Ethnologue 16/17/18th editions: A comprehensive review: Online appendices

Harald Hammarström

Language, Volume 91, Number 3, September 2015, pp. s1-s188 (Article)

Published by Linguistic Society of America

DOI: 10.1353/lan.2015.0049

For additional information about this article
http://muse.jhu.edu/journals/lan/summary/v091/91.3.hammarstrom01.html
Ethnologue 16/17/18th editions: A comprehensive review: Online appendices

Harald Hammarström
Max Planck Institute for Psycholinguistics, Nijmegen

A  Missing Languages

This section lists languages which are missing from E16/E17/E18. To be more precise, a language is listed here as missing if:

- Extant published literature can make a convincing case that the language exists (or existed, see below), and,

- Extant published literature can make a convincing case that the language is not intelligible to any language already listed in E16/E17/E18, and,

An important note is that we do not list languages which are missing solely in virtue of the interpretation of a dialect situation correctly understood (but interpreted differently) in E16/E17/E18. This matter is separately treated in Section D\(^1\). For example, if an E16/E17/E18 entry subsumes a number of varieties with borderline intelligibility, and the facts are correctly indicated (e.g., the names of the varieties as dialects and comments on intelligibility) we do not list such cases here, even if there are good reasons to interpret the same facts as yielding different entries. However, if an E16/E17/E18 entry shows signs of misunderstanding (missing the existence of a variety, has an erroneous indication of intelligibility level, or a blanket statement with no indicated basis etc.), we do list any variety that is arguably not intelligible as a missing language here.

In all cases, we provide references to the literature that support the claims made regarding the missing language in question.

A.1  Missing Languages: Australia

E16/E17/E18 Athima, a poorly but sufficiently attested extinct languages of the Upper Southwest Paman group, is missing. It is related to but different from Kunjen [kjn] and Kawarrang-Ogh Undjan [-] (Alpher 2000).

E16/E17/E18 Koko Dhawa, an extinct separate language (Sommer 1997a:3, 19) from Kok Kaper (aka Gugubera) [kkp], is missing.

\(^1\)A large number of further such cases are taken up in the list of scheduled updates to the iso-639-3 inventory, traceable via http://www.sil.org/iso639-3/changes.asp.
Barranbiya, an extinct close relative of Muruwari [zm], is missing.

Bigambal, an extinct Yugambalic language, is missing (Barlow 1873, Mathews 1902, Ridley 1873, 1875a, Turbayne et al. 1887, Wafer & Lissarrague 2008).

The extinct Bindal language recorded by Cunningham, a Greater Maric language arguably a different language from the other Bindal languages, is missing (Breen 2009, Cunningham & Gorton 1886).

The extinct Bindal language recorded by Gorton, a Greater Maric language arguably a different language from the other Bindal languages, is missing (Breen 2009, Cunningham & Gorton 1886).

The extinct Bindal language of Mount Elliot, a Greater Maric language arguably a different language from the other Bindal languages, is missing (Breen 2009, Gregory 1886).

Birrdhawal, an extinct Pama-Nyungan language, is missing (Breen 2009, Gregory 1886).


Giya, an extinct Eastern Maric language, is missing (Breen 2009, Curr 1887, Shea 1887).

Gudjal, an extinct Northern Maric language, is missing (Breen 2009, Lukin 1886, Sutton 1973a,b).

Gwar, an extinct Bandjalangic language, is missing (Jefferies 2011, Watkin & Hamilton 1887).

Guyambil, an extinct Yugambalic language, is missing (Bench of Magistrates, Queenbulla District 1887, Wafer & Lissarrague 2008).

The language of Hawkesbury, an extinct Yuin-Kuri language, is missing (Jones 2008, Mathews 1897, 1903a, 1904b, Tuckerman 1887, Wafer & Lissarrague 2008).

Ilkarranggal, an extinct Rarmul Paman language, is missing (Palmer 1884, 1886, Ray 1907, Sommer 1976, no date).

Karruwali, an extinct Central Karnic language, is missing (Anonymous 1886, Breen 1971).


The extinct language of the lower Darling and lower Lachlan (of the Lower Murray subgroup), is missing (Horgen 2004, Macdonald 1886a,b, Smyth 1878, Wafer & Lissarrague 2008).

The extinct Lower Riverland language (of the Lower Murray subgroup), is missing (Eyre 1845, Grimwade 1975, Horgen 2004, Moorhouse 1846, 1886, Pratt 1886, Taplin 1872, 1879).

The extinct Lower Southern Aranda (an Arandic language), is missing (Belt 1886, Hale 1962, London & Mueller 1886, Warburton 1886).

Minkin, a poorly attested extinct language of possible Tangkic affiliation, is missing (Coward 1886, Tryon & O’Grady 1990).

Mithaka, an extinct Central Karnic language, is missing (Bowern 2001, Breen 1971, Mathew 1899).


Ngadjuri, an extinct Thura-Yura language, is missing (Berndt & Vogelsang 1941, Simpson & Hercus 2004, Warie-Read et al. 2009).

Ngardi, a Ngumpin language, is missing (Honeyman 2005, McConvell & Laughren 2004).

Ngumbarl, an extinct Nyulnyulan language, is missing (Bowern 2010).


The extinct Northern Sunraysia language the Lower Murray subgroup, is missing (Cameron 1885, Horgen 2004, McFarlane 1886, Smyth 1878, Wafer & Lissarrague 2008).

The extinct Southern Coastal Yuin language of the Yuin subgroup is missing (Liston 1993, Mathews 1902, Wafer & Lissarrague 2008).

The extinct Southern Inland Yuin language of the Yuin subgroup is missing (Bulmer 1887a,b, Bunce 1851, Du Vé & Bulmer 1887, Eyre 1845, Hercus 1969, Koch 2011, Lhotsky 1839, Mathews 1908, Police Magistrate, Queanbeyan 1887, Wafer & Lissarrague 2008).

The extinct original Sydney language (of the Kuri subgroup), is missing (Collins 1887, Hunter 1887, Kohen 1984, Marsden 1834, Mathews 1901b, Ridley 1873, Steele 2005, Troy 1992, 1994a,b, Wafer & Lissarrague 2008).

Tagalaka, a Southwest Paman language, is missing (Sutton 1976a).

Umiida-Unggarangu, an extinct Worrorran language, is missing (Capell & Coate 1984, McGregor & Rumsey 2009).

Winjarumi, an extinct Worrorran language, is missing (Capell & Coate 1984, McGregor & Rumsey 2009).

Unggumi, an extinct Worrorran language, is missing (Capell 1940, 1940, Capell & Coate 1984, McGregor & Rumsey 2009).

The extinct Upper Riverland of the Lower Murray subgroup is missing is missing (Fulford 1886, Horgen 2004, Pegler 1886, Taplin 1879).


The long extinct and poorly attested Tasmanian languages are missing altogether. The exact number of languages in evidence is uncertain but a tradition of scholarship (Bowern 2012, Bowern & Lau 2011, Crowley & Dixon 1981, Jones 1974, Schmidt 1952) recognizes some seven languages which may be divided up as Western Coastal Tasmanian, the Northeastern Tasmanian dialect chain, Oyster Bay-Big River-Little Swanport, Port Sorell, South-Eastern Tasmanian Hinterland, Bruny Island and Ben Lomond-Cape Portland,

Yagara, an extinct language of the East Queensland Border Pama Nyungan subgroup, is missing (Holmer 1983, Lauterer 1900, Minjerribah Moorgumpin 2011, Ridley 1873, 1875a,b, 1887, Watkin & Hamilton 1887).

Yulpara, a Martuwayngkic language, is missing (Burridge 1996, Hansen 1984).

Yuru, an extinct Nyawayngic language, is missing (Breen 2009, O’Connor 1886, Wurm 2007).

The language consisting of the two dialects Kawarrang and Ogh Undjan is missing. There is a Gugu Mini [ggm] entry and the name Gugu Mini has been applied to Ogh Undjan (Sommer 1997b:1-3, Sommer 1976:133-144) as well as to Ikarrangal [ikr] (Sommer no date:1-2, Sommer 1976:133-144), but the [ggm]-entry is classed as a Rarmul Pama language, implying that it refers to Ikarrangal rather than Kawarrang-Ogh Undjan (which is not a Rarmul Pama language). The closest relative of Ogh Undjan is Ow Oykan and or possibly Kuuk Thayoore but they are not as close as to be even nearly intelligible (Sommer 1969:12, Black 2004:253).

Kalaamaya, a Pama-Nyungan language, related to the Mirning group languages but not intelligible with the varieties already included (Kalarko [kba], Ngadjunmaya [nju]) is missing (Nash 2002). A wordlist was recorded as late as 1960 so it survived beyond 1950 (O’Grady 2001).

Dharumbal, a Pama-Nyungan language surviving well beyond 1950, is missing. The Bayali [b[jy] entry has alternate names that indicate Dharumbal, but Bayali, Dharumbal’s neighbour to the south, is certainly a different language, being only 21% lexicostatistically cognate with Dharumbal (Terrill 2002:15).

Yaygir, a Pama-Nyungan language discovered only in 1973, is missing (Crowley 1979).

Anguthimri, a Pama-Nyungan language, is missing2 (Crowley 1981).

The extinct, but quite well-attested, Kulin languages are missing altogether. One of the Kulin languages was the Western Victoria language (Wafer & Lissarrague 2008), which was not necessarily extinct by 1950 (Hercus 1969).

Another extinct, but quite well-attested, Kulin languages is Warrnambool which is also missing (Blake 2003b, 2011, Blake & Reid 1995, 1998, Mathew 1899, Wilkinson 1978).

The extinct, but quite well-attested, Kulin languages are missing altogether. One of the Kulin languages was Woiwurrung (Blake 1991), which was not necessarily extinct by 1950 (Hercus 1969).

The extinct, but quite well-attested, Kulin languages are missing altogether. One of the Kulin languages was the Wathawurrung language (Blake et al. 1998a), which was not necessarily extinct by 1950 (Hercus 1969).

The extinct, but quite well-attested, Kulin languages are missing altogether. One of the Kulin languages was the Colac language (Blake et al. 1998b), which was not necessarily extinct by 1950 (Hercus 1969).

---

2I wish to thank Jean-Christophe Verstraete for discussion on this language.
E16/E17/E18 Uwinymil, a recently extinct Western Gunwinyguan language is missing ((Harvey 2003a:285), Parkhouse 1896:2)

E16/E17/E18 An extinct, presumably Western Gunwinyguan language, recorded as Wulwulam is missing (Harvey 2003b). The language is attested only in a short wordlist but this wordlist is not explainable as Warray [wrz] or any other neighbouring language (Harvey 2003b:299).

E16 Kuuk Yak (as a dialect or language) is missing (Gaby 2006).

E16/E17/E18 Gurindji Kriol, an Australian mixed language, is missing (McConville & Meakins 2005, Meakins 2011), though noted in the comments to the Gurindji [gue] entry: “Gurinji children’s language is a mixed language from Gurinji and Kriol”.

E16/E17/E18 Light Warlpiri, an Australian mixed language, is missing (O’Shannessy 2006).

