

THE PROTO-SAPIENS PROHIBITIVE/NEGATIVE PARTICLE *MA

PIERRE J. BANCEL, ALAIN MATTHEY DE L'ÉTANG

ASSOCIATION FOR THE STUDY OF LANGUAGE IN PREHISTORY (CAMBRIDGE, MA); ASSOCIATION
D'ÉTUDES LINGUISTIQUES ET ANTHROPOLOGIQUES PRÉHISTORIQUES (PARIS, FRANCE);

AND JOHN D. BENGTON

ASSOCIATION FOR THE STUDY OF LANGUAGE IN PREHISTORY (CAMBRIDGE, MA); EVOLUTION OF
HUMAN LANGUAGE PROJECT (SANTA FE INSTITUTE, NM)

We report here on a lexical root, very widespread in diverse languages worldwide, including more than 50 ancient languages, long-isolated languages, and proto-languages. Most of these rely on uncontroversial reconstructions, while others, from Proto-Nilo-Saharan to Proto-Trans-New Guinea through Proto-Austrian and Proto-Amerind, go back to far more than 10,000 years ago and cover all continents. We argue that this lexical root may only have been part of the ancestral language common to all modern humans.

1. INTRODUCTION

We will document here an ancestral word root, which is found in such a huge number of language families across all continents that it can only be a common inheritance from the original lexicon of our remote Sapiens ancestors. Following the common linguistic custom of naming the ancestral language of a family by the name of this family with the prefix *Proto-* (Proto-Germanic, Proto-Algonquian, Proto-Bantu, etc.), we call the ancestral language of our species Proto-Sapiens.

Proto-Sapiens is not a newcomer in historical linguistics: building upon the pioneering work of Trombetti (1905), about three dozen Proto-Sapiens words have recently been identified (Bengtson & Ruhlen 1994), making use of the massive linguistic materials and comparative works that have accumulated during the 20th century.

However, many historical linguists deny the validity of Proto-Sapiens etymologies, a subject which deserves a brief preliminary discussion. Their rejection basically results from an orthodoxy which has held for more than a century that languages evolve so fast that, after 5,000 to 8,000 years of evolution, nothing significant remains of an ancestor language in its descendants.

This orthodoxy is easily demonstrated to be false. We can illustrate this point with an example taken from the Indo-European family, to which most modern European languages belong and which is, for this (unscientific) reason, by far the best studied of all language families. Its ancestor language, Proto-Indo-European, is estimated to have been spoken some 5,000 to 8,000 years ago. This is close to the limit beyond which any trace of it should have vanished in modern languages;

nevertheless, thanks to comparison of its descendant languages, Indo-Europeanists have reconstructed a wealth of knowledge about it, including nearly 3,000 words and major parts of its conjugation and declension systems.

Moreover, in an unpublished study bearing on 494 Indo-European languages, we have found (Bancel & Matthey de l'Etang, ms.) that only two of them (0.4%) had lost the Proto-Indo-European 1st person singular pronominal root **m-* (found in English *me, my, mine*) and only seven (1.4%) the 2nd person singular **t-* (English, in which *thou, thee, thy, thine* subsist only in religious and other specialized uses, is counted as one of these seven cases of loss). This amounts to minuscule loss rates of 0.05% per millennium for **m-* and 0.18% for **t-*. Extrapolating these loss rates allows us to endow these roots with theoretical half-lives (Pagel 2000) of 1.39 million years and 385 000 years, respectively. These timespans are about 50 to 200 times the 5,000 – 8,000 year threshold beyond which every significant trace of an ancestor language is supposed to be lost forever, again showing the inanity of this alleged limit.

Some linguists also have attempted to demonstrate through probabilistic calculations that global or other remote etymologies could be due to chance resemblances (Ringe 2002, Boë et al. 2006). We have shown in detail elsewhere (Bancel & Matthey de l'Etang 2013) that such demonstrations were flawed by glaring mistakes. For instance, Ringe (2002), ignoring that a probability is a ratio, *i.e.* a number of chances for a given event to occur *out of* a total number of possibilities, multiplies chances as he adds parameters that obviously shrink this ratio – as if there were four times more chances to get four aces of hearts when picking a card from each of four decks than to get one when picking a card from a single deck.¹ As a consequence, the results and conclusions of these supposed demonstrations are deprived of any validity. Our study has also shown that, while several crucial parameters contributing to the validity of an etymology could not be reduced to figures, thus preventing a final probabilistic assessment, their huge distribution made some individual etymologies so obvious – like **m-* ‘1st person’ or **t-* ‘2nd person’ in the Indo-European family – that no calculation was needed.

Finally, we have studied the kinship appellative terms *papa, mama* and *kaka*, gathering kinship terminologies in over 3,000 languages covering the whole Earth. These words are so widespread that linguists never even envisioned that their convergence might be due to chance. Instead, since the mid-19th century they had elaborated an ad hoc hypothesis (Buschmann 1852, Lubbock 1889, Westermarck 1891), holding that the similarity of these words at the global level resulted from convergent innovations stemming from the phonetic limitations of babies learning to speak.

This hypothesis, refined by Murdock (1959) and Jakobson (1960), was widely accepted without any historical study having documented a single case of such innovation. However, prompted by recent publications claiming a Proto-Sapiens antiquity for *papa, mama* and *kaka* words (Ruhlen 1994a, 2000; Bengtson & Ruhlen 1994; Bancel & Matthey de l'Etang 2002; Matthey de l'Etang

¹ Actually, one has 1 chance out of 52 to get an ace of hearts (or any other card) when picking at random a card from a deck, and $(1/52)^4 = 1$ chance out of 7,311,616 to get four aces of hearts when picking a card from each of four decks.

& Bancel 2002), Trask (2004) claimed to have found such convergent innovations in a range of language families. We have shown that all of Trask's alleged innovations were inherited from the earliest stages of their respective language families (Matthey de l'Etang & Bancel 2008, Bancel & Matthey de l'Etang 2013). For instance, the allegedly “new” French words *maman* and *papa* have been inherited from Latin *mamma* and *pappa*, Welsh *tat* and *mam* have been inherited from Proto-Celtic **tata* and **mama*, all words found in classical and comparative dictionaries and apparently ignored by Trask.

The massive preservation of *papa*, *mama* and *kaka* words in an overwhelming majority of language families over the last several millennia leaves us with two theoretically possible explanations of their global distribution: (i) blatantly implausible, massive convergent innovations having originated in the Paleolithic and having inexplicably ceased in the Neolithic, a period in which these words would have started to be faithfully transmitted from one generation to another; (ii) inheritance from a common Proto-Sapiens ancestor.

As we will see, the lexical root presented here also is already known to be ancestral in a great number of language families worldwide. As a consequence, just like for *papa*, *mama* and *kaka*, any suggestion that their convergence might be due to chance would be preposterous. Let us now present the data establishing its existence.

2. THE PROTO-SAPIENS LEXICAL ROOT *MA- ‘PROHIBITIVE/NEGATIVE’

This etymology is **ma*, a prohibitive or negative particle. It was originally discovered by the Italian linguist Alfredo Trombetti (1905) and rediscovered by the first author in Trombetti's work in the late 1980s. We have entirely redesigned its empirical support, gathering reconstructions from many language families and directly adding data from language descriptions. It is found in a huge series of languages families and phyla (Table 1).

Table 1. Proto-Sapiens **ma*- ‘prohibitive/negative’²

[?? KHOISAN

SANDAWE: Sandawe **mě*: ‘not (prohibitive)’;

SOUTHERN KHOISAN: Central Khoisan: Proto-Khoe **tama*- ‘not’; Proto-Khoekhoe **tama*- ‘negative morpheme’; Nama *támá* ‘negative morpheme’; !Ora *tama* negative morpheme’; Proto-West Khoe

**-ta[ma]* ‘negative morpheme’; Naron *-tá*, - *támá* ‘negative morpheme’; //Gana *támà* ‘negative morpheme’;

(Güldemann & Elderkin 2010; G. Starostin 2007)]

NIGER-KORDOFANIAN

² Data between brackets preceded by a double question mark ([?? Khoisan ...], [?? Chadic ...] and [?? Hmong-Mien ...]) are uncertain reflexes; data between brackets preceded by a simple question mark are dubious members of the etymological series ([? Mongolic...], [? Ainu ...], [? Austronesian ...] and [? Proto-Mountain Ok ...]); they do not count as fully legitimate members of the etymology and are mentioned for the record. Sources are listed at the end of each phylum.

