedict, it is also supported by Peter Norquest (2013, 2015), Hui Li (2005) and Lawrence Reid (2006). Robert Blust (2014) has also expressed support for the theory and is in the process of working out a proof for it right now. A computerized ASJP study (Jäger 2015) found that Austronesian and Tai-Kadai formed a clade of their own on a chart of languages in Eurasia.

It is opposed by Graham Thurgood (1994), who is a specialist. It is also opposed by Campbell and Mixco, but they are not specialists. Looking at expert opinion, we have eight arguing for the theory[23] and one arguing against it (Thurgood 1994). Specialist consensus then is that Austro-Tai is a real language family.

Even the larger version of Austric, including all of Benedict's families plus Ainu and the South Indian isolate Nihali, has some supporters and there is some suggestive evidence that it may be correct (Bengtson 2006).

Campbell and Mixco, p. 202:

> **Tai-Kadai**: A large language family, generally but not universally accepted, of languages located in Southeast Asia and southern China. The family includes Tai, Kam-Sui, Kadai and various other languages. The genetic relatedness of several proposed Tai-Kadai languages is not yet settled.

Tai-Kadai is not "generally but not universally accepted." It is accepted by consensus as a valid language family (Sagart 2005; Diller *et al.* 2008). Perhaps whether some languages belong there is in doubt, but the proposal itself is not controversial. Campbell and Mixco's statement that Tai-Kadai remains controversial is a serious distortion of fact.

Campbell and Mixco, p. 133:

> **Na-Dené:** A disputed proposal of distant genetic relationship, put forward by Sapir (1915), that would group Haida, Tlingit and Eyak-Athabaskan. There is considerable disagreement about whether Haida is related to the others. The relationship between Tlingit and Eyak-Athabaskan seems more likely, and some scholars misleadingly use the name 'Na-Dené' to mean a grouping of these two without Haida.

The theory dates back a long time (Swanton 1905; Sapir 1915). Dell Hymes (1956) picked it up again in the 1950's, and Heinz-Jürgen Pinnow (1962, 1966, 1968, 1970, 1976, 1985, 2002a, 2002b) conducted major studies from the 1960's to the 2000's.

The anti-Haida reaction started much later with the publication of two papers by top Na-Dene experts Robert Levine (1979) and Michael Krauss (1978). A recent conference about Edward Vajda's Dene-Yeniseian concluded that there was no evidence to include Haida in Na-Dene (Vajda 2012). Other experts (e.g. Nikolaev 2021) say Haida is not Na-Dene.

[23] Benedict 1975; Li 2005; Sagart 2005, 2019; Ostapirat 2005, 2013; Reid 2006; Blench 2008, 2018; Norquest 2013, 2015; Blust 2014.

However, a paper by Alexander Manaster-Ramer (1996a) made an interesting case that Haida is part of Na-Dene. Further, John Enrico (2004), the scholar with the most expertise on Haida, says it is part of Na-Dene.

There have also been a number of other linguists commenting on the issue.[24]

Counting up scholarly opinion, we have a fair number of opponents of Haida as Na-Dene.[25] In addition, at the Dene-Yeniseian Symposium on February 26-29, 2008 in Fairbanks and Anchorage, Alaska, nine linguists agreed that Haida is not part of Na-Dene.[26] So we have 15 opponents to the inclusion of Haida in Na-Dene.

There is also 11 supporters of the theory.[27] The final score is 15 against 11 in opposition to including Haida in Na-Dene. This adds up to a consensus of 58-42% in opposition to Haida as Na-Dene. This means that consensus is split on the question with a slight lean to being pessimistic about the relationship. Campbell and Mixco are correct here that the subject is up in the air and has both supporters and opponents.

The pro-Haida theory started earlier and was more common in the earlier days. The anti-Haida reaction started late in the 1970's, but it seems to have picked up speed in recent years and may well have the momentum.

The statement that a relationship between Tlingit and Eyak-Athabaskan seems "more than likely" is an understatement. It is now linguistic consensus that Tlingit is related to Eyak-Athabaskan, so this part of Campbell and Mixco's statement is not true (Leer 2010).

Campbell and Mixco, pp. 9-10:

> **Amerind:** The Amerind hypothesis is rejected by nearly all practicing American Indianists and by most historical linguists. Specialists maintain that valid methods do not at present permit classification of Native American languages into fewer than about 180 independent language families and isolates. Amerind has been highly criticized on various grounds. There is an excessive number of errors in Greenberg's data.
>
> Where Greenberg stops – after assembling superficial similarities and declaring them due to common ancestry – is where other linguists begin. Since such similarities can be due to chance similarity, borrowing, onomatopoeia, sound symbolism, nursery words (the *mama*, *papa*, *nana*, *dada*, *caca* sort), misanalysis, and much more; for a plausible proposal of remote linguistic relationship, one must attempt to eliminate all other possible explanations, leaving a shared common ancestor as the most likely.
>
> Greenberg made no attempt to eliminate these other explanations, and the similarities he amassed appear to be due mostly to accident and a combination of these other factors.
>
> In various instances, Greenberg compared arbitrary segments of words, equated words with very different meanings (for example, '*excrement/night/grass*'), misidentified many languages, failed to analyze the morphology of

[24] Starostin undated; Swanton 1905; Sapir 1915; Hymes 1956, 1995; Swadesh 1987; Goddard 1996: 318s; Manaster-Ramer 1996a; Campbell 1997: 286-288; Ruhlen 1998; Rubicz *et al.* 2002; Bengtson 2002, 2004, 2008; Enrico 2004; Werner 2004.

[25] Starostin undated; Krauss 1978; Levine 1979; Goddard 1996, 318s; Campbell 1997; Werner 2004; Vajda 2012; Nikolaev 2021.

[26] Jeff Leer, Johanna Nichols, Ben Potter, James Kari, Bernard Comrie, John Ives, Yuri Berezkin, Alexander Kim, and Marie-Lucie Tarpent.

[27] Swanton 1905; Sapir 1915; Hymes 1956; 1995; Pinnow 1962, 1966, 1968, 1970, 1976, 1985, 2002a, 2002b; Swadesh 1987; Greenberg 1987; Manaster-Ramer 1996a; Ruhlen 1998; Rubicz *et al.* 2002; Bengtson 2002, 2004, 2008; Enrico 2004.

> some words and falsely analyzed that of others, neglected regular sound correspondences, failed to eliminate loanwords, and misinterpreted well-established findings.
>
> The Amerind 'etymologies' proposed are often limited to a very few languages of the many involved. Finnish, Japanese, Basque and other randomly chosen languages fit Greenberg's Amerind data as well as or better than do any of the American Indian languages in his 'etymologies'; Greenberg's method has proven incapable of distinguishing implausible relationships from Amerind generally. In short, it is with good reason Amerind has been rejected.

Greenberg's revolutionary book *Language in the Americas* (1987, henceforth *LIA*) set off a wild furor amongst Americanists and triggered a conservative reaction which then spread to all of linguistics and captured the hearts of many of our most prominent linguists. In the same sense that Napoleon's freeing of the Jews from the ghettos led to a situation where every Jew today is a "child of Napoleon," most of today's conservative historical linguists are "the children of *LIA*" or at the very least were profoundly affected by it. Greenberg threw an intellectual hand grenade into a crowded cinema full of scholars.

It's been thirty-plus years on, and they are still nursing their wounds, plotting revenge, and suffering from PTSD. Greenberg is best seen as an "intellectual terrorist" in the sense as his fellow revolutionaries Jesus, Marx, Freud. These folks drift down into our world now and again, turn everything upside down, and walk away laughing daring us to put it all together. And like the three prophets who came before him, he's long since left the room, but the ground has not stopped shaking.

According to Greenberg, the movement into the Americas came in three waves.

The first wave brought the Amerinds. It is here where the 160 language families reside. According to the reigning theory in linguistics, this group of Amerindians came in one wave that spoke *not only 160 different languages but spoke languages that came from 160 different language families, none of which were related to each other*. These being language families which, by the way, we find *scarcely a trace of in Eurasia, where even by the most conservative judgment only a few dozen language families can be found, certainly under 50 including so-called isolates like Basque, Nihali and Ainu*.

The second wave was the Na-Dene people who came along the west coast and then went inland.

The last wave were the Inuits.

Greenberg simply lumped all of the 600 languages of the Americas into a single family. The argument was certainly plausible if only on the basis of the pronouns alone, but it is less certain that he proved that every single one of those languages were all part of Amerind. But a good case can be made that a lot of them were.

The *n*-/*m*- 1st and 2nd person pronouns are found in over a hundred of those languages and scarcely exist outside of the Americas (Ruhlen 1994b, 2005a, 2005b, 2005c). I counted the presence of *n*- and *m*- in Amerind. I found 139 languages with *n*- 1st person and 157 languages with *m*- second person. So 157 languages have this paradigm in whole or in part.

The ablauted *t'ana*, *t'una*, *t'ina* word, meaning respectively human child of either sex, all females including family terms, and all males including family terms are extremely common in Amerind.

According to Greenberg, *t'ana* just means child; *t'una* means girl, woman, and includes various names for all sorts of female relatives – grandmother, cousin, aunt, niece, etc.; *t'ina* means boy or man, and includes the kinship terms grandfather, brother-in-law, uncle, cousin, and nephew. This would be considered a "submerged feature" or a grammatical paradigm. I counted 340 different languages in the Americas that had versions of this ablauted kinship form. So this form is found hundreds of times in Amerind and is nonexistent in the rest of the world (Greenberg and Ruhlen 2007), and yet this means nothing? Critics of Amerind strain credulity.

Quite probably, most if not all of those languages having that term are part of a single family. What are the other arguments? That *over 300 languages independently innovated these terms, in this precise ablauted paradigm, on their own*? What is the likelihood of that?

That these items occurring across such vast swathes of languages *is due to chance*? But this paradigm is not to be found anywhere else, so how could it be due to chance? That these core vocabulary items were *borrowed massively all across the Americas, when family terms like that are rarely borrowed*? That's not possible. None of the alternate theories make the slightest bit of sense.

Hence, the Amerind languages that have the *n-*/*m-* pronoun paradigm and the *t'ana*, *t'una*, *t'ina* ablauted names for kinship relations by sex are quite probably part of a huge language family. Yes, a few of the languages having those terms could be due to chance. An excellent case can be made that about zero of those pronouns and few, if any, of those kinship terms were borrowed.

However, not all Amerind languages have either the pronoun paradigm or the ablauted sex term. In those cases, it is less certain if those languages are all part of the same language family. But if you can put those languages in families and reconstruct to the proto-languages and end up with the pronoun paradigm or the ablauted family term reconstructed in the proto-language of that family, a good case could be made that that family would be part of Amerind. In my opinion, that's all it would take to prove a relationship in Amerind.

Both the *n-/m-* pronoun pattern and the ablauted kinship forms have been attacked. The pronoun pattern was said to be unremarkable and a point was made that it is common elsewhere in the world. Other comments were, "Pronouns are the same everywhere." Others attempted bizarre attempts at associating nasals with kinship forms, apparently based on nursery words. I looked through a chart of the pronouns in language families across the world (Babaev 2009). The pattern turns up again and again in the Americas in *proto-languages*. I could not find a single instance of it in a proto-language anywhere else on Earth. It appears that the main thing that is spurious about *LIA* are the arguments used against it.

Edward Sapir himself noticed this pattern in a letter to Alfred Kroeber, although he leaves out *m-* 'thou':

> Getting down to brass tacks, how in the Hell are you going to explain general American n- 'I' except genetically? It's disturbing, I know, but (more) non-committal conservatism is only dodging, after all, isn't it? Great simplifications are in store for us.

The argument is that *n-* and *m-* pronouns are only found in a part of Western North America, and they are attributed to an areal effect. In other words, they were massively borrowed among these families. Even if this were true, conservatives count 45 language families and isolates (by my count) in Western North America alone (not including Mexico and Central America). Conservatives say that *none* of these families have been shown to be related to each other. Thus, 1st person *n-* and 2nd person *m-* pronouns would have had to have been borrowed 45 *different times*, in an actual pronoun *paradigm*. That's never happened anywhere on Earth. The conservative argument is laughable.

Further, it's not even true that this pattern is confined to Western North America. It's actually from east to west in North America in all of the Amerind language families of North America. It's found a lot in Central America too. If we include Central America with North America, it is present in another 21 language families and isolates, for 64 families and isolates in total.

It is present to some extent in South America in Quechua, which covers much of Central West South America in Colombia, Ecuador, Peru, Bolivia, and Argentina, and in Panoan, which is found in Peru, Brazil, Paraguay, Bolivia, Uruguay and Argentina. It is found in another 16 language families and isolates, giving us 80 language families and isolates with this pattern that is found, from what I can tell, in no large language families anywhere else on Earth.

It is found in the following proto-languages (Babaev 2009):

Grouping	***1st pers. sing.***	***2nd pers. sing.***
Algic	nV / Vn	Vm / V
Keresiouan	naa-	—
Hokan	nya / nyi	mi / ma
Penutian	n-	mV
Aztec-Tanoan	neʔ	ʔeme
Chibchan	na-sV-	mue-ya
Quechuan	nu-ga	ga-m
Panoan	—	mi

Table 1. Mid-or high-level Amerind subgroups exhibiting at least one member of the pronoun root pair **n* 1sg / **m* 2sg in their proto-language.

The argument lies devastated. The pattern is found all through Amerind languages in North and Central America and a few in South America, in 80 different families and isolates. It would have had to have been borrowed 80 different times according to the conservatives. It didn't happen.

The current climate in historical linguistics is one of noncommittal conservatism or worse, angry and stubborn conservatism. And as in Sapir's day, much of this bluster is simply a dodge to avoid coming to the wrong conclusions. I should also point out that Edward Sapir was one of the

greatest linguists ever, and he was also a major long-ranger. Comments that deride the professionalism and competence of Greenberg and Ruhlen need to explain why they don't feel the same way about Edward Sapir, who did the exact same thing.

Critics have attacked *LIA* for the large number of errors in it.[28] As far as I can tell, this may be an accurate assessment. I don't think that changes the basic conclusion that there is an Amerind family that encompasses most if not all of these languages, but it's true nevertheless. Solution? Put it all up on the web. The whole thing. And let fair-minded people who believe in the reality of Amerind go at it, sort it all out, get the forms right, throw out spurious and erroneous words, etc. Greenberg himself said that *LIA* was just a beginning, a raw start, and he compared it to the first large classifications of Indo-European around 1900.

A much more important criticism is that mass comparison doesn't work.[29] The argument is that it is impossible to rule out chance, sound symbolism, nursery forms, onomatopoeia, and loans, which can all create words that appear to be cognates but are not. Examples are given of one-on-one comparisons of two languages that come up with remarkable and impossible findings.

But Greenberg doesn't do that. He only uses a short list of the most commonly used words, probably the Swadesh-100, specifically chosen to be highly resistant to borrowing. He compares languages against entire groups of other languages, and in most cases, he finds nothing. But then he hits the jackpot. Greenberg was looking over Yuki and Wappo and trying to figure out what family they belonged to.

He compared them across the board to other languages in North America and found nothing everywhere he looked. Then he compared them to the Gulf languages and hit the jackpot. Resemblances were everywhere. Greenberg concluded that Yuki and Wappo on the Northern California coast were related to the Gulf languages of the Southeastern US 2,000 miles away. Since the languages were so far away from each other, borrowing is ruled out. If the matches were due to chance, he should have found an equivalent number of matches everywhere else he looked. This is his method.

I do not think that mass comparison always comes up with the right answer, but it looks like a nice tool that often discovers some sort of language relationship, be it borrowing or genetics, and if you use enough core vocabulary, you should be able to rule out chance, often set at 7% matches (Starostin 2003). In some cases, as will be shown below, it will come up with the wrong answer. It will put an isolate into a language family – this is by far the most common error. Less commonly, it gets the entire stock wrong.

Zuni and Tonkawa may not even be Amerind, and Tonkawa may be Na-Dene instead (Manaster-Ramer 1996a). Instead of regarding the findings of mass comparison as contemptible and spurious, a better attitude is, "Good, now we can start," or, "Well, this is a good start. Let's see where it leads" – a good beginning hypothesis to be tested in coming years to see how well it holds up. I'm not an absolutist who believes mass comparison comes up with the word of God. But contempt

[28] Adelaar 1989; Berman 1992; Kimball 1992; Poser 1992; Rankin 1992; Campbell 1998; Chafe 1987; Matisoff 1990.
[29] Goddard 1987, 1990; Adelaar 1989; Berman 1992; Campbell 1998; Ringe 2000.

and scornful rejection isn't right either. On the other hand, long-rangers need to stop treating mass comparison as if it is infallible and produces some unquestionable revealed text.

Mass comparison is a tool for language comparison, as is lexicostatistics, various computer programs like ASLJP, searches for "submerged features" individual identifying morphological evidence, and last but not least, the comparative method. It's not a case of mass comparison versus the combative method. Ideally, they are both used to test out the conclusions of the others, and in a sense, the two methods are not even really contradictory.

But people in the Russian school, who reject mass comparison and claim to only use the comparative method to prove the existence of language families, also use mass comparison themselves. For instance, George Starostin (2003, 2016, 2017) is a big fan of lexicostatistics and he uses it to prove the existence of language relationships. But lexicostatistics is nothing but mass comparison. It can't possibly be anything else. Starostin also uses submerged features, individual-identifying evidence, and shared aberrancies (see the discussion at the end for more) (henceforth "submerged features") in proving the existence of language families:

> Indeed, if we assume that Khoisan languages are not related, we will be left with a great number of "similarities" between the lexicons of North, South, and Central Khoisan groups (and, to a lesser extent, of Hadza and Sandawe), all of which – including similarities in the basic lexicon and apparent similarities in the morphological inventory – will have to be explained as borrowings or chance resemblances.
>
> The number of these similarities may be overrated by some, and there is always a possibility that some of them are indeed the result of lexical diffusion through cultural exchange, but how would that explain, for instance, the use of the same unique form for the 2nd person singular pronoun (PNK **a*, PSK **a*, PCK **c-a* (masc.), **s-a* (fem.; **c-* and **s-* are gender markers)) in all the three major "subgroups" of Khoisan? (Starostin 2003)

Note that Greenberg also makes wide use of submerged features in determining his relationships. He doesn't just use mass comparison.

Another member of the Moscow school who presumably also renounces mass comparison uses it himself to prove language relationships, whether via lexicostatistics (mass comparison) (Zhivlov 2013) or submerged features (Zhivlov 2014, 2018).

Another critic of Greenberg who presumably renounces mass comparison, Terrence Kaufman (1990a), uses either it or submerged features to prove language relationships.

So, the Moscow School attacks mass comparison except when they are using it to prove the existence of language families. Once again, we see critics of mass comparison using the method itself.

To be fair, though, let us remind ourselves that this paper is about consensus in the linguistic community around long-range historical linguistic proposals and the distortion by conservatives of such.

I must admit that Amerind does not have critical support at all (see above for the hostile reviews which represent consensus little changed since *LIA* was published). In general, it is roundly and angrily rejected by the Americanist community.

Yet a few linguists support it, and some others think it stimulates ideas for further research (Munro 1995). Terrence Kaufman (1990a) was one of Greenberg's fiercest critics, but he has validated a number of *LIA*'s large groupings himself in the thirty-odd years since *LIA* was published. Furthermore, Mosan has critical consensus and specialists are optimistic about Almosan.

Penutian and Hokan are surely valid groupings, as is Macro-Gê. Campbell himself has validated part of Greenberg's Macro-Chibchan group.

All of the above groups were named in *LIA*'s rejected classification. So, we see conservatives talk out of both sides of their mouths here. First, they bash *LIA* and insist it is "spurious" and "pseudoscientific." And then, behind closed doors, they validate more and more of his proposals. They reject it only when they don't reject it.

But even where critics agreed with Greenberg's families, they often left out a number of members of the family that Greenberg had put in his classification. And much of Greenberg's South American classification has not been validated at all – to put it mildly.

Below, I will go through Greenberg's North American language families and isolates and try to give a consensus view on each of them. First of all, this view takes as a given that all or most of these languages are part of the Amerind stock, though Greenberg has not proven at all that every language he put in Amerind is actually a member of the stock. But many to most of them are. With this review, you will see that in the thirty-four years since *LIA*, Americanists have slowly come to the conclusion that, the more people looked into it, more and more of Greenberg's basic Amerind classification is correct.

Most of the larger nodes are still rejected, but a shocking number of the smaller nodes and the families contained within them have been slowly shown to be correct. Yes, he was wrong at times. In the South American section, you will see that he sometimes got whole language groups in the wrong families. But still, more of his views have been accepted than rejected. This shows the Americanist lie that Greenberg's work is "spurious." If it's so spurious, why has more and more of it slowly been proven over a few decades?

The Algic family is uncontroversial. However, Greenberg added Mosan to it to form Almosan. Almosan is not accepted, although consensus is optimistic about it. He added Kutenai to this family, and that is not accepted either. Then he created Mosan, consisting of Wakashan, Salishan, and Chimakuan. Mosan is accepted as a valid family. See the Mosan entry below for more. Only one of the three groups he put in Almosan have been validated and that is the Mosan node. Kutenai is an isolate and Algic is not Almosan, though consensus is optimistic about its membership. However, all three members of Greenberg's Mosan node have been accepted. So, 3/5 or 60% of Almosan has been validated.

His next node is Keresiouan, made up of the Caddoan, Siouan, and Iroquoian families. This is rejected as a possible though unproven family (Mithun 1999, p. 305).

None of these families have been shown to be related, so this group has failed completely.

The next is Siouan-Yuchi. Consensus does seem to support this hypothetical family. He did get Siouan and Yuchi right (Kasak 2016).

His next node is Penutian-Hokan, which is completely rejected.

Next he deals with Penutian. Most of the families he put into Penutian are now accepted, including Tsimshian, Chinookan, Plateau (including Waiilatpuan), Oregon, and California. Consensus also accepts his Gulf or Yuki-Gulf family as he calls it, consisting of Atakapa, Muskogean, Natchez, Tunica, and Yukian. And although Mexican Penutian is rejected, consensus now supports three of the four families – Huave, Mixe-Zoque, and Totonacan as part of one family – Totonozoquean, and the fourth, Mayan, is considered to be a probable member of the family. So, his Mexican Penutian is wrong and also right at the same time.

Three isolates that Greenberg put into Penutian are not accepted – Zuni, a possible isolate, Tonkawa, another isolate, which is also possibly not Amerind and may instead be Na-Dene. One member of his Gulf family, Chitimacha, is rejected, but it is accepted in a node immediately below it, Mexican Penutian, which is actually a family called Totonozoquean. New studies show that Chitimacha is actually Totonozoquean. So, Greenberg is right as well as wrong on that one too.

12 of Greenberg's 20 or 60% of his Penutian members have been accepted. However, four of the groups he put in Penutian – Chitimacha, Huave, Mixe-Zoque, and Totonacan – are part of a family called Totonozoquean. The three families he put in Mexican Penutian have not been proven to be Penutian, but all three are now thought to form a single family called Totonozoquean. So, of the 17 members, only two have been rejected, both isolates. Of the other four which are now in Totonozoquean, it is better said that he got them both right and wrong.

His next node is Hokan-Coahuiltecan, which is accepted with caveats. All of his Northern Hokan proposal is accepted. It consists of Shasta, Karok, Chimariko, Shasta-Achomawi, Yana, Pomoan, and Washo. However, his Salinan-Chumashan node is rejected. Of the three languages in the node, two, Salinan and Esselen, are accepted and one, Chumashan, is rejected. Next, Seri and Yuman are both accepted.

His Waicuri-Quinigua node is rejected, probably more on the basis of poor data than anything else. For the same reason, the one family and two isolates in the node are also rejected. It's hard to say that Greenberg is wrong here because although these languages have not been accepted, they haven't exactly been rejected either.

We simply don't have enough information to say one way or the other. Five of seven of the languages in Greenberg's Coahuiltecan node are accepted in part as Pakawan, consisting of Comecrudo, Cotoname, Mamulique, Garza, and Coahuilteco. Aranama and Solano are rejected more on the basis of poor data than anything else. Greenberg's Tequistlatecan and Jicaque are accepted as Hokan, while Yurumangui is rejected due to poor data and Subtiaba has been proven to be an Otomanguean language. However, quite a few Hokanists, including Edward Sapir, put Subtiaba in Hokan, so Greenberg is not alone in this misclassification.

Greenberg's Hokan-Coahuiltecan family has done quite well. 20 out of his 30 or 67% members have been validated, and seven of the ten of those rejected were due to poor data. Another, Subtiaba, was misclassified.

Moving on, we come to Greenberg's Central Amerind node, consisting of Tanoan-Aztecan and Otomanguean, which is rejected. The next node, Tanoan-Aztecan, combining the Tanoan and

Aztecan families, is also rejected. Otomanguean is uncontroversial, with the only error being Subtiaba is now placed there where Greenberg had it in Hokan. The next node is the very large Southern Amerind node, which is rejected. It consists of Andean-Chibchan-Paezan, Equatorial-Tucanoan, and Gê-Pano-Carib, all three of which are rejected, though in Equatorial-Tucanoan, Kariri-Tupi may have links to the Macro-Carib and Macro-Gê stocks.

The first node here is Chibchan-Paezan, consisting of Macro-Chibchan and Macro-Paezan, which is rejected. However, both Macro-Chibchan (Constenla Umaña 2005; Campbell 2012) and Macro-Paezan (Kaufman 1994, 2007; Jokelsky 2015) are now accepted as valid stocks.

Greenberg's Macro-Chibchan consists of Lencan, Chibchan, Paya, Purépecha, Yanomamo, and Yunca-Puruhan. Three of the seven or 43% of the families and isolates he put in there have been accepted and four have been rejected. Lencan (including Xincan), Chibchan, and Paya are all accepted, but Cuitlatec, Purépecha, Yanomamo, and Yunca-Puruhan are all rejected. The first three are considered isolates. Yunca-Puruhan is a small family. Greenberg also put Xincan in Lencan, which is now accepted (see below under Xincan). Betoi-Jirara (Jirajaran) and Piaroa (Piaroa-Saliban) also appear to be part of this family (Kaufman 1990b).

Greenberg's Macro-Paezan family is also now accepted, and and 12 out of 20 or 60% of possible entities included in it have also been accepted (Kaufman 1994, 2007; Jokelsky 2015); Andaqui (Jokelsky 2016); Carabayo (Seifart and Echeverri (2014); Cunza (Atacameño, Atakama, Kunza)-Capixana (Kaufman 1990b); Hotï (Hodi, Hoti, Jodï, Waruwaru) (Rosés Labrada 2015, 2016; Jokelsky 2016;); Itonama (Kaufman 1994, 2007); Kamsá (Swadesh 1959; Loukotka 1968; Suárez 1974; Jokelsky 2016); Kanoê (Kapishana) (Kaufman 1994, 2007); Paezan-Barbacoan (Kaufman 1994, 2007); Saliban (Zamponi 2017); Ticuna (Kaufman 1990b, 1994; Campbell 2012); Tiniguan (Jokelsky 2016); Warrao (Kaufman 1994, 2007), and Yurí (Kaufman 1990b, 1994; Campbell 2012) have been accepted to one degree or another.

Eight, namely Allentiac or Huarpean, Chimu-Mochita, Chocoan, Kamsá, Munichi, Mura, Timucua, and Yurumanguí are rejected.

Yurumanguí (Poser 1992) and Timucua in Florida are isolates (Campbell 1997; Mithun 1999). Munichi is better seen as an Arawakan language, either a part of the family itself or a distantly related sister to it (Gibson 1996; Jokelsky 2016, pp. 310-316). Chocoan has a close relationship with Yaruro. Both are placed with Jivaro along with Allentiac and Mura in the Cahuapana-Zaparo family in Andean which is the equivalent of Greenberg's Jivaro-Kandoshi node in Equatorial, which is invalid.

Mura is better placed with Matanawi in a small Mura-Matanawi family which is part of a larger Macro-Huarpean (Macro-Warpean) family. This family contains Mura-Matanawi and Huarpean (Alyentiyak, Huarpe, Huarpean, Warpe) (Kaufman 1994). This provisional proposal appears to be accepted (Kaufman 1990b). This family has been linked with Jibaro to a different Andean family, Cahuapana-Zaparo (Swadesh 1959; Suarez 1974), while others have left it unclassified (Migliazza 1972). The majority opinion is that Andean and Cahuapana-Zaparo are related.

Chimu-Mochita (Chimú, Chimuan, Chimúan, Chimuana, Chimuano, Mochika, Muchic, Muchik, Yunca–Puruhán, Yuncan, Yunga–Puruhá, Yungan) is a proposed family consisting of

Mochica and the Cañar-Puruhá family. The evidence for the family is good (Kaufman 1990b). Chimuan is either an isolate family or it is possibly Arawakan, which in this case would indicate a relationship with Munichi, also in this family. A proposal places it with Uru-Chipaya and Maya in a Mayan-Yunga-Chipayan family.

Chipayan is seen as either Arawakan (Greenberg 1987), or possibly Arawakan (Migliazza 1972) Mayan (Stark 1992), or part of a Quechuacho family (Swadesh 1959). Some feel it is linked to both Arawakan and Mayan (Suarez 1972). The Chipaya-Maya comparison is a poor case, and the evidence connecting anything with Chipayan is weak, although we can see that a plurality of scholars agree with Greenberg that it may be Macro-Arawakan. However, Chimuan may well be Arawakan (Migliazza 1972; Kaufman 1990b).

Coche (Camëntsëá, Camsá, Kamemtxa, Kamsá, Kamse, Mocoa, Sibundo, Sibundoy), placed in Equatorial, is actually Macro-Paezan.

Greenberg's next node, Andean, is rejected. It consists of Aymara and the Itucale-Sabela family. Aymara is theorized to possibly be part of a Quechumaran family (see below under Quechumaran, which is currently completely up in the air with regard to its acceptance). Mayna is part of an Andean group but not this one. Instead, it is are part of the Cahuapana-Zaparo group below. Itucale (Shimaku, Urarina) has been rejected.

So all three of the languages he put in this node have been rejected, and the node itself may even be invalid, so I am not certain how to classify this in terms of Greenberg's accuracy. Itucale itself is an isolate (Olawsky 2007), however the best guess for a relationship would be with the Cahuapana-Zaparo family below.

His next node is his Cahuapana-Zaparo family, consisting of Abishira (Abigira, Abiquira, Agouisiri, Auishiri, Awishiri, Avirxiri, Ixignor, Tekiraka, Tequiraca, Vacacocha) (Kaufman 2007); Candoshi (Candoxi, Kandoshi, Murato)-Shapra (Chapara) (Loukotka (1968, Tovar 1984); Cahuapanan; Chirino; Esmeraldeño (Esmeralda) (Adelaar and Muysken 2004); Jivaroan (Shuar), Peva (Peba)-Yaguan, and Zaparoan.

All except Zaparaon and Peva-Yaguan have been accepted. Zaparoan and Peva-Yaguan have not, as they are in the Carib family of Gê-Pano-Carib. Eight of the ten or 80% of the languages postulated for this family have been accepted (Kaufman 2007).

His next node is called Northern Andean, and it is also rejected. Catacao and Sechura (Sek, Talyan) are accepted (Kaufman 1990b; Campbell 2012). Kaufman 1990 adds Leco (Lapalapa, Leko) and calls the family Macro-Lecoan, which may as well be the same as Greenberg's Northern Andean. Chachapoya (Chacha) (Loukotka 1968), Cholón (Cholona, Seeptsá, Tsinganeses), Culli (Culle) (Jokelsky 2016), and Hibito (Chibito, Híbito, Hívito, Ibito, Jibito, Xibita, Zibito) also seem to be a part of this family. Hence, Greenberg's entire Northern Andean is supported.

His next node is Quechua. As a separate node this is accepted at this time, although some proposals go beyond Greenberg to posit a Quechumaran family.

His next node is Southern Andean, consisting of Araucanian (Araucano), Chimané, Chonan, Chono, Gennaken (Puelche), Moluche, Pehuenche, Querandí, Qawasqar (Alcalufan, Kawésqar,

Qawasqar), and Yagán (Háusi Kúta, Iakan, Jagan, Yaghan, Yagankuta, Yahgan, Yámana), Yuracaré, and it is rejected.

Alcalufan (Adelaar and Muysken 2004; Campbell 2012) and Yagán are generally considered to be isolates, but recent work shows that Alcalufan, Yagán, and Chono seem to be related, as Chono is a third language with words that resemble yet are quite distant from both of the others, strongly suggesting that the three make up a family (Viegas-Barros 2005). In the past, interesting arguments were presented connecting Araucanian with Alcalufan, Yagán, Chon, Mosetén, and Yuracaré (Key 1978).

This argument was interesting because it connected Araucanian to five of the 12 languages proposed for the family. A recent analysis (Viegas-Barros 2005) by the foremost scholar of these languages determined that Araucanian is not related to Southern Andean in general and specifically has no relationship to Kawésgar, Puelche, Tehuelche, and Yagán.

However, he has been overruled by a plurality of scholars. The truth here is that this entire family is eliminated and the entire family is shifted over to Macro-Panoan. So Greenberg got the whole family wrong in a sense, since it does not exist, however, his notion that all of these languages are related was ultimately correct. So he was right and wrong again and I am not sure how to analyze this situation.

The next node is the large Equatorial-Tucanoan node, which is rejected. Greenberg's Equatorial family, consisting of Macro-Arawakan, Cayuvava, and Coche has been rejected. Of these, Macro-Arawakan is accepted, while Coche and Cayuvava are misclassified, with Coche being a part of Macro-Paezan (Swadesh 1959; Loukotka 1968; Suárez 1974), and Cayuvava being close to Tupian, which is also in Equatorial, but on a different node (Suárez 1974; Kaufman 1990b, 1994).

Greenberg's small Guamo-Chapakuran family which he added to Macro-Arawakan, consists of unclassified Guamo (Wamo) and the Chapakuran family, are seen as related to each other and both are also probably Arawakan (Migliazza 1972). Three fifths or 60% of Greenberg's Equatorial has been verified. Pukina is unclassified, but it may be Arawakan (Kaufman 1990b).

The next large node in Equatorial is Jivaro-Candoshi, which is rejected as a node of Equatorial but accepted as part of an Andean node called Cahuapana-Zaparo. Jivaroan (Kaufman 2007), Candoshi-Shapra (Loukotka 1968, Tovar 1984), Chocoan, Yaruro (Jaruro, Pumé) (Pache 2016), and Esmeralda (Adelaar and Muysken 2004) are all part of this family.

So Greenberg got the fact right that 80% of these languages are related, but he put them in the wrong family, Equatorial, and he did not recognize that his Jivaro-Candoshi Equatorial family was actually part of the Cahuapana-Zaparo family in Andean. So he was wrong and right about these languages. I'm unsure what to do about this. He got it right by recognizing that the languages were related, but he put them in the wrong family. Cofán (A'ingae, Kofan, Kofane) is rejected as an isolate (Repetti-Ludlow *et al.* 2020). Besides Greenberg's Equatorial theory, the only other proposals regarding this language are Macro-Gê, Macro-Chibchan or Macro-Paezan (Migliazza 1972).

The next node is Kariri-Tupi, which is rejected. This family consists of Kariri, Katembri (Catrimbi, Kariri de Mirandela, Mirandela), Piaroa, Taruma (Taruamá), Timotean (Cuica, Loco Mutú, Timote-Cuica), Trumai, Tupi (Tupi-Guarani), Tusha (Carapató, Payacú, Rodela, Todela, Tusha), Yuracaré, and Zamucoan.

The Kariri and Tupi families are accepted as being related to each other, but they are probably misclassified, as both seem to be in the Macro-Carib family in Gê-Carib, (Swadesh 1959; Rodrigues 1966; Migliazza 1972; Kaufman 1990b). Taruma and another language close to it, Katembri, are also part of Macro-Carib (Swadesh 1959; Kaufman 1990b). Trumai is part of the Macro-Tucanoan family of Equatorial, with a close relationship to Puinave (Swadesh 1959, Kaufman 1990b).

Timote, Yuracaré, Piaroa, Zamucoan, and Tusha are also rejected. Timotean is Equatorial also, in the Macro-Arawakan branch, and Yuracaré and Zamucoan are in completely different families. Yuracaré is in the Southern Andean branch of Andean, Zamucoan is in another superstock altogether – Gê-Pano-Carib, where it is part of the Macro-Panoan family, Piaroa is in Macro-Paezan, and Tusha is unclassified due to poor data. Unfortunately, Greenberg got every one of these languages wrong, 0 out of 10 or 0% of them correct. Five of them are in Gê-Carib instead, with four of those part of the Macro-Carib family there. Only Trumai is even still in Equatorial.

The next node is Macro-Tucanoan, which has been rejected.

Greenberg's Macro-Tucanoan consists of the following: Arawan (also Arahuan, Arauan, Arauán, Arawa, Arawán), Awake (Arutani, Aoaqui, Auake, Auaqué, Oewaku, Orotani, Urutani); Auixiri (Abishira, Aewa, Aiwa, Avishiri, Ixignor, Tekiráka, Tequiraca), Canichana (Canesi, Joaquiniano), Capixana (Kanoê, Kapishana), Carabayo (Aroje, Caraballo, Yacumo), Catuquinan (Katukinan), Gamella (Acobu, Barbados, Curinsi, Gamela), Harakmbut (Arasairi, Harákmbut, Harakmbet, Tuyoneri), Huari (Aikaná, Corumbiara, Kolumbiara, Tubarão, Uari, Wari), Iranshe (Irántxe, Irántxe-Münkü, Iranxe, Mỹky, Münkü), Kaliana (Caliana, Kalianá, Sapé), Katawixi, and Kwaza (Koaiá, Kwazá).

Other languages that are part of his family include Maku (Macú, Nadahup, Vaupés-Japurá), Movima, Muniche (Munichi), Nukak-Kakwa (Cacua, Kakua), Nambikwaran (Nambiquaran), Natu (Natú, Peagaxinan), Pankaruru (Brancararu, Pancaré, Pancaru, Pankarará, Pankaré, Pankararú, Pankaroru, Pankarú, Pankaravu), Paratió, Puinave (Guaipunabi, Maku; Waipunavi, Wanse), Shukuru (Ichikile, Kirirí, Kirirí-Xokó, Xucuru, Xukuru), Ticuna (Tikuna), Tucanoan (Tukanoan, Tukánoan), Uman (Aticum, Atikum, Araticum, Huamoé, Huamoi, Huamuê, Uame, Uamué, Umã, Wamoe, Wamoé), and Yuri (Jurí).

Capixana is related to the Kunza language in a small family that is better placed in Macro-Paezan (Kaufman 1994, 2007). Kwaza (Müller *et al.* 2013) and Huari are probably related to Capixana-Kunza (van der Voort 2005) in this same family. Auixiri and Canichana have a close relationship and are part of a proposed family (Kaufman 1994). Both are better placed in the Macro-Panoan branch of Greenberg's Gê-Pano-Carib. Catuquinan and Harakmbut are part of a small proposed family (Adelaar 2000, 2007; Campbell 2012). Jokelsky (2011) concurs and adds Arawan to the group.

Katawixi also seems to be part of this group. All are better placed in Macro-Arawakan, which, while still a part of Equatorial, is a different branch. Kaliana and Maku are now thought to be a small family, so Greenberg's notion that they were together is correct. Along with Puinave, it seems to be properly placed in this Macro-Tucanoan family. Trumai, which he placed in Kariri-Tupi, is also part of this group, as is Awake, which is probably part of Kaliana-Maku, as is Nukak-Kakwa.

Iranshe is an isolate. Movima is formally an isolate, although it seems to be better placed in Gê-Carib (Suárez 1974) or the Macro-Guaicuruan family in Macro-Panoan (Swadesh 1959). In either case; it is better placed in Greenberg's Gê-Pano-Carib superstock. Uman, Gamella, and Natu are rejected as unclassified due to poor data (Kaufman 1990b).

Pankaruru seems to be an isolate, and it is also very poorly attested (Meader 1978). However, much of the known vocabulary is Tupian. Oddly, Meader concluded that this meant that Pankaruru speakers were bilingual in Tupi rather than deciding that Pankaruru is a Tupian language. The reasoning seems odd. Uman was spoken nearby but it is also an isolate, and there is no resemblance with Pankaruru. Muniche and Tikuna-Yuri are better placed in another Equatorial family, the Arawakan family.

Shukuru seems to be in a family with Paratió. However, any relationships beyond that are unknown. Nambikwaran is formally an isolate, though a relationship with Capixana, also in this purported family, is suggested. Capixana is thought to be Macro-Paezan, so a good theory is that Nambikwaran is also (Price 1978). It stands to reason that Tucanoan would be a member of a Macro-Tucanoan family.

Surprisingly, there is something left of Greenberg's postulated Macro-Tucanoan family. Much of it lies in ruins, but there are still a few batteries firing away. Kaliana, Maku, Puinave, Trumai, Awake, and Nukak-Kakwa do seem to make up a Macro-Tucanoan family in Equatorial, so six of the 25 or 24% postulated languages are actually a part of this group.

Greenberg’s next superstock in South America is Gê-Pano-Carib. It has been rejected, but two of its three large groups have been accepted, so things look better than they seem.

Greenberg’s Macro-Carib family is rejected. This family consists of Andoke (Andoque); Boran (Bora-Muinane, Bora-Muiname, Bóran, Bórano), Carib, Kukura, Witotan (Huitotoan, Uitotoan), and Yaguan (Yawa, Yawan). Kukura is spurious. It is not a real language; instead it was fabricated by an interpreter for an European anthropologist in 1901. A word list was published in 1931. In 1932, it was proven that the list consisted of one half made-up words and the other half mispronounced Guaraní words.

The last language, Yaguan, is misclassified. It is instead an Andean language, part of the Itucale-Sabela family in that group.

Bora and Uitoto are linked in a family, and some say this family is Cariban (Mason 1950; Aschmann 1993; Kaufman 1994), but others disagree (Gildea and Payne 2007; Echeverri and Seifart 2016). Consensus favors placing Bora-Uitoto in Macro-Cariban (Kaufman 1990b). Tupi and Kariri are part of this family (Swadesh 1959; Rodrigues 1966; Migliazza 1972; Kaufman 1990b). Andoke is probably related to Urequena (Arequena, Orelhudos, Uerequena) (Adelaar and Brijnen

2014; Jolkesky 2016). Some think Andoke is Witotan (Kaufman 2007), but others disagree (Aschmann (1993; Gildea and Payne 2007), and consensus is that it is not related to Bora-Uitoto.

Taruma (Taruamá) and another language close to it, Katembri (Catrimbi, Kariri de Mirandela, Mirandela), are also part of Macro-Carib (Swadesh 1959; Kaufman 1990b). Hiriharan is controversial. Some put this language in Macro-Paezan (Greenberg 1987). One paper found that Hiriharan was neither Paezan nor Chibchan (Migliazza 1972). The majority view (Kaufman 1990b, Swadesh 1959) is that it is Macro-Cariban. Yawan and Zaparoan (Auishiri, Aushiri, Huao, Huaorani, Sabela, Wao, Waorani) are linked together in a single group (Swadesh 1959; Payne 1974; Kaufman 1990b).

Peba (Peva); Yagua; Mayna (Humurana, Maina, Numurana, Omurano, Roamaina, Umurano); and Taushiro (Pinche) are also Cariban (Kaufman 1990b). Yaruro (Jolkesky 2016; Pache 2016) is related to Saparo, so it must be in the Macro-Cariban family also. Three out of five or 60% of Greenberg's Macro-Carib family have been accepted.