E16/E17/E18 Bularmu, a Pama-Nyungan language of the Southern Ngarna, Thawa group is missing (although erroneously listed as a dialect of Yindjilandji [yil] in the Southern Ngarna, Ngarru subgroup, see Breen 2004).

E16/E17/E18 Wulguru, an extinct Nyawayyic language, is missing (Donohue 2007b)

E16/E17/E18 Yirandhali, an extinct Maric language, is missing (Breen 2009)

E16/E17/E18 Pirriya, an extinct Karnic language, is missing (Barrett 2005, Breen 1990b).

E16/E17/E18 Kuungkari of the Barcoo river, another extinct Karnic language, is missing (Barrett 2005, Breen 1990a). (E16 has an entry for Kunggari [kgl] but the data in this entry matches the Bidjara-like language of Upper Nebine and Mungallala creeks and not the similarly named Kuungkari, of Barcoo river).

E16/E17/E18 Guwa, an extinct Greater Maric language, is missing (Barrett 2005, Blake & Breen 1990).

E16/E17/E18 Yanda, another extinct Greater Maric language, is missing (Barrett 2005, Breen 1990c).


E16/E17/E18 Wuthathi, an extinct northern Paman language, is missing (Hale 1966, Seligmann & Pimm 1907).

E16/E17/E18 Gudang, another extinct northern Paman language, is missing (Hale 1966, Jardine 1886).

E16/E17/E18 Mpalityan-Lutigh, another extinct northern Paman language, is missing (Hale 1966, Sommer 1976).
Yinwum, another extinct northern Paman language, is missing (Hale 1966, Sommer 1976).

The extinct northern Paman language comprising the lects Tyanngayt, Mamngayt, Ntrwangayt and Ntrangit, is missing (Hale 1966, Hey 1903).

Ngkot, yet another extinct northern Paman language, is missing (Hale 1966).

Aritningitigh, yet another extinct northern Paman language, is missing (Hale 1966).

Mbiywom, yet another extinct northern Paman language, is missing (Hale 1966).

Kok-Nar, an extinct Norman Pama language, is missing (Breen 1976).

A.2 Missing Languages: Africa

Shiwa, a Gabonese Bantu A80 language, missing (Dougère 2007, Ollomo Ella 2008, 2013). It is most closely related to Kwasio [nmg], but about equally close to it as any pair of the closest A80 languages with separate entries (Cheuclé 2014:516, Grollemund 2012).


Teshenawa, an extinct West Chadic B1 relative of Bade, is missing (Broß 1996, Schuh 2001).

Baazem, a newly discovered Yukubenic language, is missing (Lovegren 2012).

Boro, an extinct language Na-Togo language of Togo, is missing (Seidel 1898, Westermann 1922).

Gafat, a relatively recently extinct Ethiosemitic language, is mysteriously missing. It has a long history of scholarship (Adelung & Vater 1812, Beke 1846, Hetzron 1997, Leslau 1945).

Gbin, an extinct but attested language of Côte d’Ivoire is missing (Delafosse 1904, Paperno no date).


The extinct but well-known Ancient Egyptian language is missing (Loprieno 1995).

Guanche, the enigmatic poorly attested extinct original language of the Canary islands, is missing (Wölfel 1965).
E16/E17/E18 Meroitic, an enigmatic poorly attested extinct language of North Sudan, is missing (Griffith 1909, Rilly 2003, 2007, Rowan 2006).


E16/E17/E18 A series of extinct Bushman lects forming a language which may be called Lower Nosop, is missing (Bleeck 1937, 1939-1940, Guldemann 2014, Traill 1999).

E16/E17 A series of extinct Bushman lects forming a language which may be called Vaal-Orange, is missing from E16/E17 (Bleeck 1939-1940, 1956, Meinhof 1928-1929).

E16/E17/E18 Njanga, arguably a distinct language from Kwanja [knp], is missing (Connell 1998, Griffiths & Robson 2010, Robson 2011).

E16/E17/E18 Data in Klieman (2003, 1997), Sato (1992) of the language spoken by the Mikaya, Bambengangale and Baluma in the Republic of Congo are so different from other languages in the Likouala-Sangha Bantu subgroup that a separate entry is required. These lects are not covered in any form in any other entry in E16/E17/E18.

E16/E17/E18 There is a small amount of data on a now extinct enigmatic language of Ghana called Mpra (Blench 2012c, 1999, Goody 1963). It is missing from E16/E17/E18.

E16/E17/E18 The extinct but well-attested Phoenician-Punic language is missing (Friedrich 1951, Harris 1936, Kerr 2010).

E16/E17/E18 According to van der Burgt (1902), the language of the Kitwa among the Rundi speak a language of their own (in addition to Kirundi itself) which is not understood by the neighbouring Watussi and the Wahutu speakers and has the same grammar as Kirundi but is lexically different (the bulk of which is nevertheless of Rundi origin!). This language is missing from E16/E17/E18.

E16/E17/E18 Shanjo is a language of Zambia which appears as a dialect of Tonga (Zambia) [toi] in E16/E17/E18. However, it turns out that Shanjo is not intelligible to Tonga and does not even subgroup with Tonga (Bostoen 2009, de Luna 2008, 2010) and thus merits an entry of its own.

E16/E17/E18 The Nyla [nle] entry lists two dialects West Nyla and East Nyla. But the West Nyla language spoken on the shores of Lake Victoria in Busia district in Kenya is a different language from the non-adjacent East Nyla spoken inland (Kanyoro 1983). The two are not mutually intelligible (Ochwaya-Oluoch 2003:7-8) despite both having the name Nyla and being the the Abaluyia group. One of the two is thus missing from E16/E17/E18.

E16/E17/E18 Belueli, a Pygmy language spoken at Apare in the Ituri forest area of NE Congo, is missing. It is not intelligible to neighbouring Bali [bcp] and appears to the in the Nyla-Ndaka-Mbo-Budu-Vanuma subgroup (Schebesta 1953).
Nabit is listed as a dialect of Farefare [gur], but it is not intelligible with Farefare (Giffen 2013:100) and should thus be listed as a separate language.

Nterato is a language in the northern Guang area in Ghana recorded by Jack Goody in 1956 (Goody 1963), but with only two speakers still remaining at the time. The language is related to the other north Guang languages but is different from all of them (Kropp Dakubu 2013). It is missing.

Ngubi is a Sira (B.40) Bantu language in Gabon distinct from all other Sira (B.40) languages (Agadji Ayele 2002, Aleko & Puèch 1988, Puèch 1988). It is missing.

Tondi Songway Kiini, a Songhay language not intelligible to other Eastern Songhay languages listed in E16, is missing (Heath 2005).

On grounds of mutual unintelligibility, Penange Dogon merits a separate entry from other Dogon languages (Heath 2014d), and is missing from E16/E17/E18.

On grounds of mutual unintelligibility, Nanga Dogon merits a separate entry from other Dogon languages (Bertho 1953, Heath 2014c), and is missing from E16.

On grounds of mutual unintelligibility, Bankan Tey Dogon merits a separate entry from other Dogon languages (Blenc h 2005b, Heath 2014a), and is missing from E16.

On grounds of mutual unintelligibility, Bankan Tey Dogon merits a separate entry from other Dogon languages (Heath 2014b), and is missing from E16.

On grounds of mutual unintelligibility, Tirange Dogon merits a separate entry from other Dogon languages (Bertho 1953, Heath 2014e), and is this missing from E16.

On grounds of mutual unintelligibility, Ampari Dogon merits a separate entry from other Dogon languages (Blench & Douyon 2005, Hochstetler et al. 2004), and is missing from E16.

On grounds of mutual unintelligibility, Mombo Dogon merits a separate entry from other Dogon languages (Bertho 1953, Prokhoro v 2012), and is missing from E16.

Yoti, a language of the Yendang group, requires a separate entry (Blench no date). It is listed in E16 as a dialect of Yendang [yen] because most Yoti speak Yendang as well, but Yoti is not close to being a dialect of Yendang, and is, in fact closer to Bali [bcn] and Kpasham [pbn] (Blench no date).

Lebu Wolof is listed as a dialect of (Senegalese-Mauretanian) Wolof [wol]. However, Lebu Wolof is not intelligible to other Wolof speakers, e.g., Wolof speakers cannot readily understand the text specimens of Lebu in Angrand (1952), as opposed to Gambian Wolof [wof] which is intelligible to (Senegalese-Mauretanian) Wolof [wol] but uses a different written standard and source of loanwords. The distinctness of Lebu Wolof is hidden by the fact that all Lebu Wolof speakers are also bilingual in non-Lebu Wolof.
**E16** Damakawa, a now moribund or extinct language, was discovered by McGill (2008). It is missing from E16.

**E16/E17/E18** Yambe, a Congolese language of the Ndzem-Bonwali subgroup, is missing (Klieman 1997).

**E16/E17/E18** Goo, a Mande language of Côte d'Ivoire, is missing. It is closely related to Toura [neb], but the Toura do not understand Goo, and the Goo also speak Toura [neb], creating the impression that the two are intelligible, when they are not (Vydrin 2013a,b). It is missing from E16/E17/E18.

**E16/E17/E18** Siti, a Western Gurunsi language, is missing, though listed as a dialect name under Vagla [vag]. But Siti is no closer to Vagla than to, e.g., Phuie [pug]. Kleinewillinghöfer (1999) explains how the misunderstanding underlying the combining of Vagla-Siti came about.

**E16/E17/E18** Kpeego, a Mande language of the Samogo subgroup spoken by Tyurama blacksmiths, is missing (Zwernemann 1996).

**E16/E17** The Gabonese Bantu language Samay (Mokrani 2005) is missing. It is as distinct from all other Gabonese languages as any other Gabonese language included in E16/E17 (Alewijnse et al. 2007, Idiata 2007).

**E16/E17** The Gabonese Bantu language Ndambomo (Mvé 2013, Ovono Bikoro 2001) is missing. It is as distinct from all other Gabonese languages as any other Gabonese language included in E16/E17 (Alewijnse et al. 2007, Idiata 2007).

**E16/E17/E18** The Gabonese Bantu language Mwesa (Mouloungui 1999) is missing. It is as distinct from all other Gabonese languages as any other Gabonese language included in E16/E17/E18 (Alewijnse et al. 2007, Idiata 2007).

**E16/E17/E18** The Gabonese Bantu language Tombidi (Bingoumou 2005, 2008) is missing. It is as distinct from all other Gabonese languages as any other Gabonese language included in E16/E17/E18 (Alewijnse et al. 2007, Idiata 2007).

**E16/E17/E18** The Gabonese language Vili of the Ngounié (Raponda Walker 1960, Adam 1954:55) is missing. It is as distinct from all other Gabonese languages as any other Gabonese language included in E16/E17/E18 (Alewijnse et al. 2007, Idiata 2007). Vili of the Ngounié which belongs in the Njebi or B.50 group is not to be confused with the more famous Vili [vif] language, also present in Gabon, but in the H12-16c group.

**E16/E17/E18** Nzadi, a Bantu B80 language of the Democratic Republic of Congo, is missing. It is not intelligible to any of the Bantu B80 languages in E16 (nor to Lwel) (Boone 1973, Crane et al. 2011, Maes 1934).