MANDE: Proto-Mande ***maa-**, ***mee-** ‘prefixes of the negative perfective of qualificative verbs’; Western: *North-western*: Soninke **ma** ‘negative verb marker’; Bobo **má** ~ **mā** ‘negative marker’; *Central Southwestern*: Susu **mu** ‘not’; Yalunka **ami** ‘not’; Kuranko **ma** ‘not’; Vai **má** ‘not’; Konyanka **ma** ‘not’; Mandinka **maŋ** ‘negative verb marker’; Malinke **ma** ‘negative verb marker’; Bambara **ma** ‘negative verb marker (past tenses)’; Dyula **ma** ‘negative perfective marker of qualificative verbs’; Mende **mā** ‘not’;

WEST ATLANTIC: Mel: Krim **ma** ‘negative verb marker’;

SOUTH-CENTRAL NIGER-CONGO: Gur: Senúfo **mɛ** ‘negative verbal marker’; **m’a** ‘negative imperative marker’; Kwa: Baule **man** ‘negative verb marker’; Fon **ma** ‘negative verb marker’, **ma ... nɔ́** ‘prohibitive verb marker’; Abidji **mɔ́~ mu** (sg.) ‘negative verb marker’; Abron **mā ~ mī** ‘negative verb marker’; Adioukrou **-m** ‘negative verb suffix’; Attié **ma** ‘negative verb marker’; Ewe **mé** ‘negative verb marker’; Yoruba **máà** ‘negative verb marker’; Gchode **mɛ̃** ‘negative past marker’; Krachi **mɛ̃** ‘negative verb marker’; Nchumuru **mi, mā** ‘negative verb markers’; Gonja **múN-** ‘negative verb marker’; Likpe **ma-** ‘negative verb marker’; Edoid: Edo **má** ‘negative verb marker’; Ijoid: Proto-Ijo ***ma** ‘negative verb marker’; Okrika **mà** ‘negative verb marker’; Plateau: Eloyi **mô** ‘negative verb marker’; Oko: Oko **ma, mi** ‘negative verb markers’; Adamawa-Ubangian: Sango **maa**; Gbaya Kaka **ma**;

KORDOFANIAN: Talodi-Heiban: Masakin **maa**;

(Bailleul 1977; Camara 1999; Dramé 2003; Duthie 1996; Fadaïro 2001; Fofana & Traoré 2003; Girier 1996; Gregersen 1972; Hérault 1982; Kouadio N’Guessan & Kouame 2004; Kropp Dakubu 1980; Kutsch-Lojenga & Hood 1982; Long 1971; Long & Diomandé 1968; Migeod 1908; Mohammed 2001; Prost 1983; Rongier 2002; Tresbarats 1992; Williamson 2004)

NILO-SAHARAN: Proto-Nilo-Saharan ***má-** ‘negative prefix of verbs’;

KUNAMA: Kunama **-mma** ‘neg. conditional’, **-mai** ‘neg. subjunctive & imperative’, **-mme** ‘neg. aorist, optative’;

SONGHAY: **-ma-** ‘neg. with perfect’; Nara **ma-** ‘neg. of perf., imper.’; Nubian ***m-** ‘verb neg. prefix’;

EASTERN SUDANIC: Kuliak: Ik **máá** ‘verb neg. marker of perf. & imper.’; Western: Merarit **m(V)-** ‘verb. neg. pref.’; Surmic: Didinga **ma-** ‘neg. of imper.’;

NILOTIC: Eastern: Lango **ma:m** ‘preposed verb neg. particle’; Teso **mam** ‘no’; Karamojong **man** ‘no’; Proto-Maa ***m(i)-** ‘verb neg. pref.’; Maasai **m-** ‘negative verb prefix’; Southern: Kalenjin ***-ma-** ‘id.’;

(Ehret 2001; Sicard & Malherbe 2005; Kitching 1915)

AFROASIATIC: Proto-Afroasiatic ***mV** ‘prohibitive particle’;

SEMITIC: Proto-Semitic ***mā** ‘not’; Central: Arabic **mā** ‘not’; South: Harari **mɛ?** ‘not’; Amharic **al- ... -m** ‘past verbal negative suffix’;

CUSHITIC: Saho-Afar: Afar **mā-**; Somali: Somali **má-** ... **in**; Rendille-Boni: Rendille **ma-** ‘negative prefix’;

Western Omo-Tana: Dasenech **ma**; Arbore **máala**; South: Iraqw **má**;

OMOTIC: South: Hamar-Banna **-ma**;

ONGOTAN: Ongota **mi-**;

EGYPTIAN: Ancient Egyptian **m** ‘do not! (prohibitive particle)’; survives in Coptic as the initial element of the negative Sahidic Coptic particle *mpōr* and Bohairic Coptic *mp^hōr* ‘do not!’ and some negative verbal prefixes: Sahidic *mpr-*, Bohairic *mper-* ‘do not...!’, the prefix *mpe-* of the negative perfect I, the prefix *mpu-* ‘you [f. sg.] did not’, the Sahidic prefix *mpate-* (and Bohairic *mpante-* ‘not yet’), Sahidic *mare-* ~ *mere-* ~ *ma-* ~ *me-*, Bohairic *mpare-* ~ *mpa-* (prefix of the negative aorist);

[?? CHADIC: West Chad Kofyar **má** ‘negative verbal marker’; Miya **má** ‘negative verbal marker’; Bade **-m** ‘negative verbal marker’;]

(Bomhard 2008; Cohen 1936; Dolgopolsky 2008; Kropp Dakubu 1980; Militarev & Stolbova 2007; Schuh 1998, 2003)

DENÉ-CAUCASIAN: Proto-Sino-Caucasian ***ma** ‘prohibitive particle’;

NORTH CAUCASIAN: Proto-North Caucasian ***ma** ~ ***mə** ‘prohibitive particle’; Nakh: Proto-Nakh ***ma** ‘prohib. particle, do not’; Chechen **ma**; Ingush **ma**; Batsbi **ma**; Tsezian: Proto-Tsezian ***-m** ‘negative particle’; Tsez **-n-č**; Ginukh **-go-m**; Khvarshi **-b-č**; Inkhokvari **-b-**; Lak: Lak **ma**; Lezghian: Proto-Lezghian ***mV** ‘prohibitive particle’; Lezghian **-mir**; Tabasaran **m-**; Agul **m-**; Rutul **m-**; Tsakhur **m-**; Kryz **m-**; Budukh **m-**; Udi **ma**;

West Caucasian: Proto-West Caucasian *mə ‘not’ (neg. particle); Abkhaz m-; Abaza -m-; Adyghe mə-; Kabardian mə-; Ubykh -m(a)-;

SINO-TIBETAN: Proto-Sino-Tibetan *mǎ(H) ‘not’; Sinitic: Chinese 無 ‘not have, not’; Preclassic Old Chinese ma; Classic Old Chinese ma; Western Han Chinese ma; Eastern Han Chinese mwa; Early Postclassic Chinese mwo; Middle Postclassic Chinese mwo; Late Postclassic Chinese mwo; Middle Chinese mü; Beijing u¹²; Jinan u¹²; Xi'an vu¹²; Taiyuan vu¹; Hankou u¹²; Chengdu vu¹²; Yangzhou u¹²; Suzhou vu³²; Wenzhou vu¹²; Changsha u¹²; Shuangfeng əu¹²; Nanchang u³¹; Meixian vu¹²; Guangzhou mou¹²; Xiamen bu¹² (lit.), bə¹²; Chaozhou bo¹²; Fuzhou u¹²; Shanghai fu³²; Zhongyuan yinyun u¹²; Jianchuan Bai -mo¹; Dali Bai mu¹; Bijiang Bai -mu¹ (cf. also Old Chinese 勿 mət ‘don't’, 未 mətš ‘not yet’, 亡 maŋ ‘have not’); Tibeto-Burmese: Tibetan ma ‘not’; Burmese maʔ ‘verbal negative’, majʔ ‘have not’; Kachin maʔ³ ‘be exhausted, ended’, šəmat² ‘be lost’, (H) ma ‘nothing’ (cf. also mje ‘be lost, gone’); Lushai *maʔ ~ māk' (cf. ma ~ māk ‘to give up, to divorce (one’s wife)’); Lepcha ma ~ mat negative, ma ~ mā ‘negative particle’ (cf. also Tibetan min (< ma-gin) ‘is not’; Moshang mu; Namsangia ma; Kanauri ma; Kham ma ‘negative affix’; Akha mah-xv ‘no, not’; Proto-Garo *ma^L ‘be lost’; Bodo-Garo: Bodo ga-ma ‘disappear’, ka-ma ‘to lose’; Dimasa gama, kama ‘lose, disappear, perish’, khama ‘destroy’; Garo mat ‘be spent’, gimat ‘destroy’; Kham ma^L ‘become lost’; Magari hma, hmat ‘be lost’);

YENISEIAN: Proto-Yeniseian *wə- ‘not, there is not’; Ket bḥñ~bḥñ ‘not, there is not’, bḥñšəŋ¹, bḥñšəŋ¹, bḥñčəŋ¹ ‘there is not’; Yug bḥñ ‘not, there is not’, bḥšə ‘there is not’; Kottish bō ‘prohibitive particle’, mon ‘not’; Assan mon ‘not’; Arin bon ‘there is not, not’; Pumpokol a-mút, amut ‘not’;

(S. Starostin 2007; Wang 2004; McDaniel 2002)

DRAVIDIAN: Proto-Dravidian *mal- ‘negative morpheme’;

NORTH: Proto-North Dravidian *mal ‘no, not’; Kurukh mal ‘not’; Malto mala, malā ‘not, no’, malnā ‘not to be (so)’, maʔā ‘not (when the negation falls on one single word which is being opposed to another word), no’, malkā ‘deprived of, lacking’, mal- (past mall-) ‘to be not’;

SOUTH: Proto-South Dravidian *mal- ‘negative morpheme’, Tamil -mal in negative adv. suffix -āmal; (Burrow & Emeneau 1984, etym. 3883)

KARTVELIAN: Proto-Kartvelian *ma- ‘not (prohibitive)’; Svan mā-d(e), mō-de; Laz mo-t;

(S. Starostin 2005b)

EURASIATIC: Proto-Eurasian *ma ‘prohibitive particle’;

INDO-EUROPEAN:

1. Proto-Indo-European *mē ‘prohibitive particle’; Tocharian: Agnean mā ‘not, no’; Kuchean mā ‘not, no’; Indo-Iranian: Proto-Indo-Iranian mā ‘not, no’; Indic: Sanskrit (R̥gveda) mā; Pali mā; Aśokan ma ~ mā; Apabhraṃśa mā; Gypsy (Europe & Armenia dials.) ma; Waigali ma ~ mi; Dameli ma; Pashai ma; Wotapuri ma; Kashmiri mā; Sindhi ma; Gujarati mā; Kalasha (Rumbur dial.) moh; Khowar mo; with various adverbial affixes: Prakrit māia, māi, maia, maṃta; Kashmiri matā; Sindhi matā, mataṇa, matuṇi; Lahnda mat-ṭaṇ, mattā, matā, mat; Old Awali matu, mati; Hindi mat; Old Marwari mati; Old Gujarati matu, mana]; Nuristani: Ashkun m’ā ‘don’t!’, Kalasha-ala (Nishei-ala dial.) ma-a ‘don’t!’, Iranian: Avestan mā; Old Persian mā; Ossetic mā; Armenian: Classical Armenian mi; Hellenic: Proto-Greek *mā; Elean mā; Homeric mé; Attic mé; Modern Greek mí; Albanian: Albanian mos;
2. Proto-Indo-European *(s)mal- ‘mean, malicious, small’; Tocharian: Agnean smale ‘lie’, smālok ‘liar’; Iranian: Avestan mairya- ‘fraudulent’; Armenian: Classical Armenian met ‘sin’; Hellenic: Classical Greek méleo- ‘vain, unhappy’; Slavic: Proto-Slavic *māl’ ‘small’; Old Slavic mal’ ‘small’; Russian mal ‘small’; Ukrainian malij ‘small’; Bielorrussian málý ‘small’; Bulgarian mál’k ‘small’; Serbian, Croatian máli ‘small’; Slovenian máli ‘small’; Czech, Slovak malý ‘small’; Polish mały ‘small’; Upper Sorbian małki ‘small’; Lower Sorbian małki ‘small’; Baltic: Proto-Baltic *mel-a- ‘lies’; Lithuanian mėla-s ‘lie’; Lettish mēli ‘lies’, melis ‘liar’; Germanic: Proto-Germanic *smal-a- ‘small’; Gothic *smal-s ‘small, little’; Old Norse smali ‘small cattle’; Norwegian smale ‘narrow’; Swedish smal ‘narrow’; Old English smāl ‘small’; English small; Old Frisian smel ‘small, minor’; Old Saxon smal ‘small, minor, narrow’; Middle Dutch smal ‘small, minor, narrow’;

Dutch **smal** ‘small’; Middle Low German **smal** ‘small’; Old High German **smal** ‘small, minor, narrow’; Middle High German **smal** ‘small, minor, narrow’; German **schmal** ‘narrow’; Italic: Latin **malus** ‘bad, mean’; Osk **malom** ‘bad, mean’; French **mal** ‘badly, poorly; trouble, illness, pain, harm’; Portuguese **mal** ‘bad; badly, poorly; trouble, illness’; Spanish **mal** ‘bad; badly, poorly; trouble, illness’; Occitan **mal** ~ **mau** ‘bad; badly; trouble, illness’; Italian **mal** ‘bad; badly; trouble, pain, illness’; Celtic: Proto-Celtic ***mell** ‘sin’; Middle Irish **mell** ‘sin, fault’;

ALTAIC: Proto-Altaic ***ma** ‘negative particle’; Turkic: Proto-Turkic ***-ma-** ‘not’; Old Turkish **-ma-**; Karakhanid **-ma-**; Turkish **-ma-**; Tatar **-ma-**; Middle Turkish **-ma-**; Uzbek **-ma-**; Uighur **-ma-**; Sary-Yughur **-ma-**; Azerbaijani **-ma-**; Turkmen **-ma-**; Khakassian **-ma-**; Shor **-ma-**; Oyrat **-ma-**; Halaj **-ma-**; Chuvash **-ma-**; Yakut **-ma-**; Tuva **-ma-**; Tofalar **-ma-**; Kirghiz **-ma-**; Kazakh **-ma-**; Noghai **-ma-**; Bashkir **-ma-**; Balkar **-ma-**; Gagauz **-ma-**; Karaim **-ma-**; Karakalpak **-ma-**; Salar **-mi-**; Kumyk **-ma-**; [Mongolic: Proto-Mongolic ***büi**, ***bu** ‘prohibitive particle’ may be originally the same morpheme, but functioning as a separate word and thus subject to the rule ***mV** > ***bV** in monosyllabic words: Written Mongolian **bü** ‘do not’; Middle Mongolian **bu**; Monguor **bī**; Dagur **bū**; Mogol **bi**; Ordos **bū**; Khalkha **būr**; Buriat **bu**; Oyrat **bō**]; Tungusic: Proto-Tungusic ***-me** ‘prohibitive particle’; Spoken Manchu **emə** ‘don’t’; Literary Manchu **ume**; Jürchen **ume**; Nanai **em**; Oroch **em**;

NIPPO-KOREAN: Korean: Proto-Korean ***mō-t** ‘impossible (adv.), bad, wicked’; Middle Korean **mōt**, **mōtīr-**; Modern Korean **mōt** [mōs] ‘prohibitive particle’, **mōžil-**; Japonic: Proto-Japonic ***-ma-** ‘dubitative suffix’; Old Japanese **-ma-**; Middle Japanese **-ma-**; Tokyo **-ma-i**;

(Nikolayev 2007; Abaev 1970; Turner 1962–1966 [etym. 9981]; Chantraine 1968; Poppe 1955; S. Starostin 2005a; S. Starostin 2006; Strand 1997–2012)

[? **AINU**:³ Ainu **isam** ‘not to be’ (to be compared to *isu* ‘to be’);

(Greenberg 2000: 213)]

AMERIND: Proto-Amerind ***mā** ‘negative’;

PENUTIAN: Maiduan: Proto-Maiduan ***-men** ‘negative’; Maidu **-men**; Konkow **-men-te** ‘without, but not’; Nisenan **-men** ‘not’; Wintuan: Wintun **-mina**; Yokutian: Yokuts **ʔoho-m**; Yo-Yaudanchi **ʔa-m**, **kʔamu** ‘not’; Zuñi: Zuñi **-(ʔ)amme**, **-na-ʔ-ma** ‘not’; Mayan: Proto-Mayan ***(ma)-n ... ta(x)** ‘negative’; Quiche **man ... tax**; Achi **n ... tax**; Pocomchi **ma ... ta**; Cakchiquel **man ... ta**; Tzeltal **maʔ** ‘not, without’; Ch’ol **maʔ** ~ **mač** ‘not’; Chontal **maʔ** ~ **mač** ‘not’; Chorti **ma-** ‘negative prefix’, **ma ači** ‘no, not’, **maan** ‘lack, failure’;

HOKAN: Seri-Yuman: Seri **m-** ‘negative’;

CENTRAL AMERIND: Uto-Aztecan: Proto-[?Uto-Aztecan]⁴ ***ma** ‘negative verb marker’; Tetelcingo Nahuatl **a ... mo** ‘negative verb marker’; North Puebla Nahuatl **aʔmo** ‘negative verb marker’; Huasteca Nahuatl **amo** ‘negative verb marker’; Michoacán Nahuatl **amo** ‘negative verb marker’; Northern Tepehuan **mai** ‘negative verb marker’; Monachi **minoʔ** ‘prohibitive’;

CHIBCHAN-PAEZAN: Chibchan: Yanomam **ma** ‘no, negative’; Paezan: Paez **-mee**;

ANDEAN: Quechuan: Quechua **mana** ‘negative’;

EQUATORIAL-TUCANOAN: Tucanoan: Tucano **mārī** ‘negative sg.’; Macro-Arawakan: Yavitero **yama** ‘negative’; Proto-Arawakan ***ma** ‘privative’; Parecís **maha** ‘negative sg.’; Taino **mar** ‘negative’; Proto-Tupi-Guarani ***maʔe-tei** ‘negative’; Siriono **emoa** ~ **mae eā** ‘nothing’;

GE-PANO-CARIB: Ge: Apinayé **mā** ‘negative’; Macro-Panoan: Lengua **ma** ‘no, negative’; Sanapana **ama** ‘no, negative’, **ma** ‘nothing’; Pano-Tacanan: Proto-Panoan ***[-ya]ma** ‘negative’; Cashibo-Cacataibo **-ma** ‘no, negative’; Catuquina **-yama** ‘nothing, no, negative’; Chacobo **-ma** ~ **-yama** ‘no, negative’; Shipibo-Conibo **-ma** ~ **-yama** ‘no, negative’, **yamaki** ‘nothing’; Yaminahua **ma** ‘no, negative’, **aβa-ya-ma** ~ **aβa-ma** ‘nothing’; Proto-Tacanan ***-ma** ‘no, not’; Tacana **moe** ~ **mawe** ~ **aimoe ... mawe** ‘no, negative’; Araona **ma** ‘no’, **maesa** ‘no, nothing’; Cavineña **-ma** ~ **ama** ‘no, negative’, **aihama** ‘nothing’; Ese’ejja **-ama** ‘no, negative’;

³ The classificatory status of Ainu is unsure, and therefore is listed here separately, though the authors each have an opinion regarding the phylum it may belong to.

⁴ According to Langacker (1977: 34), ***ma** ‘negative’ “might be reconstructed at some level” of Uto-Aztecan.