Macro-Panoan is the next large group. It includes Charruan, Lenguan (Chaco, Enenlhet, Enlhet, Lengua-Mascoy, Mascoian, Mascoyan), Lule-Vilela (Atalalá, Chulupí, Chunupí, Uakambalelté), Mataco (Mataco-Mataguayo, Matacoano, Matacoana, Mataguayan, Mataguayo, Matákoan)-Guaicuru (Macro-Waikuruan, Macro-Waykuruan), Moseten, and Pano-Tacanan (Pano-Tacana, Páno-Takán, Pano-Takana, Pano-Takánan).

The link between Panoan (Páno, Pánoan, Panoana, Panoano) and Tacanan forming a single family suggested by Greenberg has been validated (Adelaar & Muysken 2004; Kaufman 1990b, 1994), although some still say the connection has not been proven (Loos 1999). Nevertheless, consensus supports the Pano-Tacanan family. Moseten, actually called Moseten-Tsimane (Chimané, Mosetén, Moseténan, Tsimané), and Chonan (Patagón, Tshon) are also added to the family in a group called Moseten-Chonan (Swadesh 1959; Migliazza 1972; Suarez 1974, Kaufman 1990b).

Yuracaré (Yuracare, Yurucare, Yuracar, Yurujare, Yurujuré, Yurakar, Yurakare, Yurakaré) is thought to be close to Moseten-Tsimane, while Moseten-Tsimane is close to Chon to a lesser extent (Kaufman 1990b).

Araucanian is a small family consisting of Mapudungan (Mapuche, Mapudungu, Mapuzugun), and Huilliche (Huiliche, Veliche, Williche). Araucano, Moluche, and Pehuenche are often attached to the family (Zúñiga 2006). Araucanian has been compared to many language families and is traditionally considered to be an isolate. At the moment, the proposal connecting it with Pano-Tacanan (Migliazza 1972; Loos 1973, 1999; Key 1978a, 1978b) has the most support. It has also been compared with Chon, Mosetén, and Yuracaré (Key 1978a, 1978b). A relationship with Chon and Moseten would fit in very well here.

One scholar says that a relationship between Araucanian and Kawésgar, Puelche, Tehuelche, and Yagán has been ruled out (Viegas-Barros 2005), but more seem to contradict him. At any rate, the latter four languages are related to each other, apparently in Macro-Panoan. His finding that Araucanian is not related to Puelche, etc., is in contrast to a greater number of linguists who find that Puelche is part of Chonan, which even Campbell (2012) supports.

There have also been suggestions that Alcalufan, Yagán and consequently Chono are part of this family (Key 1978a, 1978b). A plurality of linguists puts Alcalufan in Macro-Panoan (Migliazza 1972; Key 1978a, 1978b). As Yagan, Alcalufan and Chono are related, they must all be Macro-Panoan, too (Viegas-Barros 2005). Northern Tehuelche (Gününa Yajich, Pampa, Ranquelche) is also linked to this family (Viegas-Barros 2005).

Mataco-Guaicuruan (Guaicurú-Mataguayo, Macro-Guaicurú, Macro-Waikuruan, Macro-Waikurúan, Macro-Waykuruan, Mataguayo-Guaicuru), a new family postulated by Greenberg, has since been shown to be valid (Swadesh 1959; Suárez 1974; Viegas-Barros 1992-1993, 2006a, 2013; Kaufman 2007; Jolkesky 2016) (see the Mataco-Guaicuruan section below).

External relations with Mataco-Guaicuruan are more complicated. A number of scholars have linked Charruan with the family (Swadesh 1959; Suarez 1972; Kaufman 2007, p. 72), so Charruan seems properly linked to Macro-Panoan. Some scholars (Kaufman 2007, p. 72; Swadesh 1959) have added Lule-Vilela and Zamucoan. Lule-Vilela is a proposal linking the Lule-Tonocote and Vilela language families in Northern Argentina. A plurality of scholars support it (Viegas Barros 2001, 2006b; Kaufman 2007, p. 72). Tonocote (Tonocoté, Tonokoté) appears to be related to Lule.

Greenberg did very well with Macro-Panoan. All of the six or 100% of languages he put in it appear to be correct. However, he also placed many other Macro-Panoan languages in a nonexistent category called Southern Andean, all of which are actually Macro-Panoan languages. So he was wrong and right here once again.

Macro-Gê is the next large node, consisting of Bororoan (Borôroan, Borôro, Bororoano, Bororoana, Borotuke), Botocudo (Aimoré, Botocudoan, Borum, Krenakan), Caraja (Iny Rybè, Karajá), Chiquito (Bésiro, Chikitano, Chiquitano, Tarapecosi), Erikbatsa (Aripaktsa, Canoeiro, Erikbatsa, Erikpatsa), Fulnio (Carnijó, Fornió, Furniô, Iatê, Karnijó, Yahthe, Yatê), Ge (Gê, Gean, Je, Jean, Ye)-Kaingang (Kaingáng), Guató, Kamakanan (Ezeshio, Kamakã, Kamakãne), Mashakalian (Maxakali, Maxakalían), Opaie (Ofaié-Xavante, Ofayé, Opaié-Shavante, Opaye), Purian (Purían), and Yabutian (Jabutian) languages.

Botocudan, Caraja, Erikbatsa, Gê-Kaingang, Kamakã, Maxakalían (Mason 1950; Rodrigues 1966b), and Opaie are uncontroversially Macro-Gê (Kaufman 1990b).

Maxacalian is close to Botocudan and Kamakã in a Trans-São Francisco branch of Macro-Gê (Nikulin 2020).

Purian is made up of Puri, Coroado (Colorado) and Coropó (Koropó). Campbell (1997) added it to Purian, but it was later removed (Ramirez *et al.* 2015), and Hammarström (2020) feels that when Coropó is removed from Purian, there is no longer good evidence that Purian is Macro-Gê. At the moment, the consensus is that Coropó is not part of the Purian family (Kaufman 1990b).

Hammarström's position that Purian is not Macro-Gê is challenged by Swadesh (1959) who finds links between it and Maxacalian and Fulnio. Several scholars have presented evidence that Purian is Macro-Gê (Rodrigues 1966; Davis 1966, 1968; Loukotka 1968). Hence, Purian is clearly Macro-Gê (Kaufman 1990b).

Fulnio is controversial. Two scholars place it in Macro-Gê (Kaufman 1990b; Jolkesky 2016), and three others disagree (Ribiero and van der Voort 2010; Nikulin 2020). Evidence has been

presented that Fulnio is Macro-Gê (Rodrigues 1966; Davis 1966, 1968). Hence, scholarly consensus is that Fulnio is Macro-Gê and possibly has a close relationship to Purian and Maxacalian (Swadesh 1959).

Guató is also controversial. Some say it is not Macro-Gê (Ribiero and van der Voort 2010), but others say it is (Kaufman 1990b).

Bororoan (Swadesh 1959; Suarez 1972; Kaufman 1990b) is Macro-Gê. Swadesh (1959) doesn't say it is Macro-Gê, but links it to Chiquito, agreeing with Greenberg and Kaufman (1994). Chiquito is traditionally seen as an isolate. However, with Swadesh, Greenberg and Kaufman linking it to Botocudo, that makes it Macro-Gê by default. A new study has shown it to be Macro-Gê (Adelaar 2008), while another has linked it to Macro-Gê as a sister to the entire family (Nikulin 2020). In both cases, there is a clear link to Macro-Gê.

Guató is properly placed in Macro-Gê (Rodrigues 1966; Kaufman 1990b), as is Opaie (Rodrigues 1966; Davis 1966, 1968; Migliazza 1972; Kaufman 1990b). Oti is so poorly known that it is unclassified (Kaufman 1990b). Bororoan, Chiquito, Guató, and Opaie remain somewhat controversial, with first three sometimes placed with the Tupian languages and Opaie and Guató sometimes placed with the Puinave languages in Macro-Tucanoan. Some say that Bororoan is not Macro-Gê (Nikulin 2020), but more say that it is (Guérios 1939; Ribeiro and van der Voort 2010).

Tupi is in Macro-Carib and Puinave is Macro-Tucanoan. Carib is still in Gê-Pano-Carib or Gê-Carib, but Macro-Tucanoan is in Equatorial.

Other languages, including Ticuna-Yuri, Cofán and Chimuan, are postulated as part of Macro-Gê. Cofán is an isolate, with the best guesses placing it in Macro-Carib in Gê-Pano-Carib or Macro-Paezan in Andean-Chibchan-Paezan (Migliazza 1972; Kaufman 1990b). Ticuna-Yuri is in the Equatorial Macro-Arawakan family, and Chimuan may also be Arawakan.

Greenberg did very well with this one, as 13 out of his 14 or 93% of postulated members have been validated. However, to be fair to his detractors, a number of those had already been placed in the family by consensus before *LIA*. But this also must be examined. Recall that Greenberg's critics said that the entirety of Greenberg's Amerind theory must be rejected. They didn't say that some of his suggestions were correct and others were wrong. They summarily rejected the whole thing! So clearly this means that they reject all of the languages he put in Macro-Gê also.

When Greenberg's *LIA* was published, none of his proposed classifications were accepted. However, in the ensuing thirty-four years, *LIA* has become more and more vindicated and accepted. 100 out of Greenberg's 188 or 53% of Greenberg's newer classifications have been shown to be correct via more research. 88 out of 188 or 47% are still yet to be proven, but a certain number of those lack data, so they may remain forever unclassified.

When *LIA* was published, some long-rangers said, "Just give it time. In a matter of decades, much of it will be proven to be correct." It turns out that they were right after all.

The notion that *LIA* is spurious pseudo-science lies smashed and in the dust, and people who repeat this are just wrong. And it turns out that given a few decades, Greenberg, who was said to be completely wrong at first, has now been shown to be more right than wrong. That's an amazing statement and it goes against everything the Americanists and others have said about this book.

Campbell and Mixco, p. 207:

> Tonkawa: An extinct language isolate of Texas. Proposals to link Tonkawa with the languages of the Coahuiltecan or Hokan-Coahuiltecan hypotheses have not generally been accepted.

A Coahuiltecan connection with Tonkawa was denied by Manaster-Ramer (1996a), who recently proved that the family existed. That said, there are interesting parallels between Tonkawa and Coahuiltecan that seem difficult to explain (Manaster-Ramer 1996b). However, a much better case can be made that Tonkawa is in fact a Na-Dene language (Manaster-Ramer 1996c).

Campbell and Mixco, p. 158:

> **Penutian:** A very large proposed distant genetic relationship in western North America, suggested originally by Dixon and Kroeber (1913a, 1913b, 1919) for the Californian language families Wintuan, Maiduan, Yokutsan, and Miwok-Costanoan. The name is based on words for 'two', something like *pen* in Wintuan, Maiduan, and Yokutsan, and *uti* in Miwok-Costanoan, joined to form Penutian.
>
> Sapir (1929), impressed with the hypothesis, attempted to add an Oregon Penutian (Takelma, Coos, Siuslaw, and 'Yakonan'), Chinook, Tsimshian, a Plateau Penutian (Sahaptian, 'Molala-Cayuse,' and Klamath-Modoc) and a Mexican Penutian (Mixe-Zoquean and Huave).
>
> The Penutian grouping has been influential, and later proposals have attempted to unite various languages from Alaska to Bolivia with it. Nevertheless, it had a shaky foundation based on extremely limited evidence, and, in spite of extensive later research, it did not prove possible to demonstrate any version of the Penutian hypothesis and several prominent Penutian specialists abandoned it. Today it remains controversial and unconfirmed, with some supporters but with many who doubt it (see Campbell 1997: 309–20).

The statement that today it "remains controversial and unconfirmed, with some supporters but with many who doubt it," has no basis in fact. It is surely controversial, however, it has a number of supporters.

Moreover, among those who doubt it, only a few are specialists in these languages.[30] The first objection was former Penutianist William Shipley's (1957, 1961, 1966, 1969, 1970, 1973, 1978a, 1978b) classic rejoinder against the family, *Penutian amidst the Ruins* (Shipley 1980), and others followed for the next thirty-two years. In addition, the well-known Penutianist Micheal Silverstein (1967, 1969, 1972, 1974, 1977, 1978a, 1978b, 1979) later became an opponent (Silverstein 1979b).

Below I will list the specific scholars and papers that have proven that each constituent family of Penutian is valid, not just the family itself. I will also list proposed Penutian families that have been rejected as Penutian and evidence for that.

[30] Shipley 1980; Goddard 1996:290–324; Campbell 1997: 188, 273–283; Mithun 1999; Haney 2012.

A Penutian family comprising Maiduan,[31] Yok-Utian [Utian (Miwok-Costanoan)-Yokutsan],[32] Wintuan,[33] Oregon Penutian,[34] [Coosan,[35] Siuslaw,[36] Alsean,[37] Takelma,[38] Kalapuyan[39]], Chinookan,[40] Tsimshianic,[41] and Plateau Penutian,[42] Klamath-Modoc [Lutuami],[43] Cayuse and Molala {Waiilatpuan},[44] and Sahaptian[45]] has good support among specialists.[46]

The inclusion of Mixe-Zoque and Huave or Mexican Penutian (Sapir 1921a) in the family has been rejected in recent years. Mixe-Zoque may well be part of a family called Totonozoquean (see below).

Anti-Penutianism is a relatively recent movement, dating back to Shipley's famous article in 1960 (Shipley 1960). In this way, it is similar to Hoijer's important anti-Hokan piece in 1946 (Hoijer 1946) that set off anti-Hokanism. As we can see, both objections date from the middle of the last century. This is because the conservative movement itself was, oddly enough, progressive when it was started.

The first half of the century saw a lot of lumping, a fair amount of it careless, under Dixon, Kroeber, and especially Sapir, all working under the then-new framework of Boasian anthropology. So, if the conservatives are really liberals rebelling against an old lumper tendency, then long-rangers themselves are conservatives harkening back to the good old days of lumping under Kroeber and Sapir from the 1900's to the 1930's.

In this sense, the label conservative to apply to Campbell, etc. may be misleading. In the sense that they are trying to stop or slow down what they see as a radical, excessively encompassing, poorly-thought out tendency of the long-rangers, their argument for sober and deep reflection before arriving at major conclusions is indeed conservative, possible in the positive sense of the

---

[31] Dixon and Kroeber 1903, 1913a, 1913b, 1919; Sapir 1921a; Ultan 1964; Oswalt 1976; Shipley and Smith 1979; Golla 1980; Delancey and Golla 1997; Tarpent and Kendall 1998; Zhivlov 2014.
[32] Kroeber 1910; Dixon and Kroeber 1903, 1906, 1913a, 1913b, 1919; Sapir 1921a; Mason 1916; Freeland 1951; Broadbent 1960; Beeler 1961; Broadbent and Pitkin 1964; Callaghan 1967; Collord 1968; Levy 1976; Shambach 1977; Gamble 1978; Golla 1980; Whistler and Golla 1986; Delancey and Golla 1997; Okrand 1997; Tarpent and Kendall 1998; Weigel 2005; Zhivlov 2014.
[33] Dixon and Kroeber 1903, 1913a, 1913b, 1919; Sapir 1921a; Goldschmidt 1951; Pitkin 1955; Shafer 1961; Broadbent and Pitkin 1964; Schlichter 1979; Golla 1983; Delancey 1987, Delancey and Golla 1997; Tarpent and Kendall 1998; Liedtke 2007; Zhivlov 2014.
[34] Sapir 1921a; Buckley 1987; Grand 1997; Delancey and Golla 1997; Tarpent and Kendall 1998.
[35] Sapir and Swadesh 1953; Pierce 1966.
[36] Pierce 1966.
[37] Pierce 1966; Golla 1983.
[38] Sapir and Swadesh 1953; Swadesh 1965; Pierce 1966; Tarpent and Kendall 1998.
[39] Swadesh 1965.
[40] Sapir 1921a; Hymes 1956-1965; Tarpent and Kendall 1998.
[41] Sapir 1921a; Hymes 1956-1965; Delancey *et al.* 1987, Tarpent 1990, 1994, 1996, 1997, 2002b; Delancey and Golla 1997; Tarpent and Kendall 1998.
[42] Sapir 1921a; Delancey and Golla 1997; Tarpent and Kendall 1998; Liedtke 2007.
[43] Voegelin 1946; Rude 1986, 1987; Delancey 1987,1988, 1991; Delancey *et al.* 1987; Aoki 2003; Zhivlov 2014.
[44] Berman 1996; Hymes 1969; Rigsby 1969; Hunn 2000.
[45] Rude 1986, 1987; Rigsby 1969; Delancey *et al.* 1987, Aoki 2003.
[46] Dixon and Kroeber 1903, 1913a, 1913b, 1919; Sapir 1921a, 1921b; Hymes 1957; Broadbent and Pitkin 1964; Rigsby 1966, 1969; Callaghan 1967; Gursky 1969; Shevoroshkin 1980; Rude 1986; Greenberg 1987: 143-162; Delancey 1987, 1988, 1991, 1996; Tarpent 1990, 1994, 1996, 1997, 2000, 2001, 2002a, 2002b; 2014; Liedtke 1995, 2007; Berman 1996; Delancey and Golla 1997; Tarpent and Kendall 1998; Hunn 2000; Golla 2007, 2011; Zhivlov 2014.

word: "Hey, slow down! Not so fast now! Let's not go overboard. Let's think this out before we jump to conclusions."

In addition, by requiring long-rangers to dot every *I* and cross every *t* and make sure the work is at a very high standard, the conservatives are forcing the long-rangers to work harder and harder and accumulate more and more evidence for their proposals. In this sense of holding the long-rangers' feet to the fire, the conservatives are actually doing the long-rangers a favor.

If we tally up the pro-Penutian scholars, many of whom are specialists, we get a minimum of 45 in favor of Penutian, some dating back before 1950 but much more appearing in recent years as the proposal gained steam. If we count anti-Penutianists historical and modern, we get six people, half of whom are not explicitly Penutianists.

That leaves us with 88% of specialists supporting Penutian and only 12% opposing the proposal. Hokan also gets support from 90% of specialists. Further, the opponents of both families are in some cases, the same people.

In terms of papers, Penutianists have published a minimum of 71 papers. Please note that I did not include many pro-Penutian papers for reasons of time and space. Anti-Penutianists have published a minimum of six articles. We get 91% with anti-Penutianists getting only 9%. This level of support is far higher than for controversial proposals such as Altaic.

As we note below in the Hokan section, many serious scholars who do not openly consider themselves long-rangers are quite conservative in declaring that a new language family is proven, preferring instead to couch their terms in words like "unproven," "hypothesis," "optimistic," and "will eventually be proven." The major Penutianists such as Scott Delancey and Marie-Lucie Tarpent have recently written papers where they have stated outright that Penutian is a valid entity. This usually only happens when specialists feel that a critical mass of evidence has been forwarded for the proposal.

The Hokanists have been speaking the same way for thirty years. Penutianists have piled a lot of good evidence, including examples of Swadesh's "submerged features," among widely separated languages that could not have been borrowed and that could not be by chance because they are so odd (Delancey 1996; Tarpent 2002b). In an important paper, Tarpent called for a new paradigm in historical linguistics moving away from vast, unsupported proposals and overreach and obsessive focus on "close phonological and lexical correspondences," including immaculate sound correspondences.

In her former argument, she appears to be taking aim at the "Greenbergists" who confidently put forward extreme genetic language proposals dating back tens of thousands of years. In her latter argument, she is attacking Campbell and the other conservatives for excessive caution and endless moving of goalposts before a genetic relationship between languages or language families can be proven:

Tarpent (2001) calls for a balance to be struck between the unsupportable claims of the "Proto-World" hypothesis and the conviction that the only "responsible" diachronic work focuses on close phonological and lexical correspondences.

Traditionally, Penutian was always on better terms with more support than Hokan. Nevertheless, as we see below, the vast majority of specialists consider Hokan a proven language family too (see below). If it is time at last to bring Hokan in from the cold, surely now is the time at long last for Penutian to have its day in the sun also.

Campbell and Mixco, p. 27:

> **Cayuse-Molala** (also called **Waiilatpuan**): A genetic classification no longer believed that linked Cayuse (of Oregon and Washington) and Molala (of Oregon) in a single assumed family. The evidence for this was later shown to be wrong and the hypothesis was abandoned. From the beginning of the classification of the languages of Western North America by Gallatin, a relationship between Molala and Cayuse was noted, and the family was referred to as Molala-Cayuse (Gallatin 1836).

Near the turn of the century, a new name for the family was created by Brinton and Dixon (who were having a contest to see who could come up with a classification of these languages faster) – Waiilatpuan (Brinton 1891; Powell 1891). This name then continued to be used in subsequent classifications of the Indian languages of North America, including Sapir's famous classification in the *Encyclopaedia Britannica* in 1929 (Powell 1915; Rivet 1924; Sapir 1929).

The Voegelin and Voegelin classification of 1965 resulted from a consensus of linguists at a meeting in 1964 at the University of Indiana. In it, the Waiilatpuan family was abandoned, and Molala and Cayuse were listed as separate nodes within the Penutian stock.

Campbell and Mixco refer to two articles by the famous linguist Bruce Rigsby in 1966 and 1969 that examined whether the Waiilatpuan family was a real thing. He had studied the languages in this area for his PhD dissertation. I examined both articles to see if his view that Cayuse and Molala are not related was valid. Dell Hymes (1969) also wrote a manuscript around this time going over Rigsby's data and giving his own opinion on it. However, I was not able to access this article for this study.

The first article (Rigsby 1966) was a brief look at the two languages. The second article (Rigsby 1969) was simply an expanded version of the first. In the first article, Rigsby noted that the two languages had simply been declared part of a Waiilatpuan family, but the evidence for the existence of this family was never examined. Everyone was just copying someone else's classification.

In the history of research on these tribes, Indians had reported on a number of occasions that the two groups spoke the same language. Rigsby said a simple examination of the languages showed that this was not so. There was also a story told by both sides that the Molala were a former Cayuse tribe which had been expelled from the band long ago. They went westward looking for new land.

Everywhere they went was already settled. Finally, they settled in an area where the local Indians let them stay. Ever since their expulsion, the Molala had been filled with a burning hatred for the Cayuse and a desire for revenge. There were a couple of conflicts between the two tribes around 1850. Rigsby concluded that the story about the expulsion could not be verified, but he was certain it had not happened in the 1800's. However, the story fits in well with what we know about those two languages.

The languages are quite far apart, but this is not unusual in Western North America even among related tongues. Rigsby reported that he found no more than 10% possible cognates among the languages. A quick examination of the evidence he reported shows that he was wrong. I counted 46 possible cognates (19%) or what he refers to as resemblances.

The resemblances were so few that Rigsby concluded that resemblances between the languages was a result of contact or borrowings. However, in his data, he only points out a few borrowings. Everything identified as a borrowing by Rigsby was not counted towards the 19% figure. Rigsby found that the small number of matches he found could all be explained by borrowing or chance or what he calls spurious.

An examination of the evidence shows that this is dubious. The number of resemblances is more than would occur due to chance (7%). And the items he chalks up to borrowings other than the ones he points out and explains are almost all in core vocabulary that is rarely borrowed. In fact, he says that the two personal pronoun paradigm matches (see below) are due to borrowing! He then notes that the pronouns are very similar to those reconstructed for Proto-Uto-Aztecan by Kenneth Hale, as an example that the lineup may be due to chance or borrowing.

This is a common tactic used by conservatives. Pointing out what long-rangers see as a relationship between families as obviously due to chance or borrowing because it's clear that the families are not related. But it's not clear. It's a theory. Once again, they are proceeding from a foregone conclusion and actually using long-rangers' arguments against the long-rangers themselves.

The problem here is that the resemblance between Waiilatpuan and PUA pronouns is because they are *both* evidence of the very common 1st and 2nd person (and possibly beyond) *m-/n-* pronoun paradigm in the Americas.

He examines the personal pronouns at length, while acknowledging a match in the 1st, 2nd, and 3rd person independent subjective pronouns. This pronoun set in Waiilatpuan is also a good match for the set in Proto-Sahaptian.

Moving on, in the much-expanded version of the 1966 piece that was published three years later (Rigsby 1969), Rigsby assembled all of the words for Cayuse forms he could find, 330 words. Of those words, 246 also had definitions in Molala, so there were 246 lineups to examine. I went over this material carefully and found many more resemblances, though it is true that the languages are still far apart. In a number of cases, several different forms were listed for Cayuse words. Perhaps the language had dialect diversity or a lot of synonyms. In fewer cases, there was more than one Molala form for the word. This ended up expanding our 246 forms significantly, but I did not have time to do the new calculations.

At any rate, the new data turned up many new resemblances, 83 or 33%. New matches were found between the two languages for *arrow*, *arm*, *alive*, *blood vessel*, *bridge*, *brush/wood*, *buck bison and bighorn*, *cheek*, *colt*, *drink*, *face*, *flint*, *fly*, *food*, *hair*, *infant*, *meal*, *mine*, *noon*, *root*, *bread and root cake*, *our*, *shoulder*, *sturgeon*, *swim*, *that*, *thief*, *wing*, *yes*, and *your*.

There were also elaborate sets of status terms for males and females. The females showed few resemblances, but the lineup in male status terms down to specific form/function terms was impressive (C = Cayuse, M= Molala): C *man*/M *boy;* C *young man*/M *young man*, *two men;* C *boy*/M

pubescent boy; C *husband*, *old man*/M *old man;* C *old bachelor*/M *bachelor;* C *man*/M *married man;* and C *young people*, *person*, *two men*/M *young person*. Further, the matches were almost never perfect in terms of semantics, but the semantic change was often what you would expect over time.

In addition, there were good matches for both languages in other Plateau Penutian languages, cementing the notion that both are at least members of this subfamily. In a small number of cases, Rigsby found matches in far-flung parts of Penutian. Not wanting to go to bat for the larger Penutian grouping, Rigsby calls these "Penutianisms," as if they are odd, unexplained resemblances that occur in this group for no obvious reason.

Here we harken back to Malcolm Guthrie's Bantuisms, where all resemblances in Bantu languages to other African forms were called Bantuisms instead of being properly seen as probable cognates.

Similar mystical scenarios are invoked for the common *m-/t-* pronoun pattern in Eurasiatic, with the result that the match in personal pronouns is common in Eurasiatic languages – basically a Eurasiaticism – and using that to dismiss the obvious pronoun paradigm match! Here we have conservatives taking good evidence for genetic relationship and tossing it aside as some common areal resemblance of no significance, a geographical oddity, an -ism. The "Penutianisms" that Rigsby finds are forms found throughout the family, obviously old cognates.

Rigsby lists the resemblances Cayuse and Molala have for other Plateau Penutian languages, but the matches were few.

A brief look through 23 words in Penutian languages in California (Zhivlov 2014) found the following cognates: *ear* CM, *eat* CM, *I* CM, *mouth* CM, *not* M, *thou* CM, *two* C, and *we* CM. There were eight matches in total for both languages, six for both together, and one each for each one. Still, 8 out of 23 (35%) is a good score. However, there were many more matches among the California languages themselves (most if not all of the Californian Penutian languages lined up on all words) than between them and Waiilatpuan, implying that Waiilatpuan must be further from the California languages than they are from each other.

At this point, it seems clear that the announcement of the death of Waiilatpuan is premature, to say the least. This is particularly so because Cayuse and Molala share many more resemblances not explained by borrowing or chance than either do with the rest of Sahaptian. If Waiilatpuan did not exist, there would be just as many matches between the two languages as with the rest of Sahaptian or Plateau Penutian. It seems it is time to re-examine the notion that Waiilatpuan has been proven wrong and has hence been abandoned, as that statement seems to be false.

Campbell and Mixco classify Cayuse as an isolate, mostly due to sparse documentation. Molala is part of Plateau Penutian (Berman 1996; Delancey 1996; Golla 1997), and Cayuse is probably part of the same group (Rigsby 1966), although it is left out of modern Penutian proposals on the grounds of poor documentation (Delancey 1996; Golla 1997). Plateau Penutian is part of the Penutian hypothesis, which has good support (see above under Penutian).

Campbell and Mixco, p. 130:

> **Mosan:** A now abandoned proposal of distant genetic relationship that would group Salishan, Wakashan and Chimakuan together (Sapir 1929, Swadesh 1953a and b).

The Mosan proposal has absolutely not been abandoned. In fact, it has now been expanded to include Algonquian in a family called Almosan. An interesting argument is that Almosan not only is valid but is related to Nivkhi in the Old World (Nikolaev 2015a, 2015b, 2017). David Beck recently argued that Mosan is a language area or Sprachbund instead of a genetic family (Beck 1997).

Michael Fortescue argued a few years before that Mosan was a valid entity and was related to the Old-World language Nivkhi (Fortescue 1998). Recently, Murray Gell-Mann, Ilia Peiros, and George Starostin also supported Almosan and grouped it with Chukotko-Kamchatkan and Nivkhi (Gell-Man *et al.* 2009).

So far, we have five specialists arguing that Mosan is a valid family[47] and three saying it is not,[48] leading to 63% support. The consensus among specialists of Mosan is optimistic. At any rate, Campbell and Mixco's statement that this proposal is "now abandoned" is false.

For Almosan, we have four specialists saying it is valid[49] and four saying it is not.[50] Expert consensus on Almosan is evenly split.

Campbell and Mixco, p. 77-78:

> **Hokan:** A controversial hypothesis of distant genetic relationship proposed by Dixon and Kroeber among certain languages of California; the original list included Shastan, Chimariko, Pomoan, Karok, and Yana, to which they soon added Esselen, Yuman, and later Chumashan, Salinan, Seri, and Tequistlatecan. Later scholars, especially Edward Sapir, proposed various additions to Hokan. Many 'Hokan' specialists doubt the validity of the hypothesis.

It is not true that many Hokan specialists doubt the validity of the hypothesis.[51]

Below I will list all of the purported families that make up Hokan by listing specialists and publications that provided evidence that they were Hokan. In some cases, I will show that Hokan membership has been rejected and the evidence for that. In this way, we see that not only the family itself is accepted, but that the vast majority of its members have been specifically proven.

Pro-Hokan scholars usually refer to a Hokan family made up of the following languages and groups:

Coahuiltecan: Brinton 1891; Swanton 1915; Sapir 1917b, 1920a, 1921a; Bright 1956; Langdon 1974; Greenberg 1987; Kaufman 1988, 2005 (as Pakawan);

Ch'imáriko: Dixon 1910; Dixon and Kroeber 1913a, 1919; Sapir 1917b, 2001; Bright 1954, 1957; Olmsted 1956, 1957, 1959; Haas 1963; Gursky 1966,1974a, 1974b, 1988, 1989, 1990, 1995; J. M. Crawford 1976; Silver 1978; Kaufman 1988, 2005; Zhivlov 2013;

---

[47] Fortescue 1998, Gell-Man *et al.* 2009; Nikolaev 2015a, 2015b, 2017.
[48] Beck 1997; Campbell and Mixco 2007.
[49] Gell-Man *et al.* 2009; Nikolaev 2015a, 2015b, 2017
[50] Beck 1997; Fortescue 1998; Campbell and Mixco 2007.
[51] Good undated; Sapir 1929; Swadesh 1967; Silver 1976; Voegelin and Voegelin 1977; Kaufman 1988, 2005; Poser 1995; Zhivlov 2013.

Esselen: Dixon and Kroeber 1913a; Sapir 1917b; Heizer 1952; Olmsted 1956, 1957, 1959; Haas 1963; Gursky 1974a, 1974b, 1988, 1989, 1990; Beeler 1978; Shaul *et al.* 1984; Turner 1987; Kaufman 1988, 2015; Shaul 1995a, 1995b, 2019;

Jicaquean or Tolan: Swadesh and Greenberg 1953; Oltrogge 1977; Kaufman 1988, 2015;

Karuk: Dixon and Kroeber 1913a; Sapir 1917b; Harrington 1932; Haas 1963; Bright 1954, 1987; Olmsted 1956, 1957, 1959; Jacobsen 1958; Haas 1963, 1964; Silver 1964; Gursky 1966, 1995; Greenberg 1987; Kaufman 1988, 2015; Zhivlov 2013; Sandy 2017;

Pomoan: Dixon and Kroeber 1913a; Sapir 1917b, 1925; Rivet 1942; Haas 1963; Moshinsky 1964; McLendon 1964; Langdon 1974; Gursky 1974a, 1974b, 1988, 1989, 1990; O'Connor 1987; Greenberg 1987; Kaufman 1988, 2015; Gursky 1995; Zhivlov 2013;

Salinan: Kroeber 1904; Dixon and Kroeber 1913b; Sapir 1917b, 1920b, 1921a, 1925; Mason 1918; Rivet 1942; Jacobsen 1954, 1955, 1958, 1979, 1986; Olmsted 1956, 1957, 1959; Haas 1963; McLendon 1964; Gursky 1974a, 1974b; Greenberg 1987; Turner 1987; Kaufman 1988, 2015; Poser 1992; Zhivlov 2013;

Seri: Brinton 1891; Kroeber 1915; Sapir 1917b; Moser and Moser 1961; Haas 1963; J. G. Crawford 1976; Waterhouse 1976; Kaufman 1988, 2015; Müller *et al.* 2013; Zhivlov 2013;

Shasta-Palaihnihan: Dixon 1905, 1910; Dixon and Kroeber 1913a; de Angulo and Freeland 1930; Harrington 1932; Bright 1954, 1957; Bright and Olmsted 1959; Olmsted 1956, 1957, 1959; Haas 1963; Silver 1964, 1966, 1976, 1980; Gursky 1966, 1995; Silver and Wicks 1977; Greenberg 1987; Kaufman 1988, 2015; Good *et al.* 2003; Zhivlov 2013; Nevin 2019;

Tequistlatecan: Brinton 1891; Kroeber 1915; Sapir 1917b, 1925; Waterhouse and May 1950; Waterhouse 1962; Haas 1963; Gursky 1974a, 1974b, 1988, 1989, 1990; Oltrogge 1977; Greenberg 1987; Kaufman 1988, 2015; Müller *et al.* 2013; Zhivlov 2013;

Washo: Dixon and Kroeber 1919; Sapir 1921a; Jacobsen 1958, 1964, 1966, 1979: 567-70, 1986; Haas 1963; Gursky 1966, 1974a, 1974b; Waterhouse 1976; Greenberg 1987; Kaufman 1988, 2015; Gursky 1995;

Yana: Sapir 1917a, 1917b, 1920b, 1921a, 1922, 1923; Dixon and Kroeber 1919; Rivet 1942; Olmsted 1956, 1957, 1959; Jacobsen 1958; Sapir and Swadesh 1960; Haas 1963, 1964; McLendon 1964; Gursky 1966; Greenberg 1987; Kaufman 1988, 2015; Zhivlov 2013; Nevin 2019;

Yuman: Brinton 1891; Dixon and Kroeber 1913a, 1919; Sapir 1917b, 1925, 2001; Rivet 1942; Kroeber 1943; Spier 1946; Halpern 1946a, 1946b, 1946c; Winter 1957; Jacobsen 1958; Haas 1963; Gursky 1966, 1974a, 1974b, 1988, 1989, 1990, 1995; J. G. Crawford 1976; J. M. Crawford 1976; Kaufman 1986, 2015; Langdon 1990; Müller *et al.* 2013; Zhivlov 2013.

There are other languages that have been thought to be Hokan, including:

Chumashan: Dixon and Kroeber 1913b; Sapir 1917b, 1925; Haas 1963; Gursky 1974a, 1974b, 1988, 1989, 1990; Greenberg 1987;

Yurumanguí: Rivet 1942; Greenberg 1987;

Baja California or Guaicurian languages, consisting of Guaocura (Waikuri), Callejue, Cora, Huchiti, Aripe, Periúe, Pericú, and Isleño: Gursky 1966; Swadesh 1967; Waterhouse 1976.

The consensus now is that Chumashan (Kaufman 1988; Poser 1995), Yurumanguí (Poser 1992, 1995), and Guaicurian (Poser 1995) are not properly included in Hokan.

In the case of the last two, the main problem is poor documentation (Poser 1995). Subtiaba was considered to be a Hokan language for a long time,[52] but it is now no longer considered part of the family, as it is now considered to be Otomanguean. However, some now propose that Subtiaba as part of the Tlappanecan family is related to Hokan as a part of the Otomanguean family.[53] So we see that the people proposing Subtiaba as Hokan were onto something after all.

A few outliers such as Quinigua and Tlappanec (Sapir 1925; Waterhouse 1976; Greenberg 1987) have also been proposed, but there is no evidence that they are Hokan. One specialist (Zhivlov 2013) holds that Washo and Tolan are not part of Hokan, while allowing that they may be

[52] Lehmann 1920; Sapir 1925; Rivet 1942; Gursky 1968; Oltrogge 1977; Greenberg 1987.
[53] Weitlaner 1941; Radin 1944; Kaufman 1990b.

related at a very remote time-depth. Some are agnostic on Esselen (Poser 1995, Zhivlov 2018), which is only known from a 350-word list collected by friars at a California mission. However, 12 Hokanists accept Esselen[54] and two are agnostic,[55] so it seems that consensus now among Hokanists is that Esselen is Hokan.

11 Hokanists say that Seri is Hokan[56] and 15 say that Salinan is Hokan,[57] while only one Marlett (2008) says both languages are not part of the family. Consensus among Hokanists is that Seri and Salinan are Hokan.

All of the Hokan languages listed above as being part of Hokan have many scholars supporting this view. The exception is Tolan, which has five specialists[58] saying it is part of Hokan and another[59] saying it is not. Even here, consensus seems to be that Tolan is Hokan.

It is true that there are some specialists who do not believe in Hokan. Seventy-five years ago, the first opponents to Hokan appeared (Hoijer 1946, 1954). In a seminal document, Mary Haas (1963) cites a very long list of references of scholars who back up the idea of Hokan. Even sixty years ago, the list was very long and was growing rapidly at the time she wrote her paper. She takes care to point out that Hoijer opposed Hokan, but she seems shocked that anyone would oppose such an obvious concept.

There continue to be opponents of Hokan into the present era.[60] Obviously, Campbell and Mixco are opponents, and they both qualify as specialists. Campbell has also written about Hokan in other places (Campbell 1997). Marianne Mithun is a colleague of Campbell's who has collaborated with him on books on Amerindian languages. Her work deserves scrutiny because her arguments are fairly elaborate.

In one paper (Mithun 2007), she reverts to the Northern California Linguistic Area to make an areal explanation for the languages found in the area. It is true that Hokan, Penutian, Yuki-Wappo, and Ritwan languages all show certain common characteristics such as "instrumental" or lexical verbal prefixes (Delancey 1988b). It is considered an areal effect, even though the transmitted material involves intricate morphological material. In her paper, Mithun lists all of the languages in this area that have lexical prefixes and then notes that none of them are related to each other.

But she is beginning from a false assumption because she lists several Penutian and Hokan languages and claims that they show no relations to each other. But this is not proven. Then she notes that a wide variety of these occur in this area, and they rarely if ever match in form and function.

---

[54] Dixon and Kroeber 1913a; Sapir 1917b; Heizer 1952; Olmsted 1956, 1957, 1959; Haas 1963; Gursky 1974a, 1974b, 1988, 1989, 1990; Beeler 1978; Shaul *et al.* 1984; Turner 1987; Kaufman 1988, 2015; Shaul 1995a, 1995b, 2019.
[55] Poser 1995; Zhivlov 2018.
[56] Brinton 1891; Kroeber 1915; Sapir 1917b; Moser and Moser 1961; Haas 1963; J.G. Crawford 1976; Waterhouse 1976; Kaufman 1988, 2015; Müller *et al.* 2013; Zhivlov 2013.
[57] Kroeber 1904; Dixon and Kroeber 1913b.; Sapir 1917b, 1921a 19,25; Mason 1918; Rivet 1942; Jacobsen 1954, 1955, 1958, 1979, 1986; Olmsted 1956, 1957, 1959, Haas 1963; McLendon 1964; Gursky 1974a, 1974b; Greenberg 1987; Turner 1987; Kaufman 1988, 2015; Poser 1992; Zhivlov 2013.
[58] Swadesh and Greenberg 1953; Oltrogge 1977; Kaufman 1988, 2015; Poser 1995.
[59] Zhivlov 2018.
[60] Campbell 1997; Mithun 1999, 2007, 2012; Walker 2013.

But this is where she is wrong. In an intricate paper, Zhivlov (2018) points out that many of the lexical prefixes in Hokan at least are cognates and do match in form and function. None of these languages seem to have borrowed specific morphemes from each other. Instead, they borrowed a basic typology of lexical prefixes. Furthermore, these same lexical prefixes matching in both form and function are found in other Hokan languages like Salinan, Yuman, Tequistlatecan, Seri, and possibly Tol. The last four are quite distant from Northern California, and it stretches credulity that the forms were borrowed. Instead, they seem to have been inherited from Proto-Hokan.

In addition, Kaufman (2008, 2015) also demonstrates a remarkable 1st and 2nd person pronoun paradigm among Seri, Yuman, and Ch'imáriko that scarcely exists elsewhere. This remarkable resemblance is akin to *m-/t-* in Eurasiatic and *n-/m-* in Amerind and seems to clinch the case for Hokan. There are also matches for a second person singular and plural paradigm with ablauted vowels and for possessed noun marker paradigm, once again with ablauted vowels (Zhivlov 2018).

In a sense, this renders the lexical prefix argument null because Hokan also has matching paradigms in personal pronouns and possessed nominal markers. These cognates are found in Hokan languages quite distant from one another, and at any rate, personal pronoun markers are almost never borrowed, and personal pronoun *paradigms* are simply never borrowed.

One specialist has taken a neutral stance on Hokan (Jany 2014). At this point, we have five specialists who oppose Hokan[61] and one who is neutral.[62] Opponents have produced nine anti-Hokan papers. However, Hokan has had 52 supporters down through the years, and they have published 102 papers.[63]

Hokan specialists support Hokan by 52-5 and have published more papers than Hokan opponents by 102-9.

Furthermore, the hypothesis dates back over a hundred years. It shows no signs of being a formerly popular notion that has been abandoned in recent years. Instead, evidence has continued to mount from 1905 to 2019.

We can also look at the attitude of Hokanists towards their own proposal. Actually, even many long-range linguists are sanguine and conservative in their wording about whether their long-range proposal is valid or not. Many of them couch their claims with notes that the family is hypothesized

[61] Hoijer 1946; 1954; Campbell 1997; Mithun 1999, 2007, 2012; Campbell and Mixco 2007; Walker 2013.
[62] Jany 2014.
[63] Brinton 1891; Kroeber 1904, 1915, 1943; Dixon 1905, 1910; Dixon and Kroeber 1913a, 1913b, 1919; Swanton 1915; Sapir 1917a, 1917b, 1920a, 1920b, 1921a, 1922, 1923, 1925, 2001; Mason 1918; Lehmann 1920; de Angulo and Freeland 1930; Harrington 1932; Weitlaner 1941; Rivet 1942; Radin 1944; Spier 1946; Halpern 1946a, 1946b, 1946c; Waterhouse and May 1950; Heizer 1952; Swadesh and Greenberg 1953; Jacobsen 1954, 1955, 1958, 1964, 1966, 1979, 1986; Bright 1954, 1956, 1957, 1987; Olmsted 1956, 1957, 1959; Winter 1957; Bright and Olmsted 1959; Sapir and Swadesh 1960; Moser and Moser 1961; Waterhouse 1962, 1976; Haas 1963, 1964; McLendon 1964; Moshinsky 1964; Silver 1964, 1966, 1976, 1978, 1980; Gursky 1966; 1968, 1974a, 1974b, 1988, 1989, 1990, 1995; Langdon 1974, 1990; J. M. Crawford 1976; J. G. Crawford 1976; Silver and Wicks 1977; Oltrogge 1977; Beeler 1978; Shaul *et al.* 1984, 1995a; 1995b, 2019; Greenberg 1987; O'Connor 1987; Turner 1987; Kaufman 1988, 1990, 2005, 2009, 2015; Poser 1992, 1995; Berman 2001; Good *et al.* 2003; Müller *et al.* 2013; Zhivlov 2013, 2018; Sandy 2017; Nevin 2019. I did not list some supporters and papers due to time constraints and other reasons.

but not yet proven (Poser 1995, Zhivlov 2018). However, the reaction of Hokanists has been different.