**E16/E17** Palen, a third language of the Wara-Natioro subgroup of Gur (Kleinewillinghöfer 2007), is missing. Faniagara, the Jula name for the Pala village where Palen is spoken, is listed under the Wara [wbf] entry but Palen is not intelligible to any of the other
Wara dialects and may even subgroup with Natioro rather than Wara (Winkelmann 2007).

**E16/E17/E18** Jan-Awei, a scarce language of the Jukun subgroup of Central Jukunoid (Storch 2009), is missing.

**E16/E17/E18** The Ngala language of Lake Chad, as documented in a wordlist by Barth (Benton 1912, von Duisburg 1914), is missing. It is clearly of the Kotoko group of Chadic, closest to Makari (= Mpade [mpil]) and Goulfey (= Malgbé [mxf]) but is as different from both as Makari and Goulfey are from each other (Migeod 1922, Sölken 1958).

**E16/E17** The moribund Bantu language Eviya of Gabon is missing (van der Veen 1999). It possibly occurs (“Evii”) as a dialect of the Bubi [buw], but it should be separate (Alewijnse et al. 2007).

**E16/E17/E18** The enigmatic Kwadza language of Tanzania, is missing (Claus 1910, Kohl-Larsen 1943). It was still alive after the 1951 (Ehret 1980).

**E16/E17/E18** The now extinct Nubian language of Jebel Haraza in Sudan is missing (Bell 1975, Newbold 1924). It is not known exactly when the last fluent speaker died, but in the 1970s there were only rememberers (Mohamed 1974).

**E16/E17** Cuba, a Kainji language discovered in 2012 should be added (McGill 2012).

**E16/E17/E18** The Ngala language (a Zande-group language originally in the Central African Republic) discovered by Santandrea around 1950, is missing (see Santandrea 1952, data appears in Santandrea 1965). It is, in fact, missing from all listings of African, Sudanese, or CAR languages and with only some 50 refugee speakers half a century ago, it is presumably extinct now.

**E16/E17** The Tarjumo language (a Kanuri-group language in Nigeria not intelligible with any Kanuri variety listed in E16/E17) is missing. It is used as a second language mainly for purposes of interpreting the Qur’ān and was brought to the attention of Western scholars by Bondarev (2005).

**E16/E17/E18** A survey of Fadan Kanshe-Wamba-Akwanga triangle in Kaduna State, Nigeria revealed the existence of four extinct or near-extinct unknown languages Akpoudou, Babur, Nisam and Nigbo (Blench 2005a), none of which are in E16. A little data exists to confirm that Akpoudou was indeed a different language from all others (Alumu-Tesu [aab] being the closest relative), but for the others there is so far only sociolinguistic information which not sufficient to totally exclude that Babur, Nisam and Nigbo were unintelligible to themselves and all other languages. I wish to thank Roger Blench for drawing my attention to Akpoudou.

---

3I wish to thank Raymond Boyd for discussion on the classification of this language.
E16/E17/E18 The Hoan language of Botswana was discovered by Anthony Traill in the early 1970s (Traill 1973, 1974) and has recently been shown (Heine & Honken 2010) to be related to the Ju-dialect cluster (also known as North Khoisan). E16/E17/E18 has a Hua [huc] entry which has information (e.g., speaker number, grammatical characteristics) that matches the Hoan language as well as information (e.g., the classification, the name) that matches (Naumann 2014) a far eastern variety of the Taa dialect chain which is not intelligible with the westernmost end of that chain (represented in E16/E17/E18 as !Xóó [nmn]). The present-day location of both languages is the same (Gerlach & Berthold 2011) which might have lead to the confusion. Two entries are needed, one for the eastern dialect of Taa and one for the (entirely different) Hoan language, but E16/E17/E18 has a merged entry, so one is missing (and the merged entry should be corrected accordingly).

E16/E17/E18 The Mimi language of Wadai, as documented by Nachtigal (Lukas & Völckers 1938) is missing. The language has not been sighted for over a century and is therefore likely extinct. There is some evidence (body part terms) that Mimi-Nachtigal is related to the Maban group of languages (Starostin 2011).

E16/E17/E18 The Mimi language as documented by Gaston Decorse in Gaufroy-Demombynes (1907) is missing. I have never been able to find out where Decorse collected the wordlist and where his informant came from but presumably it was collected somewhere in SE Chad or NE Central African Republic on Decorse’s trip in 1902-1904 (Decorse 1906) and the wordlist shows Arabic and Maban loans (Starostin 2011), so it was probably also spoken in Wadai. The language has not been sighted for over a century and is therefore likely extinct.

E16/E17/E18 The language of the villages Kazibati and Mongoba (likely belonging to the Ngbandi group, see van Bulck 1948:178, Hackett & van Bulck 1956:74, McMaster 1988:123, 259), is missing (Costermans 1938).

E16/E17/E18 Tetserret, a recently discovered Western Berber (i.e., subgrouping with Zenaga [zen]) language, is missing (Attayoub 2001, Lux 2011).

E16/E17/E18 Ham, an endangered language of the North Masa-Musey group (Melis 2006:45), is missing (Ajello et al. 2001). Ham is not intelligible to Masa [mcn] though many Ham speak Masa (Melis 2006:44).


E16/E17/E18 Julud is mutually intelligible with Katla-Kulharong but definitely not with Katla-Cakom, so it should be listed as a separate language from Katla [kc] (p.c. Birgit Hellwig 2007, N/A 1979a,b)

E16/E17/E18 Lwel or Kelwer, a Bantu B862 language of the Democratic Republic of Congo is missing (Khang Levy 1979). It is not mutually intelligible with any of its neighbours (Koni Muluwa & Bostoen 2011).
**E16** Dahālik, an outlying Tigre language, is missing from E16 (Simeone-Senelle 2006).

**E16/E17/E18** Mangar, a Chadic language of the Ron group, is clearly different enough from Daffo-Butura (and other Ron languages) to be a separate language (Blench 2001) and is duly listed as a separate language in informed reference works (Jungraithmayr 1981).

**E16/E17/E18** Tep, previously thought to be a dialect of Nigerian Mambila [mzk], is missing though Tepo is mentioned as a potential center of a dialect (Blench 1993, Connell 2000).

**E16/E17/E18** Shang, a Hyamic language of Nigeria, is missing (Blench 2010b).

**E16/E17/E18** Kurmin Dangana, a Hyamic language, is missing (Blench 2012a). The data suggests it is sufficiently different from Hyam [jab], but reliable intelligibility judgments are difficult to get since Kurmin Dangana also speak the prestige Hyamic variety Hyam of Nok [jab]. I wish to thank Roger Blench for alerting me to this language.

**E16/E17/E18** Gwara, a close relative of Idun [ldb] of the NW Plateau of Nigeria, is missing (Blench 2009a).

**E16/E17/E18** Dyarim, a Chadic language of the West Chadic Zaar group, is missing (Blench 2007).

**E16/E17/E18** The Bapen⁴ language is (or was) not intelligible to its closest relatives Bas-sari [bsc] and Bedik [tnr] (Santos & Ferry 1975:82, Ferry 1972:353-354). It was nearly extinct in 1971.

**E16/E17/E18** The Bafotó language of Pygmoids among the Mongo (not intelligible with Mongo) is missing (Hulstaert 1960, 1978) and is highly endangered if not extinct (Motinjea Mangulu 2001-2002).

**E16/E17/E18** The now probably extinct (Cobbinah 2013:32) Northeast Bainounk varieties of northeast Casamance (once spoken in the villages Kamanka, Kansambu, Diouroul, Velingara and Samakoung) are missing in E16/E17/E18 either as part of other entries or as an entry of their own. Since they were only about 60% lexicostatistically cognate (Doneux 1991:87) to the other variety of Bainounk, they must be inferred not to have been intelligible to other Bainounk speakers⁵.

**E16/E17/E18** The Bainounk varieties of Gubēeher, Gufangor and Gubelor south of the Bainounk-Gunyamoolo [bcz] in Senegal are missing either as a separate entry or as part of the description of the other Bainounk entries. They are not intelligible to Bainouk-Gunyamoolo [bcz] and only 60-68% lexicostatistically cognate (Doneux 1991:87) with any other Bainounk and thus merit a separate entry.

---

⁴I wish to thank Guillaume Segerer for bringing this language to my attention.

⁵I wish to thank Guillaume Segerer for bringing this case my to attention.
Syan (glossonym OruSyan, Huntingford 1965) is a missing language of the North Nyanza subgroup of Bantu, which lexicostatistically too divergent to be intelligible any other language in E16/E17/E18 (Schoenbrun 1994).

A.3 Missing Languages: Pacific

Amam is a Goliolan language of Papua New Guinea. Its territory falls under the Weri [wer] entry in E16/E17/E18 but it is impossible that they are the same/intelligible languages given the data of Aki & Pennington (2014). Thus, Amam should ideally have its own separate entry as a language related to Weri [wer].

Butam is a language related to Taulil [tuh] once spoken in New Britain (Capell 1967, Filer 1994, Laufer 1959, Stebbins 2009). It has been pronounced extinct but this remains to be verified on the ground. It is missing from E16/E17/E18.

Dempwolff (1905) recorded a language called Kaniet which is so different from Kaniet [ktk] that there must have been two unintelligible Kaniet languages (Blust 1996).

Kawi, or Old Javanese, is missing (Kern 1901, 1901, 1903, von Humboldt 1836, 1838, 1839, Zoetmulder 1982, 1983).

Lemerig, a separate language from Vera’a [vra] of the Torres-Banks linkage, is missing (Codrington 1885, François 2007, Tryon 1976).

Unubahe, a Suauic language of Papua New Guinea was discovered on survey (Rueck no date). It is missing from E16.

O’oku, a presumed extinct language, Mailuan or Yareban affiliation, is missing (Ray 1938, Strong 1911).

The language of Barika, a Turama-Omatian language, is missing from E16 (Johnston 1923).

Tambora, an extinct non-Austronesian language of Indonesia, is missing (Balbi 1826, Donohue 2007a, Steinhauer 1986).

Nasvang, a separate language in Vanuatu of the Eastern Malakula linkage in the Austronesian family is missing (Lynch & Crowley 2001:75). Data appears in Charpentier (1982).

Navwien, a separate language in Vanuatu of the Eastern Malakula linkage in the Austronesian family is missing (Lynch & Crowley 2001:75). Data appears in Charpentier (1982).

Nisvai, a separate language in Vanuatu of the Eastern Malakula linkage in the Austronesian family is missing (Lynch & Crowley 2001:75). Data appears in Charpentier (1982).
The language of Marau Wawa island off the south-west coast of the Bauro district of San Cristoval was very different from, and unintelligible with, the neighbouring Bauro [bxa] language (Fox 1919:130-131). It is missing from E16/E17/E18. It is now almost certainly extinct (cf. Lanyon-Orgill 1947).

Mur Pano, an Austronesian language of the Korap subgroup was discovered on survey. It is closely related to, but not intelligible with Karnai [bbv] (Carter et al. 2012).

Molet, an Trans New Guinea language of the Warup subgroup was discovered on survey. It is closely related to, but not intelligible with Asaro'o [mtv] (Carter et al. 2012).