Carib: Proto-Carib ***-myra** ‘negative’; Eastern and Western Surinamese, Venezuelan and Guyanese Carib -**ma**, **-myn** ‘negative’; Wayana **-mna** ‘negative’; Trio **-nna**;

(Matteson 1972; Ultan 1964; Langacker 1977; Tuggy 1979; Brockway 1979; Marlett 2002; Beller & Beller 1979; Sischo 1979; Wisdom 1950; Lamb 1957; Swadesh 1967; Rowan & Burgess 2008; Key 2007a; Shell & Olive 1987; Lorient et al. 1993; Buckley & Ottaviano 1989; Kennell 2000; Prost & Prost no date; Pitman 1981; Key 2007b; Wyma & Wyma 1962; Courtz 2008)

AUSTRIC: Proto-Austrianic ***mV** ‘negative particle’;

AUSTROASIATIC: Proto-Austroasiatic ***ʔVm** ~ ***mVn** ‘not, negative’; Munda: Proto-Southern Munda ***əm** ‘negative verb particle’; Juang **ama-** ~ **am-** ~ **ma-** ‘negative verb particle’; Kharia **um** ‘negative verb particle’, ? **umbo** ‘negative verb particle’; Mon-Khmer: Proto-Mon-Khmer ***ʔam** ‘not’; Aslian: Semai **ʔama** ‘don’t (mild negative)’, **ʔamama** ‘no matter! never mind!’; Bahnaric: Mnong (Rölöm dial.) **məʔ** ‘not (preverbal)’, **ma:n** ~ **mə:** ‘do not’; Laven (Jru’ dial.) **ʔmip** ‘not possible’; Nyaheun **mmip** ‘negative’; Kasseng **ʔəhəəm** ‘not’; Yaeh **həəm** ‘not’; Tariang **həəm** ‘not’; Brao **tə- ... ʔiim** ‘not’; Ta’oih **ʔən** ‘not’; Halang **maʔ** ‘do not’; Katuic: Katu (An Diem dial.) **məʔ ʔawe:** ‘not’; Ngeq **həəm** ‘not yet’; Khasic: Proto-Khasic ***ham** ‘not’; Khasi **ʔe:m** ‘not’; War **mə ... laʔ** ‘negation of future’; Pnar (Jowai dial.) **wom** ~ **wum** ‘negative’; **ʔim** ‘negative marker’, **ham** ‘negator’; Pnar (Rymbai dial.) **ʔim** ‘negative marker’; Khmeric: Khmer **min** ‘not’; Surin Khmer **man** ‘not to, not to want to’; Khmuic: Khmu **ʔam** ‘not’; Khmu (Cuang dial.) **ʔam** ‘not’; T’in (Mal dial.) **mə:** ‘not yet’; Mlabri **met** ‘not’; Monic: Mon **mah mah** ‘nothing of importance, for nothing, gratis’; Nyah Kur (Huai Khrai, Nam Lao & Northern dials.) **ʔyym** ‘emphatic particle used with negative statements’ (Central, Klang, Southern & Tha Pong dials.), **mée** ~ (Huai Khrai, Nam Lao & Northern dials.) **mée** ‘emphatic particle used with negative, contradicting or correcting statements’; Palaungic: Proto-Palaungic ***ʔaŋ** ‘not, no’, Proto-Palaung-Wa ***mV** ‘not’, ***ʔaŋ** ‘not, no’; Plang **maŋ**³¹ ‘not’; Rumai **u⁵⁵ moh⁵¹** ‘not’; Lamet **maʔ⁵³** ‘not’; Khme **mak³¹ kəʔ⁵³** ‘not’; Proto-Waic ***ʔaŋ** ‘not’; Paraok **ʔaŋ** ‘not’; Wa **ʔaŋ** ‘not’; Wa (Kentung dial.) **āŋ** ‘not’; Wa (Southern dial.) **ang** ‘not’; U **àng** ‘not’; Pearic: Pear (Kompong Thom dial.) **miej** ‘to stop, cease; not do, be wary of’; Chong **mùuj mùuj** ‘not’; Chong (Samray dial.) **min** ‘no, not’, **maj** ‘to stop, cease; not do, be wary of’; Vietic: Vietnamese **həəm** (orthogr. *không*) ‘not’, Thavung **maʔ** ‘do not (prohibitive particle)’; Tum **nə²¹² mɛ:n²¹²** ‘not’;

TAI-KADAI: Hlai: Proto-Hlai ***ʔe:mh** ‘not’; Bouhin **ʔe:m²** ‘not’; Ha Em **ʔe:m²** ‘not’; Central Hlai (Qi) **ʔém** ‘not’; Kradai: Gelao (Judu dial.) **ma⁴²** ‘no, not’; Gelao (Niopo dial.) **ma⁵⁵** ‘no, not’; Be: Ong Be (Lincheng dial.) **mən²** ‘no, not’; Tai: Proto-Tai ***mi:A** ‘no, not’; Ningming **mi⁵** ‘no, not’; Daxin **mi²** ‘no, not’; Shangsi **mi³** ‘no, not’; Longzhou **mi⁵** ‘no, not’; Debao **mei²** ‘no, not’; Jingxi **mi²** ‘no, not’; Chongzuo **məi²** ‘no, not’; Fusui (Central dial.) **mi¹** ‘no, not’; Bouyei (Po-ai) **mei²**; Bouyei (Wangmo dial.) **mi³¹** ‘no, not’; Lianshan **mi²** ‘no, not’; Qinzhou **mai³** ‘no, not’; Yongnan **mei¹** ‘no, not’; Long’an **mui²** ‘no, not’; Proto-Zhuang-Tai ***mai¹**; Longzhou Zhuang **mi⁵**; Archaic Siamese **mi⁴** ‘no, not’; Siamese **maj³** ‘no, not’; Shan **maw²** ‘no, not’; Dehong **maau³** ‘no, not’; Kam-Sui: Southern Dong **mi³¹** ‘no, not’; Ai-Cham (Diwo dial.) **məi⁶** ‘not yet’; Mak (Yangfeng dial.) **me²** ‘no, not’; Then **me²** ‘no, not’; Biao **m⁶** ‘no, not’; Mak (Laliu dial.) **me²** ‘no, not’; Ai-Cham (Taiyang dial.) **me²** ‘no, not’;

[?? **HMONG-MIEN**: (likely Chinese borrowing) Hmongic: Dongnu Bunu (Nongjing dial.) **ma²** ‘no, not’; Younuo (Xiaozhai dial.) **mə²²** ‘no, not’; Mienic: Kim Mun **maa¹³ ~ ma³³** ‘no, not’;]

[? **AUSTRONESIAN**: East Formosan: Basai **maju** ‘not’; Basay **mia** ‘not’; Puyuman: Puyuma (Katipul dial.) **məli** ‘not’; Puyuma (Lower Pinlang dial.) **amli** ‘not’; Puyuma (Pilam dial.) **amli** ‘not’; Western Plains: Favorlang **maini** ‘no’; Malayo-Polynesian: Muna (Katobu-Tongkuno dial.) **miina** ‘no, not’; Wuna **mīna** ‘no, not’; Anakalang **da’ama** ‘no, not’; Baliledo **da’ama** ‘no, not’; Buru **mo** ‘no, not’; Paulohi **tama** ‘no, not’; Alune **mo** ‘no, not’; Selaru **léma** ‘no, not’; Levei **mole** ‘no, not’; 18th cent. Tahitian **’aima** ‘no, not’ (compare modern Tahitian *’aita*); Nanumea **mooe** ‘no, not’; Kusaie **mo[h]** ‘no, not’; Apma (Suru Kavian dial.) **ma ... nga** ‘negative sentence marker (’not)’; Nggela **mua** ‘no, not’; Kwaio **ʔamoe** ‘no, not’; Kwai **amoe** ‘no’; Dori’o **āmone** ‘no’; Oroha **mao** ‘no’; Sa’a (Sa’a village) **mao** ‘no’; ’Are’are (Ma’asupa village) **mao** ‘no’; ’Are’are (Waiahaa village) **mao** ‘no’; Santa Ana **marefa** ‘no’; Kahua (Mami dial.) **marefa** ‘no’; Kahua **mareha** ‘no’; Tawaroga **mareha** ‘no’; Santa Catalina **marefa** ‘no’; Haku **moa** ‘no, not’; Halia (Selau dial.) **moia** ‘no, not’; Numbami (Simboma dial.) **mou** ‘none’; Kove **mao** ‘not’; Kayupulau **moxa** ‘no, not’; Windesi Wandamen

moyar ‘no, not’; Marau **mao** ‘no’; Inabaknon **ma’in** ‘no, not’; Kadori **eam** ‘no, not’; Proto-Lampungic ***ma(kʔ)** ‘no, not’; Lampung **maʔwat** ‘no, not’; Komerling **maʔwat** ‘no, not’; Lampung Api (Belalau dial.) **mawat** ‘no, not’; Lampung Api (Jabung dial.) **mawat** ‘no, not’; Komerling (Kayu Agung Asli dial.) **homaʔ** ‘no, not’; Komerling (Kayu Agung Pendatang dial.) **homaʔ** ‘no, not’; Lampung Api (Kalianda dial.) **mawat** ‘no, not’; Komerling (Ulu dial., Adumanis village) **maʔwat** ‘no, not’; Komerling (Ulu dial., Darmapura village) **maʔwat** ‘no, not’; Komerling (Ulu dial., Perjaya village) **maʔwat** ‘no, not’; Komerling (Ulu dial., Perjaya village) **mawat** ‘no, not’; Komerling (Ilir dial., Palau Gemantug village) **maʔwət** ‘no, not’; Lampung Api (Kota Agung dial.) **mawat** ‘no, not’; Lampung Nyo (Abung/Kotabumi dial.) **maʔ** ‘no, not’; Lampung Nyo (Menggala/Tulang Bawang dial.) **maʔwaʔ** ‘no, not’; Lampung Nyo (Abung/Sukadana dial.) **maʔ** ‘no, not’; Lampung Api (Sukau dial.) **maweʔ** ‘no, not’; Lampung Api (Talang Padang dial.) **muwat** ‘no, not’; Lampung Api (Way Kanan dial.) **maʔwat** ‘no, not’; Lampung Api (Way Lima dial.) **maʔ** ‘no, not’; Lampung Api (Ranau dial.) **maweʔ** ‘no, not’; Lampung Api (Krui dial.) **maweʔ** ‘no, not’; Lampung Api (Pubian dial.) **mawat** ‘no, not’; Lampung Api (Sungkai dial.) **maʔwat** ‘no, not’; Modang **am** ‘no, not’; Bundu Dusun (Central Dusun) **amuʔ** ‘no, not’;]

(Anderson 2007; Blust et al. no date; Norquest 2007; Peiros & Starostin 2006; Shorto et al. 2006)

INDO-PACIFIC

TRANS-NEW GUINEA: Proto-Trans-New Guinea ***ma-** + verb ‘not’; South-East: Yareba **me**; Angan: Ankave **ma-**; Agaataha **maa-**; Finisterre-Huon: Dedua **mi**; Kâte **mi**; Ono **mi**; Komba **mâ**; Madang: Proto-Madang ***ma-**; Kalam **ma-** (-C), **m-** (-V); Waskia **me-**; Pila **me-**; Saki **me-**; Tani **me-**; Ulingan **me-**; Bepour **me-**; Wanuma **me-**; Yaben **me-**; Ukuriguma **me-**; Amaimon **me-**; Hinihon **ma-**; Abasakur **ma-**; Bilakura **ma-**; Sileibi **ma-**; Katiati **ma-**; Wadaginam **ma**; Eastern Highlands: Siane **am-**; Mid-Wahgi **ma** ‘no’ (interj.); Wiru **mo-**; Central: Samo **moi**; Kubo **moi**; Bibo **moi**; [? Proto-Mountain Ok ***ba**; Bimin **ba**; Faiwol **ba**; Telefol **ba**].