Already sixty years ago, Mary Haas (1963) seemed stunned that there was even one prominent linguist who doubted the validity of Hokan.

And as late as thirty years ago, Terrence Kaufman (1990a) was already going out of his way to state that his Hokan proposal had been frequently misconstrued as a hypothesis, whereas for him, the Hokan family was simply a proven fact. Neither Haas nor Kaufman is a notable long-ranger, and Kaufman has often allied with his Americanist colleague Campbell (Kaufman 1990a) against Greenberg's (1987) Amerind proposal. This sort of forthrightness among otherwise fairly conservative linguists is quite rare, and such judgements are typically issued only when evidence of a relationship becomes overwhelming.

Taken as a whole, Campbell and Mixco's contention that "many 'Hokan' specialists doubt the validity of the hypothesis" is at best weasel-wording and at worst a serious distortion of consensus on the subject. Looking at all of the above, there can be no other conclusion but that Hokan is a legitimate language family supported by the vast majority of specialists. *It is time to finally bring Hokan in from the cold and recognize the validity of the Hokan language family of Western North America* (Kaufman 2008 2015; Zhivlov 2018).

Campbell and Mixco, p.100:

> **Lencan:** A small language family of two languages, both now extinct, Honduran Lenca and Chilanga or Salvadoran Lenca. Hypotheses have attempted to link Lencan with various other languages in broader genetic groupings, but none has much support. The often repeated hypothesis of a connection between Lencan and Xinkan has no reliable evidence and is now abandoned.

The argument dates back a century (Lehmann 1920). Campbell's later statements contradict his statement above that no linkage of Lencan to any other family has much support. In fact, Campbell himself believes in such a linking of Lencan and two other families in a family called Macro-Chibchan or Lenmichi. Campbell traditionally regarded the argument as optimistic, but with the publication of Constenla Umaña (2005) (who also included the Chibchan family), he agreed that the case had been proven and that a Lenmichi family consisting of Chibchan, Lencan, and Misumalpan exists (Campbell 2012). At an estimated age of 9,700 years, this family is very old.

What makes this interesting is that this is the first long-range language family Campbell has accepted in forty years, which shows he is at least open to some long-range proposals. He is also optimistic about Macro-Mayan and feels it will be proven in the future. Campbell's expertise is centered on Central American languages.

This implies that where Campbell has a vast knowledge, he is open to long-range proposals, but where his knowledge is weaker, he is skeptical, which lends an interesting layer to his conservatism.

There is excellent evidence that the Lencan, Chibchan, and Misumalpan families form a Lenmichi language family.[64] An ASJP computerized phylogenetic study found links between the three

[64] Lehmann 1920; Greenberg 1987; Kaufman 1994; Campbell 1997, 2012; Constenla Umaña 2005; Müller *et al.* 2013.

families (Müller *et al.* 2013). However, some still oppose the addition of Misumalpan to Lenmichi (Hale and Salamanca 2001). Looking only at the Lenmichi argument without Xincan, we count 26 scholars in favor of it and two against it at least in part. Consensus clearly supports Lenmichi.

Campbell and Mixco, p. 225:

> **Xinca-Lenca:** Once proposed but now abandoned distant genetic relationship that sought to group the Xinkan and Lencan families. (See Campbell 1997: 961–3 for an evaluation.)

The authors are not correct that a linkage between the Lencan and Xinkan families is now abandoned is false, as 60% of specialists support a link.

Campbell and Mixco, p. 225-226:

> **Xinkan, Xincan**: A small family of four languages of Guatemala, Chiquimulilla, Guazacapán, Jumaytepeque and Yupiltepeque. The first and last became extinct in the last two decades of the twentieth century; the other two have extremely few remaining speakers, only two or three each. Hypotheses have attempted to link Xinkan with various other languages in broader genetic groupings, but none has any value. The often repeated hypothesis of a connection between Xinkan and Lencan has no reliable evidence and is now abandoned.

Others (Lehmann 1920; Greenberg 1987; Kaufman 1994) add the Xincan family to Lenmichi. Noted Americanist Terrence Kaufman (1990b) regards it as proven. The authors oppose it, and Campbell has written specifically on the proposal (Campbell 1997: 167). The argument here has three in favor and two against, so consensus does not yet support adding Xincan to Macro-Chibchan, although it is optimistic about the relationship.

There is also some suggestive recent evidence that the Lencan and Misumalpan families of South America may be Hokan. Other Hokanists have not yet commented on the proposal.[65]

Campbell and Mixco, p. 33:

> **Cochimí-Yuman:** A family of languages from Arizona, California and Baja California, with two branches, extinct Cochimí (of Baja California) and the Yuman subfamily (members of which are Kiliwa, Diegueño, Cocopa, Mojave, Maricopa, Paipai, and Walapai-Havasupai-Yavapai, among others). Cochimí-Yuman is often associated with the controversial Hokan hypothesis, though evidence is insufficient to embrace the proposed relationship.

There are two separate questions here, and I am not sure which the authors are referring to. One is, is Yuman part of Hokan? The answer to that is yes. The other question is, does Hokan itself exist? For the answer to that, see the Hokan entry. Mixco opposes the Hokan affinity of Cochimi-Yuman, and granted, he is actually a specialist on these languages (Mixco 1976, 1997).

However, the vast majority of Yuman specialists think that Yuman is a part of the Hokan family and that the Hokan family is a valid entity.[66] Müller *et al.* (2013) was a computerized phylogenetic study that found links between Yuman, Seri, and Tequistlatecan.

So, among specialists, we have 17 who say that Yuman is Hokan and two (the authors) who oppose it (Mixco 1976, 1997). The specialist consensus, then, definitely is that Yuman is a part of

[65] Poser 1995; Kaufman 1998, 2015; Zhivlov 2018.
[66] Brinton 1891; Dixon and Kroeber 1913a; Sapir 1917b, 1925; Rivet 1942; Kroeber 1943; Spier 1946; Halpern 1946a, 1946b, 1946c; Winter 1957; Haas 1963; Gursky 1974a, 1974b, 1988, 1989, 1990; J.G. Crawford 1976; J.M. Crawford 1976; Kaufman 1988, 2015; Langdon 1990; Poser 1995; Müller *et al.* 2013; Zhivlov 2013.

the Hokan family. Campbell and Mixco's conclusion that evidence is insufficient to support this theory does not represent consensus among Yumanists.

Campbell and Mixco, p. 32-33:

> **Coahuiltecan:** A hypothesis of distant genetic relationship that proposed to group some languages of south Texas and northern Mexico: Coahuilteco, Comecrudo and Cotoname, and sometimes also Tonkawa, Karankawa, Atakapa and Maratino (with Aranama and Solano assumed to be varieties of Coahuilteco).
>
> Sapir proposed a broader classification of Hokan-Coahuiltecan, joining the Coahuiltecan proposal with the broader Hokan hypothesis, and placed this in his even larger Hokan-Siouan super-stock. None of these proposals has proven sufficiently robust to be accepted generally.

This family was proposed over a hundred years ago (Powell 1891). It was later affirmed (Swanton 1915), when the proposal was expanded to include Comecrudo, Cotoname, Coahuilteco, Karankawa, Tonkawa, Atakapa, and Maratino languages into a Coahuiltecan grouping. This grouping was later reaffirmed and linked with Hokan (Sapir 1920a).

A lesser proposal of a Comecrudan language family consisting of Comecrudo, Garza, and Mamalique has recently been accepted by some who reject the larger proposal (Goddard 1979).

So here we have a situation similar to Altaic, where linguists do not accept the entire proposal but instead accept lesser versions of it. I call it "baby Altaic" in the case of that family and "baby Coahuiltecan" in the case of these languages. Presumably, the authors include all of the "baby proposals" as rejections of the hypothesis itself. But this is an argument in bad faith. A proposal need not be accepted in full to be accepted. Parts of it can be accepted too, and this is more than a null conclusion. I should also point out that Goddard is another of the Americanists whose conservatism got triggered (Goddard 1987) by *LIA* (Greenberg 1987), and he is also a close ally of Campbell's.

An interesting argument was made in a recent paper that set out to prove the proposal of a Coahuiltecan family consisting of Comecrudo, Cotoname, Aranama, Solano, Mamulique, Garza, and Coahuilteco while ruling out Karankawa, Atakapa, and Tonkawa for the time being, and saying that there was not enough data to place Maratino in the family (Manaster-Ramer 1996b). Reading the paper, the similarities in core vocabulary and even in sound correspondences were so striking that I was shocked that this was a controversial proposal.

In fact, Manaster-Ramer's grouping, now called the Pakawan family, harkened back to Powell's original proposal (Powell 1891) while rejecting the later expansions by Swanton (Swanton 1915) and Sapir (Sapir 1920a) and is considered uncontroversial consensus in the field. It is only when we attempt to expand it to Karankawa, Atakapa, and Tonkawa that we run into trouble. By tying the accepted Pakawan part of the family in with the admittedly unproven larger Coahuiltecan, the authors offer a bad faith argument that there is nothing to the whole proposal.

A recent paper stated that most linguists do not accept the existence of a Coahuiltecan language family (Logan 2001). However, this book was not written by a linguist. Instead, it was written by an archeologist.

That field has undergone severe changes in recent years and is now an outpost of political correctness and sacrifice of science to "woke" dogma. The argument Logan (2001) makes is that

the Coahuiltecans, instead of being a group of closely related tribes, were instead groups of little relation to each other who spoke mutually unintelligible languages and were linked only by place of common residence. This is a typical "Indianist" narrative that is often seen in such documents and has more to do with assuaging Indians' fears of Whites and desires to be unique than it does with actual science.

Campbell rejects Coahuiltecan in a paper published as a response to Manaster-Ramer's paper in the same journal earlier in the year (Campbell 1996). He uses the familiar arguments of irregular sound correspondences and mass borrowing of core vocabulary, two of the conservatives' go-to arguments. So, Campbell is negative about the proposal, and Goddard accepts only part of it (Campbell 1996, Goddard 1979).

On the other side, four specialists accept Coahuiltecan in whole or in the sense of Pakawan.[67] Goddard (1979) accepts a family with even fewer members than Pakawan.

So, among specialists we have four supporters, one partial supporter, and two rejectionists (Campbell and Mixco). It would seem that linguistic consensus among specialists supports at least the reduced Pakawan version. If we include Hokanists as pro-Coahuiltecan, we end up with nine supporters of Coahuiltecan versus one partial supporter and two rejectionists. If we tally it this way, consensus supports the existence of Coahuiltecan. Manaster-Ramer (1996) makes a good case that the best case for Coahuiltecan is the reduced Pakawan family.

Consensus among Hokanists also seems to accept Coahuiltecan as a member of the Hokan family.[68] Only Campbell and Mixco are against the idea. Goddard does not seem to have an opinion on the question. So, we have seven specialists in favor of some version of Coahuiltecan as Hokan and two against it. Consensus among specialists is that Coahuiltecan is part of the Hokan family.

While there is no good evidence that Atakapa and Tonkawa are part of Coahuiltecan, there are nonetheless numerous parallels between both languages and Coahuiltecan that seem hard to explain (Manaster-Ramer 1996b). Tonkawa may be Na-Dene as per another paper by Manaster-Ramer (1996c). A good case can be made that Atakapa is part of the Gulf family (Haas 1951, 1952; Greenberg 1987; Munro 1995).

Campbell and Mixco, p. 76:

> **Gulf:** Hypothesis of a distant genetic relationship proposed by Mary R. Haas that would group Muskogean, Natchez, Tunica, Atakapa, and Chitimacha, no longer supported by most linguists.

The notion that the Gulf family is no longer supported by most linguists is simply incorrect. There have only been four linguists who studied this family in depth.

The first was Mary Haas (1951, 1952), who also proposed a relationship with Yuki-Wappo as Yuki-Gulf. Haas was always dubious about Chitimacha's addition to the Gulf family.

Joseph Greenberg resurrected Yuki-Gulf in *LIA* (1987).

---

[67] Powell 1891; Swanton 1915; Sapir 1920a; Manaster-Ramer 1996.
[68] Brinton 1891; Swanton 1915; Sapir 1917b, 1920, 1921a; Bright 1956; Langdon 1974; Greenberg 1987; Kaufman 1988, 2005 (as Pakawan).

Pam Munro is an expert on these languages. She published a paper (Munro 1995) on Yuki-Gulf showing striking resemblances between the Muskogean family, Natchez, Tunica, Atakapa, and Chitimacha. Furthermore, the relationship with Yuki and Wappo, a full 2,500 miles away in Northern California, was striking. At first glance, there also appeared to be regular sound correspondences running through it. The resemblances were so blatant that I was shocked that this was a controversial proposal.

The fourth was Geoffrey Kimball (1994), who concluded that Gulf was probably a valid family but that this could not be proven.

What was shocking about Munro's paper was that Yuki and Wappo could not possibly have borrowed from Gulf because Gulf is in Louisiana, 2,500 miles away from the other two. So how did all these resemblances come in? Chance is ruled out. Borrowing could not have happened. Therefore, a relationship at least between Yuki and at least some of the Gulf languages is obvious.

Munro's paper took the position that Greenberg's Yuki-Gulf hypothesis was correct. However, there are some problems. First, Atakapa as part of the Gulf family has been controversial, in part because it has also been tied in with Coahuiltecan. Indeed, there are similarities with the latter, and the two were not spoken next to each other, so borrowing can be ruled out.

Perhaps a way of solving the matter is to posit not only Yuki-Gulf but a larger family that includes Coahuiltecan as Greenberg does in *LIA*. I have no idea how justified this is, but there are certainly surprising resemblances between Atakapa and the Coahuiltecan languages.

Furthermore, whether or not Chitimacha is part of the Gulf family has been up in the air from the beginning, when Mary Haas published her papers. Recent papers using the comparative method (Brown *et al.* 2011, 2014) have made the case that Chitimacha is related to Mesoamerican language families of Mexico such as Mixe-Zoque and Totonacan. Campbell has rejected this hypothesis.

That Tunica at the very least shows a close relationship with the Muskogean languages is not even controversial. The idea has a long pedigree (Haas 1958) and is presently supported by all experts on this family.

Geoffrey Kimball (1994) examined the data and concluded that from the evidence, it appeared that Gulf is a valid entity, but we will never be able to prove it, as he put it. However, he stated that Tunica is almost certainly related to Muskogean. A Tunica-Muskogean family at the very least is now consensus among specialists.

Kimball's paper had a number of problems, foremost among which was that he was operating with a negative stance towards the existence of the family. Further, there may be issues with his resorting to sound symbolism and borrowings to explain resemblances in the paper.

Let's evaluate Campbell and Mixco's statement that the Gulf family is no longer supported by most specialists.

We have four specialists on record about whether or not a Gulf family exists: Haas (1951, 1952) is positive, minus Chitimacha; Greenberg (1987) is positive; Pamela Munro (1995) is positive; Kimball (1994) thinks it is probably valid but it's not possible to prove it; Brown *et al.* (2014) agree that Chitimacha is a part of the Totonozoquean family, not the Gulf family, while the other members of Gulf are not members of this family.

Three out of the four specialists on the Gulf family say that the Gulf family is valid. The other feels it probably exists but cannot be proven. And there is uncertainty about whether Chitimacha is part of the Gulf family. The consensus among experts is that the Gulf family is a valid entity.

Campbell and Mixco's statement that the Gulf family is no longer supported by most linguists is simply false.

Furthermore, a good case can possibly be made for the existence of a Totonozoquean family consisting of the Mixe-Zoque and Totonacan language families. The proposal is an old one, dating back eighty years (McQuown 1942, 1956), though resemblances were noted even before then. Cecil Brown has been working on it for thirty-five years (Brown and Witkowski 1979). Brown and colleagues recently published two papers making an excellent case for Totonozoquean (Brown *et al.* 2011, 2014). Campbell and Kaufman have traditionally regarded the family as unproven (Campbell 1997; Campbell and Kaufman 1976, 1978, 1980), though neither has commented on Brown *et al.*'s powerful new arguments.

Campbell and Mixco, p. 112:

> **Macro-Gê:** A proposed distant genetic relationship composed of several language families and isolates, many now extinct, along the Atlantic coast (primarily of Brazil). These include Chiquitano, Bororoan, Botocudoan, Rikbaktsa, the Gê family proper, Jeikó, Kamakanan, Maxakalían, Purian, Fulnío, Ofayé and Guató. Many are sympathetic to the hypothesis and several of these languages will very probably be demonstrated to be related to one another eventually, though others will probably need to be separated out (Kaufman 1990b, 1994, Rodrigues 1986).

This is much too pessimistic, as Macro-Gê is not a *proposed* long-range family – it *is* a large language family in South America accepted by consensus (Jolkesky 2016; Pache 2018). It is not true that *many are sympathetic to it*; instead, the consensus is that it is *correct* (Nikulin 2020). Nor is it correct to say that it *will probably be demonstrated eventually*. In fact, it is already an *accepted reality* (Ribeiro 2012). However, it is correct that the internal structure of the family has not been completely worked out, and there is controversy about whether some of the languages proposed for membership are part of the family (Nikulin 2020; Ribeiro 2012).

Campbell and Mixco, p. 163:

> **Quechumaran:** Proposed distant genetic relationship that would join Quechuan and Aymaran. While considerable evidence has been gathered in support of the hypothesis, it is extremely difficult in this case to distinguish what may be inherited (and therefore evidence of a genetic relationship) from what may be diffused (and therefore not reliable evidence of a genetic connection) (Campbell 1995, 1997: 273–83).

It is true that there is no consensus on the existence of Quechumaran. The consensus seems to be that it is not yet proven. If we include Campbell, there are five experts who do not believe in the family (Adelaar 1992; Campbell 1997; Moulian *et al.* 2015) and three who say it is a valid entity (Kaufman 1990b; Cerrón Palomino 2000; Müller *et al.* 2003). However, the latter study was an ASJP automated computer that has a good track record on showing at the least deep relationships among families, typically either genetic, but sometimes also substrate and sometimes also mass borrowing. So, we have five against Quechumaran and four in favor of it. It sounds like it is still very much up in the air.

Those opposed to the idea propose the usual borrowing scenario, but they have had to push the large number of borrowings in core vocabulary all the way back to Proto-Aymara and Proto-Quechua (Adelaar 1992; Emlen 2017). In my opinion, "massive borrowing of core vocabulary at the proto-language level" is simply another word for genetics.

I should give Campbell a break here because he now says that there is a 50-50 chance that the language families are related (Campbell 1995). For an extreme conservative, that is a nice concession.

Campbell and Mixco, p. 112-113:

> **Macro-Guaicuruan** (also spelled **Macro-Waykuruan, Macro-Waikuruan**): A proposed distant genetic relationship that would join the Guaicuruan and Matacoan families of the Gran Chaco in South America in a larger-scale genetic classification. Grammatical similarities, for example in the pronominal systems, have suggested the relationship to some scholars, but the extremely limited lexical evidence raises doubts for others. Some would also add Charruan and Mascoyan to these in an even larger 'Macro-Waikuruan cluster' (Kaufman 1990b, 1994).

Better names for this family are Mataco-Guaicuru and Mataguayo-Guaicuru. It is not true that this is a proposed long-range family suggested by some but doubted by others. It is true that the authors doubt the proposal, and Campbell has written specifically on it (Campbell 2012). Five linguists support the proposal,[69] including Morris Swadesh, who included it in his Macro-Mapuche stock. The result is 5 against 2 or 72% of specialists in favor of the proposal.

There is debate about which families are members outside of the Guaicuruan and Mataguayo language families that make up the essence of the family. There have been suggestions to add the Lule-Vilela, Zamucoan, Charruan, and Mascoyan families to this family. Suárez (1974) included Charruan in a hypothetical Waikuru-Charrúa stock. Swadesh added Charruan and Mascoyan to the family. It is not known how much support there is for including Charruan in the proposal, so we will leave it as unproven by consensus for now. See more discussion of this family and its external relations above in the Amerind section.

Campbell and Mixco, pp. 148-149:

> **Pama-Nyungan:** A very large, widely spread language family of Australia, some 175 languages. The name comes from Kenneth Hale, based on the words *pama* 'man' in the far northeast and *nyunga* 'man' in the southwest. Languages assigned to Pama-Nyungan extend over four-fifths of Australia, most of the continent except northern areas.
>
> Pama-Nyungan is accepted by most Australianists as a legitimate language family, but not uncritically and not universally. It is rejected by Dixon; it is held by others to be plausible but inconclusive based on current evidence. Some Pama-Nyungan languages are Lardil, Kayardilt, Yukulta, Yidiny, Dyirbal, Pitta-Pitta, Arrente, Warlpiri, Western Desert language(s), and there are many more.

Pama-Nyungan was originally proposed by Kenneth Hale over fifty-five years ago (Hale 1964). Actually, consensus now is that this family of Australian languages is indeed valid (Bowern and Atkinson 2012). True, the famous Australianist R. M. W. Dixon challenged the existence of Pama-Nyungan recently (Dixon 1980, 2002), but his opposition was so outrageous that it prompted a

[69] Suárez 1974; Barros 1992-1993, 2006, 2013; Kaufman 2007; Jokelsky 2016.

quick surge of papers from Australianists defending the existence of Pama-Nyungan (Bowern and Koch 2004, O'Grady and Hale 2004).

The notion that other Australianists feel that Pama-Nyungan is possible but presently inconclusive (Kaufman 2007) is not correct. I am not aware of a single Australianist other than Dixon who feels this way. Instead, Pama-Nyungan is about as uncontroversial as Macro-Gê, Afroasiatic, or Austroasiatic.

Campbell and Mixco, p. 149:

> '**Papuan' languages**: A term of convenience used to refer to the languages of the western Pacific, most in New Guinea (Papua New Guinea and the Indonesian provinces of Papua and West Irian Jaya), that are neither Austronesian nor Australian. Papuan definitely does not refer to a genetic relationship among these languages for no such relationship can at present be shown.
>
> That is, the term is defined negatively and does not imply a linguistic relationship. While most are spoken on the island of New Guinea, some are found in the Bismarck Archipelago, Bougainville Island and the Solomon Islands to the east, and in Halmahera, Timor and the Alor Archipelago to the west.
>
> There are some 800 Papuan languages divided in a large number of mostly small language families and isolates not demonstrably related to one another.
>
> One large genetic grouping that has been posited for a number of Papuan languages is the Trans-New Guinea phylum, which is promising but not yet confirmed. Greenberg's Indo-Pacific hypothesis of a grand macro-family involving Papuan and other languages has mostly been abandoned.

As the recognized Papuan family Trans-New Guinea in a conservative sense subsumes over half of the 800 Papuan languages (Ross 2005), the notion that the Papuan languages are "divided into a large number of mostly small language families and isolates not demonstrably related to one another" (Kaufman 2007) is not a fair description of the state of these languages.

Trans-New Guinea is not "promising but not yet confirmed" (Kaufman 2007). Instead, it is an uncontroversial language family accepted by the consensus of all specialists (Ross 2005). The proposal originates with Stephen Wurm (Wurm and Laycock 1970; Wurm 1975). Much of the proposal is based on grammatical typology, which is readily borrowed, hence there are problems with the proposal (Foley 1986).

The core of the family is considered established, but its boundaries and membership are still somewhat up in the air because Ross' proposal is mostly based on a single parameter, 1st and 2nd person pronouns (Hammarström 2012). However, Ross' revised proposal (Ross 2005), which cuts out 95 languages from Wurm's original theory, is widely accepted.

CONCLUSION

A major argument in historical linguistics now is between conservatives or splitters or long-rangers or lumpers.

Conservatives say that the comparative method – described as reconstructing a proto-language based on the discovery of a set of regular sound correspondences – is necessary in order to prove that two or more languages are related. However, they also say, not entirely incorrectly, that this method is not useful beyond ~6,000 to 8,000 years (Nichols 1992). Any relationships beyond that

time frame would not be provable by the comparative method and hence could never be proven. This effectively shuts down all long-range research.

A group of long-rangers say that this method is not necessary and instead relationships can be determined by simply looking at the two or more languages through a process called multilateral comparison (formerly "mass comparison"). Multilateral comparison need not be cursory but could mean deep study of languages over 10, 15, or 20 years (Greenberg 1987).

They tend to focus on core vocabulary, numerals, kinship terms, pronouns, and deictics, in addition to small morphological particles – all things that are rarely borrowed. Once they find a number of these items that resemble one another in form and function at a rate greater than chance and rule out any borrowings, they say that the two languages are related because chance and borrowing are ruled out (Greenberg 1987).

The term *resemble* needs an explanation. "Resemble" here refers to words which resemble each other share identical or related meanings and at the various places in their sound strings and sounds which are either identical to can be related to each other through sound changes already identified by the comparative method. Hence, in this way, multilateral comparison is simply an expansion upon the admittedly powerful classical comparative method.

Long-rangers say that this is the way to prove language relatedness, not the comparative method. The comparative method, instead, is used to learn interesting things about language families that have already been discovered via multilateral comparison, such as reconstructing proto-languages and finding regular sound correspondences, including new, undocumented ones that add to the toolkit of long-range comparativists (Greenberg 1987).

Conservatives say that multilateral comparison is not a valid way of proving that languages are related and that only the comparative method can be used to prove this. However, as noted, they set a 6,000 to 8,000-year time limit on the method needed to prove this (Nichols 1992), and this walls off a lot of potential knowledge about ancient and long-range language relationships.

Some long-rangers disagree that the comparative method has a time limit on it and are attempting to use it to retrieve ancient long-range language families and find regular sound correspondences between them. Unfortunately, most of their efforts are in vain as conservatives are using increasingly strict criteria for proof of language relationship and hence are shooting down most if not all of these efforts as being "not done in the proper way."

When long-rangers try to play by the conservatives' rules to find proof, the conservatives keep moving the goalposts and using increasingly strict, petty, quibbling, and often irrational, false, or senseless arguments to say that the relationship is not proven. It feels like a game that is rigged to fail.

Hence, the current trend of extreme conservatism in historical linguistics has set up rules that seem to be designed to prevent the discovery of most if not all new language families, in particular long-range families older than a few millennia.

I am quite certain that long-range language families such as Altaic (with either three families or five), Indo-Uralic, Uralic-Yukaghir, Hokan, Penutian, Mosan, Almosan, Japanese-Korean, Gulf, Yuki-Gulf, Elamo-Dravidian, Austroasiatic-Hmong Mien, Coahuiltecan, North Caucasian,

or Na-Dene will never be accepted by conservatives in my lifetime, and that's not to mention the more extreme proposals such as Eurasiatic, Nostratic, Dene-Caucasian, Austric, and Amerind, although the evidence for the first and last of these is quite powerful.

There are simply too many emotions tied up in any of these proposals, and the conservatives do not seem to want to budge. Further, many linguists have spent a good part of their careers arguing against these proposals. It is doubtful that any amount of evidence will cause them to change their minds. Scientists, like any other humans, don't like to be shown that they're wrong.

Lyle Campbell, Marianne Mithun, Mauricio Mixco, Sarah Grey Thomason, Peter Daniels, Donald Ringe, Alexander Vovin, and Juha Janhunen are among the leaders of this new conservatism. Note that four of these eight conservatives are Americanists. The others include a general linguist, an Indo-Europeanist, a Uralicist, and an Altaicist.

At first, I was angry at what these people were doing, especially the most egregious cases such as Campbell. Then I realized that people lie and misrepresent things all day long, every single day in my life and that this behavior is fairly normal in humans, especially in a mushy area like this one, where hard truths are hard to come by, and most stated facts are more properly matters of opinion or could be construed that way.

The word *lie* is a potent word. I imagine that few if any of the scholars above are consciously lying, that is, saying something that they know to be false. Instead, I'm sure that most if not all of them fully believe in what they are saying. Nevertheless, as we saw above, quite a few things that conservatives are saying are simply objectively false based on a cursory look at the evidence.

I also realized that they are simply defending a scientific paradigm and that unfortunately, this is the rather underhanded and emotion-ridden environment that defending paradigms tends to produce.

Though to be completely honest, I should not be singling these people out because the current conservatism is simply consensus and acts as the current paradigm on the language relatedness question in historical linguistics. The people listed above are at the top of the profession and are often considered the best historical linguists. They write books on historical linguistics. A number are considered to be ultimate authorities on questions of language relatedness. They are simply the leading edge of the current conservative consensus and paradigm in the field.

Of all of them, Campbell seems to be the most extreme conservative. He is also one of the top historical linguists in the world. Mixco, Mithun (Campbell and Mithun 1979), and Poser (Campbell and Poser 2008) are about on the same level as Campbell.

Campbell (Campbell 1988), Mithun (Mithun 1999), Thomason, Poser (Poser 1992), and Mixco (Campbell and Mixco 2007) are Americanists whose conservatism was set off by the publication of Joseph Greenberg's *LIA* in 1987.

They tend to be very angry if not out and out abusive, engaging in bullying, mockery, ridicule, ostracization, and all of the usual techniques used in science against the proposers of a new paradigm.

Long-rangers are heavily disparaged in the field nowadays such that few linguists dare to be pigeonholed as a long-ranger (disparagingly also "lumper") or associated with such. However,

many historical linguists seem to be taking a more moderate fence-sitter stance in which they are neither long-rangers nor conservatives, but instead are interested bystanders potentially open to questions of new language families, including long-range families.

Among the long-range families that the moderates are open to considering nowadays are Indo-Uralic, Dene-Yeniseian, and Austro-Tai. Some of the smaller long-range families in the Americas have supporters even among the most hardline of conservatives.

Thomason took extreme umbrage after Greenberg compared conservatives' objections to his Amerind phylum with Malcolm Guthrie's objections to Greenberg's new classification of Bantu (Greenberg 1963). However, after thinking this over for some time now, I now believe that Greenberg is correct. The conservatives have their minds made up. They are going to allow few if any new families to be discovered**,** although a few of them have caved in a bit.

One motivation for conservatives may be the trajectory of one's career. If one has been arguing forcefully for thirty years that there are no known relatives to your language family, your reputation is going to take a huge hit if you have to agree that you were wrong all those years. Another reason is due to the nature of science itself. We are dealing here with a paradigm. For a good description of a scientific paradigm, see Thomas Kuhn's *The Structure of Scientific Revolutions* (1962). Kuhn holds that science is by its nature very conservative, some sciences being more conservative than others. A paradigm is set up when the field reaches a satisfactory consensus that a particular theory is correct. After a while, serious barriers go up to any challenges to overthrow the proven theory.

The challenges are first ignored, then ridiculed (often severely), then attacked (often ferociously) and then, if the challenge is successful, it is accepted (often slowly and grudgingly). Kuhn pointed out that defenders of the old theory are usually so reluctant to see the paradigm overthrown that we often must wait literally until their deaths to finally overthrow the paradigm. They defend it to their deathbeds. We are dealing with something more than pure empiricism here.

It is quite risky to challenge a paradigm in science. People's careers have suffered from it. During the 1930's, a supporter of Keynesian economics, then challenging the current paradigm in economics, could not get hired at any university in the US.

In the conservatives versus long-rangers debate, we have been in the anger phase for some time now. We seem to be settling out of it, as many are taking a fence-sitting position and arguing for attempts to resolve the debate to make it less heated.

The paradigm here involves extreme skepticism about any new language families to the point that any new families are simply going to be rejected on all sorts of grounds. Paradigms involve politics at the academic level. When a paradigm is set up in science, almost all scientists write and do research within it. Anything outside of the paradigm is derided as pseudoscience or worse.

The problem is that when a paradigm is in effect, all scholars are supposed to publish within the paradigm. Publishing outside the paradigm is regarded as evidence that one is a kook, a crank, is practicing pseudoscience, or that one is crazy or a fool. It is instructive in this debate to note that most of the prominent long-rangers are independent scholars operating outside of academia.

I have had long-rangers tell me that the only reason they can take the long-ranger position is because they are independent and don't have a university job so there are no repercussions if they are wrong. They told me that if they had a professorship, they would not be able to do this work. They have also told me that they know for a fact that certain conservatives might jeopardize their jobs, careers, and especially their funding if they took a long-ranger position. This was given as one of the reasons for their dogmatic conservatism. As an example, prominent long-ranger Merritt Ruhlen was never able to advance beyond Lecturer at Stanford University although he was an excellent scholar (Bengtson 2021).

In addition, science works according to fads, or more properly, standard beliefs. The trends for these beliefs are set by the biggest names in the field. The biggest names in linguistics are all conservatives now. They are the trendsetters, especially in whatever specialty of historical linguistics you are working in. Most of the rest of the field is dutifully following in their footsteps. As an up-and-coming young scholar, you are supposed to follow the proper trends and hypotheses of your field to uphold the consensus of scholars in your area of specialty. There is a lot more than simple empiricism going on here.

Science is no stranger to bias and emotional psychological motivations driving, or usually distorting it. We are human and humans have emotions. Emotion is the enemy of logic. Logic is the basis of empiricism. Hence, emotions are the enemy of science.

Conservatives have come up with a repertoire of reasons to shoot down proposed language relations and most are pretty poor.

They rely on overuse of the borrowing, chance, sound symbolism, nursery words, and onomatopoeia explanations for non-relatedness. There is also an overuse of the comparative method with excessively strict standards being set up for etymologies and sound correspondences. In a number of cases, linguists are going back to the etymologies of their proto-languages and reducing them by up to half.

That a relationship has irregular sound correspondences is a bad reason to throw out an etymology. Keep in mind that 50% of Indo-European etymologies have irregular correspondences (Greenberg 1987). By the logic of Uralicists we should throw out half of IE etymologies then. If Campbell finds any irregular sound correspondences in any new proposal, he automatically rejects it on those grounds alone.

In recent years, prominent long-rangers have warned that the conservatives were not just conservatives determined to stop or slow down progress as is their nature. They pointed out that there was actually a trend towards rejection and going backwards in time to dismantle families and etymologies that have already been set up on the grounds that they were not done perfectly enough. As we can see, this warning was prescient.

There are statements being made by moderates that both sides, the conservatives and the long-rangers, are being equally unreasonable. As one linguist said, the debate is between lazy long-rangers ("Just believe us, don't demand that we prove it!") and angry conservatives ("Not only is this new family false, but all new families proposed from now on will also be shot down!"). He

suggested that they are both wrong and that the solution lies in a point in the middle. I don't have a problem with this moderate centrist belief.

The conservative notion itself rests on an obvious falsehood, that there are hundreds of language families in the world that have no possible relationship with each other.

According to Campbell, there are 160 language families and isolates in the Americas. The question is where did all of these entities come from. Keep in mind, in linguistics, the standard view is that *these 160 entities are not related to each other in any way, shape, or form*. Thinking back, this means that *language would have had to have developed in humans 160 times among the Amerindians alone over the last 20,000 years at most, implying that they had invented a new language from scratch nearly every century.*

The truth is that there was no polygenesis of language.

Although it is possible that more than one human or even pre-human hominid group developed a partial language including phonetic articulation, the last stages of syntactic evolution were probably specific to *Homo sapiens* and gave them an advantage over the other contemporaneous species such as Neandertal, Denisovans, and Floresienses. Some populations may have borrowed grammatical features.

For example, Pygmies have lost their original language and now speak either Nilo-Saharan or Niger-Kordofanian languages clearly borrowed from current or former neighbors. However, the evolution of languages is a different question than the evolution of language itself. However, for the Pygmies to borrow these other languages, they clearly must have evolved the capacity for human language in their brains first as a requirement.

Children need comprehensible input to develop language. No language to hear in the environment, no language for the children to acquire on their own. With cochlear implants, formerly deaf people are now able to hear for the first time. A woman got hers at age 32. Since she missed the critical period for language development, the window of which closes at age 8, she has not, even at this late date, been able to acquire language satisfactorily. She missed the boat. No input, no language.

Obviously, language arose only once among humans. It had to. And hence, all human languages are related to each other de facto whether we can "prove" it by fancy methods or not. Those 160 language families and isolates in the Americas? All related. Now, we may not be able to prove which languages they are related to specifically and most closely, but we know they are all related to each other.

Contrary to popular rumor, everything in science does not have to be "proven" by this or that rigorous method. Many things are simply posited, as no real evidence for their existence exists, either because we were not there or because we can't see them, or in the case of pure physics, we can't even test out our theories. They exist simply because they have to according to our existing theories, and all competing theories fall down flat.

Well, the Americanists beg to disagree. Greenberg's theory was so extreme and radical that the entire field erupted in outrage. None of their alternate theories, not even one of them, make the slightest bit of sense. In order to prove that any of these 160 families in the Americas are related

to each other, we have to use a method called the comparative method whereby proto-languages and families are reconstructed and regular sound correspondences are shown between the languages being studied.

This is the only way that we can prove one language is related to another. That's simply absurd for a few reasons. I partly concur with Johanna Nichols that the comparative method does not really work on language families older than 6,000 to 8,000 years (Nichols 1992). Beyond that time, so many sound changes have taken place, semantics have been so distorted, and so many terms fallen out of use that there's not much of anything left to reconstruct. Furthermore, time has washed away any evidence of sound correspondences. However, long-rangers say that sound correspondences can still be found in language families as old as 9,000 years in the case of the hypothesized Euskaro-Caucasian family (Bengtson 2021).

Although Nichols is a conservative, I have to commend her.

She believes that the comparative method cannot reveal new families beyond 6,000 to 8,000 years. Instead of ending it there and throwing up her hands, she tried to think of new ways we could prove new families outside of the comparative method. The method she suggests is "individual-identifying evidence," which seems to be another way of saying odd morpheme paradigms that were probably not borrowed and are hardly existent outside of that family. This harkens back to Edward Sapir's "submerged features," where he says we can prove the existence of language families by these small morphemic resemblances alone:

> When one passes from a language to another that is only remotely related to it, say from English to Irish or from Haida to Hupa or from Yana to Salinan, one is overwhelmed at first by the great and obvious differences of grammatical structure.
>
> As one probes more deeply, however, significant resemblances are discovered which weigh far more in a genetic sense than the discrepancies that lie on the surface and that so often prove to be merely secondary dialectic developments which yield no very remote historical perspective. In the upshot it may appear, and frequently does appear, that the most important grammatical features of a given language and perhaps the bulk of what is conventionally called its grammar are of little value for the remoter comparison, which may rest largely on submerged features that are of only minor interest to a descriptive analysis. (Sapir 1925: 526)

Antoine Meillet had a similar idea when he spoke of "shared aberrancies" (Campbell and Mixco 2007, p. 195). The classic example is the shared paradigm of English *good*, *better*, *best* and German *gut*, *besser*, *best.*

The rest of the field are sticks in the mud. They say that we must use the comparative method to discover that languages are related because no other method exists. The problem is that as conservatives themselves note, if the comparative method fails beyond 6,000 to 8,000 years back, all attempts to prove language families beyond that time period are bound to fail.

The conservatives seem positively gleeful that according to their paradigm, few if any new language families will be discovered. This delight in nihilism seems odd and disturbing. What sort of science is gleeful that no new knowledge will be found? Even in the event that this is true, it's depressing. Why get excited about something so negative?

Many language families in the world were discovered by Greenberg's multilateral comparison or simply comparing groups of languages to each other, which should be called "comparison."

And in fact, many of the smaller language families in the world are still being posited by the means of multilateral comparison.

It need not be a shallow examination. One could examine a possible language for five, ten, fifteen, or twenty years. For instance, long-ranger John Bengtson has been working on Basque in the framework of Dene-Caucasian for more than thirty years up until the present day (Bengtson 1990, 1994, 1998, 2002, 2004, 2008, 2017; Blažek and Bengtson 1995; Dieterlen and Bengtson 2016). His main focus is Basque, but Sergei Starostin did the main multilateral comparison before him to establish the larger groups. In truth, there is no contradiction between the classic comparative method and multilateral comparison (Bancel 2021).

After studying a group of languages for some time, if one finds a group of core vocabulary items that resemble one another and are above the rate found by chance (7%), and after which borrowing has been ruled out (core vocabulary is rarely borrowed), then you have proof positive of a language family.

I fail to understand why examining a group of languages for a long period of time to find resemblances and try to rule out chance or borrowings is a ridiculous method. What's so ridiculous about that? Sure, it's nice to reconstruct and get nice sound correspondences going, but it's not always necessary, especially in long-range comparisons when such methods are doomed to failure.

There are a number of folks who have bought into the conservatives' arguments and are trying to discover long-range families by the comparative method of reconstructing the proto-language and finding regular sound correspondences between them. A number of them claim to have been successful. There have been attempts to reconstruct proto-languages and find regular sound correspondences with Altaic, Nostratic, Dene-Caucasian, Dene-Yeniseian, Austro-Tai, Totonozoquean, and Uralo-Yukaghir.

Altaic, Nostratic, and Dene-Caucasian all have proto-languages reconstructed with good sound correspondences running through them. Altaic and Nostratic have etymological dictionaries containing many words, 2,300 proto-forms in the case of Altaic in a 1,000-page volume (Starostin *et al.* 2003). Further, a considerable Nostratic proto-language was reconstructed by Aharon Dolgopolsky (1998), Vladislav Illich-Svitych (1965), and Allan Bomhard (2008).

All of these efforts claim that they have proven their hypotheses. However, the conservatives such as Campbell have rejected all of them. So, you see, even when long-rangers follow the mandated method and play it by the book the way they are supposed to, the conservatives nearly always say that the efforts come up short. It's a rigged game.

How about another question? If the comparative method is doomed beyond 6,000 to 8,000 years, why don't we use another method to discover these relationships? The conservative rejoinder is that there is no other method. It's the comparative method or nothing. But how do they know this? Can they prove that other methods can never be used to successfully discover a language relationship?

REFERENCES

Adelaar, Willem F. H. (1989). Review of *Language in the Americas*. *Lingua* 78: 249-255.

——— (1992). Quechuan Languages. In Bright, William (ed.). *Oxford International Encyclopedia of Linguistics* 3, pp. 303-10. Oxford: Oxford University Press.

——— (2000). Propuesta de un nuevo vínculo genético entre dos grupos lingüísticos indígenas de la Amazonía occidental: Harakmbut y Katukina. In Esquerre, Luis Miranda (ed.), *Actas del I Congreso de lenguas indígenas de Sudamérica*, pp. 219-236. Lima, Perú: Universidad Ricardo Palma.

——— (2007). Ensayo de clasificación del katawixí dentro del conjunto harakmbut-katukina. In Andrés Romero-Figueroa, Ana Fernández Garay and Ángel Corbera Mori (eds.), *Lenguas indígenas de América del Sur: Estudios descriptivo-tipológicos y sus contribuciones para la lingüística teórica*, pp. 159-169. Caracas, Venezuela: Universidad Católica Andrés Bello.

——— (2008). Relações externas do macro-jê: O caso do chiquitano. In Telles de Araujo Pereira Lima, Stella Virgínia, and Santos de Paula, Aldir (eds.). *Topicalizando macro-jê*, pp. 9-27. Recife: Nectar.

Adelaar, Willem F. H., and Muysken, Pieter C. (2004). *The Languages of the Andes*. Cambridge Language Surveys. Cambridge University Press.

Adelaar, Willem F. H., and Brijnen, Hélène B. (2014). Johann Natterer's Linguistic Heritage. *Archiv für Völkerkunde* 64: 162-183.

Aikio, Ante (2014a). The Uralic-Yukaghir Lexical Correspondences: Genetic Inheritance, Language Contact, or Chance Resemblance? *Finnisch-Ugrische Forschungen* 62: 7-76.

——— (2014b). On the Reconstruction of Proto-Mari Vocalism. *Journal of Language Relationship (Вопросы языкового родства)* 11: 125-157.

Anbessa, Tefera, and Unseth, Peter (1989). Toward the Classification of Shabo (Mikeyir). In Bender, Lionel M. (ed.). *Topics in Nilo-Saharan Linguistics*. Nilo-Saharan 3, 405-418. Hamburg: Helmut Buske.