E16/E17/E18 has the entry Nyelâyu [yly] merging both Balade-Arama-Tiari and Belep, but they are not intelligible to monolinguals (McCracken 2013:30) and should thus be separate entries (like they are in previous surveys, such as Leenhardt 1946).


The language spoken by the Abai Tubu and Abai Sembuak is different from all other Murutic languages, and is missing from E16/E17/E18 (Lobel 2013). I wish to thank Jason Lobel for clarifying the status of Abai Tubu-Sembuak to me.

The moribund or extinct language Orkon is lexicostatistically more different from Port Vato [ptv] and Dakaka [bpa] than they are from each other (Lynch & Crowley 2001:96). It is missing from E16/E17/E18.

The Sepik Hill language Nigilu is missing. It is not intelligible to the varieties subsumed under Bahinemo [bjh] or Berinomo [bit] (Dye & Dye 2012:38), nor to Wagu.

The Alor-Pantar language Kaera is missing (Holton et al. 2012, Klammer 2014).

Tau’t Batu, a Southern Palawanic language, is missing (Thiessen 1980).

Malaweg, an Ibanagic language (Reid 2006, no date(a),n) is missing as a separate language. It is listed as a dialect of Itawit [itv] which is not intelligible to Malaweg and belongs to Gaddangic, a different subbranch of Ibanagic (Reid 2006).

Nati, a Southwest Coastal Malekula language of the North and Central Vanuatu subgroup, is missing. It is/was not intelligible to its closest relative Navahaq (Crowley 1998:105, Dimock 2009:5).
Avok, a Eastern Malakula linkage language of the offshore island of Avok, is missing (Charpentier 1982, Tryon 1976). It is possibly missing since Tryon (1976) treat it as a dialect of Axamb [ahb]. But lexicostatistical similarity between varieties of the two are never above 75% and go down as low as 60%, leaving little room for intelligibility between the endpoints (Lynch & Crowley 2001:75).

Batanta, a possibly extinct Austronesian language of the island with the same name (northwest of the Bird's Head of Papua), is missing. Batanta was a separate language according to un-intelligibility (to the varieties subsumed under Ma'ya [slz]) reports (Remijsen 2002:42, van der Leeden 1987:83) and a little data appears in Cowan (1953).

Abu’ Arapesh, is not sufficiently intelligible with Bukiyip or other Arapeshan languages to be a dialect (Nekitel 1985). It is missing either as a dialect or a separate language in E16.

Nese-Naha, an Oceanic language of Northwest Malekula is missing (Crowley 2006b:3).

Mansim/Borai, a language related to Hatam on the northeastern Bird’s Head of Indonesian Papua, is presented as a different language (rather than dialect) from Hatam in Reesink (2002:304-305) and comparisons of old wordlists (e.g. von der Gabelentz & Meyer 1882) readily confirm this difference. Mansim has been assumed to be extinct (Reesink 2002), but persistent rumours in the Manokwari area suggest that there are still as many as 50 senior speakers (own field work, 2010). The existence of remaining speakers would also be consistent with the last known published population survey of half a century ago (Pans 1960).

Nitita, a highly endangered language of Malakula (Vanuatu), is missing (Crowley 2006a).

Vivta, a highly endangered language of Malakula (Vanuatu), is missing (Crowley 2006a).

Womo and Sumararu constitute a separate Skou language, not intelligible with any other language in E16 (Donohue & Crowther 2005, Donohue & San Roque 2004)⁶.

The Nori language as documented by Wilkes (1926) is another (presumably now extinct) missing Skou language which was not intelligible with any other Skou language, in particular not with Warapu [wra] (Baron 1983, Donohue & Crowther 2005).

---

⁶I wish to thank Mark Donohue (p.c. 2007) for clarifying the status as a separate language, rather than a dialect of some language already in the E16/E17/E18.
A.4 Missing Languages: North America

A.4.1 Adai, an extinct North American presumably isolate language, is missing (Grant 1995).

A.4.2 Alsea, an extinct North American language isolate, is missing (Frachtenberg 1918, 1920, Golla 1997).

A.4.3 Apalachee, an extinct Muskogean language, is missing (Kimball 1987).

A.4.4 Beothuk, an extinct North American isolate, is missing (Hewson 1978).


A.4.7 Biloxi, a Southeastern Siouan language, is missing (Dorsey & Swanton 1912a, Einaudi 1976, Haas 1968).

A.4.8 Ofo, a Southeastern Siouan language, is missing (Dorsey & Swanton 1912b, Swanton 1909).

A.4.9 Tutelo, a Ohio Valley Siouan language, is missing (Frachtenberg 1913, Hale 1883, Oliverio 1997).


A.4.11 Chimakum, a Chimakuan language, is missing (Boas 1892).


A.4.13 Comecrudo, an extinct poorly attested presumed language isolate, is missing (Berlandier & Chowell 1828-1829, Swanton 1915, 1940a, Uhde 1861).

A.4.14 Cotoname, an extinct poorly attested presumed language isolate, is missing (Berlandier & Chowell 1828-1829, Swanton 1915, 1940b).

A.4.15 Coahuitlaco, an extinct language isolate, is missing (Campbell 1996, de Vergara 1965 [1732], Troike 1996).

A.4.16 Cuitlatec, an extinct Mexican language isolate, is missing (Escalante 1962, McQuown 1941, R. 1939, Weitlaner 1939).

A.4.17 Waikuri, an extinct language isolate of Baja California, is missing (Baegert 1772, Golla 2011, Zamponi 2004).
Karankawa, an extinct language isolate of the coast of Texas, is missing (Berlandier & Chowell 1828-1829, Grant 1994).

E16/E17/E18 has a Yokuts [yok] entry which subsumes “many subvarieties”. However, careful analysis of all extant Yokutsan materials necessitate the division into several related languages, minimally adding three language-level entries (all extinct) to the one found in E16/E17/E18, i.e., Palewyami Yokuts, Buena Vista Yokuts and Tule-Kaweah Yokuts.

Chalon, an extinct arguably distinct Northern Costanoan language, is missing (Golla 2011, Hale 1846, Heizer 1952 [1878], Shaul 2012).

Changuena, an extinct Isthmic Chibchan language, is missing (Brinton 1891, Constenla Umaña 1985, Pinart 1890).

Dorasque, an extinct Isthmic Chibchan language, is missing (Constenla Umaña 1985, Gómez Moreno 1997, Miranda de Cabal 1974, Pinart 1882, 1890).

Huetar, an extinct Votic Chibchan language, is missing (Constenla-Umaña 1984, Quesada Pacheco 1990, 1992).

Mahican, an extinct Delawaran Algonquian language is missing from E17/E18 (Goddard 2009).

Karkin, an extinct Costanoan language is missing from E17/E18 (Callaghan 1988, Golla 2011, Okrand 1989).

Miluk, an extinct Coosan language, is missing (Doty 2012, Kroeber 2013, Pierce 1965).

Island Takic, an extinct arguably distinct Cupan Uto-Aztecan language, is missing (Golla 2011, Munro 2002).

Konomihu, an extinct poorly attested somewhat mysterious Shastan language, is missing (Dixon 1905, Golla 2011, Larsson 1987, Silver 1980).

Laurentian, an extinct northern Northern Iroquoian language, is missing (Barbeau 1961, Biggar 1924 [1534-1541], Cartier 1999, Mithun 1982).

Loup A or Nipmuck, an extinct Eastern Algonquian language, is missing (Day 1975 [1677], Gustafson 2000).

The extinct language consisting of Mohegan, Montauk and Narragansett (an Eastern Algonquian language), is missing (Aubin 1972, Cowan 1973, Edwards 1788, Granberry 2003a,b).

New River Shasta, presumably included in the Shasta [sh] entry in E16/E17/E18, was arguably a separate language (Golla 2011), thus meriting a separate entry.
Omomil, a poorly attested extinct Northern Uto-Aztecan language, is missing (Golla 2011).

Patwin, a Wintu language included in the Wintu [wit] entry in E16, is not intelligible to Wintu [wit] and thus merits a separate entry (Golla 2011, Radin no date).

Pentlatch, an extinct Central Salish language, is missing (Boas 1890, Kennedy & Bouchard 1990, Tylor et al. 1891).


Pochutec [xpo], an extinct Aztec language, is missing (Boas 1917).

Susquehannock [sqn], an extinct Northern Iroquoian, is missing (Holm 1996 [1702], Julian 2010).

Takelma [tkm], a well-attested extinct language isolate, is missing (Golla 2011, Kendall 1977, Sapir 1990).

Tongva, an extinct arguably distinct Californian Uto-Aztecan language related to Serrano, is missing (Golla 2011, Hale 1846, Kroeber 1907, Munro 2000).

Tsutsault, an extinct Northern Athapaskan language, is missing (Boas & Goddard 1924, Gallatin 1836, Golla 2011, Hale 1848).

Tualatin-Yamhill, an extinct arguably distinct Kalapuyan language, is missing (Berman 1990, Gatschet et al. 1945, Jacobs 1945, Scouler 1841, Shipley 1970).

Wampano, an extinct Eastern Algonquian, is missing (Pierson 1658, Rudes 1997).

Jicaque or Tol [jic] of Honduras had a relative in Nicaragua which is now extinct but was so different that the two could not have been intelligible (Lehmann 1920a). This Western Jicaque language is thus missing.


Yana, an extinct Californian language isolate is missing (Golla 2011, Sapir & Swadesh 1960).

Timucua, an extinct language isolate of Florida, is missing (Adam & Vinson 1886 [1614], Goddard et al. 2004).

Woccon, an extinct Catawban Siouan language, is missing (Carter 1980, Rudes 2000).

Joya, an extinct but attested Southern Uto-Aztecan language, different from all other Uto-Aztecan languages, is missing (Shaul 2001:200-205).
Xinca [xin] is listed as an unclassified language. In reality, there are at least four, perhaps five, different attested languages which may be divided as Xinca of Yupiltepeque, Xinca of Guazacapán, Xinca of Sinacantan, Xinca of Jumaytepeque and Xinca of Chiquimulilla (Sachse 2010:49-53).

Lenca [len] is listed as an unclassified language. In reality, there were (or are, given recent information about speakers remaining) two different languages, which may be called Honduran Lenca and Salvadorean Lenca (del Río Urrutia 2004, Herranz 1992).

Maratino, an extinct poorly attested language of Mexico, is missing (Campbell 1997, Swanton 1940c).

Tapachultec, an extinct Mixe language, is missing (González Casanova 1927, Lehmann 1920a, Sapper 1912).

Ulterior Mixe, a newly discovered Mixe language, is missing (Wichmann 2008).

Upper Umpqua, a poorly attested Athapaskan language that may have survived into the 1950s, is missing (Golla 2011:40-42).

Jitotolteco, a newly discovered Zoque language, is missing (Zavala Maldonado 2011).

Cusabo, an extinct nearly unattested, but nevertheless convincingly distinct language (Waddell 2004:254) of Southeast USA, is missing.

A.5 Missing Languages: South America


Careful analysis of early vocabularies of languages related to Kaweskar [alc] necessitate distinguishing a now extinct Central Alakaluf language (Viegas Barros 1990, 2005) attested by vocabularies in Borgatello (1928), Marcel (1892), Skottsberg (1913).