(Pawley 2000)

3. DISTRIBUTIONAL, PHONETIC AND SEMANTIC COMMENTS

A first remark is that **ma* ‘negative/prohibitive’ is among the most widely and firmly supported of all Proto-Sapiens etymologies published so far. Beyond the massive list of data from individual languages presented above, this is true both in terms of phyla (only Australian is not represented, perhaps because it was not investigated in any depth) and of reconstructed proto-languages, ancient languages and long-isolated languages represented (Table 2).

Table 2. Reconstructed, ancient and long-isolated languages with **ma-* ‘prohibitive/negative’

Sandawe ***mĕ:**, Pr.-Mande ***maa-**, Pr.-Ijo ***ma**, Pr.-Nilo-Saharan ***má-**, Pr.-Maa ***m(i)-**, Pr.-Afroasiatic ***mV**, Pr.-Semitic ***mā-**, Ancient Egyptian **m**, Pr.-Sino-Caucasian ***ma**, Pr.-North Caucasian ***ma**, Pr.-Nakh ***ma**, Pr.-Tsezian ***-m**, Pr.-Lezghian ***mV**, Pr.-West Caucasian ***mā**, Pr.-Sino-Tibetan ***mǎ(H)**, Preclassic Old Chinese **ma**, Pr.-Yeni-seian ***wə-**, Pr.-Dravidian ***mal-**, Pr.-North Dravidian ***mal**, Pr.-Kartvelian ***ma-**, Pr.-Eurasian ***ma**, Pr.-Indo-European ***mē**, Pr.-Indo-Iranian **mā**, Vedic Sanskrit **mā**, Avestan **mā**, Agnean **mā**, Classical Armenian **mi**, Pr.-Greek ***má**, Doric **má**, Homeric **mé**, Pr.-Altaic ***ma**, Pr.-Turkic ***-ma-**, Old Turkish **-ma-**, Pr.-Tungusic ***-me**, Middle Korean **mōt**, Pr.-Japonic ***-ma-**, Pr.-Amerind ***-mā**, Pr.-Maiduan ***-men**, Pr.-Mayan ***(ma)-n ... ta(x)**, Pr.-Uto-Aztecan ***ma**, Pr.-Arawakan ***ma**, Pr.-Tupi-Guarani ***maʔe-tei**, Pr.-Panoan ***[-ya]ma**, Pr.-Tacanan ***-ma**, Pr.-Carib ***-myra**, Pr.-Austric ***mV**, Pr.-Austroasiatic **ʔVm ~ *mVn**, Pr.-Southern Munda ***əm**, Pr.-Mon-Khmer ***ʔam**, Pr.-Khasic ***ham**, Pr.-Palaung-Wa ***mV**, Pr.-Palaungic ***ʔVm**, Pr.-Waic ***ʔaŋ**, Pr.-Hlai ***ʔe:mfi**, Pr.-Tai ***mi:**^A, Pr.-Trans-New Guinea ***ma-**, Pr.-Madang ***ma-**.

All the reconstructions in this list, based on data from their respective families, have been made by linguists operating independently from the Proto-Sapiens hypothesis (which many of them would presumably have opposed). Most or, possibly, none of them had ever heard of the particular hypothesis of a Proto-Sapiens word **ma* ‘prohibitive/negative,’ which has remained buried until today in Trombetti’s (1905) largely forgotten work. This entirely precludes the possibility that the convergence of these reconstructed words might be due to any kind of wishful thinking, much less any conspiracy, on the part of those who have postulated them.

It is to be noted that this list of ancient forms, and, for that matter, the general list of some 600 *ma*-forms, are far from exhaustive. African and Amerind languages, and still more languages of New Guinea – which represent together close to 4,000 languages – have been only superficially investigated. Most probably, both lists could be expanded to double or triple their present sizes.

An important phonetic detail is that, while the selection of putative cognates in the list was essentially made on the basis of the consonant *m-*, a strong majority of the words are also built with the vowel *-a*. This is particularly conspicuous in the more ancient putative cognates (Table 3).

Table 3. Reconstructed, ancient and long-isolated languages with an **m*-initial ‘prohibitive/negative’ + vowel *a*.

Pr.-Mande **maa-*, Pr.-Nilo-Saharan **má-*, Pr.-Semitic **mā-*, Pr.-Sino-Caucasian **ma*, Pr.-North Caucasian **ma*, Pr.-Nakh **ma*, Pr.-Tsezian **ma*, Pr.-Sino-Tibetan **mǎ(H)*, Preclassic Old Chinese *ma*, Pr.-Dravidian **mal-*, Pr.-North Dravidian **mal*, Pr.-Kartvelian **ma-*, Pr.-Eurasian **ma*, Pr.-Indo-Iranian *mā*, Vedic Sanskrit *mā*, Avestan *mā*, Agnean *mā*, Pr.-Greek **mā*, Elean *mā*, Pr.-Altaic **-ma-*, Pr.-Turkic **-ma-*, Old Turkish *-ma-*, Pr.-Japonic **-ma-*, Pr.-Amerind **mā-*, Pr.-Mayan **(ma)-n ... ta(x)*, Pr.-Arawakan **ma*, Pr.-Tupi-Guarani **maʔe-tei*, Pr.-Panoan **[-ya]ma*, Pr.-Tacanan **-ma*, Pr.-Mon-Khmer **ʔam*, Pr.-Khasic **ham*, Pr.-Waic **ʔaŋ*, Pr.-Trans-New Guinea **ma-*.

Most of these ancient forms are either words attested in writing or reconstructions supported by regular sound correspondences. Only a few of the highest-level cognates (e.g. Proto-Eurasian **ma*, Proto-Sino-Caucasian **ma*, Proto-Amerind **mā-* or Proto-Trans-New Guinea **ma-*) essentially rely on a majority of their own supporting reflexes rather than an analysis of regular sound changes. Multilateral etyma they are, and their exceptionless convergence on vowel *-a* with most regular reconstructions and attested ancient words is nonetheless striking.

Another important remark bears on the semantic side. A particular form of negation, consisting in a prohibition to act or a negation of an action, also enjoys a wide distribution in our series. It is represented by ‘prohibitive’ or ‘verbal negative’ particles or suffixes (Table 4).

Table 4. Reconstructed, ancient and long-isolated languages with **ma* ‘prohibitive’.

Proto-Mande **maa-* ~ **mee-* ‘prefixes of the negative perfective of qualificative verbs’, Proto-Nilo-Saharan **má-* ‘negative prefix of verbs’, Ancient Egyptian *m* ‘do not! (prohibitive particle)’, Proto-Nakh **ma* ‘prohibitive particle, do not’, Proto-Tsezian **-m* ‘negative particle’, Proto-Lezghian **mV* ‘prohibitive particle’, Old Chinese **mət*

‘don’t’, Proto-Kartvelian ***ma-** ‘not (prohibitive)’, Proto-Indo-European ***mē** ‘prohibitive particle’, Proto-Indo-Iranian **mā** ‘prohibitive particle’, Sanskrit **mā** ‘prohibitive particle’, Avestan **mā** ‘prohibitive particle’, Proto-Greek ***má** ‘prohibitive particle’, Proto-Turkic ***-ma-** ‘not’, Proto-Tungusic ***-me** ‘prohibitive particle’, Modern Korean **mōt** ‘prohibitive particle’, Proto-Uto-Aztecan ***ma-** ‘negative verb marker’, Proto-Southern Munda ***əm** ‘negative verb particle’, Proto-Trans-New Guinea ***ma-** + verb ‘not’.

Thus, beyond their sheer number, these ancient words and reconstructions strikingly coincide on several independent levels: (i) the root consonant *m*-; (ii) the root vowel *-a*; (iii) a negative, and, more specifically, prohibitive meaning. Taken together, these three repeatedly matching elements make random convergence an unlikely explanation.

Or do they? After all, both consonant *m* and vowel *a* are among the commonest sounds in the world’s languages (otherwise, we would not have found so many words made of them), while negation and prohibition are seemingly universal in human languages, and, like most other common shifters, are nearly always conveyed by short, typically monosyllabic words. From a sheer probabilistic viewpoint, there might be good chances that at least a good part of the *ma* words in the series presented here had arisen by random convergence.

But how do probabilities apply in this particular case? Most of the words in the list are known to descend from one or the other of the 50-odd ancestor languages mentioned in Table 2. Thus, they have been preserved over the last few millennia, at least, and cannot be said to have *recently* arisen from random convergence. Moreover, ancestral words themselves often have close parallels in related groups, e.g. Sanskrit *mā*, the ancestor language of the Indic group, matches Avestan *mā*, the ancestor language of the Iranian group, and words in various Nuristani languages, with all of which Indic constitutes the Indo-Iranian family. In turn, the Proto-Indo-Iranian form **mā* has parallels in the Tocharian, Armenian, Hellenic and Albanian families, that converge onto a still older form, namely Proto-Indo-European (PIE) **mē* ‘prohibitive particle’.