Anderson, Gregory D. S. (2017). Munda Languages. In *Oxford Research Encyclopedia of Linguistics*. Oxford: Oxford University Press.

Aoki, Haruo (1963). On Sahaptian-Klamath Linguistic Affiliations. *International Journal of American Linguistics* 29: 107-112.

Babaev, Kirill (2009). Once Again on the Comparison of Personal Pronouns in Proto-Languages. *Journal of Language Relationship* 1, 1: 37-48.

——— (2013). Joseph Greenberg and the Current State of Niger-Congo. *Mother Tongue* 18: 17-22.

Bancel, Pierre (2021). Personal communication.

Bancel, Pierre, and John D Bengtson. (2021). Pronoun roots *n* '1sg' and *m* '2sg' in the Native languages of the Americas and their historical meaning. A follow-up to Raoul Zamponi's "First-person *n* and second-person *m* in Native America: A fresh look". This volume.

Beck, David (1997). Mosan III: A Problem of Remote Common Proximity. International Conference on Salish (and Neighbo(u)ring) Languages (ms.).

Beeler, Madison S. (1961). Northern Costanoan. *International Journal of American Linguistics* 3: 191-197.

——— (1978). Esselen. *Journal of California Anthropology Papers in Linguistics* 1: 3-38.

Bender, M. Lionel (1991). Subclassification of Nilo-Saharan. In Bender, M. Lionel, (ed.). Proceedings of the Fourth Nilo-Saharan Conference, Bayreuth, Aug. 30-Sep. 2, 1989. *Nisa* 7: 1-36. Hamburg: Helmut Buske Verlag.

——— (1996). *The Nilo-Saharan Languages: A Comparative Essay*. Munich: Lincom Europa.
——— (2000). Nilo-Saharan. In *African Languages, an Introduction*, pp. 43-73. Cambridge: Cambridge University Press.
Benedict, Paul K. (1942). Thai, Kadai, and Indonesian: A New Alignment in Southeastern Asia. *American Anthropologist* 44, 4: 576-601.
——— (1975). *Austro-Thai Language and Culture, with a Glossary of Roots*. New Haven: HRAF Press.
Bengtson, John D. (1990). An End to Splendid Isolation: The Macro-Caucasian Phylum. *Mother Tongue Newsletter* 10.
——— (1994.) Edward Sapir and the 'Sino-Dene' Hypothesis. *Anthropological Science* 102, 3: 207-230.
——— (1998). Caucasian and Sino-Tibetan: A Hypothesis of S.A. Starostin. *General Linguistics* 36, 1/2: 33-49.
——— (2002). The Dene-Caucasian Noun Prefix **s-*. In Cavoto, F. (ed.). *Linguist's Linguist: A Collection of Papers in Honour of Alexis Manaster Ramer*, pp. 53-57. Munich: Lincom Europa.
——— (2004). Some Features of Dene-Caucasian Phonology (with Special Reference to Basque). *Cahiers de l'Institut de linguistique de Louvain (CILL)*: 33-54.
——— (2006). A Multilateral Look at Greater Austric. *Mother Tongue* 11: 219-258.
——— (2008). Materials for a Comparative Grammar of the Dene-Caucasian (Sino-Caucasian) Languages. *Aspects of Comparative Linguistics* 3: 45-118.
——— (2017). *Basque and Its Closest Relatives: A New Paradigm*. Cambridge, Mass.: Mother Tongue Press/Association for the Study of Language in Prehistory.
——— (2021). Personal communication.
Berman, Howard (1992). A Comment on the Yurok and Kalapuya Data in Greenberg's *Language in the Americas*. *International Journal of American Linguistics* 58, 2: 230-233.
——— (1996). The Position of Molala in Plateau Penutian. *International Journal of American Linguistics* 62 (1): 1-30.
Bisang, Walter, (1998). Structural Similarities of Clause Combining in Turkic, Mongolian, Manchu-Tungusic, and Japanese: A Typological Alternative to the Hypothesis of a Genetic Relationship. In Johanson, Lars (ed.). *The Mainz Meeting: Proceedings of the Seventh International Conference on Turkish Linguistics*, pp. 199-223. Wiesbaden: Harrassowitz.
Blažek, Václav (undated). *The First Comparative Dictionary of Nilo-Saharan* (ms.).
——— (2019). *Altaic Languages. History of Research, Survey, Classification and a Sketch of Comparative Grammar*. Brno: Masaryk University Press.
Blažek, Václav, and Bengtson, John D. (1995) Lexica Dene-Caucasica. *Central Asiatic Journal* 39, 1: 11-50; 39, 2: 161-164.
Blench, Roger (undated a). Niger-Congo: An Alternative View (ms.).
——— (undated b). Dogon Languages. On his website.
——— (2006). *The Afro-Asiatic Languages: Classification and Reference List* (ms.).
——— (2008). The Prehistory of the Daic (Tai-Kadai) Speaking Peoples. Presented at the 12th EURASEAA Meeting in Leiden, The Netherlands, 1-5 September 2008 (ms.).
——— (2015). Was There a Now-Vanished Branch of Nilo-Saharan on the Dogon Plateau? Evidence from Substrate Vocabulary in Bangime and Dogon (ms.).
——— (2017). Africa over the Last 12,000 Years (ms.).

——— (2018). Tai-Kadai and Austronesian Are Related at Multiple Levels and Their Archaeological Interpretation (ms.).

——— (2019). Shabo and Kadu: Two Orphan Branches of Nilo-Saharan. 14th Nilo-Saharan Linguistics Colloquium. Vienna: University of Vienna Department of African Studies.

Blench, Roger, and Lameen, Souag. (Undated). Saharan and Songhay Form a Branch of Nilo-Saharan (ms.).

Blench, Roger, and Ahland, Colleen (2010). The Classification of Gumuz and Koman Languages. Language Isolates in Africa Workshop. Lyons, France. December 4, 2010.

Blust, Robert (2014). The Higher Phylogeny of Austronesian and the Position of Tai-Kadai: Another Look. In *The 14th International Symposium on Chinese Languages and Linguistics (IsCLL-14)*. Taipei: Academia Sinica, Institute of Linguistics.

Bomhard, Allan R. (2008). *Reconstructing Proto-Nostratic: Comparative Phonology, Morphology, and Vocabulary*. 2 Volumes. Leiden: Brill.

——— (2021). Personal communication.

Bouda, Karl (1940). Die finnisch-ugrisch-samojedische Schicht des Jukagirischen. *Ungarische Jahrbücher* 20: 80-101.

——— (1952). *Die Verwandschaftsverhaltnisse der Tschuktschischen Sprachgruppe*. Universidad de Salamanca, Filosofia y Letras 6, 80.

Bowern, Claire, and Atkinson, Quentin (2012). Computational Phylogenetics and the Internal Structure of Pama-Nyungan: Dataset. *Language* 88, 4: 817-845.

Bowern, Claire, and Koch, Harold (eds.) (2004). *Australian Languages: Classification and the Comparative Method*. Amsterdam: John Benjamins Publishing.

Bright, William (1954). Some Northern Hokan Relationships: A Preliminary Report. In Chretien, Charles D., Madison S. Beeler, Murray B. Emeneau, and Mary R. Haas (eds.). *Papers from the Symposium on American Indian Linguistics*, pp. 63-67. Berkeley: University of California Press.

——— (1956). Glottochronologic Counts of Hokaltecan Materials. *Language* 32, 1: 42-48.

——— (1957). *The Karok Language*. University of California Publications in Linguistics 13. Berkeley: University of California.

Bright, William, and Olmsted, David L. (1959). A Shasta Vocabulary. *Kroeber Anthropological Society Papers* 20: 1-55.

Brinton, Daniel G. (1891). *The American Race: A Linguistic Classification and Ethnographic Description of the Native Tribes of North and South America*. New York: N. D. C. Hodges.

Broadbent, Sylvia M. (1960). A Grammar of Southern Sierra Miwok. Doctoral Dissertation. University of California, Berkeley.

Broadbent, Sylvia, and Pitkin, Harvey (1964). A Comparison of Miwok and Wintun. In Bright, William (ed.). *Studies in Californian Linguistics*, pp. 19-45. UCPL 34.

Brown, Cecil H., and Witkowski, Stanley R. (1979). Aspects of the Phonological History of Mayan-Zoquean. *International Journal of American Linguistics* 45: 34-47.

Brown, Cecil H., David Beck, Grzegorz Kondrak, James K. Watters, and Søren Wichmann (2011). Totozoquean. *International Journal of American Linguistics* 77: 323-372.

Brown, Cecil H., Søren Wichmann, and David Beck (2014). Chitimacha: A Mesoamerican Language in the Lower Mississippi Valley. *International Journal of American Linguistics* 80, 4: 425-474.

Buckley, Gene (1987). Coast Oregon Penutian: A Lexical Comparison. In Delancey, Scott and Tomlin, Russell S. (eds). *Proceedings of the Third Annual Pacific Linguistics Conference*, pp. 43-69. University of Oregon Department of Linguistics.

Callaghan, Catherine A. (1967). Miwok-Costanoan as a Subfamily of Penutian. *International Journal of American Linguistics* 33, 3: 224-227.

Campbell, Lyle (1988). Review of *Language in the Americas*, Greenberg 1987. *Language* 64: 591-615.

——— (1995). The Quechumaran Hypothesis and Lessons for Distant Genetic Comparison. *Diachronica* 7, 2: 157-200.

——— (1996). Coahuiltecan: A Closer Look. *Anthropological Linguistics* 38, 4: 620-634.

——— (1997). *American Indian Languages: The Historical Linguistics of Native America*, pp. 188, 273-283. Oxford: Oxford University Press.

——— (1998). Nostratic: A Personal Assessment, In Salmons, Joseph C., and Joseph, Brian D. (eds.). *Nostratic: Sifting the Evidence*. Current Issues in Linguistic Theory 142. Amsterdam: John Benjamins.

——— (2012). Classification of the Indigenous Languages of South America. In Grondona, Verónica and Campbell, Lyle (eds.). *The Indigenous Languages of South America. The World of Linguistics 2*, pp. 59-166. Berlin: De Gruyter Mouton.

Campbell, Lyle, and Kaufman, Terrence (1976). A Linguistic Look at the Olmecs. *American Antiquity* 41: 80-89.

——— (1980). On Mesoamerican Linguistics. *American Anthropologis*t 82: 850-857.

——— (1983). Mesoamerican Historical Linguistics and Distant Genetic Relationship: Getting It Straight. *American Anthropologist* 85: 362-372.

Campbell, Lyle, and Mixco, Mauricio J. (2007). *A Glossary of Historical Linguistics*. Edinburgh: Edinburgh University Press.

Campbell, Lyle, and Poser, William J. (2008). *Language Classification: History and Method*. Cambridge: Cambridge University Press.

Chafe, Wallace (1987). Review of *Language in the Americas*. *Current Anthropology* 28: 652-653.

Collord, Thomas (1968). Yokuts Grammar: Chukchansi. Ph.D. dissertation. University of California, Berkeley.

Collinder, Björn (1940). Jukagirisch und Uralisch. *Uppsala Universitets Årsskrift* 8.

——— (1957). Uralo-jukagirische Nachlese. *Uppsala Universitets Årsskrift* 12: 105-130.

——— (1965a). Hat das Uralische Verwandte? Eine Sprachvergleichende Untersuchung. *Acta Societatis Linguisticae Upsaliensis* 1: 109-180.

——— (1965b). *An Introduction to the Uralic Languages*. Berkeley and Los Angeles: University of California Press.

Comrie, Bernard (2008). Why the Dene-Yeniseic Hypothesis Is Exciting. Fairbanks and Anchorage, Alaska: Dene-Yeniseic Symposium.

——— (2010). The Dene-Yeniseian Hypothesis: An Introduction. In Kari, James, and Potter, Ben (eds.). *The Dene-Yeniseian Connection*, pp. 25-32. Anthropological Papers of the University of Alaska, New Series 5. Fairbanks: University of Alaska Fairbanks Department of Anthropology.

Constenla Umaña, Adolfo (2005). ¿Existe relación genealógica entre las lenguas misumalpas y las chibchenses? *Estudios de Lingüística Chibcha*. 23: 9-59.

Crawford, James M. (1976). A Comparison of Chimariko and Yuman. In Langdon, Margaret and Silver, Shirley (eds.). *Hokan Studies*, pp. 177-191. The Hague: Mouton.

Crawford, Judith G. (1976). Seri and Yuman. In Langdon, Margaret and Silver, Shirley (eds.). *Hokan Studies*, pp. 305-324. The Hague: Mouton.

Daniels, Peter T. (2004). Review of *Indo-European and Its Closest Relatives: The Eurasiatic Language Family*. *Language* 80, 4: 889-890.

Daniels, Peter. (2016). Personal communication.

Davis, Irvine (1966). Comparative Je Phonology. *Estudos lingüísticos* 1, 2: 10-25.

——— (1968). Some Macro-Je Relationships. *International Journal of American Linguistics* 34: 42-47.

Delancey, Scott (1987). Morphological Parallels between Klamath and Wintu. In Redden, J. (ed.) *Proceedings of the 1987 Hokan-Penutian Conference*, pp. 50-60. Carbondale, Il: Dept. of Linguistics, Southern Illinois University.

——— (1988a). Klamath Stem Structure in Genetic and Areal Perspective. In *Papers from the 1988 Hokan-Penutian Languages Workshop*. University of Oregon Papers in Linguistics 1: 31-39.

——— (1988b). Bipartite Verbs in Languages of Western North America. In Filchenko, Andrei, and Potanina, Olga (eds.). *Time and Space in Languages of Various Typology. Proceedings of the XXVth International Conference Dulson Readings*. Tomsk: Tomsk State Pedagogical University.

Delancey, Scott (1991). Chronological Strata of Suffix Classes in the Klamath Verb. *International Journal of American Linguistics* 57: 426-445.

Delancey, Scott (1996). Penutian in the Bipartite Stem Belt: Disentangling Areal and Genetic Correspondences. In Librik, David, and Beeler, Roxane, *Berkeley Linguistics Society 22: Special Session on Historical Issues in Native American Languages*, 37-54.

Delancey, Scott, C. Genetti, and Noel Rude (1987). Some Sahaptian-Klamath-Tsimshianic Lexical Sets. In Shipley, William (ed.). *In Honor of Mary Haas: From the Haas Festival Conference on Native American Linguistics*, pp. 195-224.

Delancey, Scott, and Golla, Victor (1997). The Penutian Hypothesis: Retrospect and Prospect. *International Journal of American Linguistics* 63: 171-201.

De Angulo, Jaime, and Freeland, L. S. (1930). The Achumawi Language. *International Journal of American Linguistics* 6: 77-120.

de Reuse, Willem J. (2010). Dene-Yeniseian, Phonological Substrata, and Substratic Place Names. In Kari, James, and Potter, Ben (eds.). *The Dene-Yeniseian Connection*. Anthropological Papers of the University of Alaska, New Series 5. Fairbanks: University of Alaska Fairbanks Department of Anthropology.

de Smit, Merlijn (2019). Reconsidering Uralo-Yukaghir Morphology. (draft).

Dieterlen, Florent, and Bengtson, John D. (2016). Confirmation de l'ancienne extension des Basques par l'étude des dialectes de l'Europe de l'Ouest romane. [Confirmation of the former extension of the Basques from a study of the western Romance dialects of Europe.] *Journal of Language Relationship* 14, 1: 1-27.

Diffloth, Gérard. (1974). Austro-Asiatic languages. In *Encyclopedia Britannica*, 480-4. Chicago: Encyclopaedia Britannica.

——— (2005). The Contribution of Linguistic Palaeontology and Austro-Asiatic. In Sagart, Laurent, Roger Blench, and Alicia Sanchez-Mazas, (eds.) *The Peopling of East Asia: Putting Together Archaeology, Linguistics, and Genetics*, pp. 77-80. London: Routledge Curzon.

Diller, Anthony, Jerry Edmondson, and Yongxian Luo (2008). *The Tai-Kadai Languages*. London: Routledge.

Dimmendaal, Gerrit J. (1992). Nilo-Saharan Languages, pp. 100-104. In *International Encyclopedia of Linguistics,* Volume 3. Oxford.

——— (2008). Language Ecology and Linguistic Diversity on the African Continent. *Language and Linguistics Compass* 2, 5: 840-858.

——— (2011). *Historical Linguistics and the Comparative Study of African Languages*. John Benjamins.

——— (2016). On Stable and Unstable Features in Nilo-Saharan. *The University of Nairobi Journal of Language and Linguistics*.

Dimmendaal, Gerrit, Colleen Ahland, and Angelika Jakobi (2019). Linguistic Features and Typologies in Languages Commonly Referred to as 'Nilo-Saharan'. In H. Ekkehard Wolff (ed.), *Cambridge Handbook of African Linguistics*. Cambridge: Cambridge University Press.

Dixon, Roland B. (1905). The Shasta-Achomawi: A New Linguistic Stock, with Four New Dialects. *American Anthropologist* 7: 213-217.

——— (1910). The Chimariko Indians and Language. *University of California Publications in American Archaeology and Ethnology* 5: 293-380.

Dixon, Roland B., and Kroeber, Alfred L. (1903). The Native Languages of California. *American Anthropologist* 5: 1-26.

——— (1913a). Relationship of the Indian Languages of California. *Science* 37: 225.

——— (1913b). New Linguistic Families in California. *American Anthropologist* 15: 647-55.

——— (1919). Linguistic Families of California. *University of California Publications in American Archaeology and Ethnography* 16: 47-118.

Dixon, R. M. W. (1980). *The Languages of Australia. Cambridge Language Surveys*. Cambridge: Cambridge University Press.

——— (2002). *Australian Languages: Their Nature and Development*. Cambridge: Cambridge University Press.

Dolgopolsky, Aharon (1998). *The Nostratic Macrofamily and Linguistic Palaeontology*. Oxford: The McDonald Institute for Archaeological Research.

Dunn, Michael (2012). Review of *The Dene-Yeniseian Connection*. *Language* 88, 2: 429-432.

Echeverri, Juan Alvaro, and Seifart, Frank (2016). Proto-Witotoan: A Re-Evaluation of the Distant Genealogical Relationship between the Boran and Witotoan Linguistic Families.

Ehret, Christopher (1986). Proposals on Khoisan Reconstruction. In Rottland, Franz and Vossen, Rainer, (eds.). *African Hunter-Gatherers International Symposium*, pp. 105-130. *Sprache und Geschichte in Afrika*, Special Issue 7.1. Hamburg: Helmut Buske Verlag.

——— (2001). *A Historical-Comparative Reconstruction of Nilo-Saharan*. Köln: R. Köppe Verlag.

——— (2003. Toward Reconstructing Proto-South Khoisan. *Mother Tongue* 8.

——— (2010). *History and the Testimony of Language*. University of California Press.

Emlen, Nicholas Q. (2017). Perspectives on the Quechua-Aymara Contact Relationship and the Lexicon and Phonology of Pre-Proto-Aymara. *International Journal of American Linguistics* 83, 2: 307-340.

Enrico, John (2004). Toward Proto-Na-Dene. *Anthropological Linguistics* 46, 3: 229-302.

Fleming, Harold C. (1991). Shabo: Presentation of Data and Preliminary Classification. In M. Lionel Bender (ed.). *Proceedings of the Fourth Nilo-Saharan Conference*. Bayreuth. Aug. 30, 2011.

Foley, William A. (1986). *The Papuan Languages of New Guinea*. Cambridge: Cambridge University Press.

Fortson, Benjamin W. (2010). *Indo-European Language and Culture: An Introduction*. Blackwell Textbooks in Linguistics 19. Chichester, U.K. and Malden, MA: Wiley-Blackwell.

Fortescue, Michael D. (1998). *Language Relations across the Bering Strait: Reappraising the Archaeological and Linguistic Evidence*. London and New York: Cassell.

——— (2010). Yeniseian: Siberian Intruder or Remnant? in Kari, James, and Potter, Ben (eds.), *The Dene-Yeniseian Connection*. Anthropological Papers of the University of Alaska, New Series 5. Fairbanks: University of Alaska Fairbanks Department of Anthropology.

——— (2011). The Relationship of Nivkh to Chukotko-Kamchatkan Revisited. *Lingua* 121, 8: 1359-1376.

——— (2017). Correlating Palaeo-Siberian Languages and Populations: Recent Advances in the Uralo-Siberian Hypothesis. ResearchGate (ms.)

Fournet, Arnaud (2021a). Billet d'humeur 10: Critical Innovations and the Structure of the Indo-European Family (ms.).

——— (2021b). Personal communication.

Frachtenberg, Leo J. (1918). Comparative Studies in Takelman, Kalapuyan, and Chinookan Lexicography, a Preliminary Paper. *International Journal of American Linguistics* 1: 175-182.

Freeland, L. S. (1951). *Language of the Sierra Miwok*. Publications in Anthropology and Linguistics Memoir 6. Bloomington, ID: Indiana University Press.

Gallatin, Albert (1836). *A Synopsis of the Indian Tribes within the United States East of the Rocky Mountains and in the British and Russian Possessions in North America*. Zea E-Books in American Studies.

Gamble, Geoffrey (1978). *Wikchamni Grammar*. Berkeley: University of California Press.

Gaudefroy-Demombynes, Maurice. (1907). Document sur les langues de l'Oubangui-Chari. In *Actes du XVI^e Congrès international des orientalistes*, Part II, pp.172-330. Paris: Ernest Leroux.

Guérios, R. F. Mansur F. (1939). O nexo lingüístico bororo/merrime-caiapó (Contribuição para a unidade genética das línguas americanas). *Revista do Círculo de Estudos "Bandeirantes"*. 2: 61-74.

Gell-Mann, Murray, Ilia Peiros, and George Starostin (2009). Distant Language Relationship: The Current Perspective. *Journal of Language Relationship*.

Genographic Consortium; Hui Li, Li Jin, Xingqiu Huang, Shilin Li, Chuanchao Wang, Lanhai Wei, Yan Lu, and Yi Wang (2011). Human Migration through Bottlenecks from Southeast Asia into East Asia during Last Glacial Maximum Revealed by Y Chromosomes. *PLOS ONE* 6 (8), August 31, 2011: e24282.

Georg, Stephan (2013). Review of Salmons, Joseph and Joseph, Brian (eds.), *Nostratic: Sifting the Evidence*, 1997. *Anthropological Linguistics*.

Georg, Stefan, and Vovin, Alexander (2003). From Mass Comparison to Mess Comparison: Greenberg's 'Eurasiatic' Theory. *Diachronica* 20, 2: 331-362.

Gibson, Michael Luke (1996). *El munichi: Un idioma que se extingue*. Serie lingüística peruana 42. Pucallpa: Instituto lingüístico de verano.

Gildea, Spike, and Payne, Doris. (2007). Is Greenberg's 'Macro-Carib' Viable? *Boletim do Museu Paraense Emílio Goeldi – Ciencias Humanas* 2, 2: 19-72.

Goddard, Ives (1979). The Languages of South Texas and the Lower Rio Grande. In Campbell, Lyle, and Mithun, Marianne (eds.). *The Languages of Native America*, pp. 355-389. Austin: University of Texas Press.

——— (1987). Review of *Language in the Americas*. *Current Anthropology* 28: 656-657.

——— (1990). Review of *Language in the Americas* by Joseph H. Greenberg. *Linguistics* 28: 556-558.
——— (1996). The Classification of the Native Languages of North America. In Goddard, Ives, (ed.). Languages. In Sturtevant, W. C. (general editor). *Handbook of North American Indians,* Vol. 17, pp. 290-324. Washington, D. C.: Smithsonian Institution.
Goldschmidt, Walter (1951). Nomlaki Ethnography. *University of California Publications in American Archaeology and Ethnology* 42: 303-443.
Golla, Victor (1980). Some Yokuts-Maidun Comparisons. In Klar, Kathryn, *et al.* (eds.). *American Indian and Indoeuropean Studies: Papers in Honor of Madison S. Beeler*. The Hague: Mouton, pp. 57-65.
——— (1993). Alsea-Wintu Comparisons (ms.).
——— (2007). Linguistic Prehistory. In Jones, Terry L. and Klar, Kathryn. (eds.). *California Prehistory: Colonization, Culture, and Complexity*, pp. 71-82. New York: Altamira Press.
——— (2011). *California Indian Languages*. Berkeley, Los Angeles, London: University of California Press.
Good, Jeff (undated). The Vowel Systems of California Hokan (ms.).
Good, Jeff, Teresa McFarland, and Mary Paster (2003). Reconstructing Achumawi and Atsugewi: Proto-Palaihnihan. SSILA Annual Meeting, Atlanta, January 2-5, 2003.
Grant, Anthony (1997). Coast Oregon Penutian. *International Journal of American Linguistics* 63: 144-156.
Greenberg, Joseph H. (1948). The Classification of African Languages. *American Anthropologist* 50: 24-30.
——— (1949.) Studies in African Linguistic Classification: I. Introduction. Niger-Congo Family. *Southwestern Journal of Anthropology* 5: 79-100.
——— (1949). Studies in African Linguistic Classification: II. The Classification of Fulani. *Southwestern Journal of Anthropology* 5, 3: 190-98.
——— (1949). Studies in African Linguistic Classification: III. The Position of Bantu. *Southwestern Journal of Anthropology* 5, 4: 309-17.
——— (1950a). Studies in African Linguistic Classification: IV. Hamito-Semitic. *Southwestern Journal of Anthropology* 6, 1: 47-63.
——— (1950b). Studies in African Linguistic Classification: V. The Eastern Sudanic Family. *Southwestern Journal of Anthropology* 6, 2: 143-60.
——— (1950c). Studies in African Linguistic Classification: VI. The Click Languages. *Southwestern Journal of Anthropology* 6, 3: 223-37.
——— (1950d). Studies in African Linguistic Classification: VII. Smaller Families; Index of Languages. *Southwestern Journal of Anthropology* 6, 4: 388-98.
——— (1954). Studies in African Linguistic Classification: VIII. Further Remarks on Method; Revisions and Corrections. *Southwestern Journal of Anthropology* 10, 4: 405-15.
——— (1955). *Studies in African Linguistic Classification*. New Haven: Compass Publishing Company.
——— (ed.) (1963a). *Universals of Grammar*. Cambridge, Mass.: MIT Press.
——— (1963b). The Languages of Africa. *International Journal of American Linguistics*, 29, 1, Part 2.
——— (1966a). *The Languages of Africa* (2nd ed. with additions and corrections). Bloomington: Indiana University.

——— (1966b). Language Universals, with Special Reference to Feature Hierarchies. Janua Linguarum, Series Minor 59. The Hague: Mouton.
——— (1970). Some Generalizations Concerning Glottalic Consonants, Especially Implosives. *International Journal of American Linguistics* 36: 123-145.
——— (1974). *Language Typology: An Historical and Analytic Overview*. Janua Linguarum, Series Minor 184. The Hague: Mouton.
——— (1980). Circumfixes and Typological Change. In Traugott, E. C. *et al*. (eds.) *Papers from the 4th International Conference on Historical Linguistics*, pp. 233-241. Amsterdam and Philadelphia: John Benjamins.
——— (1987). *Language in the Americas*. Stanford: Stanford University Press.
——— (2000). *Indo-European and Its Closest Relatives: The Eurasiatic Language Family*. Volume 1: *Grammar*. Stanford: Stanford University Press.
——— (2002). *Indo-European and Its Closest Relatives: The Eurasiatic Language Family*. Volume 2: *Lexicon*. Stanford: Stanford University Press.
Greenberg, Joseph H., Charles A. Ferguson, and Edith A. Moravcsik, (eds.) (1978). *Universals of Human Language*. 4 vols. Stanford: Stanford University Press.
Greenberg, Joseph H., and Swadesh, Morris (1953). Jicaque As a Hokan Language. *International Journal of American Linguistics* 19: 216-22.
Greenberg, Joseph H., and Ruhlen, Merritt (2007). *An Amerind Etymological Dictionary* (2007-09-04). Stanford: Dept. of Anthropological Sciences, Stanford University.
Gregersen, Edgar A. (1972). Kongo-Saharan. *Journal of African Languages* 4: 46-56.
Güldemann, Tom, and Elderkin, Edward D. (2010). On External Genealogical Relationships of the Khoe Family. In Brenzinger, Matthias, and König, Christa (eds.). *Khoisan Languages and Linguistics: The Riezlern Symposium 2003. Quellen zur Khoisan-Forschung 17*. Köln: Rüdiger Köppe.
Güldemann, Tom (2018). Historical Linguistics and Genealogical Language Classification in Africa. In Güldemann, Tom (ed.). *The Languages and Linguistics of Africa*, pp. 299-308. The World of Linguistics Series 11. Berlin: De Gruyter Mouton.
Gursky, Karl-Heinz (1966). On the Historical Position of Waicuri. *International Journal of American Linguistics* 32: 41-45.
——— (1968). Gulf and Hokan-Subtiaban: New Lexical Parallels. *International Journal of American Linguistics* 34: 21-41.
——— (ed.) (1969). Introduction. In Dixon, Roland B. (1924). *Zoque and Xinca Compared with Penutian* (ms.). *Abhandlungen der Volkerkundlichen Arbeitsgemeinschaft* 20. Nortorf. [German translation].
——— (1974a). Der Hoka-Sprachstamm. Eine Bestandsaufnahme des lexikalischen Beweismaterials. *Orbis* 23: 170-215.
——— (1974b). Der Hoka-Sprachstamm. Eine Bestandsaufnahme des lexikalischen Beweismaterials. *Orbis* 23: 207
——— (1988). Der Hoka-Sprachstamm: Nachtrag I *Abhandlungen der Volkerkundlichen Arbeitsgemeinschaft* 58.
——— (1989). Der Hoka-Sprachstamm: Nachtrag II *Abhandlungen der Volkerkundlichen Arbeitsgemeinschaft* 63.
——— (1990). Der Hoka-Sprachstamm: Nachtrag III *Abhandlungen der Volkerkundlichen Arbeitsgemeinschaft* 65.

——— (1995). Some Grammatical Evidence for the Hokan Stock. In: M. Dürr, E. Renner, and W. Oleschinski (eds.). *Language and Culture in Native North America: Studies in Honor of Heinz-Jürgen Pinnow*, pp. 138-153. München/Newcastle: LINCOM.

Haas, Mary R. (1951). The Proto-Gulf Word for Water (with Notes on Siouan-Yuchi). *International Journal of American Linguistics* 17: 71-9.

——— (1952). The Proto-Gulf Word for 'Land' (with Notes on Proto-Siouan). *International Journal of American Linguistics* 18: 238-240.

——— (1958). A New Linguistic Relationship in North America: Algonkian and the Gulf Languages. *Southwestern Journal of Anthropology* 14, 3: 231-264.

——— (1963). Shasta and Proto-Hokan. *Language* 39: 40-59.

——— (1964). California Hokan. In William Bright (ed.). *Studies in Californian Linguistics*, pp. 73-87. Berkeley: University of California Press.

Häkkinen, Jaakko (2012). Early Contacts between Uralic and Yukaghir. *Suomalais-Ugrilaisen Seuran Toimituksia – Mémoires de la Société finno-ougrienne* 264: 91-101.

Heggarty, P. (2013). Ultraconserved Words and Eurasiatic? The Faces in the Fire of Language Prehistory. *Proceedings of the National Academy of Sciences*. 110, 35: E3254.

Hale, Kenneth (1964). Classification of Northern Paman Languages, Cape York Peninsula, Australia: A Research Report. *Oceanic Linguistics* 3, 2: 248-264.

Hale, Kenneth, and Salamanca, Danilo (2001). Theoretical and Universal Implications of Certain Verbal Entries in Dictionaries of the Misumalpan Languages. In Frawley, William, Kenneth C. Hill, and Pamela Munro (eds.) *Making Dictionaries: Preserving Indigenous Languages of the Americas*. University of California Press.

Halpern, A. M. (1946). Yuma. *Linguistic Structures of Native America*. Viking Fund Publications in Anthropology 6: 249-288.

——— (1946). Yuma. *International Journal of American Linguistics* 12, 1: 25-33, 147-151, 204-212.

——— (1946). Yuma. *International Journal of American Linguistics* 13: 18-30, 92-107, 147-166.

Hammarström, Harald (2012). Pronouns and the (Preliminary) Classification of Papuan languages. *Journal of the Linguistic Society of Papua New Guinea*.

Hammarström, Harald, Robert Forke, Martin Haspelmath, and Sebastian Bank (eds.) (2020). Puri-Coroado. *Glottolog*.

Hamp, Eric (2004-2010). Participant in Yeniseic-Tlingit-Athabaskan Familial Proof in Tokyo, Leipzig, and Alaska.

Hamp, Eric (2010). On the First Substantial Trans-Bering Language Comparison. In Kari, James, and Potter, Ben (eds.), *The Dene-Yeniseian Connection*. Anthropological Papers of the University of Alaska, New Series 5. Fairbanks: University of Alaska Fairbanks, Department of Anthropology.

Handel, Zev (2008). What is Sino-Tibetan? Snapshot of a Field and a Language Family in Flux: Sino-Tibetan: a Snapshot. *Language and Linguistics Compass* 2, 3: 422-441.

Harms, Robert (1967). The Ugric Parallels of the Uralo-Yukaghir Focus System. *A Magyar Nyelv Története És Rendszere*: 94-103 [in Hungarian].

——— (1977). The Uralo-Yukaghir Focus System and Problem in Remote Genetic Relationship. In Hopper, Paul J. (ed.). *Studies in Descriptive and Historical Linguistics. Festschrift for Winfred P. Lehmann*, pp. 301-316. Amsterdam: John Benjamins.

Harrington, J. P. (1932). *Karuk Indian Myths*. Bureau of American Ethnology, Bulletin 107. Washington: Smithsonian Institute.

Hashimoto, Mantaro (1986). The Altaicization of Northern Chinese. In McCoy, John and Light, Timothy (eds). *Contributions to Sino-Tibetan Studies*, pp. 76-97. Leiden: Brill.

Haudricourt, André-Georges (1951). Introduction à la phonologie historique des langues miao-yao [An Introduction to the Historical Phonology of the Miao-Yao Languages]. *Bulletin de l'École française d'Extrême-Orient* 44, 2: 555-576.

Haynie, Hannah Jane (2012). *Studies in the History and Geography of California Languages*. PhD Dissertation. University of California, Berkeley.

Hayward, Richard J. (2000). Afroasiatic. In Heine, Bernd and Nurse, Derek (eds.). *African Languages: An Introduction*. Cambridge: Cambridge University Press.

Heine, Bernd (1992). African Languages. In William Bright, (ed.). *International Encyclopedia of Linguistics,* Vol. 1, pp. 31-36. Oxford: Oxford University Press.

Heine, Bernd, and Honken, Henry (2010). The Kx'a Family: A New Khoisan Genealogy. *Journal of Asian and African Studies* 79: 5-36.

Heizer, R. F. (ed.) (1952). California Indian Linguistic Records: The Mission Indian Vocabularies of Alphonse Pinart. *University of California Anthropological Records* 15: 1-84.

Hoijer, Harry (1946). Introduction. In *Linguistic Structures of Native America*. Viking Fund Publications in Anthropology 6: 9-29.

——— (1954). *Some Problems of American Indian Linguistic Research*. University of California Publications in Linguistics 10: 3-12.

Hock, Hans Henrich, and Bashir, Elena (eds.) (2016). *The Languages and Linguistics of South Asia*.

Holm, Hans J. (2008). The Distribution of Data in Word Lists and Its Impact on the Subgrouping of Languages. In Preisach, Christine, Hans Burkhardt, Lars Schmidt-Thieme, and Reinhold Decker (eds.). *Data Analysis, Machine Learning, and Applications*. 31st Annual Conference of the Gesellschaft für Klassifikation 2007, pp. 629-636. Studies in Classification, Data Analysis, and Knowledge Organization. Freiburg im Breisgau, Germany: Springer.

——— (2010). A Possible Homeland of the Indo-European Languages and Their Migrations after the Extended Separation Level Recovery Method (Separation Level Recovery under Two Distributions, SLR2D). Webpage.

Hombert, Jean-Marie, and Philippson, Gérard. (2009). 20. The Linguistic Importance of Language Isolates: The African Case. In Austin, Peter K., Oliver Bond, Monik Charette, David Nathan and Peter Sells (eds.). *Proceedings of the Conference on Language Documentation and Linguistic Theory* 2. London: SOAS.

Honken, Henry (1988). Phonetic Correspondences among Khoisan Affricates. In Vossen, Rainer (ed.). *New Perspectives on the Study of Khoisan*, pp. 47-65. Quellen Zur Khoisan-Forschung 7. Hamburg: Helmut Buske Verlag.

——— (1998). Types of Sound Correspondence Patterns in Khoisan Languages. In Schladt, Mathias, (ed.). *Language, Identity, and Conceptualization among the Khoisan*, pp. 171-193. Quellen Zur Khoisan-Forschung/Research in Khoisan Studies 15. Cologne: Rüdiger Köppe Verlag.

Hunn, Eugene S. (2000). Chapter 4: Review of Linguistic Information. In McManamon, Francis P., Jason C. Roberts, and Brooke S. Blades (eds.). *Kennewick Man Cultural Affiliation Report*. Washington DC: National Park Service, US Department of the Interior, Archeology Division.

Hymes, Dell H. (1956-65). On the Relationship of Tsimshian and Chinookan. Unpublished Paper.

——— (1956). Na-Déné and Positional Analysis of Categories. *American Anthropologist* 58, 4: 624-628.

——— (1957). Some Penutian Elements and the Penutian Hypothesis. *Southwestern Journal of Anthropology* 13: 67-87.

——— (1969). On Cayuse and Molale as Penutian. Unpublished Paper.

——— (1995). Na-Dene Ethnopoetics: A Preliminary Report: Haida and Tlingit. In Renner, Egon and Dürr, Michael (eds.). *Language and Culture in North America: Studies in Honor of Heinz-Jürgen Pinnow*, pp. 265-311. Munich: Lincom Europa.

Illich-Svitych, Vladislav M. (1965). Материалы к сравнительному словарю ностратических языков (индоевропейский, алтайский, уральский, дравидийский, картвельский, семито-хамитский). (Materials for a Comparative Dictionary of the Nostratic Languages (Indo-European, Altaic, Uralic, Dravidian, Kartvelian, Semito-Hamitic). Moscow: Etimologija.

Ives, John W. (2010). Dene-Yeniseian, Migration, and Prehistory. In Kari, James, and Potter, Ben (eds.). *The Dene-Yeniseian Connection*. Anthropological Papers of the University of Alaska, New Series 5. Fairbanks: University of Alaska Fairbanks Department of Anthropology.

Ives, John W., Sally Rice, and Edward Vajda (2010). Dene-Yeniseian and Processes of Deep Change in Kin Terminologies. In Kari, James, and Potter, Ben (eds.). *The Dene-Yeniseian Connection*. Anthropological Papers of the University of Alaska, New Series 5. Fairbanks: University of Alaska Fairbanks Department of Anthropology.

Jacobsen, William H., Jr. (1954, 1955, and 1958). Salinan Field Notes. Unpublished manuscript in possession of the author.

——— (1958). Washo and Karok: An Approach to Comparative Hokan. *International Journal of American Linguistics* 24: 195-212.

——— (1964). The Washo language. Berkeley: University of California at Berkeley Department of Linguistics. Ph.D. dissertation.

——— (1966). Washo Linguistics Studies. In D'Azevedo, Warren L., *The Current Status of Anthropological Research in the Great Basin,* Desert Research Institute Technical Report Series S-H. Social Sciences and Humanities Publications 1, 113-136. Reno, Nevada: Desert Research Institute.

——— (1979). Hokan Inter-branch Comparisons. In Campbell, Lyle, and Mithun, Marianne (eds.). *The Languages of Native America: Historical and Comparative Assessment*, pp. 545-591. Austin: University of Texas Press.

——— (1986). Washo Linguistic Prehistory. In James Redden (ed.). *Papers from the 1983, 1984, and 1985 Hokan-Penutian Languages Conferences*. Occasional Papers on Linguistics 13: 33-58. Carbondale, Illinois. Department of Linguistics, Southern Illinois University.

Jany, Carmen (2014). Chimariko Grammar: Areal and Typological Perspective. University of California Publications in Linguistics 142. Berkeley: University of California Press.

Jasanoff, Jay H. (2003). *Hittite and the Indo-European Verb*. Oxford: Oxford University Press.

Jolkesky, Marcelo Pinho de Valhery. (2011). Arawá-katukína-harakmbet: correspondências fonológicas, morfológicas e lexicais. Paper presented at the Encontro internacional: arqueologia e linguística histórica das línguas indígenas sul-americanas, Brasília. October 24-28, 2011.

——— (2015). Semejanzas léxicas entre el páez, el andakí y el tinígua. (ms.)

——— (2016). *Estudo arqueo-ecolinguístico das terras tropicais sul-americanas*, 2nd edition). Ph.D. dissertation. Brasília: Universidade de Brasília.

Kallio, Petri, and Koivulehto, Jorma (2017). Beyond Proto-Indo-European. In Klein, Jared; Brian Joseph and Matthias Fritz (eds.) *Handbook of Comparative and Historical Indo-European Linguistics* 3, pp. 2280-2291. Walter de Gruyter.

Kari, James, and Potter, Ben A. (2010). The Dene-Yeniseian Connection, Bridging Asia and North America. In Kari, James, and Potter, Ben (eds.). *The Dene-Yeniseian Connection*, pp. 1-24. Anthropological Papers of the University of Alaska, New Series 5. Fairbanks: University of Alaska Fairbanks Department of Anthropology.

Kari, James (2010). The Concept of Geolinguistic Conservatism in Na-Dene Prehistory. In Kari, James, and Potter, Ben (eds.). *The Dene-Yeniseian Connection*. Anthropological Papers of the University of Alaska, New Series 5. Fairbanks: University of Alaska Fairbanks Department of Anthropology.

Kasak, Ryan (2016). A Distant Genetic Relationship between Siouan-Catawban and Yuchi.

Kassian, Alexei, Michael Zhivlov, and Georgiy Starostin. (2013). Proto-Indo-European-Uralic Comparison from the Probabilistic Point of View. *Journal of Indo-European Studies* 43: 301-347.

Kaufman, Terrence (1988). A Research Program for Reconstructing Proto-Hokan: First Groupings. In *Papers from the 1988 Hokan-Penutian Languages Workshop*. University of Oregon Papers in Linguistics: Publications of the Center for Amerindian Linguistics and Ethnography 1: 50-168. Eugene, Oregon: University of Oregon Department of Linguistics.

——— (1990a). Tlapaneko-Subtiaba, OtoManguean, and Hokan: Where Greenberg Went Wrong (ms.).

——— (1990b). Language History in South America: What We Know and How to Know More. In Payne, Doris L. (ed.). *Amazonian Linguistics*, pp. 13-74. Austin: University of Texas Press.

——— (1994). The Native Languages of South America. In Moseley, Christopher, and Asher, Ronald E. (eds.). *Atlas of the World's Languages*, pp. 46-76. London: Routledge.

——— (2007). South America. In: Asher, Ronald E., and Moseley, Christopher (eds.). *Atlas of the World's Languages* (2nd ed.), pp. 59-94. London: Routledge.

——— (2015). Some Hypotheses Regarding Proto-Hokan Grammar. Project for the Documentation of the Languages of Mesoamerica.

Kaye, Alan S., and Daniels, Peter T. (1992). Comparative Afroasiatic and General Genetic Linguistics. *Word* 43, 3: 429-458.

Key, Mary Ritchie (1978a). The History and Distribution of the Indigenous Languages of Bolivia. Paper presented at the Annual Meeting of the American Anthropological Society. Los Angeles, California.

——— (1978b). Araucanian Genetic Relationships. *International Journal of American Linguistics* 44, 4: 280-293.