Allentiac, one of the two extinct Huarpean languages, is missing (de la Grasserie 1901, de Valdivia 1894 [1607]).

Millcayac, one of the two extinct Huarpean languages, is missing (Márquez Miranda 1943, Tornello et al. 2011).

Araicu, an extinct Northern Arawakan language attested with a vocabulary in von Martius (1867b), is missing (Ramirez 2001).
E16/E17/E18 Uaimuma-Mariaate, an extinct Inland Northern Maipuran language attested with vocabularies in von Martius (1867b) and Wallace (1853), is missing (Ramírez 2001).

E16/E17/E18 The extinct Arawán language Aruá at the Jurua river attested with a vocabulary in Chandless (1869), is missing.

E16/E17/E18 Chaná, an Charrúan language which may have survived longer than previously assumed (Barros 2006, Barros & Jaime 2013, Viegas Barros 2008, 2009), is missing (Larrañaga 1923).


E16/E17/E18 Atacame or Esmeraldeño, an extinct language of unknown affiliation spoken, when attested, by black people in Ecuador, is missing (Franch 1996, Jijón y Caamaño 1919, Lehmann 1920b, Seler 1902, Wolf 1892).

E16/E17/E18 Atshahuaca, an extinct Panoan languages of the Madre de Dios area, is missing (Fleck 2013, Nordenskiöld 1905).

E16/E17/E18 Bonari, an extinct Cariban language most closely related to Waimiri-Atroari [atr], is missing (de Sousa 1875, Girard 1971).

E16/E17/E18 Cañari, an extinct language of Ecuador whose existence can be verified mainly through toponyms, is missing (Howard 2010, Jijón y Caamaño 1919, Paz y Miño 1961).

E16/E17/E18 Capoxo, an extinct Western Maxakali language attested with a wordlist in von Martius (1867b), is missing.

E16/E17/E18 Cariay, an extinct Northern Arawakan (Ramírez 2001) language attested with a wordlist in von Martius (1867b), is missing.

E16/E17/E18 Chané, the original Arawakan language of the Bolivian Chané (Rocca & Rossi 2004), attested with only a short wordlist in Nordenskiöld (1910, no date), is missing.

E16/E17/E18 Payagua, a poorly attested extinct language isolate of Paraguay, is missing (Boggiani 1900, Lafone Quevedo 1910, Mansfield 1856, Parodi 1880, Peña 1898, Schmidt 1949).


E16/E17/E18 Guamo, an extinct language isolate attested only with the wordlist in Anónimo (1928 [1778]), is missing.
Mure, an extinct Bolivian language of uncertain affiliation attested only with the Pater Noster in Teza (1868) and a few sporadic mentions (Hervás 1800:251, Bueno 1770), is missing. It is not to be confused with the more well-known Chapacuran Bolivian language called Moré or Itene [ite] (Angenot-de Lima 2002).


Chapakura, an extinct Chapacuran language of Bolivia, is missing (Chamberlain 1912, d’Orbigny 1839, Montaño Aragón 1987).

Charrúa, an extinct Charruan language of Uruguay, is missing (d’Orbigny 1839, Ramos Boerr [ca. 2007], Rona 1964).

Chiriva, an extinct probably Panoan language of Bolivia attested only with the minuscule wordlist in Palau & Saiz (1989 [1794]), is missing. I wish to thank David Fleck and Willem Adelaar for discussion on this language.

Chono, an extinct isolate language of Southern Chile, is missing (Bausani 1975, Viegas Barros 2005) unless it is the language (unsuccessfully, see below) intended by the Kakauhua [kbf]-entry.


Manao, an extinct Northern Arawakan language, is missing (de Goeje 1948, de Lourdes Joyce 1951, Métraux 1940, Ramirez 2001, von Martius 1867b).

Urina, an extinct Northern Arawakan language (Ramirez 2001) attested only in the wordlists of Kann (1989), von Martius (1867b), is missing.

Coeruna, an extinct Huitotoan language attested with a vocabulary in von Martius (1867b), is missing.

Andoquero, an extinct Huitotoan language attested with a vocabulary in von Martius (1867b), is missing.

Gauchi, a poorly attested of Southern Brazil of uncertain affiliation (Viegas Barros 2004), is missing (de Castelnau 1851, Koch-Grünberg 1902, von Martius 1867b).

Maipure, a Northern Maipuran Arawakan language, is missing (Gilij 1780, 1781, 1782, 1784, Hervás y Panduro 1971 [1799](a),, Zamponi 2003b).

Otomaco, an extinct poorly attested presumed language isolate once spoken along the Orinoco river, is missing (de Luzena no date, Rosenblat 1936, Veigl 1785).

Jumana, an extinct Northern Arawakan (Ramirez 2001) language attested with a wordlist in von Martius (1867b), is missing.
E16/E17/E18 Kaishana, a presumed extinct Northern Arawakan (Ramirez 2001) language, is missing (Hanke 1960, Nimuendajú 1955b, von Martius 1867b).

E16/E17/E18 Passe, an extinct Northern Arawakan (Ramirez 2001) language attested with a wordlist in von Martius (1867b), is missing.


E16/E17/E18 Cotoxó, an extinct Brazilian Kamakanan language, is missing (Loukotka 1932, Silva Martins 2007, von Martius 1867b).

E16/E17/E18 Masacará, an extinct Brazilian Kamakanan language, is missing (Silva Martins 2007, von Martius 1867b).

E16/E17/E18 Menien, an extinct Brazilian Kamakanan language, is missing (Métraux 1930, Silva Martins 2007, von Martius 1867b).

E16/E17/E18 Duit, an extinct language closely related to Chibcha [chb], is missing (Lehmann 1920b, Uricoechea 1871a).

E16/E17/E18 Güenoa, an extinct Charrúan language, is missing (Hervas 1787, Lafone Quevedo 1897, Vignati 1940).

E16/E17/E18 Geicó, an extinct Jéan language (Loukotka 1955b) attested with a wordlist in von Martius (1867b), is missing.

E16/E17/E18 Guerén, an extinct close relative of of Krenak [kqq], is missing (Ignace 1909, Loukotka 1955a).

E16/E17/E18 Guináu, an extinct Inland Northern Maipuran Arawakan language (Ramirez 2001), is missing (Koch-Grünberg 1928a, Schomburgk 1848, von Martius 1867b).

E16/E17/E18 Haush, an extinct insular Chonan language, is missing (Coiazzi 1914, 1914, Furlong 1917, Viegas Barros 2005).

E16/E17/E18 Teushen, an extinct Continental Chonan language, is missing (Viegas Barros 2005, von Martius 1867b).

E16/E17/E18 Ingain, a poorly attested Southern Je language, is missing (Jolkesky 2010, Vogt 1904).

E16/E17/E18 Kimda, a poorly attested Southern Je language, is missing (Lista 1883).

E16/E17/E18 A number of extinct lects subsumed under the name Jirajaran constitute an isolated language which is missing from E16/E17/E18 (Jahn 1927, Oramas 1916, Querales 2008).

E16/E17/E18 Yurí, an extinct language related to Ticuna (Orphão de Carvalho 2009), is missing (Goulard & Montes 2013, Ortiz 1965, von Martius 1867a, Wallace 1853).
Kitemoka, an extinct Chapacuran language of Bolivia, is missing (de Créqui-Montfort & Rivet 1913c).


Kustenau, a Central Maipuran Arawakan language, documented in a wordlist by von den Steinen (1894), is missing.

Lule, an extinct but well-attested language of present-day Argentina, is missing (Maccioni 2008 [1732], Viegas Barros 2001a,b).

Magiana, an extinct Bolivia-Parana Arawakan language of Bolivia attested only with the minuscule wordlist in Palau & Saiz (1989 [1794]), is missing. I wish to thank Swintha Danielsen for discussion on this language.

Macuni, an extinct Western Maxakali language attested with a wordlist in von Martius (1867b), is missing.

Maragua, an extinct Northern Maipuran Arawakan language (Ramirez 2001, Tastevin 1919, Verneau 1921) attested with a wordlist in von Martius (1867b), is missing.

Marawan, an extinct poorly attested extinct Eastern Maipuran Arawakan language (the closest relative of Palikūr [plu]), is missing (Rivet & Reinburg 1921).

Antioquian, an extinct poorly attested Core Chibchan language, is missing (Duque & Espinosa 1994, Rivet 1946).

Apiaka of the Tocantins river, an extinct Cariban language of the Xingú subgroup attested with wordlists in Ehrenreich (1895), Krause (1936), is missing. It is not to be confused with the more well-known Apiaká [api] language of the Tapajos river in Tupi-Guarani Subgroup VI.

Matanawí, an extinct isolated language of the Madeira river, attested with a wordlist in Nimuendajú (1925), is missing.

Morike, an extinct poorly attested extinct Western Maipuran Arawakan language (the closest relative of Chamicuro [ccc]), is missing (Fleck 2007b, Jakway 1975, Tessmann 1930).

Napecañ, an extinct Chapacuran language of Bolivia, is missing (Cardús 1886, de Créqui-Montfort & Rivet 1913c, Pauly 1928, Riester 1965).

Opón-Carare, an extinct far western Cariban language, is missing (Durbin & Seijas 1973a, Pardo et al. 1998, von Lengerke 1878).

Palmela, an extinct displaced Cariban once spoken in Rondônia, is missing (Becker-Donner 1956, da Fonseca 1881, Nordenskiöld 1915).
Ramanos, an extinct seemingly isolated Bolivian language of Bolivia attested only with the minuscule wordlist in Palau & Saiz (1989 [1794]), is missing.

Rocorona, an extinct Chapacuran language of Bolivia attested only with the Pater Noster in Teza (1868), is missing (Birchall 2013).


Coroad-Puri, an extinct language of Brazil related to Puri [prr], is missing (da Neto 2007, Loukotka 1937, von Martius 1867b).


Pimenteira, an extinct Kuikuroan Cariban language, is missing (Loukotka 1955b, von Martius 1867b).

Urupa, an extinct Chapacuran language, is missing (Nimuendajú 1925, Nimuendajú & do Valle Bentes 1923).

Wanyam, an extinct Chapacuran language, is missing (Hanke 1975, Loukotka 1963, Nordenskiöld 1915, Ribeiro 1998).

Wanyam, an extinct Chapacuran language, is missing (Loukotka 1963).

Yao, an extinct Guianan Cariban language, is missing (de Laet 1633, Girard 1971).

Wayumara, an extinct Guianan Cariban language most closely related to Maquiritare [mch], is missing (Farabee 1924, Girard 1971, Koch-Grünberg 1928b, Schomburgk 1849, von Martius 1867b).

Paravilhana, an extinct Venezuelan Cariban language, is missing (Adam 1893, d’Almada 1861 [1787], von Martius 1867b).

Sapara, a presumed extinct Venezuelan Cariban language, is missing (d’Almada 1861 [1787], Girard 1971, Koch-Grünberg 1928b).

Pauxiana, a presumed extinct Venezuelan Cariban language, is missing (d’Almada 1861 [1787], Meyer 1956, Salathé 1932).
Demushbo, a recently extinct separate Panoan language of the Matses subgroup, is missing (Fleck 2010, 2013).