4. CHANCES FOR INDO-EUROPEAN *M-* NEGATIVES TO BE RECENT

How likely is it that PIE **mē* had resulted from a sudden innovation in this language, while its survival for 6 to 8 millennia in many of its daughter languages testifies to its resistance? Wait a minute, there also is a good number of daughter languages that lost it! Would not those languages having lost PIE **mē* reveal that negation is a fruitful soil for linguistic innovation, so that PIE could after all plausibly have created a new negation **mē*?

This is not the case, however. Most Indo-European languages that lost PIE **mē* did not replace it by a new word, but generalized, instead, the other PIE negative particle **nē*.⁵ Consequently,

⁵ PIE **nē* ‘neg. particle’; Anatolian: Hittite *natta* ‘not’, *nawi* ‘not yet’; Lydian *ni* ‘not’; Tocharian: Agnean *a(n)-* ‘neg. pref.’; Kuchean *e(n)-* ‘idem’; Indo-Aryan: Sanskrit *na*, *nā* ‘not’, *ned* ‘not, not indeed’, *a-* ‘un-’; Avestan *na* ‘not’, *nōit* ‘not at all’, *a-* ‘un-’; Armenian: Classical Armenian *an-* ‘neg. pref.’; Hellenic: Old Greek *ne-* ‘neg. pref.’, *a- ~ an-* ‘un-

many new negative words in Indo-European languages, like Latin *non* ‘no, not’ (< *noenum* < Archaic Latin *ne oinom* ‘no one’) or English *not* (< *naught* < Old English *na wiht* ‘no thing’) are compounds of an ancestral PIE negation with a reinforcing word.

Nevertheless, negatives not bearing a clear trace of an inherited PIE negative are found in the Indo-European family. Classical examples are Greek *ouk* ‘no, not’, Armenian *oč* ‘no, not’, Old Norse *eigi* ‘not’, *enngi* ‘no one’, *hvergi* ‘never’, and French *pas* ‘not’. All are explained by Indo-Europeanists as former complex expressions made of a descendant of **nē* with a reinforcing word, in which the negative meaning was transferred to the latter, whereafter the descendant of **nē* disappeared more or less completely, on the model of French⁶ *ne ... pas* (Cowgill 1960).

The most frequent emergence of new negatives from a pre-existing one certainly are the reasons why the Albanian negative *mos* is considered by Indo-Europeanists to be a legitimate descendant of PIE **mē*. Albanian *mos*, however, like any other Albanian word, is attested no earlier than five centuries ago and is thus separated from PIE by a factual night of 5,500 to 7,500 years – apparently, more than enough time for Pre-Albanian to lose PIE **mē* and recreate a new negative *mos*. Moreover, its ending *-os* is impossible to explain without postulating still another compounding of **mē* with a reinforcing word – for which there is no evidence except by analogy with compounds on **nē*. According to Joseph (2002), Albanian *mos* would derive from PIE **mē-k^wid* ‘not anything’ or **mē-k^we* ‘not anyone.’

However, it does not seem likely to Indo-Europeanists that the lineage of Albanian might have lost PIE **mē* and recreated *mos* during the 5 to 7 millennia before Albanian was first put in writing. Indeed, the idea that **mē* was preserved in Albanian, like in many other Indo-European languages, and underwent at some point a compounding by a process known in several other languages, is simpler than the idea that it was lost and then re-created, which did *not* happen in a single language of the IE branches whose respective ancient written languages had lost **mē*, like Italic, Celtic, Germanic or Slavic.⁷

Thus, preservation certainly is the most likely hypothesis for Albanian *mos*, but this goes with an important consequence. There being a slim chance that Albanian independently recreated an *m*-initial negative particle in 5,500 to 7,500 years (as well as for Tocharian, Indo-Iranian, Greek and

; Slavic: Proto-Slavic **ne* ‘not’, Baltic: Proto-Baltic **nē* ‘neg. particle’; Germanic: Proto-Germanic **nē* ‘neg. particle’, **un-* ‘neg. pref.’; Italic: Latin *ne-* ‘neg. pref.’, *nē* ‘not, that not’, *in-* ‘neg. pref.’; Osk *ne* ‘no, that not’, *an-* ‘un-’; Celtic: Old Irish *ne-* ‘neg. pref.’, *nō* ~ *no* ~ *nū* ‘no’, *ní* ‘not’; Cornish *ny* ‘not’; Breton *ne* ‘not’. By the way, **nē* is itself another widespread root, with correspondences at least in Uralic, Altaic, Japonic, Eskimo-Aleut, Kartvelian, Afroasiatic and Austric, showing again that it was not a new word in PIE, either.

⁶ Or, rather, Gallo-Romance, as *pas* ‘not’ is also general in Occitan and Franco-Provençal, both of which have, even more completely than French, eliminated the original *ne* from which *pas* ‘step’ originally drew its negative meaning (*je ne marche pas* ‘I do not walk a step’).

⁷ Actually, prohibitive *mi* does occur in Southeast Macedonian and Eastern Bulgarian, two South Slavic dialects in close contact with Greek, which still maintains PIE **mē* under a form *mí*, a regular evolution of Proto-Greek **mā* > Attic *mē*. Joseph (2002) rightly states that Southeast Macedonian and Eastern Bulgarian *mi* words have “clearly [been] borrowed from Greek [...], given [their] form and [their] absence from Slavic languages not in intimate contact with Greek.”

Armenian, in the shorter but by no means negligible timespans separating their first written attestations from their Indo-European origin) logically entails that chances for **mē* ‘prohibitive’ to have been recent in PIE itself are small as well.

At first sight, this improbability that **mē* had been recent in Proto-Indo-European may seem not to lead anywhere: it may or may not have *actually* been recent – from the Indo-European viewpoint, there simply is no way to tell.

5. CHANCES FOR ANCIENT **MA* NEGATIVES TO HAVE BEEN RECENT

But the PIE case is not unique. The same reasoning applies to each of the 40-odd other written or reconstructed ancestor languages, whose respective descendants have preserved an ancestral negative **ma* over the last millennia – for many of them, for 3,000 to 6,000 years or more.⁸ Just as with Albanian or Proto-Indo-European, it is also not very likely that their respective **ma* negatives were recent. And, just like for PIE, we cannot tell *for each* of these ancestor languages, seen from inside their respective families, whether its own **ma* word was or was not recent.

However, their number now allows us to make a general inference: as all had a tiny chance to have emerged randomly in a recent past, we may be sure that, *taken together*, most (and possibly all) of them were not recent. A few of them may have been, but the chances for more than a few to have appeared independently quickly drop to infinitesimality.⁹ As a consequence, while any of

⁸ Such estimated ages hold for low-level, easily reconstructible families. Ages of remote phyla like Khoisan, Nilo-Saharan, Dené-Caucasian or Eurasiatic may not be assessed, except sometimes through non-linguistic means, e.g. Proto-Amerind must be more or less isochronic with the genetic clock data for Amerind and the first human remains found by archeologists in North America, pointing to a period comprised between some 12,000 and 20,000 years ago, and consistent with the existence of an ice-free corridor in Beringia during this period.

⁹ It is possible to give a rough idea of the probability of the 40-odd reconstructed *m*-initial ancestral negatives (leaving aside those found in isolates or posited in remote macrofamilies and phyla) to be the result of convergent innovations. Each of the five Indo-European language groups with an *m*-initial negative is believed to have preserved PIE **mē* over the 3 to 7 millennia separating their first respective attestations from PIE times. Their *m*-initial negatives must thus be considered as having less than 1 chance out of 5 to have been a recent innovation in the group they are found in, a maximal probability which may indeed be much lower.

Even if roughly estimated, this individual probability of innovation of less than 1/5 in, say, 5 millennia makes possible, thanks to Bernoulli’s Binomial Law, to calculate the probability for any number of 40 ancestral *m*-initial negatives to have been recent in the proto-languages they belong to.

We have used the Microsoft Excel function BINOMDIST (*w*; *x*; *y*; *z*), which calculates the probability to get *w* winning trials out of *x* trials with an individual probability *y* for each trial to be a winning one;

these ancestral **ma* words, considered individually, may or may not be recent, we can safely assume that an overwhelming majority has been inherited from a remoter ancestor over several millennia before the time they were spoken.

Being mostly inherited from remoter ancestral languages, these 40 low-level ancestral *ma* words draw us a good bit farther back in time – say, some 10,000 to 15,000 years ago. What may have happened in the millennia having preceded them? Just like in the interval between PIE times and today, there must have occurred many language splits. Since the **ma* words of some 30 to 40 of our recent ancestor languages have to be several millennia older than the ancestors in which they are reconstructed, it would be a kind of miracle if none of them had resulted from the splits of ancestor languages spoken some millennia earlier, and did not descend from the same ancestral word, exactly like Albanian *mos*, Sanskrit *mā*, Avestan *mā* and Classical Greek *mē* descend from PIE **mē*. With so many ancestor languages having inherited the same word worldwide, the conclusion seems inescapable that at least some – and probably not a few – of them ought to have descended from common ancestors. This seems to be a strong argument in support of negative **ma* words posited by various authors (mostly independently of each other) in such ancient phyla as Nilo-Saharan, Afroasiatic, Eurasiatic, Dené-Caucasian, Amerind and Trans-New Guinea.

Transitively, this near universal presence in high-level linguistic phyla strongly supports an inheritance from a common origin, namely, in the language of the ancestral population of all modern humans, namely those who, some 100,000 years ago, left their African cradle and conquered the whole world.

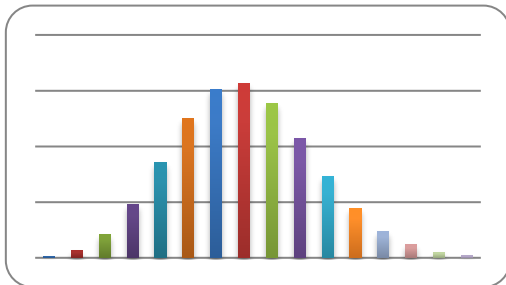


Table 5. Chances in percentages for 1 to 16 out of 40 ancestral *m*-initial negatives to be recent, calculated with the Microsoft Excel BINOMDIST function, based on an individual probability of 1/5. The total chances for 17 to 40 of them to be recent are of 0.1%, or 1/1,000.

argument *z* is a Boolean one; if set to “False,” the function returns the probability to get the exact given number *w* of winning trials; if set to “True,” it returns the total probability to get from 0 to *w* winning trials.