Kimball, Geoffrey (1992). A Critique of Muskogean, 'Gulf,' and Yukian Materials in *Language in the Americas*. *International Journal of American Linguistics* 58: 447-501.

——— (1994). Comparative Difficulties of the 'Gulf' Languages. In Langdon, Margaret (ed.). *Proceedings of the Meeting of the Society for the Study of the Indigenous Languages of the Americas July 2-4, 1993, and the Hokan-Penutian Workshop July 3, 1993*, Ohio State University, Columbus, Ohio). Survey of California and Other Indian Languages Report 8. Berkeley: University of California Press.

Kiparsky, Paul (2015). New Perspectives in Historical Linguistics. In Bowern, Claire, and Evans, B. (eds.). *The Routledge Handbook of Historical Linguistics*, pp. 64-102. London and New York: Routledge.

Kortlandt, Frederik (1989). Eight Indo-Uralic Verbs? *Münchener Studien zur Sprachwissenschaft* 50: 79-85.

——— (2002). The Indo-Uralic Verb. In *Finno-Ugrians and Indo-Europeans: Linguistic and Literary Contacts: Proceedings of the Symposium at the University of Groningen*, Groningen, The Netherlands, November 22-24, 2001, 217-227. Maastricht: Shaker Publishing.

Krauss, Michael E. (1979). Na-Dene and Eskimo-Aleut. In Campbell, Lyle and Mithun, Marianne (eds.). *The Languages of Native America: Historical and Comparative Assessment*, pp. 803-901. Austin: University of Texas Press.

Krejnovich, Erukhim (1978). On Some Uralo-Yukaghir Language Parallels. *Sovetskoe Finnougrovedenie* 4: 241-249 [in Russian].

Kroeber, Alfred L. (1906). The Dialectic Divisions of the Moquelumnan Family in Relation to the Internal Differentiation of the Other Linguistic Families of California. *American Anthropologist* 8: 652-663.

——— (1904). Languages of the Coast of California South of San Francisco. U*niversity of California Publications in American Archaeology and Ethnology* 2: 2.

——— (1910). The Chumash and Costanoan Languages. *University of California Publications in American Archaeology and Ethnology* 9: 259-263.

——— (1915). Serian, Tequistlatecan, and Hokan. *University of California Publications in American Archaeology and Ethnography* 11: 279-290.

——— (1943). Classification of the Yuman Languages. *University of California Publications in Linguistics* 1: 21-40.

Kuhn, Thomas S. (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Langdon, Margaret (1974). *Comparative Hokan-Coahuiltecan Studies: A Survey and Appraisal*. Janua Linguarum Series Critica 4. The Hague: Mouton.

——— (1990). Morphosyntax and Problems of Reconstruction in Yuman and Hokan. In Baldi, Philip (ed.). *Linguistic Change and Reconstruction Methodology*, pp. 57-72. Berlin: Mouton de Gruyter.

Lecarme, Jacqueline (2003). *Research in Afroasiatic Grammar* 2. Amsterdam: John Benjamins Publishing.

Leer, Jeff (2010). The Palatal Series in Athabascan-Eyak-Tlingit, with an Overview of the Basic Sound Correspondences. In Kari, James, and Potter, Ben (eds.). *The Dene-Yeniseian Connection*. Anthropological Papers of the University of Alaska, New Series 5. Fairbanks: University of Alaska Fairbanks Department of Anthropology.

Leer, Jeff, James Kari, and Ben Potter (eds.) (2010). *The Dene-Yeniseian Connection*, Anthropological Papers of the University of Alaska, New Series 5: 33-99, 168-193. Fairbanks: University of Alaska Fairbanks Department of Anthropology.

Lehmann, Walter (1920). *Zentral-Amerika*. Berlin: Reimer.

Levine, Robert D. (1979). Haida and Na-Dene: A New Look at the Evidence. *International Journal of American Linguistics* 45: 157-70.

Levy, Richard S. (1976). *Costanoan Internal Relationships*. Berkeley: Berkeley Archaeological Research Facility, University of California, Berkeley.

Lewy, Ernst (1928). Possessivisch und Passivisch. Bemerkungen zum Verbalausdruck in der Sprachlichen Typenlehre. *Ungarische Jahrbücher* 8: 274-289.

Li, Hui (李辉) (2005). Genetic Structure of Austro-Tai Populations. Doctoral Dissertation. Fudan, China: Fudan University.

Liedtke, Stefan (1995). *Wakashan, Salishan, and Penutian and Wider Connections Cognate Sets*. Linguistic Data on Diskette Series 9. Munich: Lincom Europa.

——— (2007). The Relationship of Wintuan to Plateau Penutian. *Lincom Studies in Native American Linguistics* 55. Munich: Lincom Europa.

Lindsay, Robert (2021). Altaic Consensus among Specialists. (draft). Academia.edu website.

Logan, Jennifer L. (2001). Reassessing Cultural Extinction: Change and Survival at Mission San Juan Capistrano, Texas. College Station: Texas A&M University: Center for Ecological Archaeology.

Loos, Eugene (1973). *Algunas implicaciones de la reconstrucción de un fragmento de la gramática del proto-pano. Estudios panos II*. Serie lingüística peruana 11, pp. 263-282. Yarinacocha: Instituto lingüístico de verano.

——— (1999). Pano. In *The Amazonian Languages*. Dixon, R. M. W., and Aikhenvald, Alexandra Y., (eds.), pp. 227-49. Cambridge: Cambridge University Press.

Lukas, Johannes, and Völckers, Otto (1938). G. Nachtigals Aufzeichnungen über die Sprache der Mimi in Wadai. *Zeitschrift für Eingeborenensprachen* 29, pp. 145-154.

Loukotka, Čestmír (1968). *Classification of South American Indian Languages*. Los Angeles: UCLA Latin American Center.

Luthin, Herbert W. (ed.) *Surviving Through the Days: Translations of Native California Stories and Songs*. Berkeley: University of California Press.

Manaster Ramer, Alexis (1996a). Sapir's Classifications: Haida and the Other Na Dene Languages. *Anthropological Linguistics* 38: 179-215.

——— (1996b). Sapir's Classifications: Coahuiltecan. *Anthropological Linguistics* 38, 1: 1-38.

Manaster-Ramer, Alexis (1996c). Tonkawa and Zuni: Two Test Cases for the Greenberg Classification. *International Journal of American Linguistics* 62: 264-288.

Marlett, Stephen A. (2008). The Seri and Salinan Connection Revisited. *International Journal of American Linguistics* 74, 3: 393-99.

Martins, Andérbio Márcio Silva. (2011). Uma avaliação da hipótese de relações genéticas entre o guató e o tronco macro-jê. PhD dissertation. Universidade de Brasília.

Mason, J. Alden (1916). The Mutsun Dialect of Costanoan Based on the Vocabulary of de la Cuesta. *University of California Publications in American Archaeology and Ethnology* 11: 399-472.

——— (1918). The Language of the Salinan Indians. *University of California Publications in American Archaeology and Ethnology* 14: 1-154.

Matisoff, James (1990). On Megalo-Comparison: A Discussion Note. *Language* 66: 106-120.

Matthews, Peter H. (2007). *Oxford Concise Dictionary of Linguistics* (2nd ed.). Oxford: Oxford University Press.

McLendon, Sally (1964). Northern Hokan (B) and (C): A Comparison of Eastern Pomo and Yana. In William Bright (ed.). *Studies in Californian Linguistics*, pp. 126-144. Berkeley: University of California Press.

McQuown, Norman A. (1942). Una posible síntesis lingüística macro-mayance, mayas, y olmecas. Sociedad mexicana de antropología, Reunión de mesa redonda sobre problemas antropológicos de México y Centro América 2: 37-38.

——— (1956). Evidence for a Synthetic Trend in Totonacan. *Language* 32: 78-80.

Meader, Robert E. (1978). Índios do Nordeste: Levantamento sobre os remanescentes tribais do Nordeste brasileiro. Brasília: SIL International.

Melchert, H. Craig (2012). The Position of Anatolian (ms.).

Migliazza, Ernesto (1972). The Languages of South America. (ms.).

Mikkola, Pertti (1999). Nilo-Saharan Revisited: Some Observations Concerning the Best Etymologies. *Nordic Journal of African Studies* 8, 2: 108-138.

Mithun, Marianne (1999). The Languages of Native North America. Cambridge: Cambridge University Press.

——— (2007). Grammar, Contact, and Time. *Journal of Language Contact* THEMA 1.

——— (2012). Core Argument Patterns and Deep Genetic Relations: Hierarchical Systems in Northern California. In Suihkonen, Pirkko, Bernard Comrie, and Valery Solovyev (eds.). *Argument Structure and Grammatical Relations. A Crosslinguistic Typology*, pp. 257-294. Amsterdam: John Benjamins Publishing Company.

Mixco, Mauricio J. (1976). Kiliwa Texts. *International Journal of American Linguistics Native American Text Series* 1: 92-101.

——— (1977). The Linguistic Affiliation of the Ñakipa and Yakakwal of Lower California. *International Journal of American Linguistics* 43: 189-200.

Moshinsky, Julius (1964). Southeastern Pomo Grammar. Ph.D. dissertation. University of California, Berkeley.

Moser, Edward, and Mary B. (1961). Vocabulario Seri. Mexico, D. F.

Mous, Maarten (2012). Cushitic. In Frayzingier, Zygmunt and Shay, Erin (eds.). *The Afroasiatic Languages*, pp. 342-422. Cambridge: Cambridge University Press.

Müller, André, Viveka Velupillai, Søren Wichmann, Cecil H. Brown, Eric W. Holman, Sebastian Sauppe, Pamela Brown, Harald Hammarström, Oleg Belyaev, Johann-Mattis List, Dik Bakker, Dmitri Egorov, Matthias Urban, Robert Mailhammer, Matthew S. Dryer, Evgenia Korovina, David Beck, Helen Geyer, Pattie Epps, Anthony Grant, and Pilar Valenzuela. ASJP World Language Trees of Lexical Similarity: Version 4 Website. (October 2013).

Munro, Pamela (1995). Gulf and Yuki-Gulf. *Anthropological Linguistics* 36: 125-222.

Nevin, Bruce (2019). Why Proto-Palaihnihan is Neither. (ms.)

Nevskaya, Irina (2010). Inclusive and Exclusive in Altaic Languages. In Johanson, Lars and Robbeets, Martine (eds). *Transeurasian Verbal Morphology in a Comparative Perspective: Genealogy, Contact, Chance*, pp. 115-128. Turcologica Vol. 78. Wiesbaden: Harrassowitz.

Newman, Paul (1977). Chadic Classification and Reconstructions. *Afroasiatic Linguistics* 5, 1: 1-42.

——— (1980). The Classification of Chadic within Afroasiatic. Leiden: Universitaire Pers Leiden.

Nichols, Johanna (1992). Linguistic Diversity in Space and Time. Chicago: University of Chicago Press.

——— (2012). Selection for M : T Pronominals in Eurasia. In Johanson, Lars and Robbeets, Martine (eds). *Copies Versus Cognates in Bound Morphology*, pp. 47-70. Brill's Studies in Language, Cognition, and Culture, Vol. 2. Leiden: Brill.

Nikolaev, Sergei L. (2015a). Toward the Reconstruction of Proto-Algonquian-Wakashan. Part 1: Proof of the Algonquian-Wakashan Relationship. *Journal of Language Relationship* 13, 1: 23-61.

——— (2015b). Toward the Reconstruction of Proto-Algonquian-Wakashan. Part 2: Algonquian-Wakashan Sound Correspondences. *Journal of Language Relationship* 13, 4: 289-328.

——— (2017). Toward the Reconstruction of Proto-Algonquian-Wakashan. Part 3: The Algonquian-Wakashan 110-item Wordlist. *Journal of Language Relationship* 15, 4: 250-278.

——— (2021). Personal communication.

Nikolaeva, Irina (1988a). Проблема урало-юкагирских генетических связей (The Problem of Uralo-Yukaghir Genetic Relationship). Moscow: Institute of Linguistics. PhD dissertation. [in Russian].

——— (1988b). On the Correspondence of Uralic Sibilants and Affricates in Yukaghir, *Sovetskoe Finnougrovedenie* 2: 81-89 [in Russian].

——— (2006). A Historical Dictionary of Yukaghir. Berlin/New York: Mouton De Gruyter.

Nikulin, Andrei (2020). *Proto-Macro-Jê: Um Estudo Reconstrutivo*. Ph.D. dissertation. Brasília: Universidade de Brasília.

Norquest, Peter K. (2013). A revised inventory of Proto Austronesian consonants: Kra-Dai and Austroasiatic evidence. *Mon-Khmer Studies* 42: 102-126.

——— (2015). *A Phonological Reconstruction of Proto-Hlai*. Languages of Asia 13. Leiden: Brill.

O'Connor, Mary Catherine (1987). *Topics in Northern Pomo Grammar*. Ph.D. dissertation. University of California, Berkeley.

O'Grady, Geoffrey, and Hale, Kenneth (2004). The Coherence and Distinctiveness of the Pama-Nyungan Language Family within the Australian Linguistic Phylum. In Bowern, Claire and Koch, Harold (eds.). *Australian Languages: Classification and the Comparative Method*, p. 69. Amsterdam: John Benjamins Publishing.

Okrand, Marc (1977). Mutsun Grammar. Ph.D. dissertation. University of California Press, Berkeley.

Olawsky, Knut J. (2007). Introduction. In *A Grammar of Urarina*. Berlin, New York: De Gruyter Mouton.

Olmsted, David L. (1956). Palaihnihan and Shasta I: Labial Stops. *Language* 32, 1: 73-77.

——— (1957). Palaihnihan and Shasta II: Apical Stops. *Language* 33, 2: 136-138.

——— (1959). Palaihnihan and Shasta III: Dorsal Stops. *Language* 35, 4: 637-644.

Olson, Ronald D. (1964). Mayan Affinities with Chipaya of Bolivia I: Correspondences. *International Journal of American Linguistics* 30, 4: 313-324.

Ostapirat, Weera (2005). Kra-Dai and Austronesian: Notes on Phonological Correspondences and Vocabulary Distribution. In Sagart, Laurent, Roger Blench and Alicia Sanchez-Mazas (eds.). *The Peopling of East Asia: Putting Together Archaeology, Linguistics, and Genetics*, pp. 107-131. London: Routledge Curzon.

——— (2013). Austro-Tai Revisited. Paper Presented at the 23rd Annual Meeting of the Southeast Asian Linguistics Society, 29-31 May 2013, Chulalongkorn University, Thailand.

Oswalt, Robert L. (1976). Switch-Reference in Maiduan: An Areal and Typological Contribution. *International Journal of American Linguistics* 42: 297-304.

Paasonen, Heikki (1907). Zur Frage von der Urverwandtschaft der Finnisch-Ugrischen und Indoeuropäischen Sprachen ('On the Question of the Original Relationship of the Finnish-Ugric and Indo-European languages'). *Finnisch-Ugrische Forschungen* 7: 13-31.

Pache, Matthias J. (2016). Pumé (Yaruro) and Chocoan: Evidence for a New Genealogical Link in Northern South America. *Language Dynamics and Change* 6: 99-155.

——— (2018). Contributions to Chibchan Historical Linguistics Ph.D. dissertation. Leiden: Universiteit Leiden.

Pagel, Mark, Quentin D. Atkinson, Andreea S. Calude, and Andrew Meade (2013). Ultraconserved Words Point to Deep Language Ancestry across Eurasia. *Proceedings of the National Academy of Sciences* 110, 21 : 8471-8476.

Pedersen, Holger (1931). Linguistic Science in the Nineteenth Century: Methods and Results, Translated by John Webster Spargo. Cambridge, Massachusetts: Harvard University Press.

——— (1933). Zur Frage nach der Urverwandtschaft des Indoeuropäischen mit dem Ugrofinnischen. *Mémoires de la Société finno-ougrienne* 67: 308-325.

Pedersen, Holger (1938). Hetitisch und die anderen indo-europäischen Sprachen. Copenhagen: Munksgaard.

Pierce, Joe E. (1966). Genetic Comparisons and Hanis, Miluk, Alsea, Siuslaw, and Takelma. *International Journal of American Linguistics* 32: 379-387.

Pinnow, Heinz-Jürgen (1962). Two Problems of the Historical Phonology of Na-Dene Languages. *International Journal of American Linguistics* 28: 162-166.

——— (1966). Grundzüge einer historischen Lautlehre des Tlingit. Wiesbaden: Otto Harrassowitz (in German).

——— (1968), Genetic Relationships versus Borrowing in Na-Dene. *International Journal of American Linguistics* 34, 3: 194-203.

——— (1970). Notes on the Classifiers in the Na-Dene Languages. *International Journal of American Linguistics* 36, 1: 63-67.

——— (1976). *Geschichte der Na-Dene-Forschung*. Indiana Beihefte 5. Berlin: Mann.

——— (1985). Das Haida als Na-Dene Sprache. *Abhandlungen der Völkerkundlichen Arbeitsgemeinschaft*, 43-46. Nortorf, Germany: Völkerkundliche Arbeitsgemeinschaft.

——— (2006a). *Die Na-Dene-Sprachen im Lichte der Greenberg-Klassifikation* [The Na-Déné Languages in Light of Greenberg's Classification], 2nd revised ed. Bredstedt: Lempfert.

——— (2006b). *Sprachhistorische Untersuchung zur Stellung des Haida als Na-Dene-Sprache*. Unveränderte Neuausgabe aus Indiana 10, Gedenkschrift Gerdt Kutscher. Teil 2, Berlin 1985. Mit einem Anhang: Die Na-Dene-Sprachen im Verhältnis zum Tibeto-Chinesischen. Bredstedt: Lempfert.

Piispanen, Peter S. (2013). The Uralic-Yukaghiric Connection Revisited: Sound Correspondences of Geminate Clusters. *Journal de la Société finno-ougrienne* 94: 165-197.

——— (2021). Personal communication.

Pitkin, Harvey (1955). *Wintu Grammar*. Ph.D. dissertation. Berkeley: University of California.

Price, David P. 1978. The Nambiquara Linguistic Family. *Anthropological Linguistics* 20, 1: 14-37.

Pokorny, Julius (1957/1969 1st ed.), (5th ed. 2005). *Indogermanisches etymologisches Wörterbuch*, 2 vols. Tübingen, Bern, Munich: A. Francke.

Pooth, Roland (2021). Personal Communication.

Poser, William J. (1992). The Salinan and Yurumanguí Data in *Language in the Americas*. *International Journal of American Linguistics* 58, 2: 202-229.

——— (1995). Binary Comparison and the History of Hokan Comparative Studies. *International Journal of American Linguistics* 61, 1: 135-144.

——— (2008). The Languages of the Caucasus. *Language Log* website. August 25, 2008.

Powell, John W. (1891). Indian Linguistic Families of America North of Mexico. Seventh Annual Report, Bureau of American Ethnology, pp. 1-142. Washington, D.C.: Government Printing Office. Reprinted in P. Holder (Ed.), 1966, Introduction to Handbook of American Indian languages by Franz Boas and Indian linguistic families of America, north of Mexico, by J. W. Powell. Lincoln, Nebraska: University of Nebraska.

——— (1915). Linguistic Families of American Indians North of Mexico. Revised by Members of the Staff of the Bureau of American Ethnology with Map. Washington, D.C.: Government Printing Office.

Radin, Paul (1944). The Classification of the Languages of Mexico, *Tlalocan* 1, 3: 259-265.

Ramstedt, Gustaf J. (1924). A Comparison of the Altaic Languages with Japanese. *Transactions of the Asiatic Society of Japan Second Series* 7: 41-54.

Ramirez, Henri, Valdir Vegini and Maria Cristina Victorino de França (2015). Koropó, puri, kamakã e outras línguas do Leste brasileiro. LIAMES: Línguas indígenas americanas 15, 2: 223-277.

Rankin, Robert (1992). Review of *Language in the Americas* by J. H. Greenberg. *International Journal of American Linguistics* 58, 3: 324-351.

Ratliff, Martha (2010). *Hmong-Mien Language History*. The Australian National University, College of Asia and the Pacific, Pacific Linguistics School of Culture, History and Language.

——— (undated). Grammatical Features Shared by Hmong-Mien and Austronesian. Powerpoint Presentation. Detroit: Wayne State University.

Rédei, Károly (1990). Zu den uralisch-jukagirischen Sprachkontakten. In Elek Bartha, László Keresztes, Sándor Maticsák, and Zoltán Ujváry (eds.), *Congressus septimus internationalis fenno-ugristorum*. *Pars 1 A. Sessiones Plenares*. Debrecen.

——— (1999). Zu den uralisch-jukagirischen Sprachkontakten. *Finnisch-Ugrische Forschungen* 55: 1-58.

Reid, Lawrence A. (2006). Austro-Tai Hypotheses. In Brown, Keith (ed.). *The Encyclopedia of Language and Linguistics*, 2nd ed., pp. 609-610. Amsterdam: Elsevier Science.

Repetti-Ludlow, Chiara, Haoru Zhang, Scott AnderBois, Hugo Lucitante, and Chelsea Sanker (December 2020). A'ingae (Cofán). *Journal of the International Phonetic Association* 50, 3: 431-444.

Ribeiro, Eduardo Rivail (2012). *A Grammar of Karajá*. Ph.D. dissertation. Chicago: University of Chicago.

Ribeiro, Eduardo Rivail, and van der Voort, Heinz (2010). Nimuendajú Was Right: The Inclusion of the Jabutí Language Family in the Macro-Jê stock. *International Journal of American Linguistics* 76, 4: 517-70.

Rice, Keren (2011). Dene-Yeniseian. *Diachronica* 28, 2: 255-271.

Rigsby, Bruce J. (1965). Continuity and Change in Sahaptian Vowel Systems. *International Journal of American Linguistics* 31: 306-311.

——— (1966). On Cayuse-Molala Relatability. *International Journal of American Linguistics* 32, 4: 369-378.

——— (1969). The Waiilatpuan Problem: More on Cayuse-Molala Relatability. *Northwest Anthropological Research Notes* 3: 68-146.

Ringe, Donald (2000). Some Relevant Facts about Historical Linguistics. In Renfrew, Colin (ed.) *America Past, America Present: Genes and Languages in the Americas and Beyond*, pp. 139-62. Cambridge, UK: McDonald Institute for Archaeological Research.

——— (2015). Response to Kassian *et al.*, Proto-Indo-European-Uralic Comparison from the Probabilistic Point of View. *Journal of Indo-European Studies* 43: 348-363.

Rivet, Paul (1924). Langues américaines III: Langues de l'Amérique du Sud et des Antilles. In Meillet, Antoine and Cohen, Marcel (eds.). Les Langues du Monde 16: 639-712. Paris: Champion.

——— (1942). Un dialecte hoka colombien: le yurumangui. *Journal de la Société des américanistes de Paris* 34: 1-59.

Robbeets, Martine (2004). Swadesh 100 on Japanese, Korean, and Altaic. *Tokyo University Linguistic Papers TULIP* 23: 99-118.

——— (2005). *Is Japanese related to Korean, Tungusic, Mongolic, and Turkic?* Wiesbaden: Harrassowitz.

——— (2015). *Diachrony of Verb Morphology: Japanese and the Transeurasian Languages*. Trends in Linguistics, Studies and Monographs 291. Berlin: Mouton de Gruyter.

Robbeets, Martine, Remco Bouckaert, Matthew Conte *et al*. (2021). Triangulation Supports Agricultural Spread of the Transeurasian Languages. *Nature* 559: 616-621.

Rodrigues, Aryon D. (1966a). Classificação da língua dos cinta-larga. Revista de Antropologia 14: 27-30.

——— (1966b). Línguas brasileiras. São Paulo: Loyola.

Rosés Labrada, Jorge Emilio (2015). Is Jodï a Sáliban Language? Paper presented at the Workshop on Historical Relationships among Languages of the Americas. Leiden, The Netherlands, September 2-5, 2015.

——— (2019). Jodï-Sáliban: A Linguistic Family of the Northwest Amazon. *International Journal of American Linguistics* 85, 3: 275-311.

Ross, Malcolm (2005). Pronouns as a Preliminary Diagnostic for Grouping Papuan Languages. In Pawley, Andrew, Robert Attenborough, Robin Hide, and Jack Golson (eds.). *Papuan Pasts: Cultural, Linguistic and Biological Histories of Papuan-speaking Peoples*, pp. 15-66. Canberra: Pacific Linguistics.

Rubicz, Rohina, Kristin L. Melvin, and Michael H. Crawford (2002). Genetic Evidence for the Phylogenetic Relationship between Na-Dene and Yeniseian Speakers. *Human Biology* 74, 6: 743-760.

Rude, Noel (1987). Some Klamath-Sahaptian Grammatical Correspondences. *Kansas Working Papers in Linguistics* 12: 67-83.

Ruhlen, Merritt (1987). *A Guide to the World's Languages*, Volume 1: *Classification*. Stanford, California: Stanford University Press.

——— (1994). *The Origin of Language: Tracing the Evolution of the Mother Tongue*. New York: Wiley & Sons.

——— (1994b). First- and Second-Person Pronouns in the World's Languages. In Ruhlen, Merritt, On the Origin of Languages: Studies in Linguistic Taxonomy. Stanford, California: Stanford University Press, pp. 252-60.

——— (1995a). Proto-Amerind Numerals. *Anthropological Science* 103, 3: 209-225.

——— (1995b). A Note on Amerind Pronouns. *Mother Tongue* 24: 60-61.

——— (1995c). On the Origin of the Amerind Pronominal Pattern. In Chen, Matthew Y. and Tzeng, Ovid J. L. (eds.). *In Honor of William S-Y. Wang*, pp. 405-407. Taipei: Pyramid Press.

——— (1998). The Origin of the Na-Dene. Proceedings of the National Academy of Sciences 95: 13994-13996.

Sagart, Laurent (2005). Tai-Kadai as a Subgroup of Austronesian. In Sagart, Laurent, Roger Blench, and Alicia Sanchez-Mazas (eds.). *The Peopling of East Asia: Putting Together Archaeology, Linguistics, and Genetics*, pp. 177-181. London: Routledge.

——— (2019). A Model of the Origin of Kra-Dai Tones. *Cahiers de linguistique – Asie orientale* 48, 1: 1-29.

Sands, Bonny (1998). Eastern and Southern African Khoisan: Evaluating Claims of Distant Linguistic Relationships. Cologne: Rüdiger Köppe Verlag.

——— (1998). Comparison and Classification of Khoisan Languages. In Maddieson, Ian, and Hinnebusch, Thomas J. (eds.). *Language History and Linguistic Description in Africa*, pp. 75-85. Trenton, NJ: Africa World Press.

Sandy, Clare Scoville (2017). Prosodic Prominence in Karuk. Doctoral Dissertation. University of California, Berkeley.

Sapir, Edward (1915). The Na-Dené Languages: A Preliminary Report. *American Anthropologist* 17 (3): 765-773.

——— (1917a). The Position of Yana in the Hokan Stock. *University of California Publications in American Archaeology and Ethnology* 13: 1-34.

——— (1917b). The Hokan and Coahuiltecan Languages. *International Journal of American Linguistics* 1: 280-290.

——— (1920). Review of Mason, J. Alden. The Language of the Salinan Indians. *University of California Publications in American Archaeology and Ethnology*, 14, 1: 1-154. *International Journal of American Linguistics* 1, 4: 305-309.

——— (1921a). A Supplementary Note on Salinan and Washo. *International Journal of American Linguistics* 2: 68-72.

——— (1921b). A Characteristic Penutian Form of Stem. *International Journal of American Linguistics* 2, 1/2: 58-67.

——— (1921c). A Bird's-Eye View of American Languages North of Mexico. *Science* 54: 408.

——— (1922). The Fundamental Elements of Northern Yana. *University of California Publications in American Archaeology and Ethnology* 13: 215-334.

——— (1923). Text Analyses of Three Yana Dialects. *University of California Publications in American Archaeology and Ethnology* 20: 261-294.

——— (1925). The Hokan Affinity of Subtiaba in Nicaragua. *American Anthropologist* 27: 402-435, 491-527.

——— (1926). The Hokan Affinity of Subtiaba in Nicaragua (Conclusion). *American Anthropologist* 27, 3: 402-435.

——— (1929). Central and North American Indian Languages. *Encyclopaedia Britannica* 14th ed. 5: 138-141.

——— (1984). Letter to A. L. Kroeber (1918), The Sapir-Kroeber Correspondence: Letters between Edward Sapir and A. L. Kroeber, 1905-1925, Berkeley: University of California at Berkeley, Survey of California and Other Indian Languages.

Sapir, Edward, and Swadesh, Morris (1953). Coos-Takelma-Penutian Comparisons. *International Journal of American Linguistics* 19: 132-137.

——— (1960). *Yana Dictionary*. Berkeley: University of California Press.

Sapir, Edward, and Berman, Howard (2001). Chimariko Linguistic Material. In Golla, Victor, and O'Neill, Sean (eds.). *The Collected Works of Edward Sapir XIV: Northwest California Linguistics*, pp. 1039-1076. Berlin: Mouton de Gruyter.

Schlichter, Alice (1979). Wintu Internal Variation. *International Journal of American Linguistics* 45: 236-244.

Schnoebelen, Tyler (2009). Classifying Shabo. Presentation at the 40th Annual Conference on African Linguistics (ACAL 40). University of Illinois at Urbana-Champaign. April 9-11, 2009.

Seifart, Frank, and Echeverri; Juan Alvaro (2014). Evidence for the Identification of Carabayo, the Language of an Uncontacted People of the Colombian Amazon, as Belonging to the Tikuna-Yurí Linguistic Family. *PLoS ONE* 9, 4:e94814.

Shafer, Robert (1961). Tones in Wintun. *Anthropological Linguistics* 3, 6: 17-30.

Shambach, Catherine (1977). A Preliminary Reconstruction of Proto-Costanoan. Unpublished paper.

Shaul, David (1995a). The Huelel (Esselen) Language. *International Journal of American Linguistics* 61: 191-239.
——— (1995b). The Last Words of Esselen. *International Journal of American Linguistics* 61, 2: 245-249.
——— (2019). *Esselen Studies: Language, Culture, and Prehistory*. Munich: Lincom Europa.
Shaul, David, Katherine Turner, and James Collins (1984). *Esselen Linguistic Materials*. Kansas Working Papers in Linguistics.
Shevoroshkin, Vitaly (1980). On Sonorants in Penutian and Some Other NA Languages. Unpublished paper.
Shipley, William (1957). Some Yukian-Penutian Lexical Resemblances. *International Journal of American Linguistics* 23: 269-274.
——— (1961). Maidu and Nisenan: A Binary Survey. *International Journal of American Linguistics* 27: 46-51.
——— (1966). The Relation of Klamath to California Penutian. *Language* 42: 489-498.
——— (1969). Proto-Takelman. *International Journal of American Linguistics* 35: 226-230.
——— (1970). Proto-Kalapuyan. In Swanson, Earl H., Jr (ed.). *Languages and Cultures of Western North America*, pp. 97-106. Pocatello, Idaho: Idaho State University.
——— (1973). California. In Sebeok, Thomas A. (ed.). *Current Trends in Linguistics* 10: 1046-1078.
——— (1978). Native Languages of California. In Heizer, Robert F. (ed.). *Handbook of North American Indians* 8: *California*, 80-90. Washington, D.C.: Smithsonian Institution.
——— (1980). Penutian among the Ruins: A Personal Assessment. In *Proceedings of the Sixth Annual Meeting of the Berkeley Linguistics Society*, pp. 437-441.
Sicoli, Mark A., and Holton, Gary (2014). Linguistic Phylogenies Support Back-Migration from Beringia to Asia. *PLOS ONE* 9, 3: e91722.
Sidwell, Paul (2009). *Classifying the Austroasiatic Languages: History and State of the Art*. Lincom Studies in Asian Linguistics 76. Munich: Lincom Europa.
Silver, Shirley (1964). Shasta and Karok: A Binary Comparison. In Bright, William (ed.). *Studies in Californian Linguistics*, pp. 170-181. Berkeley: University of California Press.
——— (1966). *The Shasta Language*. Ph.D. dissertation. University of California, Berkeley.
——— (1976). Comparative Hokan and the Northern Hokan languages. In Langdon, Margaret and Silver, Shirley (eds.). *Hokan Studies*, pp. 170-181. The Hague: Mouton.
——— (1978). Chimariko. In Heizer, Robert F. (ed.) *Handbook of North American Indians* 8, *California*, 203-210. Washington, D.C.: Smithsonian Institution.
——— (1980). Shasta and Konomihu. In Kathryn Klar, Margaret Langdon, and Shirley Silver (eds.), *American Indian and Indoeuropean Studies: Papers in Honor of Madison S. Beeler*, pp. 245-263. The Hague: Mouton.
Silver, Shirley, and Wicks, Clara (1977). Coyote Steals Fire (Shasta). In Golla, Victor, and Silver, Shirley (eds.). *Northern California Texts*, pp. 121-131. Chicago: University of Chicago Press.
Silverstein, Michael (1967). Penutian Implications of Takelma Sibilant Phonemes. Term Paper for MIT Course, Fall 1966.
——— (1969). Tsimshian and Penutian Possessives. Unpublished paper read at the American Anthropological Association Meeting, New Orleans, Nov. 1969.
——— (1972). Review of Callaghan 1970. *General Linguistics* 12: 128-134.
——— (1974). Dialectal Developments in Chinookan Tense-Aspect Systems, an Areal-Historical Analysis. *International Journal of American Linguistics* Memoirs 29.

——— (1975). On Two California Penutian Roots for 'Two'. *International Journal of American Linguistics*: 41: 369-380.
——— (1977). Person, Number, Gender in Chinook: Syntactic Rule and Morphological Analogy. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 3: 143-156.
——— (1978a). Deixis and Deducbility in a Wasco-Wishram Passive of Evidence. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 4: 238-54.
——— (1978b). Yokuts: Introduction. In Heizer, Robert F. (ed). *Handbook of North American Indians* 8: *California*, pp. 446-447.
——— (1979a). Two Bis. *International Journal of American Linguistics* 45: 187-205.
——— (1979b). Penutian: An Assessment. In Campbell, Lyle, and Mithun, Marianne (eds.). *The Languages of Native America*. Austin: University of Texas Press, pp. 650-691.
Spier, Leslie (1946). Comparative Vocabularies and Parallel Texts in Two Yuman Languages of Arizona. *University of New Mexico Publications in Anthropology* 2: 1-150.
Stark, Louisa R. (1972). Maya-Yunga-Chipayan: A New Linguistic Alignment. *International Journal of American Linguistics* 38, 2: 119-135.
Starostin, Sergei A., Anna V. Dybo, and Oleg A. Mudrak, with the assistance of I. Gruntov and V. Glumov (2003). *Etymological Dictionary of the Altaic Languages*. Leiden: Brill.
Starostin, George (2003). A Lexicostatistical Approach towards Reconstructing Proto-Khoisan. *Mother Tongue* 8: 81-126.
——— (2008). From Modern Khoisan Languages to Proto-Khoisan: The Value of Intermediate Reconstructions. *Aspects of Comparative Linguistics* 3: 337-470.
——— (2011). On Mimi. *Journal of Language Relationship* 6: 115-140.
——— (2013). *Languages of Africa: An Attempt at a Lexicostatistical Classification*. Volume I: *Methodology. Khoesan Languages*. Moscow.
——— (2016). The Nilo-Saharan Hypothesis Tested through Lexicostatistics: Current State of Affairs (ms.).
——— (2017). Языки Африки. Опыт построения лексикостатистической классификации. Т. 3. Нило-Сахарские Языки (Languages of Africa: An Attempt at a Lexicostatistical Classification. Volume 3: Nilo-Saharan Languages). Moscow: State University of the Humanities.
Storch, Anne, and Dimmendaal, Gerrit J. (2016). Niger-Congo: a Brief State of the Art. *Oxford Handbooks Online* website.
Sturtevant, Edgar H. (1931). Hittite Glossary: Words of Known or Conjectured Meaning, with Sumerian Ideograms and Accadian Words Common in Hittite Texts. Language Monograph 9. *Language* 7, 2: 3-82.
——— (1932). The Development of the Stops in Hittite. *Journal of the American Oriental Society* 52, 1: 1-12.
——— (1933). Comparative Grammar of the Hittite Language. New Haven: Yale University Press.
——— (1942). The Indo-Hittite Laryngeals. Baltimore: Linguistic Society of America.
——— (1940). Evidence for Voicing in Indo-Hittite γ. *Language* 16, 2: 81-87.
Sturtevant, Edgar H., and Bechtel, George (1935). *A Hittite Chrestomathy*. William Dwight Whitney Linguistic Series. Philadelphia: Linguistic Society of America.
Suárez, Jorge A. (1974). South American Indian Languages. In *Encyclopaedia Britannica,* 15th ed., Macropaedia 17: 105-112.
Swadesh, Morris (1959). Mapas de clasificación lingüística de México y las Américas. México: UNAM.

——— (1965). Kalapuya and Takelma. *International Journal of American Linguistics* 31: 237-240.

——— (1967). Lexicostatistical Classification. In McQuown, Norman A. (ed.). *Linguistics*. In Wauchope, Robert (general editor). *Handbook of Middle American Indians* Vol. 5, pp. 79-115. Austin: University of Texas Press.

Swadesh, Mauricio (1987). Tras la Huella Lingüística de la Prehistoria (in Spanish).

Swanton, John R. (1905). *Haida Texts and Myths. Skidegate Dialect*. Bureau of American Ethnology Bulletin 29. Washington, D.C.: Smithsonian Institution.

——— (1915). Linguistic Position of the Tribes of Southern Texas and Northeastern Mexico. *American Anthropologist* 17: 17-40.

Tarpent, Marie-Lucie (1990). Nisgha Vowel Alternations. In *Papers from the 25th International Conference on Salishan and Neighbouring Languages*, pp. 313-334. Vancouver: University of British Columbia Dept. of Linguistics.

——— (1994). Tsimshianic Stress-Assignment and the Words for 'Meat, Bear'. In *Papers from the 29th International Conference on Salishan and Neighbouring Languages*, pp. 243-264. Pablo, Montana: Salish Kootenai College.

——— (1996). Reattaching Tsimshianic to Penutian. *Survey of California and Other Indian Languages* 9: 91-112.

——— (1997). Tsimshianic and Penutian: Problems, Methods, Results, and Implications. *International Journal of American Linguistics* 63: 65-112.

——— (2000). The Original Structure of Sapir's 'Characteristic Penutian Form of Stem'. Paper Presented to the SSILA (Society for the Study of the Indigenous Languages of the Americas) Chicago, January 6-9, 2000.

——— (2001). On the Eve of a New Paradigm. In J. Brinton (ed.) *Historical Linguistics 1999: Selected Papers from the 14th International Conference on Historical Linguistics*, Vancouver, 9-13 August 1999. Amsterdam: John Benjamins, Studies in the Theory and History of Linguistic Science 4.

——— (2002). A Pan-Penutian Database of Materials for Comparison and Reconstruction: Its Organization, Uses and Current Results. In Conathan, Lisa (ed.), *Report 12 Survey of California and Other Indian Languages*, pp.119-130.

——— (2002). Tsimshianic L-Initial Plurals: Relics of an Ancient Penutian Pattern. In Buszard-Welcher, Laura (ed.) *Proceedings of the Meeting of the Hokan-Penutian Workshop*, June 17-18, 2000. Report 11 *Survey of California and Other Indian Languages*: 88-108.

——— (2014). Some Pan-Penutian Lexical Elements in Names of Small Animals. Paper read at the Meeting of the Society for the Study of the Indigenous Languages of the Americas.

Tarpent, Marie-Lucie, and Daythal, Kendall (1998). On the Relationship between Takelma and Kalapuyan: Another Look at 'Takelman'. Paper presented to the Society for the Study of the Indigenous Languages of the Americas Annual Meeting, Linguistic Society of America, New York.

Taye, Kibebe Tshay. (2015). Documentation and Grammatical Description of Chabu. Doctoral dissertation. Addis Ababa: Addis Ababa University.

Theil, Rolf (2006). Is Omotic Afro-Asiatic? (draft).

Thomason, Sally (May 8, 2013). Ultraconserved Words? Really?? *Language Log* website.

Thurgood, Graham (1994). Tai-Kadai and Austronesian: The Nature of the Relationship. *Oceanic Linguistics* 33: 345-368.

Tovar, Antonio, and Larrucea, Consuelo (1984). Catálogo de las lenguas de América del Sur: Con clasificaciones, indicaciones tipológicas, bibliografía y mapas. Madrid: Gredos.

Traill, Anthony (1986). Do the Khoi Have a Place in the San? New Data on Khoisan Linguistic Relationships. In Rottland, Franz, and Vossen, Rainer (eds.). African Hunter-Gatherers International Symposium, pp. 407-430. *Sprache und Geschichte in Afrika* Special Issue 7.1. Hamburg: Helmut Buske.

——— (March 1998). Linguistics 112 Lecture, Department of Linguistics, University of the Witwatersrand, South Africa.

Turner, Katherine (1987). *Aspects of Salinan Grammar*. Ph.D. dissertation. University of California, Berkeley.

Ultan, Russell (1964). Proto-Maidun Phonology. *International Journal of American Linguistics* 30: 355-370.

Vajda, Edward J. (2010a). Siberian Link with Na-Dene Languages. In Kari, James, and Potter, Ben (eds.), *The Dene-Yeniseian Connection*, pp. 33-99. Anthropological Papers of the University of Alaska, New Series: Vol. 5. Fairbanks: University of Alaska Fairbanks Department of Anthropology.

——— (2010b). Yeniseian, Na-Dene, and Historical Linguistics. In Kari, James, and Potter, Ben (eds.), *The Dene-Yeniseian Connection*, pp. 100-118. Anthropological Papers of the University of Alaska, New Series 5. Fairbanks: University of Alaska Fairbanks Department of Anthropology.

——— (2012). Geography, Demography and Time Depth: Explaining How Dene-Yeniseian is Possible. Presentation at the 2012 Dene-Yeniseian Workshop. Alaska Native Language Center, University of Alaska, Fairbanks, Alaska, March 24, 2012.

——— (2020). In Robbeets, Martine, and Savelyev, Alexander (eds.). *The Oxford Guide to the Transeurasian Languages*, pp. 726-734. Oxford: Oxford University Press.

van Driem, George (2018.) The East Asian Linguistic Phylum: A Reconstruction Based on Language and Genes, *Journal of the Asiatic Society* 60, 4: 1-38.

Viegas-Barros, J. Pedro (1992). La familia lingüística tehuelche. *Revista Patagónica* 54: 39-46.

——— (1992-1993). ¿Existe una relación genética entre las lenguas mataguayas y guaycurúes? In Braunstein, José A. (ed.), *Hacia una nueva carta étnica del Gran Chaco,* vol. V, pp. 193-213. Las Lomitas (Formosa): Centro del hombre antiguo chaqueño (CHACO).

——— (1994). La clasificación de las lenguas patagónicas. Revisión de la tesis del grupo lingüístico "andino meridional" de Joseph H. Greenberg. *Cuadernos del Instituto nacional de antropología* 15: 167-184.

——— (2001). *Evidencias del parentesco de las lenguas lule y vilela*. Colección folklore y antropología 4: 15-21. Santa Fe: Subsecretaría de cultura, Dirección provincial de gestión cultural.

——— (2005). *Voces en el viento: Raíces lingüísticas de la Patagonia.* Colección El Suri. Buenos Aires : Mondragón Ediciones.

——— (2006a). La hipótesis macro-guaicurú. Semejanzas gramaticales guaicurú-mataguayo. Revista UniverSOS 3: 183-212. Valencia: Universidad de Valencia.

——— (2006b). Proto-lule-vilela: Una reconstrucción fonológica preliminar. Comisión Lenguas chaqueñas. Paper presented at the 52 Congreso internacional de americanistas, July 17-21, 2006. Sevilla (España): Universidad de Sevilla.