Arazaire, an extinct Panoan language of the Madre de Dios area, is missing (Fleck 2013, Llosa 1906).

Cashinawà of the Tarauacá river, an extinct Panoan language in the Pano Nawa subgroup, is missing (Fleck 2013, Rivet & Tastevin 1927, 1929). It is not be confused with the more well-known Kashinawà of the Ibaçu River which is covered in the Cashinahua [cbs] entry.

Careful analysis of early vocabularies of languages surfacing under the name Remo (Fleck 2013) necessitate distinguishing another two extinct languages Remo of the Blanco river (Philips Leuque 2002 [1927]) and Remo of the Jaquirana river (Braulino de Carvalho 1931, de Oliveira Carvalho et al. 2006), in addition to the extant Remo (of the Moa river) [rem] entry (Loos 1973-1974). The three are closely related but fall into the Pano Nawa subfamily with different languages between them (Fleck 2013).

Careful analysis of early vocabularies of languages surfacing under the name Mayoruna (Fleck 2013) necessitate distinguishing another three extinct languages Mayoruna of the Amazon river (de Castelnau 1851, von Martius 1867b), Mayoruna of the Jandiatuba river (de Alviano 1957), Mayoruna of the Tabatinga river (von Martius 1867b), in addition to the extant Matsés-Mayoruna [mcf] entry. All of them belong to one subgroup of Pano (called Mayoruna) (Fleck 2013).

Yurumanguí, a poorly attested extinct language isolate of Colombia, is missing (Arcila Robledo 1951, Elias Ortiz 1945, Rivet 1942).

The extinct Chibchan language Chibcha, is missing (Anónimo no date, Uricoechea 1848, 1871b).

Aruan, an extinct Northern Maipuran language, is missing (Ferreira Penna 1879, Nimuendajú 1926).


Nomuya, a Colombian Huitotoan language not intelligible to any other Huitotoan language that survived long beyond 1950, is missing from E16 (Echeverri & Landaburu 1995, Landaburu et al. 2009).

Taruma, an isolated language of the Brazil/Guyana borderlands, is missing. While often presumed extinct, there are three Taruma left, living among the Wapishana (Carlin 2006:315). One of the three is fluent enough to produce coherent texts in Taruma while the other two have a much weaker knowledge of the language (p.c. Sergio Meira 2015).
Pisamira, a Tucanoan language of the Eastern Eastern Tucanoan II subgroup (Chacon 2014), is missing as a separate language (de Pérez 2000). Pisamira is listed under Tucano [tuo] which is of the Eastern Eastern Tucanoan I subgroup and not at all intelligible to Pisamira (Chacon 2014, de Pérez 2000).

Kankuamo, an Eastern Arhuacic language (different from, but most closely related to Damana [mbp]), is missing (Trillos Amaya 2005:70).

Kontanáwa (Panoan, thought to be extinct until 2002) is missing (de Aguiar 2007:40). However, no linguistic data has been published to contradict or verify that Kontanáwa is a different language (as opposed to merely a dialect of a different ethnic group, or a different ethnic group who no longer speak their ancestral language) from other Nawa-languages (de Aguiar 2007, Tastevin 1919). If not a distinct language, Kontanáwa should at least be included as a dialect of another language.

In 1995, a small group of isolated Indians were contacted by the department of isolated Indians (of the FUNAI) in Rondônia in the Brazilian Amazon. They turned out to speak Kanoé, an already known language. However, this contact lead to the contact of a second group of Indians who speak a thenherto unknown variety labeled Akuntsun by the Kanoé (Gabas 2005). This variety this very close to Mekéns (a Tupí language of the Tupari group) but arguably deserves an entry (Aragon 2008, Aragon & Cabral 2005).

E16/E17/E18 has two entries Karirí-Xocó [kzw] and Xukurú [xoo], with the following information:
The are (at least) three attested languages relevant for these two entries. The first, which we may call Kipeá-Dzubukuá, is a dialect complex consisting of Dzubukuá, Kipeá, Kamurú (Camuru), Sabujá (Pedra Branca). Other dialects or languages are even less well attested.

Region: Alagoas. Pernambuco, Serra de Urubá (Arobá) near Cimbres City, Bahía.

Classification: Unclassified Unclassified

Language development: Grammar. Grammar.

country: Brazil Brazil

Alternate names: Karirí, Kariri Xoxó, Kipeá, Xoxó-Kariri, Xoxó-Kariri, Kamurú, Kamu Kariri


Language use: Shifted to Portuguese. Shifted to Portuguese.
by 1951), and the names and locations of the Kipeá-Dzubukuá [kzw] and Xukurú [xoo] entries should be changed to reflect the names and locations of the (past) speakers of the language of the entry, rather than the names and locations used for descendants who do not speak the language of the entry.


**E16/E17/E18** Yároamë, the recently characterized fifth Yanomami language, is missing (Ferreira 2011).

**E16/E17/E18** The highly endangered Chapacuran language Kujubi is missing (Duran 2000).

**E16/E17/E18** The highly endangered Arawakan Paunaca-Paiconeca language of Bolivia is missing (Montaño Aragón 1989:175-205, Danielsen 2013).

**E16/E17/E18** The Arawakan Apolista language of Bolivia is missing (de Créqui-Montfort & Rivet 1913a, Payne 2005). Although it has been endangered for a long time, Montaño Aragón (1989) was able to gather data in the 1970s and it cannot be ruled out that there are still some very old speakers (p.c. Simon van de Kerke 2011).

**E16/E17/E18** Yarumá, a Cariban language of the Xingu subgroup, is missing (Galvão & Simões 1965, Lévi-Strauss 1948a). Yarumá is listed as a dialect of Suyá [suy] (a Jê language), presumably because the ethnic group assimilated to the Suyá (Villas Boas & Villas Boas 1972), but the original Yarumá, which is attested (Galvão & Simões 1965), is certainly not a Suyá dialect.

**E16/E17/E18** The Flecheiros is one of the uncontacted peoples in the Javari region (Riccardo 1986). If anything, Ethnographic considerations would point to the language being Kanamarí [knum]. But since Wallace (2011) recounts one meeting between a Kanamarí [knum] and the Flecheiros, showing that the two have different languages, they arguably speak a language not counted point to the language being Kanamarí [knum] already in in E16/E17/E18.

### A.6 Missing Languages: Eurasia


**E16/E17/E18** Mator, an extinct Samoyedic language, is missing (Adelung 1806, Helimski 1997).

---

\(^7\)It became extinct sometime between 2000 and 2002 (p.c. Raoul Zamponi 2005).
E16/E17/E18 Aghwan or Caucasian Albanian, an extinct but amply attested Lezgic language, is missing (Gippert et al. 2008, Greppin 1982).

E16/E17/E18 Akkadian, an extinct ancient East Semitic language whose existence is well-known, but not found in E16/E17/E18 (Huehnergard 1997, Ungnad 1964).

E16/E17/E18 Ugaritic, an extinct ancient Northwest Semitic language whose existence is well-known, but not found in E16/E17/E18 (Schniedewind & Hunt 2007).

E16/E17/E18 Sumerian, an extinct ancient language isolate whose existence is well-known, but not found in E16/E17/E18 (Edzard 2003).


E16/E17/E18 Arin, an extinct Yeniseic language, is missing (Werner 2005).

E16/E17/E18 Assan, an extinct Yeniseic language, is missing (Werner 2005).

E16/E17/E18 Pumpokol, an extinct Yeniseic language, is missing (Werner 2005).

E16/E17/Yugh, a recently extinct Yeniseic language, is missing (Werner 1997).

E16/E17/E18 Kott, an extinct Yeniseic language, is missing (Verner 1990).

E16/E17/E18 Pyu, an extinct Sino-Tibetan language (possibly of the Sak subgroup) known from inscriptions, is missing (Beckwith 2002a,b, Luce 1937).

E16/E17/E18 Dalecarlian or Elfdalian is listed as a dialect in the Swedish [swe] entry, but Dalecarlian (of which there are some varieties) is not intelligible to Swedish [swe] and thus merits a separate entry (Äkerberg 2012, Levander 1909).

E16/E17/E18 Chuvan, an extinct Yukaghir language, is missing (Nikolaeva 2008, Tailleur 1962). The lexicostatistical figures in (Nikolaeva 2008:327) make it more likely that Chuvan was a language not intelligible with Kolyma or Tundra Yukaghir, but clearly lexically closer to Kolyma than to Tundra.

E16/E17/E18 Some lexical data recorded as Omok likely represent an extinct Yukaghir language which is missing (Nikolaeva 2008, Tailleur 1959). There is some uncertainty as to whether the lexical data recorded as Omok represent a Yukaghir variety not intelligible to the other Yukaghiric languages or whether the puzzling items represent misunderstanding and poor transcriptions (Nikolaeva 2008:327-328, Nikolaeva 2006:24-25). It is only by assuming the former and following the lexical distances lexical distances in Nikolaeva (2008:327) that we interpret the data as representing a distinct language.

E16/E17/E18 A large array of classical or old literary forms of languages have a timespan such that they are not intelligible to their more modern variants which have E16/E17/E18 entries. They are:
E16/E17/E18 Classical Armenian (Meillet 1913)
E16/E17/E18 Classical Tibetan (DeLancey 2003)
E16/E17/E18 Old Malay (Mahdi 2005)
E16/E17/E18 Old Aramaic (Fitzmyer 1995, Greenfield 1966)
E16/E17/E18 Old Burmese (Luce 1985, Yanson 1990, 2002)
E16/E17/E18 Old Chinese (PeYraube 2008, Starostin & Starostin 2011)
E16/E17/E18 Old Dutch (Quak & van der Horst 2000)
E16/E17/E18 Old English (Hogg 2011)
E16/E17/E18 Old French (Einhorn 1974)
E16/E17/E18 Old Frisian (Bremmer 2009)
E16/E17/E18 Old Georgian (Fähnrich 1994)
E16/E17/E18 Old Irish (Thurneysen 1909)
E16/E17/E18 Old Japanese (Kupchik 2011, Vovin 2005)
E16/E17/E18 Old Kannada (Gai 1946)
E16/E17/E18 Old Khmer (Sidwell 2015b)
E16/E17/E18 Old Marathi (Master 1964)
E16/E17/E18 Old Mon (Jenny & McCormick 2015)
E16/E17/E18 Old Norse (Faarlund 2004)
E16/E17/E18 Old Ossetic (Bielmeier 1989)
E16/E17/E18 Old Persian (Skjærvø 2002)
E16/E17/E18 Old Tamil (Lehmann 1994)
E16/E17/E18 Old Uighur (Tekin 1968)
E16/E17/E18 Older Runic (Antonsen 1975)

E16/E17/E18 A number of ancient Indo-European languages of Anatolia are not listed. They are:

E16/E17/E18 Hittite (Hoffner & Melchert 2008)
E16/E17/E18 Palaic (Carruba 1970, Kassian & Šatskov 2013, Melchert 2008c)
E16/E17/E18 Cuneiform Luwian (Melchert 2003)
E16/E17/E18 Cuneiform Luwian (Payne 2004)
E16/E17/E18 Lycian (Melchert 2008b)
E16/E17/E18 Lydian (Carruba 1969, Ševoroškin 1967)
E16/E17/E18 Milyan (Sherevoshkin 2008)
Phrygian (Brixhe 2008, Orel 1997)
Pisidian (Brixhe et al. 1987, Kassian 2013a)
Sidetic (Melchert 1995:2157-2158, Kassian 2013b)

A number of ancient European and Ancient Near Eastern languages. They are:

Elamite (Khačikjan 1998)
Etruscan (Wallace 2008a)
Faliscan (Bakkum 2009)
Hattic (Girbal 1986, Kassian 2010)
Hurrian (Wagner 2000)
Urartian (Benedict 1958)
Iberian (Bergua Camón 1994, Untermann 1990)
Gothic (Wright 1954)
Oscan (Tikkanen 2004, von Planta 1897)
Umbrian (von Planta 1897)
Lusitanian (Untermann 1997)
Messapic (Pisani 1953, Simone & Marchesini 2002)
Venetic (Lejeune 1974, Wallace 2008b)
Noric (Eichner et al. 1994, Eska & Evans 2010)
Galatian (Eska & Evans 2010, Freeman 2001, Weisgerber 1931)
Gaulish (Transalpine) (Dottin 1918, Lambert 2003b)
Celtiberian (Wodtko 2000, 2003)
Bolgarian, a Turkic language ancestral to or subgrouping with Chu-vash [chv] (Benzing 1959, Erdal 1993, Pritsak 1955, Zimonyi 1992)

E16/E17/E18 has an entry Nenets [yrk] combining both Tundra and Forest Nenets, but the two are not intelligible thus require separate entries (Salminen 1997:13). I wish to thank Matthew Dryer for alerting me to this fact.