Using this BINOMDIST function, we have calculated the probability for each number of ancestral *m*-initial negatives to have resulted from recent innovations, based on a individual probability of 1/5 for each of them to be recent (Table 5).

The most likely numbers of innovations range between 6 (BINOMDIST (6; 40; 1/5; False) = 12.5%) and 10 (BINOMDIST (10; 40; 1/5; False) = 10.7%), while the total probability of any number over 16 drops close to zero (1 – BINOMDIST (16; 40; 1/5; True) = 0.01%).

As a consequence, even calculated based on the grossly over-rated probability that each 1 out of 5 of them might be recent, 23 ancestral *ma* negatives at least have to

have been inherited from an earlier ancestral language several millennia older.

Recall that the individual probability of 1/5 is an absolute limit, and that the actual one may only be lower, perhaps much lower, which would entail that much less ancestral *ma* words might be innovations. If it is set to 1/20 (instead of 1/5), there would be 1 out of some 1,400 chances that 8 or more of them be recent (1 – BINOMDIST (7; 40; 1/20; True) = 0.07%); if set to 1/100, there would be 1 out of some 1,400 chances that 4 or more be recent (1 – BINOMDIST (3; 40; 1/100; True) = 0.07%).

Whether their language was relatively unified or highly diversified, and what its degree of evolution may have been remain questions that have hardly been posed, much less answered. But we believe that the most reasonable hope of shedding some light into this darkness is to continue the job interrupted for nearly a century and unearth more Proto-Sapiens words – only they and the problems they will pose will tell us something about the evolution of language ability in humans.

REFERENCES

- ABAEV V. I. 1970. *Osetinsko-Russkij Slovar'—Iron-Uyryssag Dzyrduat* (Ossetic-Russian, Russian-Ossetic Dictionary). Ordžonikidze: IR.
- ANDERSON G. D. S. 2007. *The Munda Verb. Typological Perspectives*. Berlin/New York: De Gruyter Mouton.
- BAILLEUL C. 1999. *Cours pratique de bambara*. Faladjè: Centre d'études linguistiques.
- BANCEL P. J. & MATTHEY DE L'ETANG A. 2002. Tracing the Ancestral Kinship System: The Global Etymon *Kaka*. Part I: A Linguistic Study. *Mother Tongue* 7, 209–43.
- BANCEL P. J. & MATTHEY DE L'ETANG A. 2013. Brave New Words. In C. LEFEBVRE, B. COMRIE & H. COHEN, *New Perspectives on the Origin of Language*, Amsterdam: John Benjamins, pp. 333–78.
- BANCEL P. J. & MATTHEY DE L'ETANG A. ms. Reconstructing the Early History of Languages. The Eurasian Person Markers **m* and **t*.
- BELLER R. & BELLER P. 1979. Huasteca Nahuatl. In R. W. LANGACKER (ed.), *Studies in Uto-Aztecan Grammar*, vol. 2: *Modern Aztec Grammatical Sketches*. Dallas/Arlington: Summer Inst. of Linguistics/U. of Texas.
- BENGTSON J. D. & RUHLEN M. 1994. Global Etymologies. In M. RUHLEN (1994b).
- BLUST R. et al. No date. *Austronesian Basic Vocabulary Database*. Retrieved Jan. 12, 2012, <http://language.psy.auckland.ac.nz/austronesian/classification.php?node=Austronesian>.
- BOË L.-J., BESSIERE P., LADJILI N. & AUDIBERT N. 2006. Simple Combinatorial Considerations Challenge Ruhlen's Mother Tongue Theory. In B. L. DAVIS & K. ZAJDO (eds.), *The Syllable in Speech Production*. New York: Erlbaum.
- BOMHARD A. R. 2008. *Reconstructing Proto-Nostratic. Comparative Phonology, Morphology, and Vocabulary*. 2 vols. Charleston: Signum.
- BROCKWAY E. 1979. North Puebla Nahuatl. In R. W. LANGACKER (ed.), *Studies in Uto-Aztecan Grammar*, vol. 2: *Modern Aztec Grammatical Sketches*. Dallas/Arlington: Summer Inst. of Linguistics/U. of Texas.
- BUCKLEY DE OTTAVIANO A. & OTTAVIANO J. 1989. *Diccionario tacana-castellano, castellano-tacana*. Dallas, TX: Summer Inst. of Linguistics.
- BURROW T. & EMENEAU M. B. 1984 (2nd ed.). *A Dravidian Etymological Dictionary*. Oxford, UK: Clarendon. Retrieved December 31, 2011, from <http://dsal.uchicago.edu/dictionaries/burrow/>.
- BUSCHMANN J. C. E. 1852. Über den Naturlaut. *Philol. u. hist. Abhandl. kgl. Akad. Wiss. Berlin* 3, 391–423.
- CAMARA M. 1999. *Parlons malinké*. Paris: L'Harmattan.

- CHANTRAINE P. 1968. *Dictionnaire étymologique de la langue grecque. Histoire des mots*. Paris: Klincksieck. Retrieved May 7, 2010, from <http://www.archive.org/details/Dictionnaire-Etymologique-Grec>.
- COHEN M. 1936 [3rd ed. 1995]. *Traité de langue amharique (Abyssinie)*. Paris: Musée de l'Homme.
- COURTZ H. 2008. *A Carib Grammar and Dictionary*. Ph.D. diss., Leiden. Toronto: Magoria. Retrieved online Dec. 11, 2011, <<https://openaccess.leidenuniv.nl/handle/1887/12578>>.
- COWGILL W. 1960. Greek *ou* and Armenian *oč*. *Language* **36**–3/1, 347–50.
- DOLGOPOLSKY A. 2008. *Nostratic Dictionary*. Cambridge, Eng.: Cambridge U. Retrieved January 1, 2012, from <http://www.dspace.cam.ac.uk/handle/1810/196512>.
- DRAME M. F. 2004. *Parlons mandinka*. Paris: L'Harmattan.
- DUTHIE A. S. 1996. *Introducing Ewe Linguistic Patterns. A Textbook in Phonology, Grammar, and semantics*. Accra: Ghana U. Pr.
- EHRET C. 2001. *A Historical-Comparative Reconstruction of Nilo-Saharan*. Cologne: Rüdiger Köppe.
- FADAÏRO D. 2004. *Parlons fon. Langue et culture du Bénin*. Paris: L'Harmattan.
- FOFANA A. T. & TRAORE M. 2003. *Bamanankan Learners' Reference Grammar*. Madison, WI: NALRC.
- GIRIER C. 1996. *Parlons soninké*. Paris: L'Harmattan.
- GREENBERG J. H. 2000-2002. *Indo-European and Its Closest Relatives. The Eurasiatic Language Family*, vol. 1 (2000): *Grammar*; vol. 2 (2002): *Lexicon*. Stanford: Stanford U. Pr.
- GREGERSEN E. 1972. Kongo-Saharan. *J. Afr. Lang.* **11**–1, 69–89.
- GÜLDEMANN T. & ELDERKIN E. D. 2010. On External Genealogical Relationships of the Khoe Family. In M. BREZINGER & C. KÖNIG (eds.), *Khoisan Languages and Linguistics*. Proceedings of the 1st International Symposium January 4–8, 2003, Riezlern/Kleinwalsertal. Quellen zur Khoisan-Forschung 24. Köln: Rüdiger Köppe, 15–52.
- HERAULT G. (ed.). 1982. *Atlas des langues kwa de Côte d'Ivoire*, vol. I: *Monographies*. Abidjan: Université d'Abidjan.
- JAKOBSON R. 1960 [1971]. Why *Mama* and *Papa*? In B. KAPLAN & S. WAPNER (eds.), *Perspectives in Psychological Theory. Essays in Honor of Heinz Werner*. New York: International Universities Pr. Repr. in JAKOBSON R. 1971. *Selected Writings*, Vol. I: *Phonological Studies*, 538–45. The Hague: Mouton.
- JOSEPH B. D. 2002. Balkan Insights into the Syntax of **mē* in Indo-European. In M. SOUTHERN (ed.), *Indo-European Perspectives*. Washington, DC: Institute for the Study of Man, 103–20.
- KENNEL G. R., Jr. 2000. *Catuquina*. Intercontinental Dictionary Series. Online linguistic database, <http://lingweb.eva.mpg.de/ids/>.
- KEY M. R. 2007a. *Sanapana*. Intercontinental Dictionary Series. Online linguistic database, <http://lingweb.eva.mpg.de/ids/>.
- KEY M. R. 2007b. *Cavineña*. Intercontinental Dictionary Series. Online linguistic database, <http://lingweb.eva.mpg.de/ids/>.
- KITCHING A. L. 1915. *A Handbook of the Ateso Language*. London: Society for Promoting Christian Knowledge.