——— (2013). La hipótesis de parentesco guaicurú-mataguayo: Estado actual de la cuestión. *Revista brasileira de linguística antropológica* 5, 2: 293-333.

Voegelin, Charles F. (1946). Notes on Klamath-Modoc and Achumawi Dialects. *International Journal of American Linguistics* 12: 96-101.

Voegelin, Charles F., and Voegelin, Florence M. (1965). Classification of American Indian languages. Languages of the World, Native American Facsimile 2, 1.6. *Anthropological Linguistics* 7, 7: 121-150.

——— (1977). *Classification and Index of the World's Languages*. New York: Elsevier.

Waterhouse, Viola Grace (1962). *The Grammatical Structure of Oaxaca Chontal*. Indiana University Publication of the Research Center in Anthropology, Folklore, and Linguistics 19.

——— (1976). Another Look at Chontal and Hokan. In Langdon, Margaret and Silver, Shirley (eds.). *Hokan Studies*, pp. 325-343. The Hague: Mouton.

Waterhouse, Viola Grace, and Morrison, May (1950). Chontal Phonemes. *International Journal of American Linguistics* 16: 35-39.

Weigel, William (2005). *Yowlumne in the Twentieth Century*. Ph.D. dissertation. Berkeley: University of California.

Werner, Heinrich K. (2004). Zur jenissejisch-indianischen Urverwandtschaft (On the Yeniseian-[American] Indian Primordial Relationship). Wiesbaden: Harrassowitz.

Weitlaner, Roberto J. (1941). Los Pueblos no nahuas de la historia tolteca y el grupo lingüístico macro-otomangue. *Revista mexicana de estudios antropológicos* 5, 2-3: 249-269.

Westermann, Diedrich (1912). The Shilluk People, Their Language and Folklore. Philadelphia: Board of Foreign Missions of the United Presbyterian Church.

Westphal, Ernst O. J. (1962). A Re-Classifcation of Southern African Non-Bantu Languages. *Journal of African Languages* 1:1-8.

——— (1963). The Linguistic Prehistory of Southern Africa: Bush, Kwadi, Hottentot, and Bantu Linguistic Relationships. *Africa* 33: 237-265.

——— (1965). Linguistic Research in S.W.A. and Angola. In *Die ethnischen Gruppen in Südwestafrika* 3, pp. 125-144. Windhoek: S.W.A. Wissenschaftliche Gesellschaft.

——— (1971). The Click Languages of Southern and Eastern Africa. In Sebeok, Thomas A. (ed.). *Current Trends in Linguistics*, Volume 7: *Linguistics in Sub-Saharan Africa*, pp. 367-420. Berlin: Mouton.

Whaley, Lindsay J., and Li, Fengxiang (2000). Emphatic Reduplication in Oroqen and Its Altaic Context. *Linguistics* 38, 2: 355-372.

Whistler, Kenneth W. and Golla, Victor (1986). Proto-Yokuts Reconsidered. *International Journal of American Linguistics* 52: 327-358.

Winter, Werner (1957). Yuman Languages I: First Impressions, *International Journal of American Linguistics* 23: 18-23.

Wurm, Stephen A., and Laycock, Donald C. (eds.) (1970). *Pacific Linguistic Studies in Honour of Arthur Capell*. Canberra: The Australian National University: Pacific Linguistics.

Wurm, Stephen A. (ed.) (1975). *Papuan Languages and the New Guinea Linguistic Scene: New Guinea Area Languages and Language Study* 1. Canberra: Australian National University, Dept. of Linguistics, Research School of Pacific Studies.

Zamponi, Raoul. (2008). Sulla Fonologia e la Rappresentazione Ortografica del Lule. In Maccioni, Antonio (ed.), *Arte y Vocabulario de la Lengua Lule y Tonocoté* XXVI-LVII. Cagliari: Centro di Studi Filogici Sardi.

——— (2017). Betoi-Jirara, Sáliban, and Hodi: Relationships among Three Linguistic Lineages of the Mid-Orinoco Region. *Anthropological Linguistics* 59, 3: 263-321.

Zhivlov, Mikhail (2013). The Hokan Family and Lexicostatistics. Comparative-Historical Linguistics of the 21st Century: Issues and Perspectives. Moscow: Institute for Oriental and Classical Studies, Russian State University for the Humanities, March 20-22, 2013.

——— (2014). Калифорнийские Пенути: Группа, Семья Или Макросемья?. Paper read at the 9th Annual Sergei Starostin Memorial Conference on Comparative Historical Linguistics. Moscow: Russian State University for the Humanities.

——— (2018). Some Morphological Parallels between Hokan Languages. *Journal of Language Relationship* 16, 1-2: 138-161.

——— (2020). Personal communication.

Zúñiga, Fernando (2006). *Mapudungun. El habla mapuche*. Santiago, Chile: Centro de estudios públicos.

RESONANT VARIATIONS ON IMMORTALITY

GREGORY HAYNES[1]

Both the Greek word *ambrosia* and the Sanskrit *amrita* signify a divine elixir that confers immortality on the gods. These words are composed of the negative prefix **a-* and a root **mer-(t)-* 'to die,' hence 'undying, immortal, without death,' and with wider meanings that include: 'imperishable, eternal, unfading, undamaged, unwounded, undestroyed.'

This root shares semantic value and phonetic form with other roots that differ phonetically only in the attested resonant. The form common to these resonant-variants can be generalized as ****n̥-me(R)-(t)-*** , where *(R)* can represent any resonant (or a zero grade with no resonant). The majority of the attestations cited show a suffix (or final) in **t-* or **th-*, which is represented below by the symbol *(t)*. The following table lists four primary resonant-variants in which attestations can be found showing a consistent semantic value.[2]

****n̥-me(R)-(t)*** **'Undying, immortal'**

PIE Root	Neg. Pref.	Initial	R1	R2	Final	Ref.	Semantic Value
**n̥-mer-(t)-*	n̥	m		r	(t)	1	Undying, immortal, imperishable, eternal, unfading
**n̥-mel-(t)-*	n̥	m		l	(t)	2	Undying, unharmed, unfading, undamaged, unwounded
**n̥-men-(t)*	n̥	m		n	(t)	3	Undying, immortal, unharmed, undestroyed (Proposed Root)
**n̥-me-(t)-*	n̥	m		ø	(t)	4	Undying, immortal, undestroyed, unhurt (Proposed Root)

[1] Correspondence can be addressed to haynes@sonic.net.

[2] This study is based upon my article "Resonant Variation in Proto-Indo-European," which appeared in the journal *Mother Tongue* (Volume 22, 2020), published by the Association for the Study of Language in Prehistory. That article is fundamental to the thesis presented here and must be consulted along with the present work. See https://www.aslip.org/MT/Resonant_Variation_in_Proto-Indo-European.pdf for an on-line version of that article in a revised format.

1. **n̥-mer-(t)-* 'Undying, immortal' (**mer-*)

Lat *immortālis* 'not subject to death, immortal, imperishable, everlasting, eternal, unending,' *immortālitās* 'immunity from death, immortality,' Skt *amṛ́ta* 'immortal; also the nectar conferring immortality, produced at the churning of the ocean, ambrosia, also called *soma* or *mádhu*,' Av *aməša-* 'immortal,' Grk *ἄμβροτος* 'immortal,' *ἀμβρόσιος* 'ambrosia, the drink of immortality, elixir of life' *ἀμάραντος* 'unfading, imperishable,' *ἀμάραντον* 'the plant amaranth,' Lat *amarantus* 'the plant amaranth, literally: never-fading flower' (see Skt *amalātaka* 'amaranth' below). Two additional Skt words, *ámartya* and *amára*, both signify 'undying, immortal, imperishable.' A compound of the latter, *amára-ratna*, signifies 'crystal' (literally, imperishable jewel). See also *amala-ratna* 'crystal' below.

Examples of forms without the negative particle would include: Grk *βροτός* 'mortal, mortal man, human,' Skt *mṛ́tás* 'dead,' *márta* 'a mortal, man,' *mártavya* 'liable to die,' *mrit, mrityáti* 'decay, be dissolved,' *marmartu* 'kill,' OPers *martiya-* 'a mortal man,' Arm *mard* 'man, mortal,' Hitt *merta* 'disappear, be lost,' OCSlav *mrěti* 'to die,' Lith *mirtìs* 'death,' Lat *mors, mortis* 'death,' ON *morð* 'murder,' OE *morþ* 'murder, death, destruction, anything horrible.' —LIV 439; IEW 735; Watkins (2011) 56; Mallory and Adams (2006) 198; Monier-Williams 80, 791; LSJ 8-9, 77; OLD 112, 838; De Vries 392; Onions 30; Mayrhofer II 319; AHD 55; NIL 488-91; Bomhard 873. (A list of abbreviations used appears at the end of this article.)

Note that in Tocharian B the word for 'man' (*eṅkwe*) is formed from a different root (**nek̂s* 'death') but also signifies 'mortal'.

2. **n̥-mel-(t)* - 'Immortal, without harm, unfaded, undamaged, unwounded' (mel-*)**

OIr *millid* 'harms,' Toch B *mäl* 'wound, damage,' —Mallory and Adams (2006) 279.

Mallory and Adams recognize the root **mel-* 'harm,' as supported by the Old Irish and Tocharian B attestations. To these I would add the forms that they note under the root **melo/*méles*, and which IEW lists under *2.*mel-*:

Arm *melk'* 'sin,' MIr *mell* 'mistake,' Lat *malus* 'bad,' Grk *μόλη* 'abortion,' *ἀμβλίσκω*, *ἀμβλόω* 'cause a miscarriage, cause an abortion' (the initial *ἀ* is not privative; the medial *β* is not original but arises as *μβλ-* < *μλ-*), MIr *mell* 'error, failure,' *maile* 'evils,' Cymr *mall* 'spoil, damage, ruin.' —Mallory and Adams (2006) 197; IEW 719-20; LSJ 79; DELG 70.

To these, I propose the following four additions (all with negative prefix):

- Grk *Ἀμάλθεια* (*Ἀ-μάλθεια* 'without death, without harm'). *Amaltheia* 'Amalthea' is the mythical goat that suckled Zeus in a Cretan cave and who was made immortal in the

heavens. She is additionally associated with 'protection from harm' in that the magic mantle of Athena, the *aegis*, which protected its possessor from all injury, was, according to some authorities, fashioned from her hide. Her horns were miraculous, flowing with nectar and ambrosia. One of those horns was broken off and was endowed by Zeus with the power to refill itself inexhaustibly with whatever food or drink was wished for—the original cornucopia[3], Grk *ἀμαλθεύω* (= *τρέφω* per LSJ) 'to nurture, keep, cause to grow or increase, bring up, rear, tend, cherish, foster, maintain, support, feed.' —LSJ 76, 1814; OLD 112; DELG 67.

It has been suggested (LSJ, DELG) that the etymology for the name of this mythical goat can be found in Grk *ἀμαλακία* (cf. *μαλθακός*) 'hardness, incapability of being softened, unmitigated, harsh,' but it is not clear how these concepts would appropriately apply to the being who nurtured Zeus. Given her associations with immortality, nectar, ambrosia, inexhaustible sources of sustenance, protection and nurturing, it seems more correct to analyze the name as 'without death, without harm.'

- Lat *amulētum / amolētum* 'an object used as a charm to avert evil or injury, an amulet, the power of averting evil,' —AHD 62; AHD (1975) 45; OLD 124, 125; Kluge 20; Onions 33.

This is a word of obscure origin, supposed by some to stem from Lat *amylum* < Grk *ἄμυλον* 'fine meal, starch, gruel,' but the semantic divergence is great. More probably the source of this word (*a-muletum*) is the concept, "without harm, without death," because warding off harm and death is the function of the object or charm used as an amulet.

- Skt *mlai* 'fade, wither, decay, vanish, to be exhausted, enfeeble,' Skt *amalātaka, amilātaka* 'amaranth' (amaranth is literally "the unfading, imperishable flower"), *amala-ratna* 'crystal' (literally: 'the undying jewel' because crystals grow, but they do not perish), *amlāna* 'unwithered.' —Mayrhofer II 388; Monier-Williams 80-84, 837-838.

Mayrhofer suggests a possible connection of Skt *mlai* to PIE **melh₁* 'grind,' but semantically this concept is distant. He also suggests a possible connection to Grk *μαλακός* 'soft, mild, gentle,' but here again the semantics are significantly different. Monier-Williams places *amala-ratna* 'crystal,' with an adjective *a-mala* 'spotless, stainless, clean, pure, shining,' rather than with the concept "undying, imperishable, immortal," but this is unlikely. The forms *amala* and *amára* are evidently parallel since they share widely divergent semantic values. For example, when compounded with *-ratna* 'jewel,' both signify 'crystal,' and both share the unexpected signification "umbilical cord." *Amára* is clearly *a-mára* 'undying, immortal, imperishable,' as its innumerable compounds attest

[3] For the myth of Amaltheia, see *New Larousse Encyclopedia of Mythology*, Guirand, Felix, ed., (London: Hamlyn, 1959), 91, Andrews, E. A., *Latin–English Lexicon*, (New York: Harper and Brothers, 1858), 89, and *The Oxford Classical Dictionary*, Cary, M. et al., eds., (Oxford: Clarendon Press, 1964), 41.

(i.e. *amarī-bhū* 'to become immortal,' *amara-sadas* 'the assemblage of the gods,' *amara-vat* 'like an immortal,' *amara-pura* 'the residence of the immortals, paradise,' etc.).

With respect to my reading of Skt *amala-ratna* and *amára-ratna* 'crystal,' as reflecting their imperishable, durable form, (as opposed to "gem of the immortals") compare the Skt compound *amṛitêshṭakā́* "a burnt or baked (and therefore imperishable) brick (used for the sacrificial altar)."[4] It is evident that Sanskrit does not hesitate to employ terms signifying *immortal, imperishable* to physical objects that possess an enduring form.

- ON *Amlóði* 'the Old Icelandic name for a character who appears in the Norse sagas, Latinized by Saxo as *Amlethus*, and who appears again in the work of William Shakespeare as *Hamlet*.' —De Vries 8; Vigfusson 19, 771.

Various etymologies for this word have been proposed, but none are convincing, most authorities stating that it is of uncertain origin.[5]

It should be noted that the Indo-European myth of an elixir of life, or of immortality (see *amrita, ambrosia* above) is closely related to the widespread axis mundi / Tree of Life myth. The axis mundi myth-complex involves an enormous tree, mountain, or pillar located at the center of the earth, and which is also the axis of stellar motion. It is often believed to be located on an island in the ocean somewhere west of Europe. The most defining characteristic of this central axis of the world is that it always involves axial motion in some form that reflects the mystery of the apparent movement of the stars across the sky from east to west and around the pole star. This mythical element can be represented by notions of spinning, churning, making fire with a fire drill, making pots on a potter's wheel, mills, whirlwinds, whirlpools, etc.[6]

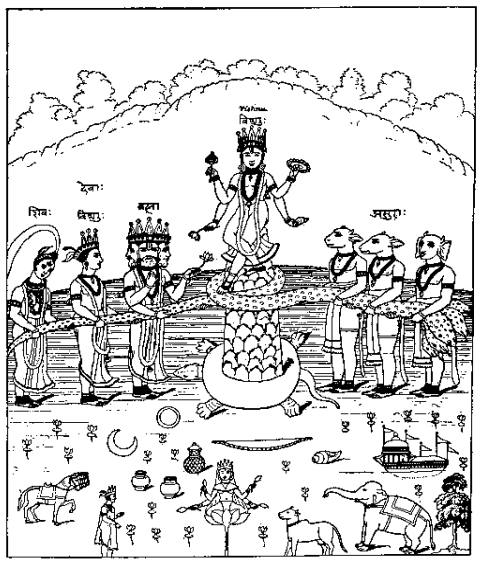


Figure 3. The Indian Myth of the Churning of the Ocean of Milk, which produced *Amrita*.

In Indian legend, for example, *amrita* was created when the gods *churned* the sea of milk using a giant snake as a rope to *twist* the axis mundi mountain, Mount Mandara. In Norse myth the goat, Heidrun, eats the leaves from Yggdrasill, the Scandinavian Tree of Life, and from her udder flows an endless supply of mead.[7] The axial motion element in Norse myth is supplied by the three Norns who sit

[4] Monier-Williams, Monier, *A Sanskrit–English Dictionary*, (Oxford: Oxford University Press, 1899), 82.
[5] See De Vries, Jan, *Altnordisches Etymologisches Wörterbuch*, (Leiden: E. J. Brill, 1977), 8.
[6] For an extensive treatment of the Axis Mundi / Tree of Life myth, see Haynes, Gregory, *Tree of Life, Mythical Archetype* (San Francisco: Symbolon Press, 2009). See also Michael Witzel, "Sur le chemin du ciel," *Bulletin des Etudes Indiennes* 2 (1984): 213-279, http://www.people.fas.harvard.edu/~witzel/CheminDuCiel.pdf.
[7] Sturluson, Snorri, *Gylfaginning*, in *Prose Edda,* Young, Jean, trans. (Berkeley and Los Angeles: University of California Press, 1954), 64. It is interesting to compare the two myths of Heidrun and Amaltheia: Both are female goats, both produce an inexhaustible supply of an exhilarating or ambrosial drink, and both inhabit a lofty place on the axis mundi tree or mountain.

in the shade of Yggdrasill and *spin out* the fates of human beings. In another Norse myth, Odin steals the mead of poetic inspiration by *drilling* into the mountain cave where it is stored, illustrating yet another form of axial motion.[8] In Greece, Zeus is fed ambrosia by the goat *Amaltheia* in a cave on Mount Ida. Mount Olympus and Mount Ida are the typical representations of the axis-mundi mountain in Greek myth, and Zeus is connected to axial motion by his associations with fire drills and to whirlwinds.[9]

Another Norse myth describes the *mill* of *Grotti* where two maids, *Fenya* and *Menya*, *grind out* gold for King *Fróthi*. At the place where the ship carrying the mill and the maids was sunk, a *whirlpool* now exists in the sea. The grinding of the mill and the whirlpool are two additional examples of the axial motion typically associated with axis mundi representations.[10]

This mill of Grotti is mentioned again, this time in the Skáldskaparmál section of the *Prose Edda*.[11] The Brodeur translation of the verse is quoted here:

> They say nine brides of skerries
> Swiftly move the Sea-Churn
> Of Grótti's Island-Flour-Bin
> Beyond the Earth's last outskirt –
> They who long the corny ale ground
> Of Amlódí; the Giver
> Of Rings now cuts with ship's beak
> The Abiding-Place of boat-sides.
> Here the sea is called Amlódi's Churn.

This churn, now a whirlpool, is associated with an island in the sea beyond the last outskirt of Europe, and is named after one *Amloði*, a name which appears again, much later, as Shakespeare's *Hamlet*. Shakespeare's play follows the story line as described by Saxo Grammaticus in his *Gesta Danorum* completed at the beginning of the 13th century, where the name of the Hamlet character is *Amleth*. But even earlier than Saxo is the 12th century *Chronicon lethrense,* where the Hamlet character is called *Amblothæ.*

Here then, is a name strongly associated with axis mundi myth and with the production of an exhilarating drink: "They who long the corny ale ground of Amlódí." In Indian legend, churning the sea produces soma, and in the Norse legends, churning the sea produces the ale of Amlóði. Soma or amrita is the elixir of immortality. The "ale of Amlóði" would appear to represent the same ancient mythological motif.

---

[8] Sturluson, Snorri, *Prose Edda*, Brodeur, Arthur G., trans. (Oxford: Oxford University Press, 1929), 93-94.
[9] For the association of Zeus with the fire drill, see Graves, Robert, *Greek Myths,* (Wakefield, Rhode Island & London: Moyer Bell, 1988), article 39.8; and Cook, A. B., *Zeus*: *A Study in Ancient Religion*, Vol. 1, (Cambridge: Cambridge University Press, 1914), 330. For the association of Zeus with whirlwinds, see Cook, A. B., *Zeus*: *A Study in Ancient Religion*, Vol. 3, 162–165.
[10] Sturluson, Snorri, *Prose Edda*, Brodeur translation, 140.
[11] Sturluson, Snorri, *Prose Edda*, Brodeur translation, 162.

3. **n̥-men-(t)-* 'Immortal, unharmed' (**men-*) Proposed Root

- Grk *ἀμανῖται* (*ἀ-μανῖται* 'un-dying, immortal') 'a fungus or mushroom; a tribe of mostly poisonous mushrooms, the most famous being the hallucinogenic *Amanita muscaria*, the fly agaric. —LSJ 76; Monier-Williams 1249; Daniélou 66; DELG 67; AHD (1975) 40.

R. Gordon Wasson has argued that this hallucinogenic mushroom is the *soma* of the *Ṛg Veda*, and that it was called *amrita* (the beverage of immortality) for its powerful psychotropic effects.[12] Other Vedic scholars question this association,[13] while still others believe that Wasson was mistaken in this identification.[14] Even if the *Amanita muscaria* was not the Vedic soma, still it was unquestionably employed in a religious context by other cultures for its psychotropic effects and may have derived its name (*a-manita* 'im-mortal') from those hallucinogenic properties. The Grk word, *ἀμανῖται,* has no known PIE origin.

- Skt. *ámānusha* (*á-mānusha* 'not-human, not-mortal') 'superhuman, divine, celestial,' but also 'inhuman, brutal,' *ámānusha-loka* 'the celestial world, heaven' (literally "abode of immortals"). —Monier-Williams 80.

The following are additional proposed attestations of the same root. They lack the negative prefix and so signify simply 'death, harm, deprivation,' or beings who are subject to the same (i.e. mortals).

- Skt *mánu* 'man, mankind,' ON *maðr*, *mannr* 'man, human,' Goth *manna* 'man, human,' SCr, Czech, Russ *muž* 'man, human,' Pol *mąž* 'man, human,' NE *man* 'human being, mankind, male.' —Monier-Williams 784; De Vries 374-75; Buck (1949) 79-80; IEW 700; Mayrhofer II 309.

Mallory and Adams (2006) write,

> Perhaps more conjectural is the derivation of *mVnus, which rests on a not entirely clear Germanic-Indic isogloss (e.g. NE *man*, Skt *mánu-* 'man, person'), which many claim to go back to **men-* 'think', presumably under the illusion that man is a cognitive creature.[15]

C.D. Buck writes,

> The principal source of words for 'man' [as human being], so far as their etymology is clear, is the notion of 'earthly' or 'mortal', thus distinguishing men from the gods. But a few are derivatives of words for 'man' [the male gender], and the ultimate semantic source of one important group (NE *man*, etc.) is uncertain. ... These point to an IE word for 'man', but its root connection and so its ultimate semantic source are wholly uncertain. It has been derived from IE **men*' 'think', or, since the designation 'thinker' seems too sophisticated, from **men-* in an assumed earlier meaning 'breath' or the like; or by some connected with Lat. *manus* 'hand'...[16]

---

[12] Wasson, R. Gordon, *Soma: Divine Mushroom of Immortality*, (Harcourt Brace Jovanovich, 1968).
[13] Stephanie W Jamison and Joel P Brereton, *The Rigveda : the earliest religious poetry of India*, 3 vols., (Oxford: Oxford University Press, 2014).
[14] Michael Witzel, personal communication.
[15] Mallory, J. P., and Douglas Q. Adams, *The Oxford Introduction to Proto-Indo-European and the Proto-Indo-European World*, (Oxford and New York: Oxford University Press, 2006), 203-204.
[16] Buck, Carl Darling, *A Dictionary of Selected Synonyms in the Principal Indo-European Languages*, (Chicago & London: University of Chicago Press, 1949), 80.

Pokorny writes, "...perhaps connected with 3.*men- 'think.'

I propose that these attestations parallel other IE words for 'man, human being' that signify 'mortal,' or 'subject to death,' as seen in the forms listed above (i.e., Grk *βροτός* 'mortal, mortal man, human,' Skt *márta* 'a mortal, man,' OPers *martiya-* 'a mortal man,' Arm *mard* 'man, mortal,' Toch B *eṅkwe* 'man, mortal').

- Lat *mānēs* 'the spirits of the dead, the shade of a particular person, mortal remains, corpse, death, doom, fate,' —OLD 1072-73, 1078; IEW 693; AHD 1063; OCD 533; Onions 551; Watkins 51.

AHD suggests that this word is "perhaps from Lat *mānis* 'good'`." Onions writes, "plausibly (but not certainly) referred to *mānis, mānus* 'good'..." Watkins includes these Latin forms with the root **mā-* 'good' (IEW**mā-* 693), but that proposed root also contains the Grk word *μῆνις* 'wrath' which would presumably be the opposite of "good." This appears to be a contradictory "catch-all" root that attempts to combine too many disparate concepts. Perhaps the true source for "mortal remains, corpse, and death" is simply **men-* 'to die, to be mortal.'

- Grk *μῆνις* 'the dead worshiped as heroes, wrath' —LSJ 1128; IEW 693; DELG 670-71.

This is a word with no known PIE etymology. IEW suggests a possible connection with 2. **mā* 'good,' but DELG states that the etymology is unknown.

- Grk *μαντιχώρας* 'man-eater,' Lat *mantichōrās* 'man eater,' (*manti* 'man' + *chōrās* 'eater') —LSJ 1080-81; Onions 553; OLD 1075; AHD 1066; DELG 644; Monier-Williams 785.

This Greek word is the source of NE manticore, which is either a tiger, or else a fabulous monster from India or Africa having a lion's body, the head of a man, and the tail of a dragon or scorpion. All of the authorities agree that the etymology of this word signifies "man-eater" but they also uniformly opine that Grk *μαντιχώρας* is a misreading of an original *μαρτιχόρας* 'eater of mortals' (i.e. eater of men). In light of the present analysis, this may or may not be the case. Compare Skt. *manushya-hārin* 'man stealer.'

- Lat *menta, mentha* 'mint,' Grk *μίνθη* 'mint.' If correct, this would be in the sense of the mint, pennyroyal (*Mentha pulegium*), well-known to contain an abortifacient oil, hence the association with death.[17] —OLD 1100, 1113, 1517; Bailey 2035; Galen III.xii.183; AHD 1121.

Lat *menta* and Grk *μίνθη* have no known PIE origin.

- Toch B *mänt-* 'remove (utterly) from its place, destroy, disturb, fall into misfortune, be irritated, be evil-minded, malicious,' *mäntarṣke* 'evil,' *mänts-* 'be sorrowful, afflicted, lament.' —Adams s.v. *mänt-*, *mänts-*; Mayrhofer 298-99; LIV 438-39.

[17] Riddle, John M., *Contraception and Abortion from the Ancient World to the Renaissance*, (Cambridge, Mass: Harvard University Press, 1994); Aristophanes, *Peace*, Rogers, B. B., trans., in *The Complete Plays of Aristophanes* (New York: Bantam, 1988), 207.

LIV would derive *mänt-* from **manth₂-*, 'whirl, stir,' but the concepts "stir" and "destroy" are significantly divergent (see below Skt *math-*). Adams says of Toch B *mänts-* that the etymology is unclear, but with a possible connection to **menth₂*.

*4. *n̥-me-(t)-* 'Undying, immortal, undestroyed, unhurt' (**me-t-*) Proposed Root

- Grk *ἀμέθυστος*, traditionally glossed as: 'not drunken, without drunkenness, not intoxicating, remedy against drunkenness, amethyst,' Lat *amethystus* 'a precious stone of violet-blue color, the amethyst, a kind of vine supposed to yield a non-intoxicating wine.' — LSJ 79, 1091; Monier Williams 82; Onions 31; OLD 117; AHD 59; IEW 707; DELG 650-51.

Of *amethyst*, AHD gives, "*a-*, 'not' + Grk **methúskein* 'to intoxicate,' from *methuein* 'to be drunk,' from *méthu* 'wine'; see **medhu-*."

The crystal amethyst was given this name—so the story goes—because it was supposed to prevent intoxication. It was said that wine, drunk out of a cup made of amethyst, would not intoxicate.

Accepting this etymology (which goes back to classical authors), most authorities ascribe the derivation of *amethyst* to PIE **méd^{h}u-* 'honey, mead.'

I would argue, however, that such a development is more likely to be an instance of folk etymology. Amethyst can have no effect against the intoxicating property of wine, so it is very unlikely that such a belief would have been part of its original designation. Neither is there any reality to the idea that a grape-vine exists that produces a type of wine that will not intoxicate. Probably there were two other factors that led to the stone being associated with wine.

First the light purple color of the crystal is reminiscent of the color of diluted wine. Since the custom in ancient Greece was to mix wine with an equal part of water before drinking, this similarity in color would have been commonly noted. Second, by taking the initial *α-* of *ἀμέθ-* (the ancient name for the crystal) as the alpha-privative, the remaining word certainly does resemble the initial element in *μεθυστάς* 'drunken.' For that reason, it would have seemed logical, in a later period, to somehow associate the crystal with the concept, *non-intoxication*. It is probable, however, that the form *ἀμέθυστος* 'amethyst' has been remodeled (by the suffixation of *υ-* to the stem) to align it with *μέθυ* in its mistaken identity with that word.

But, as has been observed above in the examples from Sanskrit (*amára-ratna, amala-ratna*), a crystal was sometimes associated with the concepts "undying, imperishable," due to its durable, unfading, non-perishing properties. It is more likely that the word *amethyst* was originally derived from the zero resonant-grade *ameth-* 'undying, imperishable,' rather than the fanciful and implausible *α* + *μεθύειν* 'not drunk.'

Adding to the confusion is the fact that, in the Ṛg Veda, the word *mádhu* 'mead' is frequently associated with soma; in many cases the two words are used interchangeably. It is likely that this association goes back to PIE, since in the Norse myths *mead* is used to indicate the exhilarating

drink of gods and heroes. In Greek, the reflex of PIE **med^{h}u-* 'honey, mead' was *μέθυ*, which came to signify *wine*. The Vedic scholar Arthur Macdonell writes,

> "The Soma juice, which is intoxicating, is frequently termed *mádhu* or sweet draught... The exhilarating power of Soma led to its being regarded as a divine drink bestowing immortal life. Hence it is called amrita *draught of immortality*. All the gods drink Soma; they drank it to gain immortality. ...It is possible that the belief in an intoxicating divine beverage, the home of which was in heaven, goes back to the Indo-European period. It must then have been regarded as a kind of honey mead (Skt. *mádhu*, Grk. *μέθυ*, Anglo-Saxon *medu*)."[18]

Additional candidates for inclusion with this root are (without negative prefix):

- Skt *maṭaka* 'a dead body, corpse,' *matá* 'man.' —Monier-Williams 774, Mayrhofer 296.
- Skt 1. *math* 'to trouble, disturb, afflict, distress, hurt, destroy, kill (with an additional group of meanings including: to stir or whirl around, to produce fire by rapidly whirling around or rotating a dry stick in another dry stick prepared to receive it, to churn milk into butter, to churn nectar out of the ocean of milk),' 2. *math* 'destroying, a destroyer,' *mádhu-math* 'destroyer of *Madhu* (a name of Indra),' *úrā-máthi* 'killing sheep,' *pra-math* 'destroy, lay waste, assault violently,' *pra-mathana* 'harassing, tormenting, hurting, injuring, destroying, killing,' *havir-máthi* 'destroying sacrifices.' —Monier-Williams 217, 685, 777, 780, 787, 1293; Mayrhofer II 298-99; Adams s.v. *mänt-*; Mallory and Adams (2006) 258-59; LIV 438-39; IEW 732.

Mayrhofer and Monier-Williams suggest that Skt *math* may be connected to the concepts "rub, rotate, or steal."

I would suggest that two separate roots have been combined into the lemma 1. *math*. Those attestations signifying "whirl, stir, producing fire by rotating a stick, churn, mix, mingle" belong with PIE **menth$_2$* 'churn, stir, rotate, etc.' Those with semantic value 'hurt, destroy, afflict,' belong with 2. *math* 'destroyer.' It is this second group that is cognate to *a-methyst*.

LIV and Adams (but not IEW) attempt to derive both groups from a common source via a complex semantic development that would run: *'stir' > 'disturb' > 'irritate' or 'harm' > 'destroy.' Such a development may be possible but is not likely.

- NE *matador* 'bullfighter who performs the final passes and kills the bull,' —AHD 1079; Onions 561.

Of *matador*, AHD says, "Spanish, from *matar*, to kill, possibly from Vulgar Latin **mattāre*, to beat senseless, perhaps from Latin *mattus*, stupefied."[19] Onions derives the word from Persian *māt* 'dead.' AHD notes the Arabic word *māt* as it appears in NE *checkmate* 'the king is dead,' but does not connect this with Spanish *matar* 'to kill.'

[18] Macdonell, Arthur Anthony, *A Vedic Reader for Students* (First publ. by: Oxford University Press, 1917), 152-55.
[19] This is probable as it has parallels in French *mater* 'to suppress (by violence),' e.g. *mater une rébellion*, Portuguese, Catalan and Occitan *matar* 'to kill,' Rumantsch *mazzar* 'kill' (Pierre Bancel, personal communication).

COMPARISONS TO OUTSIDE LANGUAGE FAMILIES

Bomhard 16.55 cites negative particles in **na*, **ni*, and **nu* that occur in: Afrasian *n, nn, ny, nw* 'not,' Elamite *in-, inni, ani* 'negative particle,' Kartvelian **nu, numa, numu, nəmə, nōma, nōm-* 'no, not,' Uralic **ne, në, mëm, ni, neem* 'not, nothing, nobody, never,' Altaic *an, ana* 'no, not,' Eskimo-Aleut **na-, naa, naami, nalaa, naakka, naaxxa* 'no, don't!'.

1. Bomhard 873 cites Afrasian **mar-, marāṣu, māraṣ, mrṣ, mariḍa, etc.* 'to be weakened, wither away, decay, be sick, die, perish, disease, death, with pain, etc.,' Dravidian *maṛuṅku, *mṛakku, maṛgu, margu āpini*, etc. 'be dull, blunt, obscured, to fade, disappear, be lost, die, perish, decay, molder, dead,' Eskimo **məʁnuʁ-* - 'to be tired, to be without strength.'

3. Bomhard 862 cites Afrasian **man-, mann, méena, manna, man-t, etc.* 'man, male, progenitor, person, people, woman,' Dravidian *māntar, mañja, etc.* 'human beings, male persons, man,' Uralic **man^{y}t^{y}ɜ* 'man, male.'[20]

4. AHD (Appendix II, Semitic Roots, 2065) cites the root *mt.* "Common Semitic noun **mut-* man." Dolgopolski 1493 cites Afrasian: Old Akkadian *mutum* 'husband, man, warrior,' Ugaritic *mt* 'husband, man, hero,' Biblical Hebrew (pl.) *mə'tīm* 'men, people, human beings,' East Chadic *máti* 'man,' Egyptian *mt* 'man,' Dravidian: **mut* 'aged man, old.' AHD (Appendix II, Semitic Roots, 2065) cites the root *mwt* 'to die,' NE *checkmate* 'the king must die' (check < Persian *shah* 'king,' and *mate* < Arabic *māt* 'he has died' < earlier *māta* 'to die'). Dolgopolski 1494 cites Afrasian: Common Semitic *mwt* 'die,' Egyptian *mt, mwt* 'death,' Akkadian *mūtu(m)* 'death,' Biblical Hebrew *māwet* 'death,' Somali *mōtan* 'dead,' Dravidian: Common Dravidian root **muɻ* 'die,' Tamil *muṭi* 'die.'[21]

These four roots appear to have strong affiliations to the outside language groups, suggesting that the differentiation of the resonant-variants occurred while PIE was still in contact with the linguistic community that later split up into the language families cited.

CONCLUSIONS

In my previous paper (Haynes, 2020), I have cited numerous cases where a series of PIE roots, all sharing a common semantic value, differ only in the variation of the non-initial resonant. These families of roots form an ablaut series that must have shared a common linguistic ancestor in the pre-Proto-Indo-European stage of the language.

In this paper, I have proposed two new PIE roots (****n̥-men-(t)-*** ,****n̥-me-(t)-***) and the serious expansion of one other previously accepted root (***n̥-mel-(t) -***). The impetus for this proposal arises

[20] For a wider outside comparison of this root, see Ruhlen, Merritt *Amerind Etymological Dictionary*, etymology 472 MAN$_7$ 'man' (pp. 146-147), with reflexes in Almosan, Penutian, Hokan, Paezan, Andean, Macro-Tucanoan, Equatorial, Macro-Carib, Macro-Panoan and Macro-Ge; and John Bengtson and Merritt Ruhlen's chapter 14 "Global Etymologies," in Ruhlen, Merritt, 1994, *On the Origin of Languages—Studies in Linguistic Taxonomy* (Stanford, Calif.: Stanford University Press), etymology 15 MANO 'man' (pp. 310-312), with reflexes in ?Niger-Congo, Nilo-Saharan, Afro-Asiatic, Indo-European, Uralic, Dravidian, Japanese-Ryukyuan, Ainu, Caucasian, Yeniseian, Indo-Pacific, Nahali, Miao-Yao, and Amerind.

[21] Compare Austronesian *mate* 'to die', as shown on Robert Blust's Austronesian database. Austronesian: 'to die, be dead,' whose reflexes include Bunun and several other Taiwanese Austronesian languages and a host of languages from elsewhere. https://abvd.shh.mpg.de/austronesian/word.php?v=75 (Thanks to Pierre Bancel for these references.)

from the observation that resonants often vary in what appears to be an ablaut series, and that many words with no adequate etymologies can be explained by this linguistic feature.

Each of these roots is semantically equivalent and phonetically parallel to the well-established PIE root ***n̥-mer-(t)-*** 'undying, immortal,' differing phonetically only in the variation of the non-initial resonant. If the final *t- is intrinsic to this root-family, as may very well be the case, then it would be appropriate to say that the roots differ phonetically only in the variation of the medial resonant.

The evidence suggests that these four roots must have shared a common archaic ancestor (***n̥-me(R)-(t)*** 'undying, immortal') and that their differences are due to either grammatical ablaut or to dialectical variation.

Recognition of the variable nature of PIE resonants allows us to assign etymologies with a fair degree of confidence to words of otherwise obscure origin. When corresponding forms are observed in external language families, then our degree of confidence is significantly increased.

Ancient religious narratives (myths) preserve linguistic terminology that would often otherwise have become obsolete and lost to posterity. The relationship between historical linguistics and mythology is then all the more important the farther back in time that both of these cultural artifacts are investigated.

ABBREVIATIONS USED

Adams: Adams, Douglas Q. *A Dictionary of Tocharian B*. Amsterdam: Rodopi, 1999.

AHD: *American Heritage Dictionary of the English Language*. 4th edition. Boston, New York: Houghton Mifflin Company, 2000.

AHD (1975): *American Heritage Dictionary of the English Language*. Morris, William, ed. Boston, New York: Houghton Mifflin Company, 1975.

Bailey: Bailey, L. H. *The Standard Cyclopedia of Horticulture*. Three volumes. New York, The Macmillan Company, 1935.

Bomhard: Bomhard, Allan R. *A Comprehensive Introduction to Nostratic Comparative Linguistics*: *With Special Reference to Indo-European*. First Edition. Charleston, South Carolina: 2014.

Buck (1949): Buck, Carl Darling. *A Dictionary of Selected Synonyms in the Principal Indo-European Languages*. Chicago & London: The University of Chicago Press, 1949.

CLL: Melchert, Craig H. *Cuneiform Luvian Lexicon*. Chapel Hill, North Carolina, 1993.

Daniélou: Daniélou, Alain. *The Myths and Gods of India*. Rochester, Vermont: Inner Traditions, 1991. Originally published under the title *Hindu Polytheism*. New York: Bollingen Foundation, 1964.

DELG: Chantraine, Pierre. *Dictionnaire étymologique de la langue grecque: Histoire des mots.* New Edition. Klincksieck, 2009.

De Vries: De Vries, Jan. *Altnordisches Etymologisches Wörterbuch*. 2nd edition. Leiden: E. J. Brill, 1977.

Galen: Galen. *On the Natural Faculties*. Brock, Arthur John, trans. Cambridge, Massachusetts: Harvard University Press (Loeb Classical Library), 1963.

Haynes (2020): Haynes, Gregory, "Resonant Variation in Proto-Indo-European," *Mother Tongue Journal* (Volume 22, 2020) 151-222.

IEW: Pokorny, Julius. *Indogermanisches Etymologisches Wörterbuch*. Bern: Francke Verlag, 1959.

Kluge: Kluge, Friedrich. *Etymologisches Wörterbuch der Deutschen Sprache*. 19th edition. Berlin: Walter De Gruyter & Co., 1963.

LIV: Rix, Helmut, et al. *Lexicon der Indogermanischen Verben*. 2nd edition. Wiesbaden: Dr. Ludwig Reichert Verlag, 2001.

LSJ: *A Greek–English Lexicon*, Liddell, Scott, and Jones, Oxford: Clarendon Press, 1968.

Mallory and Adams (2006): Mallory, J. P. and D. Q. Adams. *The Oxford Introduction to Proto-Indo-European and the Proto-Indo-European World*. Oxford: Oxford University Press, 2006.

Mayrhofer: Mayrhofer, Manfred. *Etymologisches Wörterbuch des Altindoarischen*. Two volumes. Heidelberg: Carl Winter Universitätsverlag, 1992.

Monier-Williams: Monier-Williams, Sir Monier. *A Sanskrit–English Dictionary*. New Edition. Delhi: Motilal Banarsidass, 1981. Originally published Oxford: Oxford University Press, 1899.

NIL: Wodtko, Dagmar S., Britta Irslinger, and Carolin Schneider. *Nomina im Indogermanischen Lexikon*. Heidelberg: Universitätsverlag Winter, 2008.

OCD: *Oxford Classical Dictionary*, Cary, M. et al., eds. Oxford: Clarendon Press, 1964.

OLD: *Oxford Latin Dictionary*, Glare, P. G. W., ed. Oxford: Clarendon Press, 1982.

Onions: *Oxford Dictionary of English Etymology*. Onions, C. T., ed. Oxford: Clarendon Press, 1966.

Vigfusson: *Icelandic–English Dictionary*, Cleasby-Vigfusson, Oxford: Clarendon Press, 1874.

Watkins (2011): Watkins, Calvert. *American Heritage Dictionary of Indo-European Roots*. 3rd edition. Boston, New York: Houghton Mifflin Harcourt, 2011.

THE CHINESE PRIMORDIAL GIANT PANGU AND HIS POSSIBLE INDO-EUROPEAN ORIGIN

VÁCLAV BLAŽEK
MASARYK UNIVERSITY

In the present contribution a hypothetical relation between the Chinese primordial giant Pangu and his Indo-European counterparts is studied on the basis of both their mythic descriptions and etymological analysis.

1. PANGU

1.1. Documentation

Traditional: 盤古; Simplified: 盘古; Modern transcription in Pinyin: *Pángǔ*; Wade-Giles: *P'an ku*. Several reconstructions have been proposed for the historical pronunciation of the signs reflecting various stages in the development of Chinese:

盤 - Beijing *pán*, Min dialects: Xiamen, Chaozhou *puã*, Fuzhou *puaŋ*, Jian'ou *puiŋ* (Starostin 1989, 66); Cantonese *p'ūn* < 'Ancient' Chinese *ⱪb'uân "*vessel; tub, tray, dish, plate" (Karlgren 1923, 213; GSR 0182e) = Late Middle Chinese [AD 900] **phuan* < Early Middle Chinese [AD 600] **ban* "basin, dish, tray, plate; to turn round; joy" (Pulleyblank 1991, 231) ~ Late Old Chinese [3rd cent. BC - 5th cent. AD] **bwân "*поднос, блюдо", i.e. "tray, dish, bowl, pan, basin" < Preclassic Old Chinese [11-8th cent. BC] **bān* (Starostin 1989, 143; 129: cf. the Siamese loan *bhān*) = **[b]ˤan* (Baxter & Sagart, 2014) = **baːn* (Zhengzhang 2003[1]).