Careful examination (Gates 2012) of Jiarong speech varieties filed under Jiarong [ʃja] in E16/E17/E18 necessitate the recognition of separate languages divisible as Situ (Nagano 2003), Zbu (Sun 2003a), Tshobdun (Gates 2012) and Japhug (Jacques 2008). Thus, E16/E17/E18 has three entries too few.

The Ancient South Arabian languages are missing. They are:
A number of ancient languages of the Middle East and Central Asia are attested, but missing from E16/E17/E18. They are:

- Tumshuqese-Khotanese (Emmerick 1989)
- Khwarezmian (Henning 1955)
- Kitan (Jin 2008, Kane 2009)
- Bactrian (Davary 1982)
- Pahlavi (Nyberg 1964-1974)
- Parthian (Durkin-Meisterernst 2014, Rastorgueva & Molčanova 1981)
- Sogdian (Skjærvø 2003)
- Tokharian A (Krause & Thomas 1960)
- Tokharian B (Krause & Thomas 1960)

A number of ancient Sino-Tibetan languages of various levels of attestation and decipherment are missing. They are:

- Nam (very poorly understood) (Thomas 1951)
- Pai Lang (poorly understood) (Beckwith 2008, Ding 1942)
- Tangut (imperfectly understood) (Jacques 2014, Sofronov 1968)
- Zhangzhung (imperfectly understood) (Haarh 1968, Takeuchi et al. 2001, van Driem 2001b)

Andro, an extinct Sak language of Myanmar documented by a wordlist in McCulloch (1859), is missing (unless counted in the Lui [lba] entry, see above).

Sengmai, an extinct Sak language of Myanmar documented by a wordlist in McCulloch (1859), is missing (unless counted in the Lui [lba] entry, see above).

Chairel, an extinct Brahmaputran language of Myanmar documented by a wordlist in McCulloch (1859), is missing (unless counted in the Lui [lba] entry, see above).

Dotyali, a close relative of Nepali, is missing from E16 (Eichentopf 2014, Upreti 1900).

Molise Croatian, a variety of Croatian spoken in Italy which is not intelligible to Serbian-Bosnian-Croatian of Yugoslavian, is missing from E16 (Breu & Piccoli 2000, Sammartino 2004).
E16/E17/E18 Recent fieldwork reveals that there are at least three mutually unintelligible languages falling under the E16/E17/E18 Biao [byk] entry, which may be divided as Shidong Biao, Yonggu Biao and Dagang-Baoshan-Liangcun Biao. Two entries are thus missing (Hsiu 2013a).

E16/E17/E18 Cosao, a newly discovered language on the border of Laos and China, is missing (Bai & Xu 2015). It is probably in the Bi-Ka subgroup of Sino-Tibetan but mutually unintelligible with any of them. I wish to thank Andrew Hsiu for bringing this language and the work of Bai & Xu (2015) to my attention.

E16/E17/E18 Seb Seliyer, a Gypsy variety not known to be immediately related to any other Gypsy language (Windfuhr 2002), in the mountains of the Firuzkuh area east of Tehran (Sotuda 1341 AHS [1962]), is missing.

E16/E17/E18 E17/E18 has one entry for Newari [new] but the dialects around Dolakha are not mutually intelligible to those in the Kathmandu Valley (Genetti 2007:1). Thus the entry should be split into two entries, at least.

E16/E17/E18 Bânà, a Hmongic language of China, most closely related to She [shx], is missing (Chen 2001).

E16/E17/E18 Two enclaves of Greek speakers are found in Calabria and Apulia of present-day Italy. They are subsumed in E17/E18’s entry on Modern Greek [ell], noting, however that the “Greek of Italy and that of Corsica are probably distinct languages”. Indeed, especially the Apulian variety of Otranto is not intelligible to Greek speakers of Greece (Tozer 1889:37), and thus merits a separate entry.

E16/E17/E18 Homshetsma, the Western Armenian variety of Hamshen, is missing from E16/E17/E18 either as a separate language or a dialect of Armenian. It is not mutually intelligible with other Armenian (Vaux 2007:257) and should thus merit a separate entry.

E16/E17/E18 Bih, a language closely related to Ede (Rade [rad] in E16/E17/E18), but not intelligible to Ede (Nguyen 2013:4-11), is missing. It was presumably filed with Ede [rad] because nearly all Bih also speak Ede.

E16/E17/E18 Basum, a Bodish but non-Tibetic language of Tibet, is missing (Tournadre 2014). Data appears in Qu et al. (1989). I wish to thank Guillaume Jacques for clarifying the status of Basum to me.

E16/E17/E18 There are attestations of an earlier Northwestern Iranian language in Azerbaijan (e Mortazavi 1384 [2005], Sadeqi 2002, Yarshater 1987) which we may call Adhari. Though modern languages likely descendant from Adhari are in E16/E17/E18, Adhari itself is missing.

8I owe this observation of Raoul Zamponi.
There are two Southeast Asian Tai-Kadai languages surfacing under the name Mulao, one in the Kam-Sui subgroup (Jun & Guoqiao 1980) and one (nearly extinct) the Gelao subgroup (Bo 2003). E16 accurately reflects the Kam-Sui one in the Mulam [mlm] entry, which even contains the correct clarificatory statement “Different from the nearly extinct Ge-Yang Tai-Kadai language also called Mulao”. However, there is no entry in E16 for the other Mulao language.

An entry for Northern Hua Miao in Southeast Yunnan is missing. It is not quite intelligible with Southern Hua Miao, represented by E16’s Hmong Njua [hnj]; this impression probably represents acquired intelligibility due to radio broadcasts (McLaughlin 2012:9-10).

Hezhou, a far eastern high-contact Uyghur-origin language of China, is missing (Lee-Smith 1996, Wurm 1997).

Tangwang, a high-contact Mandarinic language in the Gansu area of China which is unintelligible with Mandarin, is missing (Yiblaxam 1985).

Dao a high-contact Mandarinic language unintelligible to Mandarin, Wutunhua and Tangwang speakers (Acuo 2004), is missing.

Pinghua, a Sinitic language in the south of China of uncertain precise subclassification, is missing (Li 2000).

Syriem, a Kuki-Chin language closely related to Aimol [aim], is missing (Haokip 2011).

The language of Jabal Räzih, a Sayhadic language or Arabic language with Sayhadic influence, which is famously not intelligible with any other Arabic language, is missing (Watson et al. 2006).

The Bangru language, a close relative of Miji [sjl], is missing (Sun 1993:159, Ramya 2012, Blench 2012b). Bangru and Levai (another name for Bangru) is listed as a dialect of Hruso [hru], presumably because of the confused idea that Hruso [hru] is a dialect of Miji [sjl], going back to Shafer (1947). Miji and Hruso share some borrowed vocabulary but are otherwise vastly different – nowhere near being dialects – and probably not even deep-level related (Blench & Post 2011). If anything, Bangru is a dialect of Miji, not of Hruso, and though the amount of data so far available is small, it suggests that Miji and Bangru are different enough not to be intelligible.

Cäijia, a newly discovered Sinitic (or sister branch to Sinitic) language in China, is missing (Bo 2004). I wish to thank Hilário de Sousa for informing me of this language.

Longjia, a poorly attested extinct Sino-Tibetan language in Western Guizhou, is missing (Guizhou Province Ethnic Classification Commission 1984, 1986, Hsiu 2013). I wish to thank Andrew Hsiu for bringing this language to my attention and for passing on the data.
Sadu, a language discovered in 1958 and filed under the Bai nationality, is missing. It is a Loloish language, not a Bai language (Bai 2012, Bai & Wang 2012, Bradley 2014).

Lawu, a newly discovered Ngwi language of Yunnan province in China, is missing (Yang 2012).

Lizu is an Ersuic language not intelligible to Ersu [ers] (or Duoxu) and should be a language separate from the Ersu [ers] entry (Yu 2012:2, Chirkova 2014:107).

Douxu is an Ersuic language not intelligible to Ersu [ers] (or Lizu) and should be a language separate from the Ersu [ers] entry (Yu 2012:2, Chirkova 2014:107).

Samatu, Central Loloish language, is missing (Bradley 2002, Bradley 2007a:399). The failure to include it was possibly catalyzed by a confusion with Samatao [ysd], a similar sounding name for a different Northern Loloish language.


Khammigan Mongol, a Oirat-Khalkha Mongolic language, is missing (Janhunen 2003).

The Koro-Aka language, a Siangic language of Arunachal Pradesh, is missing from E16 (Anderson & Murmu 2010). This is the language arrogantly purported to be a “newly discovered” language by National Geographic’s "Enduring Voices" project and picked up by many major news outlets, e.g., BBC. In fact, language data on Koro (as Miri-Aka) was published by Grewal (1993, 1997) and the existence of the language (as a separate language) has been mentioned (without data) in publications several times earlier, e.g., Sinha (1962:13-15).

Nyisu, a Southeastern Ngwi language, is missing (Wu 1997, Yunnan Institute of Nationalities 1984).

Zhongu, a Tibetic language not intelligible with the Tibetic languages listed in E16, is missing (Sun 2003b).

Kundal Shahi, an Indo-Aryan language of the Shina group, is missing (Rehman & Baart 2005) from E16 in spite of being published in an SIL publication series.

Old Wanj, an Iranian language of the Shugni-Yazgulami subgroup that went extinct some time after 1924, is missing (Edelman 2000).