- KOUADIO N'GUESSAN J. & KOUAME K. 2004. *Parlons baoulé. E kan bawle. Langue et culture de Côte d'Ivoire*. Paris: L'Harmattan.
- KROPP DAKUBU M. E. (ed.). 1980a. *West African Language Data Sheets*, 2 vols. West African Linguistic Society.
- HOOD E. & KUTSCH-LOJENGA C. 1983. L'attié. In HÉRAULT 1983.
- LAMB S. M. 1957. *Mono Grammar*. PhD dissertation. Berkeley: University of California.
- LANGACKER R. W. 1977. *Studies in Uto-Aztecan Grammar*, vol. 1: *An Overview of Uto-Aztecan Grammar*. Dallas/Arlington: Summer Inst. of Linguistics/U. of Texas.
- LONG R. W. 1971. *A Comparative Study of the Northern Mande Languages*. Ph.D. diss. Bloomington: Indiana U.
- LONG R. W. & DIOMANDE R. S. 1999. *Basic Dyula*. Bloomington: Indiana U. Pr.
- LORIOT J., LAURIAULT E. & D. DAY. 1993. *Diccionario shipibo-castellano*. Lima: Min. de Educación/Inst. Lingüístico de Verano.
- LUBBOCK J. 1889. *The Origin of Civilisation and the Primitive Condition of Man. Mental and Social Condition of Savages*. (5th ed.). London: Longmans & Green.
- MARLETT S. A. 2002. Reanalysis of Passive and Negative Prefixes in Seri. *Ling. Disc.* 1–1.
- MATTESON E. 1972. Toward Proto-Amerindian. In E. MATTESON, A. WHEELER, F.L. JACKSON, N. E. WALTZ & D. R. CHRISTIAN, *Comparative Studies in Amerindian Languages*. *Janua Linguarum. Studia Memoriae Nicolai Van Wijk Dedicata*. The Hague/Paris: Mouton.
- MATTHEY DE L'ETANG A. & P. J. BANCEL. 2002. Tracing the Ancestral Kinship System: The Global Etymon *Kaka*. Part II: An Anthropological Study. *Mother Tongue*, 7, 245–258.
- MATTHEY DE L'ETANG A. & P. J. BANCEL. 2008. The Age of *Mama* and *Papa*. In J. D. BENGTSON (ed.), *In Hot Pursuit of Language in Prehistory. Papers in the Four Fields of Anthropology in Honor of Harold C. Fleming*, 417–438). Amsterdam/New York: John Benjamins.
- MCDANIEL M. 2002. *Akha to English Handbook. Two Thousand Words*. Maesai, Chiangrai: The Akha Heritage Foundation.
- MIGEOD F. W. H. 1908. *The Mende Language. Containing: Useful Phrases, Elementary Grammar, Short Vocabularies, Reading Materials*. London: Kegan Paul, Trench & Trübner.
- MILITAREV A. & STOLBOVA A. 2007. *Afroasiatic Etymology*. Online etymological database, Tower of Babel Project. Retrieved January 1, 2012, from <http://starling.rinet.ru/cgi-bin/main.cgi?flags=eygtnnl>.
- MOHAMMED M. A. 2001. *Modern Swahili Grammar*. Nairobi/Kampala/Dar es Salaam: East African Educational Publ.
- MURDOCK G. P. 1959. Cross-Language Parallels in Parental Kin Terms. *Anthr. Ling.* 1, 1–5.
- NIKOLAYEV S. 2007. *Indo-European Etymology*. Online etymological database, Tower of Babel Project. Retrieved January 1, 2012, from <http://starling.rinet.ru/cgi-bin/main.cgi?flags=eygtnnl>.
- NORQUEST P. K. 2007. *A Phonological Reconstruction of Proto-Hlai*. Ph.D. diss. University of Arizona.

- PAGEL M. 2000. Maximum Likelihood Models for Glottochronology and for Reconstructing Linguistic Phylogenies. In C. RENFREW, A. MCMAHON & L. TRASK (eds), *Time Depth in Historical Linguistics*, vol. I, 189–207. Cambridge (Eng.): McDonald Inst. for Arch. Res.
- PAWLEY A. 2000. *Some Trans New Guinea Etymologies*. Ms. Canberra: Australian National U.
- PEIROS I. & STAROSTIN S. 2006. *Austric Etymology*. Online etymological database, Tower of Babel Project. Retrieved January 1, 2012, from <http://starling.rinet.ru/cgi-bin/main.cgi?flags=eygttnl>.
- PITMAN Mary. 1981. *Diccionario araona y castellano*. Riberalta, Bolivia : Inst. Lingüístico de Verano/Min. de Educación y Cultura.
- POPPE N. 1955 [1987, 2nd ed.]. *Introduction to Mongolian Comparative Studies*. Helsinki: Suomalais-Ugrilainen Seura.
- PROST A. 1983. *Essai de description grammaticale du dialecte bobo de Tansila (Haute-Volta)*. Paris: Mandenkan.
- PROST G. & PROST M. No date. *Unpublished Chacobo dictionary file*. Intercontinental Dictionary Series. Online linguistic database, <http://lingweb.eva.mpg.de/ids/>.
- RINGE D.A. 2002. Review of Greenberg (2000). *J. of Linguistics*, **38**–2, 415–20.
- RONGIER J. 2002. *Parlons sénoufo*. Paris: L'Harmattan.
- ROWAN O. & BURGESS E. 2008. *Gramática Parecis*. Anápolis, GO: Associação Internacional de Lingüística.
- RUHLEN M. 1994a. *The Origin of Language. Tracing the Evolution of the Mother Tongue*. New York: John Wiley.
- RUHLEN M. 2000. Why *Kaka* and *Aya*? In D. G. LOCKWOOD, P. H. FRIES & J. E. COPELAND (eds). *Functional Approaches to Language, Culture and Cognition*. Amsterdam/Philadelphia: John Benjamins.
- SCHUH R. G. 1998. *A Grammar of Miya*. Berkeley/Los Angeles: University of California Pr.
- SCHUH R. G. 2003. The Functional Unity of the Hausa and West Chadic Subjunctive. *UCLA Working Papers in Linguistics, Papers in African Linguistics* **3**, 17–42.
- SHELL R. & OLIVE A. 1987. *Vocabulario cashibo-cacataibo. Revisado por Marlene Ballena Dávila*. Yarinacocha: Min. de Educación/Inst. Lingüístico de Verano.
- SHORTO H. L., SIDWELL P., COOPER D. & BAUER C. (eds.). 2006–. *A Mon-Khmer Comparative Dictionary*. Canberra: Pacific Linguistics. Retrieved January 21, 2012, from <http://sealang.net/monkhmer/dictionary/>.
- SICARD G. M. & MALHERBE M. 2005. *Parlons massai*. Paris: L'Harmattan.
- SISCHO W. R. 1979. Michoacán Nahuatl. In R. W. LANGACKER (ed.), *Studies in Uto-Aztecan Grammar*, vol. 2: *Modern Aztec Grammatical Sketches*. Dallas/Arlington: Summer Inst. of Linguistics/U. of Texas.
- STAROSTIN G. 2007. *Central Khoisan Etymology*. Online etymological database, Tower of Babel Project. Retrieved January 10, 2012, from <http://starling.rinet.ru/cgi-bin/main.cgi?flags=eygttnl>.
- STAROSTIN S. 2005a. *Uralic Etymology*. Online etymological database, Tower of Babel Project. Retrieved January 1, 2012, from <http://starling.rinet.ru/cgi-bin/main.cgi?flags=eygttnl>.

- STAROSTIN S. 2005b. *Kartvelian Etymology*. Online etymological database, Tower of Babel Project. Retrieved January 1, 2012, from <http://starling.rinet.ru/cgi-bin/main.cgi?flags=eygtnnl>.
- STAROSTIN S. 2006. *Altaic Etymology*. Online etymological database, Tower of Babel Project. Retrieved January 1, 2012, from <http://starling.rinet.ru/cgi-bin/main.cgi?flags=eygtnnl>.
- STAROSTIN S. 2007. *Sino-Caucasian Etymology*. Online etymological database, Tower of Babel Project. Retrieved January 1, 2012, from <http://starling.rinet.ru/cgi-bin/main.cgi?flags=eygtnnl>.
- STRAND R. 1997–2012. *Nuristân. Hidden Land of the Hindu-Kush*. Retrieved January 22, 2012, from <http://nuristan.info/>.
- SWADESH M. 1967. Linguistic Classification in the Southwest. In D. HYMES, *Studies in Southwestern Ethnolinguistics*, pp. 281–309. The Hague: Mouton.
- TRASK L. 2004. Where Do *Mama/Papa* Words Come From? *U. Sussex Work. Pap. Ling. & Eng. Lang.* Retrieved online 10 Apr. 2010, <http://www.sussex.ac.uk/linguistics/documents/where_do_mama2.pdf>.
- TRESBARATS C. 1992. *Esquisse linguistique de l'abidji*. Paris: Agence de coopération culturelle et technique.
- TROMBETTI A. 1905. *L'Unità d'Origine del Linguaggio*. Bologna: Luigi Beltrami.
- TUGGY D. H. 1979. Tetelcingo Nahuatl. In R. W. LANGACKER (ed.), *Studies in Uto-Aztecan Grammar*, vol. 2: *Modern Aztec Grammatical Sketches*. Dallas/Arlington: Summer Inst. of Linguistics/U. of Texas.
- TURNER R. L. 1962–1966. *A Comparative Dictionary of Indo-Aryan Languages*. London: Oxford University Press. Retrieved January 1, 2012, <http://dsal.uchicago.edu/dictionaries/soas/>.
- ULTAN R. 1964. Proto-Maidun Phonology. *Int. J. Am. Ling.* 30: 355–70.
- WANG W. S.-Y. 2004. *Chinese Dialects*. Online etymological database, Tower of Babel Project. Retrieved 26 Jan. 2012, <http://starling.rinet.ru/cgi-bin/main.cgi?flags=eygtnnl>.
- WESTERMARCK E. 1891. *The History of Human Marriage*. London: Macmillan.
- WILLIAMSON K. 2004 *Proto-Ijoid Reconstructions*. Retrieved online 9 Jan. 2012, <<http://www.nostratic.ru/index.php?page=books&lang=en>>.
- WISDOM C. 1950. *Chorti Dictionary*. Transcr. and translit. by B. STROSS. Austin: U. of Texas.
- WYMA L. & WYMA R. 1962. *Ese 'ejja y castellano*. Cochabamba: Inst. Lingüístico de Verano.