The external cognates confirm this reconstruction: Written Tibetan *ban* "beer-jug, pitcher", *ben* "large pitcher, jug, beer-pot", Jingpo *bàn* "tray, waiter, salver", Burmese *panh-kan* "deep basin", *laŋ-panh* "metal tray" (Starostin 1989, 143) < Sino-Tibetan **pān* ~ **bān* (CVST I, 51).

古 - Beijing *gǔ*, Cantonese *ku* < 'Ancient' Chinese *'*kuo* "ancient; strange" (Karlgren 1923, 145; GSR 0049a) = Late Middle Chinese *kuǎ´* < Early Middle Chinese **kɔ*' "olden times, former times" (Pulleyblank 1991, 111) ~ Middle Chinese [7th-10th AD] *kó* < Postclassic Old Chinese [3rd-5th cent. AD] **kǒ* < East & West Han Chinese **kǎ* [3rd cent. BC - 3rd cent. AD] = Early Classic Old Chinese **kǎ* [5th-3rd cent. BC] < Preclassical Old Chinese [8-11th cent. BC] *kā*'

[1] <https://en.wiktionary.org/wiki/盤>

(Starostin 1989, 272, 521, 687) = **ka:ʔ* (Zhengzhang 2003[2]) = Archaic Chinese **kā*ʾ "antiquity, of old" (Schuessler 2007, 259) = **kˁaʔ* (Baxter & Sagart, 2014). The word and its sign are already known from the oracle inscriptions on bones and bronzes from the Shang and Western Zhou dynasties respectively[2]:

It is also attested in the text *Shījīng* (*c*. 600 BC). The external cognates confirm the Archaic Chinese reconstruction **kā*ʾ: Written Tibetan *rga-ba* "to be old, aged", Jingpo *ləga* "to be old, hardened, stiff" (Benedict 1972, #445; Schuessler 2007, 259) < Sino-Tibetan **kāH* (CVST V, 42).

1.2. Sources

The following five source-texts including their translations are taken from Kósa (2009, 171, 179-181). The first secure source recording the myth of Pangu are two treatises, which have been ascribed to the Daoist writer Xu Zheng (徐整) living AD 220-265, i.e. during the Three Kingdoms (三國) period. *Xú Zhěng* (Chinese: 徐整; Wade-Giles: *Hsü Cheng*).

1.2.1. ***Sānwǔ lìjì*** **三五曆紀** "Three Five Historic Records", literally: "Three Five[3] Calendar", as preserved in the *Yiwen leiju* **藝文類聚** ("Classified Collection of Literary Writings") 1.4a–b. (compiled by Ouyang Xun **歐陽詢** et alii, AD 624)

> "[1] Xu Zheng's *Sanwu liji* says: [2] "The sky and the earth were chaotic like an egg, and **Pangu** was born in it. [3] It took 18,000 years for the separation of the sky and the earth: the *yang* and the pure became the sky, the *yin* and the impure became the earth. [4] **Pangu** was in-between, nine times did he transform daily, he became more spiritual than the sky, more saintly than the earth. [5] The sky grew one *zhang* higher every day, the earth grew one *zhang* thicker every day, **Pangu** also grew one *zhang* bigger every day. 18,000 years passed like this. [6] The sky reached its utmost height, the earth reached its utmost depth, **Pangu** (reached) his utmost size, and then the Three Sovereigns were born."[4]

> [1] 徐整《三五曆紀》曰: [2]「天地渾沌如雞子，盤古生其中。**[3]** 萬八千歲，
> 天地開闢，陽清為天，陰濁為地。[4] 盤古在其中，一日九變，神於天，聖於地。
> [5] 天日高一丈，地日厚一丈，盤古日長一丈。如此萬八千歲。[6] 天數極高，
> 地數極深，盤古極長，後乃有三皇。

---

[2] <https://en.wiktionary.org/wiki/古>

[3] The numerals "3" & "5" refer to the "Three August Ones and Five Emperors" (三皇五帝).

[4] Cf. the alternative translation by Mair 1998, 14:
"Heaven and earth were in chaos like a chicken's egg, and **Pangu** was born in the middle of it. In eighteen thousand years Heaven and earth opened and unfolded. The limpid that was Yang became the heavens, the turbid that was Yin became the earth. **Pangu** lived within them, and in one day he went through nine transformations, becoming more divine than Heaven and wiser than earth. Each day the heavens rose ten feet higher. each day the earth grew ten feet thicker, and each day **Pangu** grew ten feet taller. And so it was that in eighteen thousand years the heavens reached their fullest height, earth reached its lowest depth, and **Pangu** became fully grown. Afterward, there were the Three Sovereign Divinities. Numbers began with one, were established with three, perfected by five, multiplied with seven, and fixed with nine."

1.2.2. ***Wǔyùn lìniánjì*** **五運曆年紀** "Chronicle of the Five Cycles of Time" in Ma Su's **馬驌** (1621–1673) *Yishi* **繹史** "Unravelling History" 1.2a.

> "[1a] When the first-born **Pangu**'s death was approaching, he transformed his body, [1b] his breath became the wind and the clouds, his voice became the thunder, [1c] his left eye became the sun, his right eye became the moon, his four limbs and the five parts of his body became the four extremities and the five (sacred) mountains, his blood became the rivers, his arteries became the arteries of the earth, his flesh became the soil, his hair became stars and constellations, his skin became grasses and trees, his teeth and bones became minerals and stones, his marrow became precious stones and jades, his sweat and his fluids became rain and marshes, and the insects on his body, when they were touched by the wind, were transformed into the common people."[5]
>
> [1a] 首生盤古垂死化身，[1b] 氣成風雲，聲為雷霆，[1c] 左眼為日，右眼為月，
> 四肢五體為四極五嶽，血液為江河，筋脈為地里，肌肉為田土，發為星辰，皮膚為草木，
> 齒骨為金石，精髓為珠玉，汗流為雨澤，身之諸蟲，因風所感，化為黎甿。

1.2.3. ***Wǔyùn lìniánjì*** **五運曆年紀** (see 1.2.2.) in Dong Sizhang **董斯張**: *Guang bowu zhi* **廣博物志** "The Enlarged Treatise on Research into Nature", 9.2b. Also quoted as *Diwang Wuyun linian ji* **帝王五運歷年紀** in the earlier *Shiwu jiyuan* **事物紀原** "The Origins of Affairs and Things" (1.4a−7a) by Gao Cheng **高承**.

> "[1] Lord **Pangu** had a dragon head and a serpent body, [2] his breathing became wind and rain, his breathing out became thunder. When he opened his eyes, it was daylight, when he closed his eyes, it became night. [3a] After his death [3b] his bones and joints became mountains and forests, his limbs became the rivers and seas, his blood became the Huai river, his hair became grasses and trees."
>
> [1] 盤古之君龍首蛇身，[2] 噓為風雨，吹為雷電，開目為晝，閉目為夜。
> [3a] 死後 [3b] 骨節為山林，體為江海，血為淮瀆，毛髮為草木。

1.2.4. ***Shùyìjì*** **述異記** "A Record of Accounts of Marvels" by Ren Fang **任昉** (460–508 CE) in *Han-Wei congshu* **漢魏叢書** "Han and Wei Dynasties Collectanea", compiled by Cheng Rong **程榮** (fl. 1573−1620), preface from 1592.

> "[1] In former times when **Pangu** died, his head became the four (sacred) mountains, his eyes became the sun and the moon, his fat became the rivers and the seas, the hair on his body and on his head became the grasses and the trees. [2] In Qin and Han, people said: '**Pangu**'s head became the eastern mountain, his stomach became the central mountain, his left arm became the southern mountain, his right arm became the northern mountain, his legs became the western mountain.' [3] Former literati said: '**Pangu**'s weeping became the rivers, his breath became wind, his voice became the thunder, the pupils in his eyes became the lightning.' The ancients said: '**Pangu**'s joy became good weather, his anger became cloudy weather.' [4] In Wu and Chu they say: '**Pangu** and his wife are the beginning of *yin* and *yang*.' [5] Nowadays at the southern sea there is **Pangu**'s tomb, which is 300 *li* long. The folk say that his descendants buried his soul here. In Guilin [today Guangxi] there is a temple dedicated to **Pangu**, where people pray and offer sacrifice to him. At the southern sea there is a land called

[5] Cf. the alternative translation by Mair 1998, 15:
"When the firstborn, **Pangu** was approaching death, his body was transformed. His breath became the wind and clouds; his voice became peals of thunder. His left eye became the sun; his right eye became the moon. His four limbs and five extremities became the four cardinal points and the five peaks. His blood and semen became water and rivers. His muscles and veins became the earth's arteries; his flesh became fields and land. His hair and beard became the stars; his bodily hair became plants and trees. His teeth and bones became metal and rock; his vital marrow became pearls and jade. His sweat and bodily fluids became streaming rain. All the mites on his body were touched by the wind and were turned into the black-haired people."

Pangu, nowadays everybody there has **Pangu** as a surname. (Ren) Fang's commentary: '**Pangu** is the ancestor of the sky, the earth and the myriad things, living creatures originate from **Pangu**.'"

[1] 昔盤古氏之死也，頭為四岳，目為日月，脂膏為江海，毛髮為草木。[2] 秦漢間俗說：盤古氏頭為東岳，腹為中岳，左臂為南岳，右臂為北岳，足為西岳。[3] 先儒說：盤古氏泣為江河，氣為風，聲為雷，目瞳為電。古說：盤古氏喜為晴，怒為陰。[4] 吳楚間說：盤古氏夫妻，陰陽之始也。[5] 今南海有盤古氏墓，亘三百里，俗云後人追葬盤古之魂也。桂林有盤古祠，今人祝祀。南海有盤古國，今人皆以盤古為姓。昉按：» 盤古氏，天地萬物之祖也而生物始於盤古。«

1.2.5. Chengguan 澄觀 (738−839): ***Dafangguang fo Huayanjing suishu yanyi chao* 大方廣佛華嚴經隨疏演義鈔** "Subcommentary on The Flower Garland Sūtra", T36, n1736, on Śikṣānanda's translation of *Avataṃsaka Sūtra* or "The Flower Garland"

"[1] The *History of the Three Kings* says: [2] "The sky and the earth were chaotic, **Pangu** was born in it. [3] Nine times did he transform daily, he became more spiritual than the sky, more saintly than the earth, more powerful than the sky and the earth. [4] The sky grew one *zhang* higher every day, the earth grew one *zhang* thicker every day, **Pangu** also grew one *zhang* bigger every day. 18,000 years passed like this, and afterwards the sky and the earth got separated. [5] **Pangu** had a dragon body and a human head. [6] His head reached the edge of the east-west direction, his legs reached the edge of the east-west direction, his left hand reached up to the southernmost end, his right hand reached up to the northernmost end. [7] When he opened his eyes, it was daylight, when he closed his eyes, it became night. When he breathed out, it was summer, when he breathed in, it was winter. The air he blew became wind and clouds, his shouting voice became thunder. [8] When **Pangu** died, his head became the *jia*, his throat became the *yi,* his shoulders became the *bing,* his heart became the *ding,* his bladder became the *wu,* his spleen became the *ji,* his ribs became the *geng,* his lungs became the *xin,* his kidney became the *ren,* his feet became the *gui.* [9] His eyes became the sun and the moon, his beard became the stars and constellations, his brow became the pole of the Big Dipper, the nine holes became the nine continents, his chest became the Kunlun (mountain), his knee became the southern mountain, his thigh became the Taishan, his buttocks became fish and tortoises, his hands became flying birds, his claws became turtles and dragons, his bones became gold and silver, his hair became grasses and trees, his hair on his body became mallards and ducks, his teeth became gems and rocks, his sweat became rain and water, his large intestines became rivers and seas, his small intestines became the Huai and the Si rivers, his blisters became the hundred rivers, his face became caves."

[1] 案《三王曆9》云: [2]「天地渾10沌，盤古生其中。[3] 一日九變，神於天，聖於地，主於天地11。[4] 天日高一丈，地日厚一丈，盤古亦長一丈，如此萬八千年 12，然後天地開闢 13。[5] 盤古龍身人首。

[6] 首極東西，足極東西 14，左手極南，右手極北。[7] 目成晝 15，合目成夜。呼為暑，吸為寒。吹氣成風雲，叱 16。聲為雷霆 17。

[8] 盤古死，頭為甲，喉為乙，肩為丙，心為丁，膽為戊，脾為己，脊為庚，肺 為辛，腎為壬，足為癸，

[9] 目為日月，髭為星辰，眉為斗樞，九竅為九州，乳為崑崙，膝為南嶽 18，股為太 19 山，尻 20。為魚鱉，手為飛鳥，爪為龜龍，骨為金銀，髮為草木，毫毛為鳧鴨，齒為玉石，汗為雨水 21，大腸為江海，小腸為淮泗，膀胱為百川，面輪為洞庭。」

1.3. Interpretation of the name

1.3.1. Paul Carus (1907[1974][6]) offered the following explanation:

P'an Ku is written in two ways: one means in literal translations, "**basin ancient**", the other "**basin solid**". Both are homophones, i.e., they are pronounced in the same way; and the former may be preferred as the original and correct spelling. Obviously the name means "**aboriginal abyss**" or in German, **Urgrund**, and we have reason to believe it to be a translation of the Babylonian Tiamat, "**the Deep**". But this explanation is rather problematic. The metaphor "basin, vessel" instead of "abyss" is not quite natural. The goddess *Tiamat* according to the Akkadian myth *Enuma eliš* was a personification of the primordial salt-water ocean. It is a common Semitic term: Akkadian *tiamtu(m)* "the deep, sea"; Eblaite *ti-'a-ma-tum*, *ti-'à-ma-tum* /*tihām-at-um*/; Ugaritic m. *thm* "ocean; the Ocean", f. *thmt* "primordial ocean; abyss"; Hebrew *t^{e}hōm* "primaeval ocean, primaeval flood; flood, deluge" > Jewish Aramaic *t^{e}hōmā*, Syriac *t^{e}hōmā*, pl. *t^{e}hōmātā*; Arabic *Tihāmā* 'the coastal plain along the southwestern and southern shores of the Arabian Peninsula' (Koehler & Baumgartner 2001, 1690-91; del Olmo Lete & Sanmartín 2003, 864).

1.3.2. It seems more probable that the Chinese compound represents an adaptation of a foreign theonym (this conclusion is supported by the strange syntax, as George Starostin mentions). There are significant details connecting the myth of Pangu with one of the most important Indo-European myths of the primordial giant, whose body served as a building material for both the heavens and earth and their inhabitants.

1.3.2.1. The most suggestive parallels appear in the so-called *Púruṣa*-hymn [RV X, 90]

13. *candrámā mánaso jātáś cákṣo sryo ajāyata*
a) The Moon was gendered from his mind, and from his eye the Sun had birth,
b) The moon was born from his mind. From his eye the sun was born.

múk^{h}ād índraś cāgníś ca prāṇā́d vāyúr ajāyata
a) Indra and Agni from his mouth were born, and Vāyu from his breath.
b) From his mouth Indra and Agni, from his head the heaven developed.

14. *nā́b^{h}yā āsīd antárikṣaṃ śīrṣṇó dyaú sám avartata*
a) forth from his navel came mid-air; the sky was fashioned from his head,
b) From his navel was the midspace. From his head the heaven developed.

padb^{h}yā́m b^{h}ū́mir díśa śrótrāt tát^{h}ā lokā́m̐ akalpayan
a) Earth from his feet, and from his ear the regions. Thus they formed the worlds.
b) From his two feet the earth, and the directions from his ear. Thus they arranged the world.

[Translated by a) Griffith 1889, 559-60; b) Jamison & Brereton 2014, 1540]

1.3.2.2. A similar motif appears in the *Poetic Edda* in the description of cutting up the body of the primordial giant Ymir in portions:

§40. *Ór Ymis holdi*	§ 40. Out of Ymir's flesh
var iǫrð um skǫpoð,	was fashioned the earth,
en ór sveita sær,	And the ocean out of his blood;
biǫrg ór beinom,	Of his bones the hills,

[6] <http://en.wikipedia.org/wiki/Pangu>.

baðmr ór hári, *en ór hausi himinn*;	of his hair the trees, Of his skull the heavens high.
§41. *en ór hans brám* *gørðo blíð regin* *miðgarð manna sonom*; *en ór hans heila* *vóro þau in harðmóðgo* *ský ǫll of skǫpoð.*	§41. Mithgarth the gods from his eyebrows made, And set for the sons of men; And out of his brain the baleful clouds They made to move on high.

[Translated by Henry Adams Belows (1936)- see <http://www.sacred-texts.com/neu/poe/poe06.htm>]

1.3.3. Although the correspondences between the Chinese and Indo-European traditions are not one-to-one, they are so apparent that it is legitimate to think about a cultural influence. The first candidate could be the Old Indic cultural circle. The cultural vector directed from India to China brought Buddhism into China. But it is not very probable that the Buddhist missionaries propagated the Vedic deity. It is also impossible to explain the name *Pangu* from *Púruṣa-*. But there were two Indo-European branches which were in a direct contact with the Chinese civilization: Iranian and Tocharian. In this contribution both of these hypothetical mediators are discussed.

2. ETYMOLOGICAL INTERPRETATION OF *PÚRUṢA-*

To date there is no unambiguous etymology of ***Púruṣa-***. A survey of proposed etymologies was summarized by Mayrhofer (KEWA II, 312-13; III, 760; EWAI II, 149-50), cf. also Bailey 1979, 230. The present new etymology is based on two synonymous rhyme-words: *púruṣa-*, *pū́ruṣa-* "man, male, human being, person", pl. "people, mankind" & *mánuṣa-* "man" (but also the athematic form *mánuṣ-* "man; Manu - father of men"), all attested already in RV. With regard to the forms *pūrú-* "man"(?) and *mánu-* "man, mankind; Manu - father of men" it is apparent that *-ṣ(a)-* is a derivational suffix. Brugmann (1906, 535) speculated about the primary abstract function of the nouns extended in *-ṣ-*, namely *mánuṣ-* *"Menschtum, Menschheit" (cf. the Avestan proper name *Manuš-čiθra-*). A key to the origin of this suffix could be found in the early Anatolian suffix *-hsu-* forming numerous personal names attested in the Old Assyrian tablets from Kültepe in Cappadocia. Analyzing the anthroponyms in their internal structure, it is possible to conclude that the primary function of the suffix *-hsu-* was "born". In this case the proper names can be interpreted as follows (see Laroche 1966, 297-303):

Apiziahsu- 'Latter-born' : *appezi-* "last(-born), backmost, hindmost",
Arawahsu 'Born as free' : *arawa-* "free",
Arzanahsu- 'Born in brothel' : *arzana-* "inn, hostel, brothel",
Nakiahsu 'Born as hard' : *nakki-* "hard, heavy",
Supiahsu 'Born as holy' : *suppi-* "pure, holy",
Udniahsu : *udne-* "country, land";

anthroponyms formed from toponyms:

Pesahsu 'Born in Pisa' : ^{URU}*Pisa*;

or personal names formed from divine names:

Ilaliahsu 'Born by Ilali' : *Ilali*,
Peruahsu 'Born by Perua' : *Peru(a)*,
Taruhsu- 'Born by Taru' : *Taru*, i.e. the 'Storm-god'.

If the determination of the function of the suffix *-hsu-* is correct, it is possible to seek its origin in the primary appellative attested in Hittite *has(s)-* "to beget, procreate, engender, produce, bear, give birth, bring to birth", extended by the *-u-* with a passive function (Laroche 1966, 301 following Goetze; more problematic is the hypothesis of Hamp quoted also by Laroche, seeking here the verb **seu̯H-* "to give birth"). Puhvel (3, 245, 212-18) adds other Anatolian comparanda: Hieroglyphic Luwian *has(a)-* "to beget", *hasmi-* "progeny, issue" = Milyan *qezm̃mi*?

In the light of the Anatolian anthroponyms it is possible to interpret the sigmatic designations of "man" in Vedic in two ways, (i) passive and (ii) active:

mánuṣ(a)- = (i) "born by *mánu-*" or (ii) "giving birth to *mánu-*", where *mánu-* = "man; Manu"; cf. also epic *manuja-* "born by Manu" = "man" [MBh];

pū́ruṣa- = (i) "born by *pūrú-*" or (ii) "giving birth to *pūrú-*", where *pūrú-* = "man" or "men" (EWAI II, 156);

púruṣa- = (i) "born by *púru-*" or (ii) "giving birth to *púru-*", where *púru-* has nothing in common with men, but it means "many, much, abundant", cf. also *purú víśva* "one and all, every".

With respect to this analysis it seems there were two different terms:

pū́ruṣa- = "born by *pūrú-*", where *pūrú-*[7] = "man", parallel to *mánuṣ(a)-* = "born by *mánu-*", where *mánu-* = "man"[8] or 'the first man Manu';
púruṣa- = "giving birth to many".

The interpretation "manifold, multiple" represents a truthful description of the primordial giant *Púruṣa-*, who is described in RV X, 90.1 as

sahásraśīrṣā púruṣa sahasrākṣá sahásrapāt
"thousand-headed is Puruṣa, thousand-eyed, thousand-footed".

7 According to Kuiper 1991, 7 a tribal name of non-Aryan origin.
8 This idea is based on the metaphor "man / human being" = "son of mankind", wide-spread in the ancient languages of the Near East (see Blažek 2008, 57-58): Sumerian DUMU.NAM.LÚ.ULÙ^{LU}, Akkadian *mār awīlūti* and *mār nišī* "man" in the sense "human being", lit. "son of mankind"; Ugaritic *bn nšm* "men", i.e. "sons of mankind", and the compound *bnš* "man, an individual, someone, person; people, personnel; service personnel; farmhand, labourer", pl. *bnšm*, cf. also *bnš bnšm* "every man", lit. "a man of men"; plus syllabic *bu-nu-šu* < **bun-nōš-*; similarly Hebrew *bɛn-ʔĕnōš* (Ps. 144:3), Aramaic *bar ʔĕnāš*. Ugaritic *bn ảdm* "man" corresponds with Hebrew *bɛn-ʔadam* (Nu. 23, 19). Similar formations are also known in IE: Vedic *púmāṁs-*, nom. sg. *púmān*, voc. sg. *púmas* etc. "man", originally perhaps a compound "child of man", where the first component could correspond with Greek παῖς "child", Gothic pl. *fawai* "few", Latin *paul(l)us* "little" etc., and the second component is connected with Vedic *mánu-* and / or Latin *mās*, gen. *maris*, *masculus* (Eichner 1974, 39-40; Čop 1976, 25-28). Similarly Slavic **čelověkъ* "man, human being" is analyzable as German *Menschenkind* "human being", lit. "child of men or people" (Ivanov 1975, 20).

With regard to his role in the process of creation, when his body was used as a building material for many (better *all*) objects, from the earth and heavens to animals and men, it is possible to use the epithet "multigenitor" or even "omnigenitor". The latter interpretation is explicitly expressed in RV X, 90.2:

púruṣa evédáṃ ***sárvaṃ*** *yád b*^{h}*ūtáṃyác ca b*^{h}*áviyam*
"This Puruṣa is **all** that yet hath been and **all** that is to be"
[translated by Ralph Griffith]

"The Man alone is this **whole** (world): what has come into being and what is to be."
[translated by Stephanie W. Jamison & Joel P. Brereton].

These conclusions agree with the information recorded in the Chinese text in §1.2.4. that ... "living creatures originate from Pangu", while the text in §1.2.2. says that ... "insects on his body ... were transformed into the common people."

3. TOCHARIAN DATA

Thanks to the Buddhist conversion of the Tocharians no information on pre-Buddhist deities is preserved in Tocharian texts. But the existence of the parallel creation myths in Germanic and Indo-Aryan traditions guarantees their common Indo-European origin. It implies their possible existence in other branches too, including Tocharian, before the spread of new religions. If there are no pre-Buddhistic theonyms in Tocharian texts, it is necessary to seek a hypothetical source of Chinese *Pangu* among appellatives. With respect to the semantic analysis of *Púruṣa-* as "multigenitor" or "omnigenitor" it is legitimate to seek in the semantic field "many, much, all". In Tocharian there are good candidates A *puk*, *pont/c/ś*^{o} and B *po*, *pont/c*^{o} "each, every, all"[9], derivable from the ablaut opposition zero vs. *o*-grade and following levelling, which is comparable with the paradigm of Greek πᾶς, ntr. πᾶν, f. πᾶσα, Arcadian πάνσα, gen. sg. παντός, nom. pl. πάντες, acc. pl. πάντας, levelled from the primary opposition nom. sg. m. **pónts* : gen. sg. *patós* < **pŏnts* : **pn̥tós* (cf. van Brock 1972, 276; Hilmarsson 1986, 110, 214-15 added initial **H*$_1$- seeing here a participle from the verb **H*$_1$*ep-* "to take"; but in this case one would expect $^{+}$*epant*^{o} in Greek).

[9] Tocharian A: Poucha 1955, 181-83: "omnis, cunctus, quisque"; Tocharian B: Adams 1999, 402: "all, every, each, complete"; Tocharian A & B: Pinault 2008, 522-24: sg. "totus", pl. "omnes"; he reconstructs **pānt-* < **peH*$_2$*-nt-*.

case	Tocharian A	Tocharian B	internal reconstruction
nom. sg. m.-f.	*puk*		**pn̥k^{w}(t)-s*[10]
nom.-obl. m.-f.		*po*	**pŏn(k^{w})t-s*
obl. sg. m.	*poñcäṃ*	= nom.	**pn̥(k^{w})t-en-m̥*
obl. sg. f.	*pontsāṃ*	= nom.	**pn̥(k^{w})t-iH$_2$-n-m̥*
nom. pl. m.	*poñś^{ä}*, *poṃś*, *poś*[11]	*poñc*	**pn̥k^{w}-es / *pn̥(k^{w})t-es*
nom.-obl. pl. f.	*pont*	*ponta*	*pn̥(k^{w})t-H$_2$*
obl. pl. m.	*poñcäs*	*pontäṃ*	**pn̥(k^{w})t-n̥s*
gen. pl. f.	*pontāśśi*		**pn̥(k^{w})t-H$_2$-n̥s-Tei*
		pontäṃts	**pn̥(k^{w})t-H$_2$-n̥s-ōm*

4. IRANIAN DATA

At least two literary Middle Iranian languages, namely Sogdian and Sakish, were in direct contact with Chinese. But in their known lexical corpora there are no promising candidates for both a counterpart of Vedic *Púruṣa-* and a source of Chinese *Pangu*[12]. In the remnant Iranian languages from Pamir and Hindukush there are more promising parallels: Wakhi *pak* "each, every(body)" (Paxalina 1975, 231; Steblin-Kamenskij 1999, 256), Bartangi *fuk & fúk-aθ* "every, all" (Sokolova 1960, 111-12; *-aθ* is an intensifier), Shughni *fuk & fuk-aθ*, Sarikoli *fyk* "all" (Tomaschek 1880, 818), maybe also Munjan *po*, pl. *pówi* id. (Paxalina 1983, 25, 202; Grjunberg 1972, 343), derivable from Iranian **pak(u)-/*fak(u)-/*bak(u)- < *p$^{(h)}$/b$^{(h)}$n̥k(u)-* (see Paxalina 1983, 156-57, 173 about the phonological development). Their relation to Persian *pāk* "omnis, totus" and Ossetic Iron *fag*, Digor *fagæ* "enough, sufficient(ly)" (Abaev I, 416), implying the initial **p-*, is not clear. An alternative cognate in Young Avestan *fånku-*, attested in nom. pl. *fånkauuō* "(mountain) peaks" [Yašt 19.3; see Bartholomae 1904, 973] would indicate the initial **p^{h}-*. Bailey (1979, 190) connected the Avestan word with designations of various monsters, as Persian *nihang* "monster", Syriac *nhng-*' "crocodile" or "hippopotamus", Armenian *nhang* "water beast", reconstructing their Middle Iranian source in the form **ni-fanka-*. According to him, Sanskrit *pra-pañca-* "expansion" [Kāvya; Kathāsaritsāgara, etc.], *pra-pañcaka-* "multiplying" [Harṣa carita] can also be related, but these words have been derived from *pañca* "five" (cf. EWAI III, 297). Let us mention that in the various Chinese versions of the myth about Pangu there appear both the mountains originating from the

[10] Reconstruction of van Brock (1972, 272-73) and Schwartz (1992, 423). Otherwise Pinault (2008, 523) who separates the final *-k* in A *puk* and explains it as a secondary extension influenced by the word *māk* (= B *māka*) "many; numerous". On the other hand, Hilmarsson (1986, 110) saw in *puk* a suppletive base. In principle, it could be a continuant of the synonym **b^{h}n̥ĝ^{h}u-*.

[11] Tocharian *ś* has its origin in palatalized **d* and velars, not **t*, changing into *c* in the palatal context (Pinault 2008, 423). Tocharian A nom. pl. m. *poñś^{ä}*, *poṃś*, *poś* reflects **pn̥k^{w}-es* or **pn̥kw-es*, parallelly the nom. sg. m. *puk* is derivable from **pn̥k^{w}-s* or **pn̥kw-s*. It seems that the *t*-extension appears in other cases than the nominative. The reason could consist in suppletion, proposed already by Hilmarsson.

[12] A hypothetical compound of Khotanese *pana-* "each, every", pl. "all" (with Osset *fæjnæ* "every" have been derived from Iranian **patina-*; see Bailey 1979, 209) and the derivative of the verb **gau-* "to grow" (> Khotanese *gvāna-* "growth", Young Avestan *gūnaoiti-* "verschafft" - see Bailey 1979, 96; Bartholomae 1904, 504) looks too artificial.

limbs of Pangu (§§1.2.3., 1.2.4.) and the information about the dragon-like head and serpent-like body of Pangu (§1.2.3.).

5. OTHER POSSIBLE INDO-EUROPEAN COGNATES (CF. BLAŽEK 1999, 228-29 ABOUT POSSIBLE ETYMOLOGY OF IE **penkᵘe* "FIVE")

Hittite *panku-* adj. "all, entire, complete; every; general", n. c. "multitude, the people, the masses; assembly", v. *pankuēšš-* "to become plentiful" (Kloekhorst 2008, 624-25; he connects it with Vedic *bahú-* "many, much, frequent, numerous" etc., from **bʰn̥ǵʰu-*); ?Cuneiform Luwian *pūna-* "all", perhaps via vowel metathesis or from **Pn̥Kwo- > *pn̥wa- > *pwn̥(w)a-*.

Greek ἅπαξ "(only) once, once for all", if it was syncopated from *ἅπακυς < **sm̥-pn̥kʷu-s* "one-all" (Meyer 1993, 44-45).

Latin *cunctus* "all, whole" < **kwonkto- < *kwenkwto- < *pₑnkwto- < *pn̥Kwto-* (Hamp 1973, 169-70); Umbrian nom.pl. *puntes*, abl.pl. *puntis*, referring to a group of priests, attested in the Iguvinian Tables (III.9,10/III.4), is usually translated "quiniones", but Polomé established perhaps a preferable meaning "all" or "the whole group" (1966, 233; 1968).

6. CONCLUSION

6.1. Middle and older Chinese transcriptions of foreign words, usually proper names, reflect the opposition between voiced and voiceless labials:

(i) ***b***

Middle Chinese 浮屠 *bi̯u-dou* = personal name/epithet *Buddha* (Pulleyblank 1962, 213)
Middle Chinese 梵 *bi̯am`* < **blāmh* = title *bramma* < *brahma* (Pulleyblank 1962, 114, 231)
Middle Chinese 蒲盧 *bou-lou* & 婆羅 *bα-lα* = place name *Bālā* (Pulleyblank 1962, 214)
Middle Chinese 高附 *kau-bi̯ou`* = place name *Kabul*, Κάβουρα [Ptolemy, VI, 18.5] (Pulleyblank 1962, 223)

(ii) ***p***

Middle Chinese 邠祁 *pi̯in-nα* < **plə̂n-nα = Pūrṇa-* in personal name *Pūrṇamaitrāyaṇīputra-* (Pulleyblank 1962, 124)
Middle Chinese 賓頭盧 *pyin-du-lou* = place name *Piṇḍola* (Pulleyblank 1962, 214)
Middle Chinese 襆達 *puk-dαt* < **phōk-ljāw* = place name *Puśkalāvatī-*, Πευκελαῶτις (Pulleyblank 1962, 101; Starostin 1989, 452)
Middle Chinese 虎魄 *hou´-phak* "amber" < **hα-phlαk* < **hā-phrāk* = Greek *ἄρπαξ "amber" [recorded by Pliny] (Pulleyblank 1962, 124; Starostin 1989, 453).

6.2. These examples of Middle and earlier Chinese transcriptions are based on proper names and titles of Indo-Aryan and Iranian origin. If it is also possible to generalize them for adaptation of foreign words from other sources, the expected protoform in a hypothetical donor-language should look like ± **bānkå*, if the adoption was realized in the 3rd cent. AD, i.e. at the time of the first attestation, or ±**bānkā*, if this process happened earlier. In the first case it is possible to think about Tocharian A *puk*, pl. *poñśᵃ̈*, *poṃś*, *poś* "all", derivable also from **bʰn̥ǵʰwᵒ*, perhaps via **bānkwᵒ*

with a loss of the nasal known e.g. from the Tocharian B verb *tsāk-* "to pierce, bite" (cf. van Brock 1971, 290). It is also necessary to accept that at least in the initial position the originally voiced aspirated labial still preserved its voice. From the point of historical phonetics of Tocharian this could be important information for the chronology of the process of devoicing.

6.3. George Starostin (p.c. July 22 and Sept 29, 2010) drew my attention to the variant 盤瓠 *pán hù*, where the second member of this compound is the word designating "bottle-gourd, melon, calabash", in Late Middle Chinese *xɦuǎ`* < Early Middle Chinese **γɔ^{h}* (Pulleyblank 1991, 128) = Middle Chinese [7th-10th AD] **γo*, Postclassic Chinese [3rd-5th cent. AD] and West & East Han Chinese [3rd cent. BC - 3rd cent. AD] **gā ~ *γā*, Classic Old Chinese [5th-3rd cent. BC] **g(h)wā*, Preclassic Old Chinese [11-8th cent. BC] **g(h)^{w}ā*, with the alternative reconstruction **g(h)^{w}ā-s* (Sergei Starostin 2005) = **[ɢ]^{wś}a-s* (Baxter & Sagart, 2014), while Zhengzhang[13] reconstructed **g^{w}la:s*. The Old Chinese reconstruction of the signs 盤瓠 **bān g(h)^{w}ā(s)* supports the solution formulated above, i.e. the source of the type Tocharian A *puk*, pl. *poñś^{äṃ}*, *poṃś*, *poś* "all", derivable from **b^{h}n̥g^{h}w^{o}*, via **bā^{n}kw^{o}* with the labial umlaut and a loss of the nasal in the nom. sg. In this case, when the labialized velar was preserved in the Old Chinese period, the adaptation could already have been realized during the first millennium BC. This conclusion is in agreement with the idea of Hilmarsson (1986, 110; see fn. 4) who proposed the suppletivism in Tocharian A nom. sg. *puk*, nom. pl. *poñś^{äṃ}*, *poṃś*, *poś* "all", vs. obl. *pont^{o}* / *poñc^{o}*, reflecting two synonymous, but different bases **b^{h}n̥g^{h}w^{o}* vs. **pn̥t^{o}* (**H$_{1}$pn̥t^{o}* according to Hilmarsson 1986, 110, 214-15).

6.4. In Iranian the most promising candidates for a source are Young Avestan pl. *fånkauuō* "(mountain) peaks" with hypothetical cognates in Middle Iranian **ni-fanka-* "monster" (→ "giant"; cf. the dragon-like head of Pangu in the text in §1.2.3.) and / or in the word **p/f/bak(u)-* "each, every, all" in the minor languages from the Pamir-Hindukush borderland. Iranian prevocalic **f-* is an extremely rare phoneme. Its origin is more probable from **b^{h}-* than from **p^{h}-*, cf. also Persian *faγfūr* "divine son" < Iranian **baga-putra-* (Bailey 1936, 1054 = 1981, 302) or Gilaki *faγ / fiγ* "hornbeam" < Iranian **bāga-* < **b^{h}āgo-* "beech" (Henning 1963, 68-69, fn. 2). Both solutions remain so only on a hypothetical level.

6.5. The Chinese tradition itself connected Pangu with the West or Northwest of the Chinese world (cf. Mair 1998, 12-14).

APPENDIX: ON ETYMOLOGY OF *YMIR*

The name of Ymir, the Scandinavian counterpart to *Púruṣa-*, has usually been projected in proto-Germanic as **Jumijaz*, derivable from **i̯n̥mi̯o-*, and further compared with the Indo-Iranian first mortal and ruler of the dead **Yama-* whose name also meant "twin" (Puhvel 1975, 146-57). Although this solution is attractive and generally accepted, it has no etymological support in Germanic itself. The present alternative solution is based on both direct cognates in Germanic and external typological parallels. Gothic *iumjo* "crowd" (only in nom. pl. *iumjons managos* 'ὄχλοι πολλοί',

[13] <https://en.wiktionary.org/wiki/瓠>

i.e. “large crowds”), Old Swedish *ympnian* “abundant” (Lehmann 1986, 208), if related, locate *Ymir* in the same semantic field as *Púruṣa-*, whose name is derivable from *púru-* “many, much, abundant”. Like *Púruṣa-*, *Ymir* was also characterized as he from whom all come, as recorded in The Prose Edda of Snorri Sturluson:

***Gylfaginning* “Beguiling of Gylfi”**

§5.7.	§5.7.
Eru völur allar	All the witches
frá Viðolfi,	spring from Witolf,
vitkar allir	All the warlocks
frá Vilmeiði,	are of Willharm,
en seiðberendr	And the spell-singers
frá Svarthöfða,	spring from Swarthead;
jötnar ***allir***	**All** the ogres
frá Ymi komnir.	of Ymir come.
§5.8.	§5.8.
Ór Élivágum	Out of the Ice-waves
stukku eitrdropar	issued venom-drops,
ok óx, unz ór varð jötunn;	Waxing until a giant was;
þar órar ættir	Thence are our kindred
kómu ***allar*** *saman;*	come **all** together;
því er þat æ allt til atalt.	So it is they are savage forever.

[Text: <http://www.heimskringla.no/wiki/Gylfaginning>
Translated by Arthur Gilchrist Brodeur (1916):
<http://www.sacred-texts.com/neu/pre/pre04.htm>]

The epithet of ‘omnigenitor’ or ‘all-father’ was also ascribed to Odhin, e.g. in the Poetic Edda:

***Grímnismál* “Ballad of Grímnir”**

§46	§46.
Hétomk Grímir, hétomk Gangleri,	Grim is my name, | Gangleri am I,
Herian ok Hiálmberi	Herjan and Hjalmberi
Þekkr ok Þriði, Þuðr ok Uðr	Thekk and Thridhi, | Thudh and Udh,
Helblindi ok Hár;	Helblindi and Hor;
§47	§47.
Saðr ok Svipall ok Sanngetall,	Sadh and Svipal | and Sanngetal,
Herteitr ok Hnikarr,	Herteit and Hnikar,
Bileygr, Báleygr, Bǫlverkr, Fiǫlnir,	Bileyg, Baleyg, | Bolverk, Fjolnir,
Grímr ok Grímnir, Glapsviðr ok Fiǫlsviðr;	Grim and Grimnir, | Glapsvidh, Fjolsvidh.
§48	§48.
Síðhǫttr, Síðskeggr, Sigfǫðr, Hnikuðr,	Sidhhott, Sidhskegg, | Sigfather, Hnikudh,
Alfǫðr*, Valfǫðr, Atríðr ok Farmatýr:*	**All-father**, Val-father, | Atridh, Farmatyr:
eino nafni hétomk aldregi,	A single name | have I never had
síz ek með fólkom fór.	Since first among men I fared.
....	....

§54
Óðinn ek nú heiti, Yggr ek áðan hét,
hétomk Þundr fyr þat,
Vakr ok Skilfingr, Váfuðr ok Hroptatýr,
Gautr ok Iálkr með goðom,
Ofnir ok Sváfnir, er ek hygg at orðnir sé
***allir** af einom mér.*

§54.
Now am I Odhin, | Ygg was I once,
Ere that did they call me Thund;
Vak and Skilfing, | Vofudh and Hroptatyr,
Gaut and Jalk midst the gods;
Ofnir and Svafnir, | and **all**, methinks,
Are names for none but me.

[Text: Neckel 1936, 64-66;
Translated by Henry Adam Bellows (1936)
<http://www.sacred.texts.com/neu/poe/poe06.htm>]

An analogous and etymologically even identical epithet (see Pokorny 1959, 25) characterizes Dagda, the highest god of the Old Irish pantheon:

> *Aed Abaid Essa Ruaid misi .i. Dagdia druidechta Tuath De Danann ₇*
> *in Ruad Rofhessa ₇ Eochaid **Ollathair** mo tri hanmanna*
> "I am Aed Abaid Essa Ruaid; i.e., the Good God of druidic wisdom of the Túatha dé Danann, the Mighty One of Great Knowledge, and Eochaid **All-father** are my three (other) names"
>
> [Yellow Book of Lecan, fol. 176; see Olmsted 1994, 43]

Note 1: The sign 盤 *pán* < Middle Chinese *bân* < Old Chinese **bān* (see §1 above) was used in the Chinese word *shǔ-pán* "counting board, tally, abacus", whose Middle Chinese predecessor *ṣü-bân* together with *gōngju* "instrument" were adapted in the Tocharian B word *ṣipāṅkiñc* "abacus" (Lubotsky & Starostin 2003, 263).

Note 2: George Starostin (p.c., July 22, 2010) also drew my attention to the Hmong Dog-God *Pan Hu* who had married a princess and their offspring became ancestors of the Hmong people (see also Mair 1998, 27f; White 1991 - *non vidi*). With regard to the Old Chinese reading of the sign 瓠 by Zhengzhang, namely **g^{w}la:s* (see §6.3.), it is possible to see the connection with Proto-Hmong (Miao) **klə^{C}* "dog" > "dog": Hmu *hla*^{3}, Qo-Xiong *qwɨ*$^{3/7}$, Hmong *tle*^{3}, Bunu *tle*^{3}, She *kja*^{3}, besides Mien *tcu*^{3}, Taipan *klu*^{3}, Kimmun *klo*^{9} < proto-Mien (Yao) **klu^{C}* < proto-Hmong-Mien (Miao-Yao) **klu^{C}* (Peiros 1998, 152).

Note 3: Yuri Berezkin informed me that the motif of the primordial being, whose body became a building material in the creation of the world, is relatively very frequent in South and Southeast Asia, Oceania and America. He prefers to seek in the Vedic *Púruṣa*-myth a South Asian substratum influence and in the Scandinavian *Ymir*-myth an import from the Eurasian steppe-cultures. It is a logical solution, but without any concrete material support.

BIBLIOGRAPHY

Abaev, Vasilij I. 1959-73-78-89. *Istoriko-ėtimologičeskij slovaŕ osetinskogo jazyka*, I-IV. (Moskva-) Leningrad: Nauka.

Adams, Douglas Q. 1999. *A Dictionary of Tocharian B*. Rodopi: Amsterdam - Atlanta.

Bailey, Harold W. 1936[81]. Persia II: Language and dialects. In: *Encyclopaedia of Islam* III, 1050-58, reprinted H.W. Bailey's *Opera Minora*. *Articles on Iranian Studies*, Vol. I, ed. by M. Nawabi. Shiraz: Forozangah Publishers 1981, 299-307.

Bailey, Harold W. 1979. *Dictionary of Khotan Saka*. Cambridge: University Press.

Bartholomae, Christian. 1904[61]. *Altiranisches Wörterbuch*. Berlin: Walter de Gruyter.

Baxter, William H. 1992. *A Handbook of Old Chinese Phonology*. Berlin-New York: Mouton de Gruyter.

Baxter, William H. & Sagart, Laurant, 2014. *Old Chinese reconstruction*, version 1.1 (20 September 2014) <http://ocbaxtersagart.lsait.lsa.umich.edu/BaxterSagartOCbyMandarinMC2014–09–20.pdf>

Benedict, Paul. 1972. *Sino-Tibetan: A Conspectus*. Cambridge: University Press.

Blažek, Václav. 1999. *Numerals.* Brno: Masarykova univerzita.

Blažek, Václav. 2008. Egyptian *rmṯ* "Man": An attempt at an Afroasiatic Etymology. In: *Semito-Hamitic Festschrift for A.B. Dolgopolsky and H. Jungraithmayr*. Berlin: Reimer, 57-62.

van Brock, Nadia. 1971. Le traitement des nasales voyelles en tokharien. *Zeitschrift für vergleichende Sprachforschung* 85/2, 280-295.

van Brock, Nadia, 1972. De πύξ à πᾶς. In: *Mélanges de linguistique et de philologie grecques offerts à P. Chantraine*, ed. A. Ernout. Paris: Klincksieck (Etudes et Commentaires 79), 263-276.

Carus, Paul. 1974. *Chinese Astrology*: *Early Chinese Occultism*. LaSalle (IL): Open Court Publishing Company, from an earlier book by the same author, *Chinese Thought* (Chicago: Open Court Publishing Company 1907), chapter on 'Chinese Occultism'.

CVST *A Comparative Vocabulary of Five Sino-Tibetan Languages*, I-VI, by Ilya Peiros & Sergei Starostin. Melbourne: University of Melbourne - Department of Linguistics and Applied Linguistics 1996.

Čop, Bojan. 1976. Nochmals ai. *púmắṅs-*. *Die Sprache* 22, 25-28.

Eichner, Heiner. 1974. Zu Etymologie und Flexion von vedisch *strī́* und *púmān*. *Die Sprache* 20, 26-42.

EWAI *Etymologisches Wörterbuch des Altindoarischen*, I-III, by Mayrhofer, Manfred. Heidelberg: Winter 1986f.

Griffith, Ralph T.H. 1889. *Hymns of the Ṛgveda*. New Delhi: Munshiram Manoharlal.

Grjunberg, Aleksandr L. 1972. *Jazyki vostočnogo Gindukuša: mundžanskij jazyk.* Leningrad: Nauka.

GSR = Grammata Serica Recensa by Bernhard Karlgren. Stockholm: Museum of Far Eastern Antiquities, no. 29, Stockholm 1957.

Hamp, Eric P. 1973. On the phonology and morphology of Lat. *cunctus*. *American Journal of Philology* 94, 169-170.

Henning, W.B. 1963. The Kurdish elm. *Asia Major* 10, 68-72.

Hilmarsson, Jörundur. 1986. *Studies in Tocharian Phonology, Morphology and Etymology with special emphasis on the* o-*vocalism*. Reykjavík: Author.

Horowitz, Franklin E. 1992. On the Proto-Indo-European etymon for 'hand'. *Word* 43, 411-19.

Ivanov, Vjačeslav V. 1975. K tipologičeskomu analizu vnutrennoj formy praslav. **čelověkъ* 'čelovek'. *Ėtimologija* 1973, 17-22.

Jamison, Stephanie W. & Joel P. Jamison (translators). 2014. *The Rigveda. The Earliest Religious Poetry of India*. Oxford: University Press.

Karlgren, Bernhard. 1923. *Analytic Dictionary of Chinese and Sino-Japanese*. Paris: Geuthner.

KEWA *Kurzgefasstes etymologisches Wörterbuch des Altindischen*, I-IV, by Mayrhofer, Mannfred. Heidelberg: Winter 1956f.
Kloekhorst, Alwin. 2008. *Etymological Dictionary of the Hittite Inherited Lexicon*. Leiden-Boston: Brill.
Koehler, Ludwig & Baumgartner, Walter. 2001. *The Hebrew and Aramaic Lexicon of the Old Testament*. Leiden-Boston-Köln: Brill.
Kósa, Gábor. 2009. Pangu's Birth and Death as Recorded in a Tang Dynasty Buddhist Source. *Oriental Archive* 77, 169-192.
Kuiper, Franciscus B.J. 1991. *Aryans in the Rigveda*. Amsterdam-Atlanta: Rodopi.
Laroche, Emmanuel. 1966. *Les noms des Hittites*. Paris: Klincksieck.
Lehmann, Winfred P. 1986. *A Gothic Etymological Dictionary*. Leiden: Brill.
Lubotsky, Alexander & Starostin, Sergei. 2003. Turkic and Chinese loan words in Tocharian. In: *Language in Time and Space. A Festschrift for Werner Winter on the Occasion of his 80th Birthday*, ed. by Brigitte L.M. Bauer & Georges-Jean Pinault. Berlin - New York: Mouton de Gruyter, 257-269.
Mair, Victor H. 1998. Canine Conundrums: Eurasian Dog Ancestor Myths in Historical and Ethnic Perspective. *Sino-Platonic Papers* 87, 1-74.
Meyer, Denise P. 1993. A Reexamination of Indo-European **pn̥Ku-* 'all, whole'. *Indogermanische Forschungen* 98, 40-47.
Neckel, Gustav. 1935. *Edda. Die Lieder des Codex Regius nebst verwandten Denkmälern*3. Heidelberg: Winter.
del Olmo Lete & Sanmartín 2003. *A Dictionary of the Ugaritic Language*. Leiden-Boston: Brill.
Olmsted, Garrett S. 1994. *The Gods of the Celts and the Indo-Europeans*. Innsbruck: Innsbrucker Beiträge zur Kulturwissenschaft, Sonderheft 92.
Paxalina, Tat'jana N. 1983. *Issledovanie po sravniteľno-istoricheskoj fonetike pamirskix jazykov*. Moskva: Nauka.
Peiros, Ilia. 1998. *Comparative Linguistics in Southeast Asia*. Canberra: Pacific Linguistics, Series C, Vol. 142.
Pinault, Georges-Jean. 2008. *Chrestomathie tokharienne. Textes et grammaire*. Leuven-Paris: Peeters.
Pokorny, Julius. 1959. *Indogermanisches etymologisches Wörterbuch*. Bern-München: Francke.
Polomé, Edgar. 1966. Les numéraux indo-européennes. *Revue belge de philologie et d'histoire* 44, 229-33.
Polomé, Edgar. 1968. The Indo-European numeral for 'five' and Hittite *panku-* 'all'. In: Pratidānam. Fs. F.B.J. Kuiper. Ed. J.C. Heesterman et al. The Hague-Paris: Mouton, 98- 101.
Poucha, Pavel. 1955. *Institutiones Linguae Tocharicae*, Pars I: *Thesaurus Linguae Tocharicae Dialecti A*. Praha: Státní pedagogické nakladatelství.
Puhvel, Jaan. 1975. Remus et Frater. *History of Religion* 15, 146-157.
Puhvel, Jaan. 1991. *Hittite Etymological Dictionary*, Vol. 3: *Words beginning with* H. Berlin-New York: Mouton de Gruyter.
Pulleyblank, Edwin G. 1962. The Consonant System of Old Chinese. *Asia Major* N.S. 9, 58-144, 206-265.
Pulleyblank, Edwin G. 1991. *Lexicon of Reconstructed Pronunciation in Early Middle Chinese, Late Middle Chinese, and Early Mandarin*. Vancouver: UBC Press.
Schuessler, Axel. 2007. *ABC Etymological Dictionary of Old Chinese*. Honolulu: University of Hawai'i Press.

Schwarz, Martin. 1992. On Proto-Indo-European **penk*^{w}- 'hand'". Word 43, p. 421-27.
Sokolova, Valentina S. 1960. *Bartangskie teksty i slovaŕ*. Moskva-Leningrad: Izdateľstvo Akademii nauk.
Starostin, Sergei. 1989. *Rekonstrukcija drevnekitajskoj fonologičeskoj sistemy*. Moskva: Nauka.
Starostin, Sergej. 2005. *Chinese Database*.
<http://starling.rinet.ru/cgi-bin/response.cgi?root=config&morpho=0&basename=\data\china\bigchina&first=1>
Steblin-Kamenskij, Ivan M. 1999. *Ėtimologičeskij slovaŕ vaxanskogo jazyka*. Sankt-Peterburg: Peterburgskoe Vostokovedenie.
Tomaschek, Wilhelm. 1880. Centralasiatische Studien II.: Die Pamir-Dialekte. *Sitzungsberichte der philosophisch-historischen Classe der kaiserlichen Akademie der Wissenschaften* 96, 735-900.
White, David Gordon. 1991. *Myths of the Dog-Man.* Chicago: The University of Chicago Press.
Zhengzhang, Shangfang. 2003. 上古音系 *Shànggǔ yīnxì* (Old Chinese phonology). Shanghai: Educational Publishing House.

SUMMARY

In this contribution the name of the Chinese primordial giant Pangu is studied from two points of view: (i) internal reconstruction, proceeding to evaluation of (ii) external etymological relations, especially in the Indo-European perspective. Two variant records of the Chinese theonym 盤古 *Pángǔ*, in Early Middle Chinese **ban* '*kuo* (Pulleyblank) and in Late Old Chinese **bwân kó́* < Preclassic Old Chinese **bānkā*' (Starostin), plus 盤瓠 *pán hù*, Old Chinese **bān g(h)*^{w}*ā(s)*, imply various scenarios:

(a) Adaptation of Early Tocharian A **ba*^{n}*kw*-, continuing in historically attested Tocharian A nom. sg. *puk*, nom. pl. *poñś*^{äṃ}, *poṃś*, *poś* "all".
(b) Adaptation of Iranian **fanka-/*fankā*- "monster".
(c) Adaptation of proto-Tocharian **bangwas/*bangwā* "all" < IE **b*^{h}*ŋ̍ǵ*^{h}*wos/*b*^{h}*ŋ̍ǵ*^{h}*wā*.

The idea that the primordial giant was an *multi*- or *omnigenitor* was explicitly expressed in both Vedas and Eddas and can be supported etymologically: *Púruṣa*- : Vedic *púru*- "many, much, abundant" and *Ymir* : Gothic *iumjo* "crowd", Old Swedish *ympnian* "abundant". This conclusion agrees with the information about Pangu from the text from §1.2.4.

ACKNOWLEDGEMENT:

The present study was prepared under the auspices of the Specific Research Fund at Masaryk University, nr. 2817. The author would like to express his gratitude to John D. Bengtson for help with correction of English. In the final version I have projected the valuable comments of George Starostin and Yuri Berezkin.

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

THE ENDS OF THE EARTH[1]

JOHN M. SAUL
ORYX, PARIS

The appearance of cave paintings thirty to forty thousand years ago at the two extremities of the Eurasian continent (Spain/France and Sulawesi/Borneo[2]) *but not in between* appears as though it might be the consequence of incomplete data, sampling error, or similar, as may indeed be the case. Yet this is the data available to us (as of late 2019), and we should also consider that it might actually represent the reality of the matter. If so, an explanation would be required.

In speculating on this matter, let us first assume that rock-paintings were already being made by our ancestors before they left Africa. This has not been demonstrated, but it seems safer than assuming otherwise.

Early groups of our ancestors may have left Africa and migrated north in an effort to learn the "immortalizing secret" of the storks or cranes who would annually leave Africa and then return, seemingly born again, as I have proposed elsewhere.[3] On reaching southern Europe or the Middle East, our ancestors would have discovered the disorienting ornithological truth. At that point, our ur-religion, if we may call it such, may have undergone a schism, as have all religions since, with some of our ancestors following the stars to the west and others heading east toward the birthplace of the life-giving Sun. Driven by hope and *ideas*, and moving through lands lacking other *sapiens*,

[1] Dr. Saul has contributed to earlier issues of *Mother Tongue* (Journal and Newsletters), *e.g*. "Was the First Language Purposefully Invented?" in MT VII (2002). Besides his proficiency in mythology and cosmology, Saul is fully conversant in historical linguistics and other fields of anthropology (see Saul, John M. [2019]: *What the Stork brought: African click-speakers and the spread of humanity's oldest beliefs,* Old Africa Books, Naivasha, Kenya, ISBN 978-9966-7575-2-4). The present essay is consistent with ASLIP founder Harold Fleming's emphasis on four fields anthropology. Saul's suggestion of some kind of connection between the Indus civilization's undeciphered symbols and the glyphs found on Easter Island is a reasoned hypothesis that is testable. Special attention should be given to Saul's warning in footnote 7 [Ed.].

[2] Aubert, Maxime, A. Brumm, M. Ramli, et al. (2014): "Pleistocene cave art from Sulawesi, Indonesia," *Nature*, 514: 223-227. https://doi.org/10.1038/nature13422; Aubert, Maxime, P. Setiawan, A.A. Oktaviana, et al. (2018): "Palaeolithic cave art in Borneo," *Nature*, 564: 254-257. https://doi.org/10.1038/s41586-018-0679-9; Aujouat, Norbert, J.-J. Cleyet-Merle, J. Gaussen, et al. (1998): "Approche chronologique de quelques sites ornées paléolithiques du Périgord par datation carbone 14 en spectrométrie de masse par accélérateur de leur mobilier archéologique," *Paléo, Revue d'Archéologie Préhistorique*, 10: 319-323 ; Brumm, A., A.A. Oktaviana, B. Burhan et al. (2021): "Oldest cave art found in Sulawesi," *Science Advances,*7(3): eabd4648.

[3] Saul (2019), *cit.*

they could have reached the "ends of the Earth" within a very few years (as did Marco Polo), and there left their mark.

On approaching or reaching the ends of their journeys, our ancestors, east and west, painted cave walls with depictions of animals. We do not know why they painted animals, but the enormously insightful Africanist and ethnologist Leo Frobenius (1873–1938) conjectured that the painted animals depicted in the European caves were stars.[4]

Matters in which a sequence of events, i.e., *history*, is important are commonly highly sensitive to the conditions that prevailed at their outset, i.e., to the initial event or to initial conditions. Familiar cases are the health of a child at birth and the preparation of a sports team or an army before going into action. So even if we know nothing for certain about the matter, and whether my migrating-bird hypothesis is correct or not, we are entitled to suspect that the extremities of the world, or of the continents, may have had an important role in human affairs from the outset of the human adventure ... perhaps because these "ends of the Earth" were seen as correspondents Below of the ends of the Milky Way, or even as its entrances.

On their way to westernmost continental Europe, pilgrims to the "Field of Stars" at Compostela (Galicia, Spain) would kick up dust that by tradition would become the stars of the Milky Way, much as a Bushman girl had formed the Milky Way by throwing Earthy white wood ashes into the heavens,[5] As Below, So Above. From Compostela itself, pilgrims might continue westward to the nearby Chapel of Nuestra Señora at Cape Finisterre (*finis terre,* "World's End"), the last finger of land jutting into the ocean, said to mark the end of the Milky Way. (Islands "don't count": well into the 19th century, the idea that there are islands that had never been connected to the continental mainland was not acceptable.[6])

"World's End" at Nuestra Señora at Cape Finisterre, at 42°52'N, 9°15'W (*Times Atlas* coordinates) designates the western end of the Milky Way. Where, then, is the eastern end? The question is awkward, because we are in unfamiliar intellectual territory. But the Sri Lankan cliff-site called "World's End," situated at 6°48'N, 80°47'E (*Times Atlas*), seems a proper candidate because it is located *exactly* 90° of longitude east of Cape Finisterre. (Yet, as put by one tourist brochure, "nobody really knows how the escarpment got its name.")

Vestiges of the Indus River Valley civilization, c.2600–1900 BC, include seals and other objects marked with 400 to 600 different "symbols" of unknown significance. Many of the seals are identical, mass-produced using molds. On the opposite side of the world, approximately 60° of latitude and 180° of longitude away, is Easter Island, believed to have been first settled, by Polynesians, c.1200 AD. On Easter Island, undeciphered characters called rongorongo provide what may (or may not) be vestiges of a writing system with "glyphs," some of which are strikingly

[4] Von Dechend Hertha (1979): unpublished teaching notes at the Massachusetts Institute of Technology, p. 13, citing Frobenius, *Kulturgeschichte Afrikas* (1934) 46, 187.
[5] Bleek and Lloyd Archive: http://lloydbleekcollection.cs.uct.ac.za/, 6879–6884.
[6] Oosterzee, Penny van (1997): *Where Worlds Collide: The Wallace Line*, Cornell University Press, p. 104. Callimachus, *Hymn to Delos,* 28–54, has the islands cut off from the continent by a mighty god, with Delos, as "the goddess Asteria" as an exception, once flung into the sea from heaven.

similar to the symbols from the Indus Valley. Approximately 120 different glyphs are known, plus numerous variants and combinations.

Most scholars intrigued by these matters have concluded that the undeciphered symbols from the Indus Valley and the undeciphered glyphs from Easter Island, half a world and 2500 years away, are unrelated. I am not convinced, however, that scholars are the right people to determine the matter; they know too much. ("It ain't what you don't know that gets you into trouble, it's what you know for sure that ain't so.") I would prefer to give the problem to students or to puzzle addicts who have never heard of the Indus Valley seals or of rongorongo.

A first hypothesis would have the undeciphered "symbols" and the undeciphered "glyphs" as unrelated to one another, which would leave us with two unresolved historical puzzles, which is the present situation.

A second hypothesis would assign a common origin to the two systems and since the Indus system is older, a means to export it to Easter Island would have to be proposed.

In attempting to join heaven Above, the Indus Valley people may have sought a terrestrial correspondent for the "End of the Milky Way." A voyage undertaken by inexperienced and poorly equipped explorers from the valley of the Indus to the region of Easter Island would have little chance of success even in our own times. But many expeditions may have been mounted from the Indus Valley just as many pilgrims go to Compostela, many stone circles incorporating astronomical orientations were built, many Gothic cathedrals constructed, and many astrological systems concocted, all in the same quest for immortality.

The Indus Valley "symbols" appear to have been of importance to the Indus Valley people so in a subsidiary hypothesis, I propose that these explorers carried the Indus Valley symbols with them, whether physically, or as part of their intellectual baggage.

Many of these hypothesized voyagers would have perished. But with great good fortune, one individual or small group could have arrived on Easter Island. Before he, she, or they died out, or attempted to return, or perhaps attempted to push on, they left objects with their symbols in a cave, whether to mark burials or for some other reason.

Polynesians arrived on the island many centuries later, most likely sometime after the invention of the double-hulled canoe (catamaran), somewhat before 1000 AD. To their astonishment they may have found inscribed symbols that, for them, were as extraordinary as would be the discovery these days of evidence of a past visit to Earth of extraterrestrials.[7]

In any case, a time difference of 2500 years need not negate the possibility of a connection, as shown by the next example.

The extreme westernmost point of the Asian continent is marked by the walled city of Troy, "Earth's most famous town."[8] At Troy, there had been trouble from the beginning when Poseidon sent a sea monster to ravage the city and its coasts, complaining that he had not been paid his

[7] Anyone following up on this hypothesis is urged to use photographs of the symbols and glyphs or the actual objects, and to avoid hand-sketched copies. It may also be necessary to identify frauds and replicas made for tourists.
[8] Morley, Christopher (1937): *The Trojan Horse: a contemporary drama laid in 1185 B.C.*, J.B. Lippincott, Philadelphia.

wages for building the city's sea-walls. The beast was subsequently killed by Hercules but when he too was refused payment, he returned with a band of warriors and killed most of Troy's royal line, vengeance for the city's "twice perjured walls."[9] Thus, the war whose end is recited in Homer's *Iliad* is actually the story of the city's third sacking. Trojan and Greek refugees, men and gods, then set out to found new cities in Europe, a new continent for a new "Age," with Trojan refugees asking Apollo for "a walled city that shall endure" (*Aeneid*, III).

According to Eratosthenes (276?–196 BC), the final fall of Troy occurred in 1184 BC. Some believe the fall was of an historical nature. Others see it as mythological. What is certain is that the year 1184 BC was the year the North Celestial Pole (the point about which the northern heavens appear to turn) had moved (precessed) through three stars of the constellation of *Draco* to a position that was for the first time closer to a bear-star of an *Ursa*-constellation than to the ancient pole star Thuban (*Alpha Draconis*): 1184 BC was the year the North Celestial Pole definitively abandoned the stars of its original home- or mother-constellation, a consequence of the Precession of the Equinoxes.

The pole star and other nearby stars are central to the beliefs of peoples living in the northern hemisphere because they never set or "die" at any time of the night or year. The Spring Equinox incorporates rebirth, the Pole Star incorporates immortality itself. But neither is immune to damage wrought by the Precession of the Equinoxes.

The change of pole stars in 1184 BC, which was observable throughout the northern hemisphere, appears to have been recorded in the Chinese account of "Qi, the Abandoned One," abandoned three times by his mother.[10] Qi had been conceived after his mother, a consort of the emperor, had stepped into the divine footprint of Shangdi,[11] in which *dì,* 帝, represented "the supreme God as the celestial pole."[12] The footprint represents the most familiar stars of the constellation of *Ursa Major*; **Figure 1**.

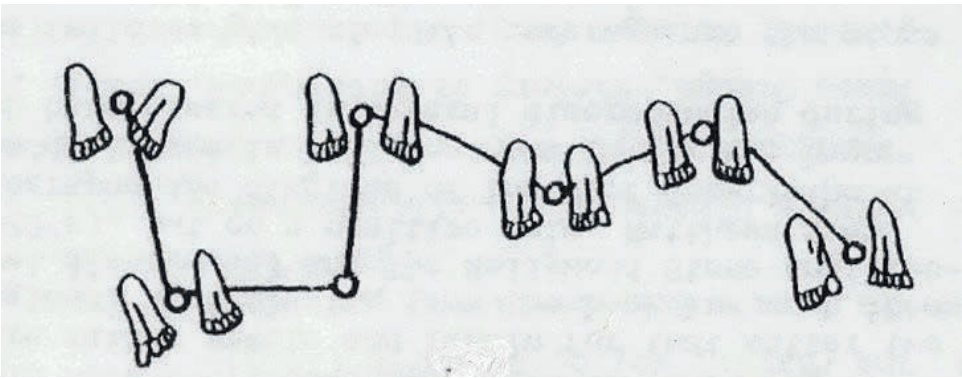


Figure 1: Steps in the pattern of the stars of *Ursa Major,* danced by the legendary Chinese king Yu the Great.

[9] Ovid, *Metamorphosis,* 11:215.
[10] Wikipedia, Zhou dynasty.
[11] Wikipedia, Zhou dynasty.
[12] Wikipedia, Shangdi.

The lineage of the Zhou Dynasty (1046–256 BC) was said to have begun with the birth of thrice-abandoned Qi.[13] Also established at this time was the "Ji [the Walled] City" district of Beijing,[14] a matter linked to Zhou affairs.

Tales of Troy, even if they then existed, are highly unlikely to have reached the other end of the Asian continent at the inception of the Zhou dynasty in 1046 BC. But there were two much later periods when accounts of Earth's most famous town could, *or would*, have reached China. One was between c.1250 and 1350 AD when it was possible for a European to walk from the Black Sea or the eastern Mediterranean all the way to the Pacific Coast,[15] passing through the khanates into which the Mongol Empire had fragmented. It was during these times that Kublai Khan established his capital at Beijing (Khanbaliq) in accordance with the *Rites of Zhou*. The year was 1271 AD but the details are unavailable. Writing in *Discover Mongolia* (14 April 2018), "Enkhzul" asked "Why did Kubilai Khan moved (*sic*) the capital of his dynasty to Beijing," concluding that the move down to Beijing was an "under-discovered and uncertain event."[16]

Then in the early 1400s, the Yongle Emperor, 永樂帝,third emperor of the Ming Dynasty, moved his capital from Nanjing, a capital since "about 212 BC," when a dragon had been assuaged by the burial of a treasure on its neck.[17] The Yongle Emperor re-established the capital in Beijing. Construction began at the Forbidden City in 1406 and was completed 28 October 1420, with the city then officially designated the Ming capital. A painting at the National Museum of China shows the great imperial march from Nanjing to Beijing but inquiries on the spot in 2010 indicated that the reason for the change of capitals is unclear.

Warring with the Mongols had closed direct access to the West during the early Ming Dynasty, yet much about the world had by then become known to the scholarly Emperor Yongle, a consequence of the compilation of his great encyclopedia, put together between 1403 and 1408 by a team that came to include 2169 scholars, not all of whom were Confucian. Yongle's encyclopedia was designed to include the entire Confucian canon and its commentaries, as well as Buddhism, Daoism, history, philosophy, astronomy, geography, medicine, and the arts.[18] At the rate of 72 years per degree, the year 1408 AD corresponds to a precessional advance of 36° – one-tenth of the celestial circle – from 1184 BC, date of the fall of Troy.

---

[13] Wikipedia, Zhou dynasty.
[14] Wikipedia, Jicheng.
[15] Berthon, Simon and Andrew Robinson (1991): *The Shape of the World: The mapping and discovery of the Earth*, George Philip, London, p. 64.
[16] Enkhzul: https://www.discovermongolia.mn/blogs/why-did-kubilai-khan-moved-the-capital-of-his-dynasty-to-beijing (posted 14 April 2018, accessed 16 Feb. 2020).
[17] Geil, William Edgar (1911): *Eighteen Capitals of China,* Constable & Co., London, pp. 192, 194.
[18] Yongle's compilation was the world's largest general encyclopedia until overtaken by the Wikipedia on September 9, 2007. Only approximately 3.5% of Yongle's encyclopedia survives.

Kublai Khan and the Yongle Emperor both appear to have had knowledge of something highly specific concerning the site of Beijing. The same can perhaps be said of the Zhou author of the story of Qi. We can easily imagine that this was related to the configuration of lakes and waterways at the site, and their unmistakable evocation of the "dipper stars" of the constellation of *Ursa Major*; **Figure 2**. But the sources consulted all describe the lakes as "man-made." If true, it must have been the *site* of Beijing that was initially important, with the bear-constellation lakes later added as an acknowledgement of something already recognized.

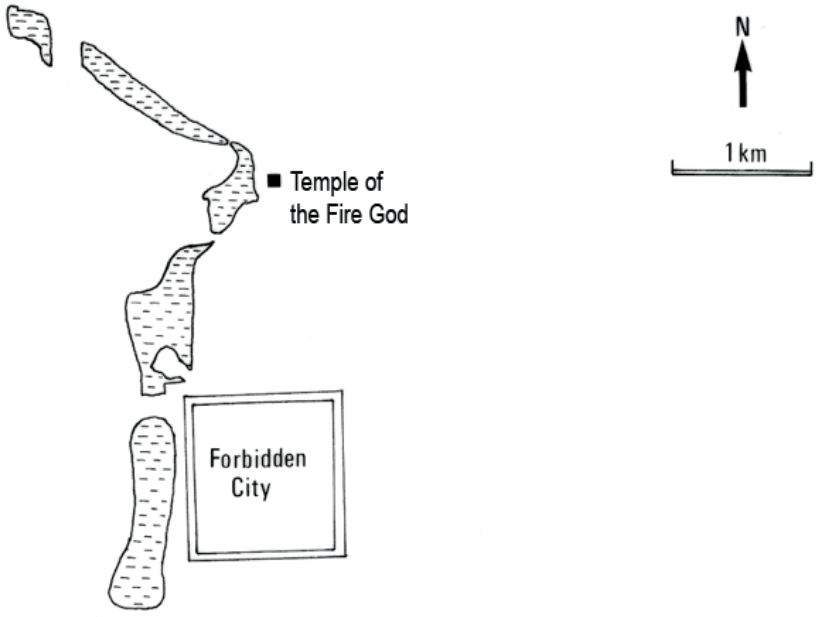


Figure 2: The Forbidden City of Peking and nearby lakes after a map by Father Hyacinthe Bitchurin (1829). The shapes of the lakes were substantially modified *circa* 1650 and again *circa* 1950. Yet the overall configuration still resembles that of the dipper-stars in the constellation of *Ursa Major.*

There is another difficulty with this argument, too, because the precessional abandonment of the three stars in *Draco* had *not* produced a pole star in *Ursa Major.* Instead, the new pole star was Kochab (*Beta Ursae Minoris*) in the constellation of *Ursa Minor.*

Ursa Major or *Ursa Minor*? What are we to do about this? Apparently nothing. The disagreement stands, another ancient schism. For with splendidly succinctness, Seneca (c.4 BC – 65 AD) signaled disagreement elsewhere on this exact same matter: *Maior Pelasgis apta; Sidoniis minor* (*Medea,* 697).

Ursa Major is suitable for the [pre-Hellenic] Greeks,

Ursa Minor for the Phoenicians of Sidon.

With this background, we can return to the "famous town" of Troy, at the extreme western edge of the Asian continent, whose third and final fall initiated a new Age. The *Times Atlas* coordinates for Troy are 39°55' North, 26°17' East.

Beijing, located hard by the eastern extremity of the Asian continent, was founded after Qi had been abandoned the third and final time at the beginning of a new Mandate of Heaven. At 39°55' North, 116°26' East (*Times Atlas*), Beijing lies *exactly* 90° of longitude *exactly* due east of Troy.

A GEOGRAPHICAL ALIGNMENT SCHOLARS MIGHT LOVE

Scholars and laymen alike may prefer to treat the geographical position of Easter Island, half a world away from the Indus Valley, as inconsequential and to discount the similarity in their symbols and glyphs as coincidence. They might also see the naming of "World's End" in Sri Lanka, 90° of longitude east of Compostela, and the positioning of Beijing 90° of longitude east of Troy as occult matters dating to recent centuries and lacking truly archaic roots. Yet arguments along such lines cannot be applied to the matter of Kish, now taken up.

The ruined Mesopotamian city of Kish is of prime historical importance because it was there, according to the Sumerian King List, that "kingship was lowered from heaven." Kish was thus the site of the world's first kings. Thereafter, one man could *legitimately* rule over others, even those he did not know, a most astonishing development.

"King of Kish," as a title, came "to signify or symbolize imperial, even universal, dominion" in times "long after Kish had ceased to be the seat of kingship."[19]

At Kish, the largest mound among the ruins is believed to be that of the temple of the goddess Inanna,[20] Sumerian counterpart-precursor of Aphrodite-Venus. To the west of Kish, temples of Aphrodite-Venus at Aphaca and at Byblos, both in modern-day Lebanon (ancient Phoenicia), were associated with stones supposedly fallen from heaven.[21] Another temple to Aphrodite-Venus with a stone (incorrectly) identified as a meteorite is located at "Aphrodite's birthplace" at Palaepaphos, near Paphos, on "Aphrodite's Isle" of Cyprus.[22]

From Kish to the Temple of Aphrodite at Aphaca, where Aphrodite Ourania had fallen as "fire from a star," the geographical bearing is 11°51' – virtually 12° – North-of-West, and from Kish to Byblos the bearing is 11°55' North-of-West.[23] Kish to Palaepaphos, site of the third of these meteorite-sanctuaries of Aphrodite, has a bearing of 12°11' North-of-West, but it was also said that Aphrodite had actually risen from the Ouranian "foam" (*paphos*) at Aphrodite's Rock, a nearby sea rock whose bearing from Kish is precisely 12° North-of-West.

Kish plus the three meteorite-temples form part of a group of sites aligned at 12° North-of-West. Other members of the group include the Poseidon-Aphrodite Cave on Karpathos at 11°36' North-of-West from Kish and the temple, now destroyed, of Aphrodite Ourania on Cythera, the "Isle of Love," formerly located near Paleopoli, whose bearing from Kish is 11°47' North-of-

[19] Hallo, W.W. and W.K. Simpson (1971): *The Ancient Near East: A History* (2nd ed.), Harcourt Brace College Publishers p. 39.
[20] Gibson, McGuire (1972): *The City and Area of Kish.* Coconut Grove, Miami, Florida, Field Research Projects.
[21] Burke, John G. (1986): *Cosmic Debris: Meteorites in History*, University of California Press, p. 221.
[22] Burke (1986) *cit.,* p. 221.
[23] Bearings were calculated using Movable Type Scripts.

West.[24] Another Venus-site is Monte Venera in eastern Sicily, at 12°30' North-of-West of Kish, and yet another Aphrodite-Venus site was at Erice in western Sicily where the bearing from Kish to the Temple of Venus, notorious for its temple prostitution, is 11°53' North-of-West; **Figure 3**.[25]

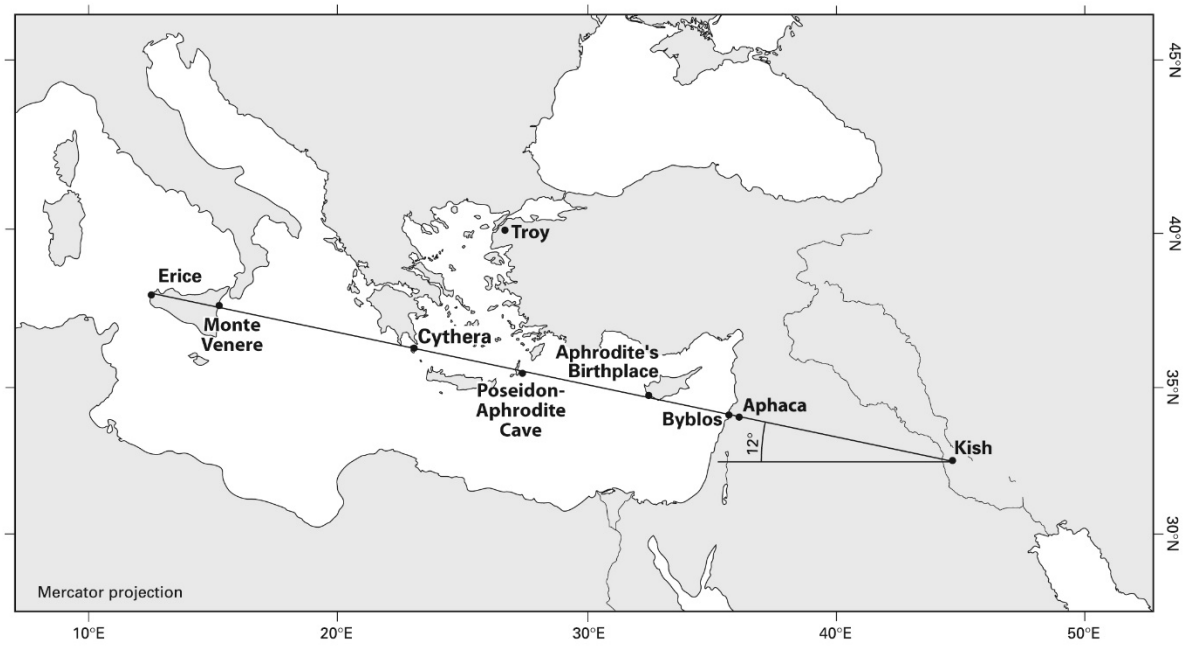


Figure 3: Ancient sites dedicated to Aphrodite-Venus aligned at close to 12° North-of-West from Kish.

[24] The temple on Cythera was known as the oldest temple of Aphrodite in Greece.

[25] Geographical coordinates of sites mentioned in text:

Site	Coordinates	Source of information for coordinates
Kish, where "Kingship was lowered from heaven"	32°32'25"N, 44°36'17"E	Wikipedia
Aphaca (Afqa); meteorite or supposed meteorite	34°04'05"N, 35°53'34"E	http://imperium.ahlfeldt.se/places/37121.html
Byblos, Phoenician site; meteorite or supposed meteorite	34°07'25"N, 35°39'04"E	Wikipedia
Palaepaphos, Cyprus; false meteorite	34°42'26"N, 32°34'28"E	Dept. of Antiquities, Republic of Cyprus
Aphrodite's Rock (*Petra tou Romiou*), Cyprus	34°39'49"N, 32°37'38"E	GPS
Poseidon-Aphrodite Cave, Karpathos; Venus statuettes found near Pigadia nearby	35°30'04"N, 27°11'49"E	Cave and Pigadia are both 11°36' N-of-W of Kish. Karpathos Guide Mobile Application; https://www.karpathos.org.uk/sights.html
Paleopoli, Cythera, Greece, ruins near Agios Kosmas	36°15'30"N, 22°59'54"E	GPS coordinates of nearby lodging
Monte Venere, eastern Sicily	37°51'N, 15°17'E	*Times Atlas*
Erice, western Sicily, Temple of Venus on Phoenician	38°02'13"N, 12°35'11"E	Wikipedia

In Akkadian, a Semitic language predominantly used at the Sumerian city of Kish, Kish was *kiššatu,* which translates as "3600" or "Totality,"[26] or perhaps as "Center," and might, I propose, signify Kish as a counterpart in the world Below of the Ecliptic Pole. Angular positions measured using Kish as pole might then correspond to dates, whether historical or mythological, with one degree representing 72 years.

By tradition, twelve is "the number of Jupiter," keyed to the orbital period of 11.9 years of the planet with the same name.[27] The angular displacement of 12° from due East-West, as measured with Kish as center, might thus indicate the beginning or the end of a 12°-Age ruled by Jupiter-Zeus. But, if so, which Age would that be, and what date?

The role of Aphrodite, goddess of love, provides the answer. The Age, which was described by Hesiod in the 8th century BC, was the "Age of the godlike Race of *heroes*," "the Age before our own," in which men, subject to *eros,* "crossed the open seas in ships for fair-haired Helen's sake" to be "killed at Troy."[28]

The "Age before our own" ended in 1184 BC, date of the third and final destruction of Troy. Dialing back 12° from 1184 BC using the rate of 72 years per degree, produces the year 2048 BC, 864 years earlier.

In myth, 2048 BC would have been *the beginning of the Age that ended at Troy,* or more exactly, because a destruction is involved, it would have been the end of the Age that preceded the Age of Troy. It was the year Poseidon sent a sea monster to ravage Troy and its coasts, the year Qi was abandoned for the first time, and the year in which Beowulf the Geat overcame Grendel, prowler of the borderlands, in his first of three battles with the dragon brood.

In the heavens, 2048 BC was the year the North Celestial Pole first came closer to the star *Kappa Draconis* than to the ancient pole star Thuban (*Alpha Draconis*), i.e., the year *Kappa Draconis* became the pole star.

Historically, 2048 BC should thus correspond to the beginning of the Minoan-Mycenean Palace Civilization, destroyed at Troy 864 years later, which indeed it does. For according to one standard academic reckoning, the year 2050 BC (which is structurally identical to the year 2048 BC) was the beginning of the "Middle Minoan I," an interval marked by "the foundation of the Palace proper at Knossos."[29]

This Minoan-Mycenean Age-before-our-own began with the star *Kappa Draconis* at the Pole in 2048 BC. It ended in 1184 BC with the abandonment of both *Kappa Draconis* and of *Draco* itself and with the final Fall of Troy on Earth Below. When the Minoan-Mycenean Palace Civilization ended, its celestial Palace abandoned, the scattered survivors, Greeks and Trojans alike, set out to found new cities for a new Age on the new Continent of Europe.

---

[26] Labat, René (1976): *Manuel d'Épigraphie Akkadienne: Signes, Syllabaire, Idéogrammes*, new edition reviewed and corrected by F. Malbran-Labat, Librairie Orientaliste Paul Geuthner, Paris, entries 396, 425.

[27] Philolaos of Croton (c.470 - c.385 BC) "says that the angle of the dodecagon is the angle of Zeus because Zeus holds together in a single unity the whole duodecimal number. In Plato likewise Zeus leads 'the twelve' and has absolute domination over all things" (from von Dechend, 1979 *cit.,* p. 35).

[28] Hesiod, *Works and Days,* 156–170.

[29] Pendlebury, J.D.S. with a foreword by Sir Arthur Evans (1969): *Handbook to the Palace of Minos at Knossos with its Dependencies.* Macdonald, London, p. 28.

As Above, So Below, the Minoan-Mycenean Age before our own, which began in 2048 BC and ended at Troy in 1184 BC, was the period of historic time during which the star Kappa Draconis *was pole star*.[30]

BEFORE THE MINOANS

Jupiter's father was Saturn, whose number is 30.[31] Thirty degrees multiplied by 72 years per degree gives 2160 years, and 2160 years before 2048 BC takes us back to 4208 BC. This was approximately 36 years after the Spring Equinox had exited the Milky Way near the star *Zeta Tauri,* point of the Bull's horn. This retrodicted 30°-Age preceded the Minoans and Myceneans in what may have been a Taurine Time, though not an Age of Gold because the Equinox had already departed the Milky Way, thereby prohibiting direct unhindered access to heaven. It would have been a "gold-like" Age, perhaps of amber (electrum), or perhaps an Age suitable for a Golden Calf. It ended when Minoans jumped over the ecliptic Bull of Heaven and established palaces in the image of the pole star.

These numbers "work." *Of course they do*: they were *forced* to work. In truth, they are numerical concoctions. Ingenious theological conventions. Elements from a wish-list. Sacred mumbo jumbo. Lies. The precessional rate is actually closer to 71.6 years per degree than to 72; Jupiter's period is approximately 11.86 Earth-years, not 12; Saturn's period is 29.46 Earth-years, not 30; and *Zeta Tauri* as exit from the Milky Way and point of a bull's horn is three or four or five times removed from reality. All that said, this is our heritage.

LATER TIMES

Troy fell with the defeat of the stars of *Draco. Draco* was then replaced by *Ursa* and for some peoples, cosmological attention shifted to Jerusalem, built on hills whose original topography evoked the pattern of the stars of *Ursa Major*; **Figure 4**.

[30] Saul, John M. (2013): *The Tale Told in All Lands*, Edite, Paris. ISBN 978-2-37081-005-2. The ecliptic was implicated in Minoan belief as well; according to von Dechend (1979) *cit.,* p. 65, the Minoan "double-axe labrus" … "marks the equinox."

[31] "*Kyrios triakontaeteridon*, Lord of the thirty-year period."

Figure 4: Photograph of a three-dimensional model of the original topography of the site of Jerusalem (Museum of the History of Jerusalem, Jerusalem) showing the resemblance of the site to the outline of the most familiar stars of the constellation of *Ursa Major*. The "bucket" of the Big Dipper corresponds to the Temple Mount.

Thereafter several parallel calendars were maintained in order to mark progress toward a New Age. One such calendar depended on periods of 216 years, equivalent to precessional advances of three degrees, one-tenth of a zodiacal sector:

-12°	2048 BC: Beginning of Minoan-Mycenean civilization.
0°	1184 BC: End of Minoan-Mycenean civilization; Fall of Troy.
+3°	968 BC: Building of First Temple, Jerusalem.
+6°	752 BC: Foundation of Rome as a New Troy.
+9°	536 BC: Foundation of Second Temple Jerusalem.
+12°	320 BC: Possible target date for universal empire of Alexander (died 323 BC).
+15°	104 BC: Eagle replaced Bull in Rome. Grand Inception calendric reform in China, designed to render the Emperor immortal.
+21°	328 AD: Constantinople founded as Second Rome, 326–330 AD, although Constantine had originally intended to establish his city on the site of Troy.
+36°	1408 AD: Ming capital moved to Beijing.

Among the calendars established for times after the Fall of Troy, one of the strangest had a gap of 72 years for the time, or absence thereof, between the destruction of the First Temple of Jerusalem in 587 BC and the completion of the Second Temple of Jerusalem in 515 BC.

Work on these matters showed that the Nazi ideology emerged from the same nexus of archaic ideas.[32] Nazi concepts of "Europe," "Aryan," and "Race" were all of a mythological nature. The Nazis mistook myth for reality. So too did Alexander, the Aztecs, and certain Rwandans in 1994. So do we all on occasion. This too is our heritage.

john.saul@wanadoo.fr
www.tomebook2.com
ORCID 0000-0002-2282-2430

[32] Saul (2013) *cit.*