The Hruso [hru] entry mentions Koro (as Aka) with the incoherent claim that it is a ‘dialect’ that shares 9% of its vocabulary with Hruso [hru]. This information is presumably a misinterpretation of the unpublished 2005 survey (Abraham et al. 2005), which correctly identified Aka as a separate language and contains a vocabulary and the lexicostatistical comparison with the 9% figure.

E16/E17/E18 Theen, a Khmuic language of Laos that subgroups with Ksingmul, is missing (Proshan 1996).

E16/E17/E18 Muangphe is an attested separate Mondzish language missing from E17/E18 (Hsiu 2014).

E16/E17/E18 Maang is an attested separate Mondzish language missing from E16/E17/E18 (Hsiu 2014, Zhou 2014).

E16/E17/E18 Mango is an attested separate Mondzish language missing from E16/E17/E18 (Hsiu 2014).

E16/E17/E18 Maza is an attested separate Mondzish language missing from E16/E17/E18 (Hsiu 2014, 2013b).

E16/E17/E18 Mondzi, a Lolo-Burmese language close to Mantsi [nty], but not intelligible to it (Lama 2012:140, Hsiu 2014 Hsiu 2013b) is missing.

E16/E17/E18 Luren is an ethnic group of Guizhou province of China which is typically classified as Manchu. The data in (Guizhou Province Ethnic Classification Commission 1984) shows that it is a new Tibeto-Burman language which is missing from E16/E17/E18. I wish to thank Andrew Hsiu for bringing this language to my attention.

E16/E17/E18 Milang, a Tibeto-Burman language of the Siangic subgroup, is missing as a distinct language (Grewal 1997, Post & Modi 2011, Tayeng 1976), first identified as such by Duff-Sutherland-Dunbar (1916:10-11) with a wordlist from an L2 speaker. The name occurs under the hodge-podge entry Adi [adi] (a highly ambiguous term, see Singh 1994:17-20, Post 2013a:86-88) which unites the names of several very different languages. Milang is not intelligible with any of them (Sun 1993:473).

E16/E17/E18 Bokar, a Tibeto-Burman language of the Western Tani subgroup, is missing as a distinct language (Grewal 1997, Megu 1990, Ouyang 1985). Bokar occurs under the entry Adi [adi] along with a number of Western and Eastern Tani leets. Bokar should be taken out of “Adi” since it is far too divergent from both the other Western and Eastern Tani languages (Sun 1993).

E16/E17/E18 Tangam, a Tibeto-Burman language of the Tani subgroup is missing as a separate language, though is incorrectly listed under in the Adi [adi] entry. It is not intelligible to any other Tani language (Badu 2004, Post 2013b).

E16 Lautu, a Chin language different from all other Chin languages (Bareigts 1981:33-34), is missing from E16 (Luce 1959, Thang 2001).

E16/E17/E18 Laze (also known as Muli Shuitian), a Tibeto-Burman language of the Naish group (Jacques & Michaud 2011), is missing (Huang 2009).

11I wish to thank Frank Proshan for clarifying the status of Theen to me.
Sorbung (also known as Southern Tangkhul), a Tibeto-Burman language organizable as part of the Old Kuki group, is missing (Brown 1837, Mortenson & Keogh 2011).

Xuzhang Lalo, is a variety of Lalo recently shown to be unintelligible to all other Lalo varieties with entries in E17/E18 (Yang 2010).

Eka is a variety of Lalo recently shown to be unintelligible to all other Lalo varieties with entries in E17/E18 (Chan & Yang 2008, Yang 2010).

Mangdi or Southern Lalo is a variety of Lalo recently shown to be unintelligible to all other Lalo varieties with entries in E17/E18 (Yang 2010).

Yangliu is a variety of Lalo recently shown to be unintelligible to all other Lalo varieties with entries in E17/E18 (Yang 2010).

Cypriot Greek is not intelligible to Greek of Greece (Arvaniti 2006:26), hence Cypriot Greek is missing as a separate language.

Yilan, a Japanese-lexified creole of Taiwan, is missing (Chien & Sanada 2010, Sanada & Chien 2012).
Appendix B lists entries in E16/E17/E18 which are spurious. To be more precise, an entry is listed here as spurious if:

- If it duplicates another extant E16/E17/E18 entry, or,
- It cannot be asserted that the entity denoted in the entry was a different language from that of every other entry in E16/E17/E18

Again, we do not list languages which are spurious solely in virtue of the interpretation of a dialect situation correctly understood (but interpreted differently) in E16/E17/E18 and, in all cases, we provide references to the literature that support the argument made regarding the spurious language in question.

B.1 Spurious Languages: Africa

E16/E17/E18 E16/E17/E18 lists two languages Ngombe [nmj] (in Central African Republic) and Bangandu [bgf] (in Cameroon and Congo-Brazzaville). There exist two such populations with those names and locations. They were separated approximately fifty years before the arrival of Europeans and are mutually aware of each other (Moñino 1995:16). They speak the same language, as comparison with actual data confirms (Moñino 1988, Syssa & Pagbanda 2013). The classification of Ngombe [nmj] (but not location or any other information in the entry) in the Baka-Gundi group suggests that there is confusion with the clan of Pygmies called Bangome (where Ba- is arguably a prefix) encountered in Congo-Brazzaville (Pepper 1955, Verhille 1948, 1948, 1948, 1949, Klieman 1997:290, Sato 1992, Boyeldieu & Cloarec-Heiss 1986, Ballif 1992:81-82). But the language of the Bangome clan (Pepper 1955, Verhille 1948, 1948, 1948, 1949, Boyeldieu & Cloarec-Heiss 1986, Ballif 1992:212) as far as it is documented, is indistinguishable from that of the vocabularies in Ouzilleau (1911) which represent the Gundi [gdi] entry, cf. Klieman (1997:277). Thus, even if Ngombe [nmj] were to refer to the Bangome clan of Pygmies, the entry is spurious as a separate language.

E16 Beti [btb] is listed as an unmapped language of Cameroon. However, Beti is a term encompassing Eton [eto], Ewondo [ewo], Fang [fak], Mengisa [mct] which already have their own E16 entries (Alexandre 1965, Dugast 1949:60).

E16/E17 The Forest Maninka (also known as Wassulunka) [myq] entry is described as an ethnic Fulani group who now speak an East Manding language. This variety, however, scores over 90% lexicostatistical similarity with several other East Manding (e.g., Eastern Maninkakan [emk]) languages which have separate entries in the E16/E17 (Long 1971:56) and should thus not have a separate entry.

E16/E17/E18 Bubia [bbx] is listed as a separate language of Cameroon. If the entry refers to their ancestral language, it is simply a variety of Bube [bvb], needing no separate entry, and if it refers to the language they are shifting to, it is simply a variety of Wumboko [bqm], equally not in want of a separate entry (Ardener 1956:10-12, 2013, 31, 35, Blench & Martin 2010).
E16/E17/E18 E16/E17/E18 list Idesa [ids] as a separate language, but Idesa is a section of Otuo town and its people were speakers who shifted from Ghotuo [aaa] to Okpe [oke], languages which already have entries (Wente-Lukas 1985:280).

E16/E17/E18 Kofa [kso] is listed as a separate language, with the comment that it is “Reportedly a separate language from Bura-Pabir [bwr]”. It is correct that it is distinct from Bura-Pabir [bwr] but already (Meek 1931:118-121) had data from Kofa and shows that they speak the same language as Malabu, i.e., Bata [bta].

E16/E17 E16/E17 lists Imeraguen [ime] as an unclassified language of Mauretania spoken by an ethnic group traditionally subsisting on fishing. Presumably based on a blanket statement by Gerteiny (1967:56), E16/E17/E18 reports that it is “Reportedly a variety of Hassaniyya structured on an Azer (Soninke) base”. However, it is unclear what Gerteiny (1967:56)’s statement is based on and no data or other evidence for the Azer-Hassāniyya mix is adduced. Other ethnographers with field experience, notably Fortier (2004), say they speak Ḥassāniyya with some special vocabulary related to fishing. The specific forms of the fishing vocabulary are cited and they clearly derive from a Berber (not Azer) source, and there is no ethnohistoric evidence for an Azer origin either (Fortier 2004). Since there is insufficient evidence to posit a past or present separate language for the Imraguen, the entry is spurious.

E16/E17/E18 E16/E17/E18 lists Shempire Senoufo [seb] as a separate Senoufo language located in the extreme North of Northwest Côte d’Ivoire. This area is marked as Supyire Senoufo [spp]-speaking by Carlson (1994:xx). The same are is marked as West Senari, i.e., a dialect of Senoufo-Cebaara [sef], in the detailed meta-survey by Mills (1984:xii-xx) (also followed by von Roncador & Miehe 1998:264). West Senari is reported to have 84% intelligibility to the more central dialects, but with the explicit caveat that the transition of the Supyire dialects needs more investigation (Mills 1984:xiv, xvi). Thus Shempire Senoufo [seb] is either a dialect of Supyire [spp] or of Cebaara [sef], not a separate language.

E16/E17/E18 E16/E17/E18 lists Syenara Senoufo [shz] as a separate Senoufo language in the Southeast corner of Mali. But this variety is mutually intelligible with the variety across the border in Burkina Faso (Garber 1987:6) subsumed under Sícité Sénoufo [sep].

E16/E17/E18 E16/E17/E18 lists Dek [dek] as an unmapped language of Cameroon. It goes back to a name Dek on a map by Tessmann (1928), but no data has surfaced to assert it is/was a language of its own (Elders 2006).

E16/E17 Daza [dzd] is a presumed Chadic language of the Bole-Tangale group, but is actually unattested and thus cannot be ascertained to be separate language (Schuh 1978b).

E16/E17 Sara Dunjo [koj] is listed as a separate language, but our only information on Dunjo (Tucker & Bryan 1956:57, Caprile 1981:239) in the past century claims that it is a dialect (of the city Njounjou) of the language in the Sara Kaba Nàà [kwv] entry.
where indeed Dunje is listed as a dialect). This is confirmed in the more recent detailed sociolinguistic survey of Marianne (2012).


E16/E17 E16/E17 lists Gbati-ri [gti] as a separate language, but the only source on this language (Hackett & van Bulck 1956:74) has it as a dialect of Nyanga-li [nyc].

E16/E17 Bemba [bmy] is listed as an Unclassified Bantu language of South Kivu province in the DRC. A more precise location is not given. It is difficult to interpret this entry as anything other than a duplication of the Bembe [bmb] entry – a relatively well-known Bantu language of the Lega complex in South Kivu province (van Bulck 1952:48, Hulstaert 1950:31, Biebuyck 1973). The population of Bemba [bmy] is given as 296,000 which is a plausible figure if it refers to Bembe [bmb], and there is absolutely no sign of a different Bemba language of such a large population in Kivu province (Kadima 1983:34-37, Moeller 1936). The alternative name Kinyabemba for Bemba [bmy] suggests that the information on Bemba [bmy] came from transient Bembe in Rwanda and therefore the equality with Bemba [bmb] may have been overlooked.

E16/E17 Similarly, Songa [sgo], alternative name Kisonga, is listed as an Unclassified Bantu language of South Kivu province without a precise location or a population estimate. There is no sign of a language with that name in South Kivu province (Kadima 1983:34-37, Moeller 1936) unless it is Kisanga (Botne 2003:423). If it is Kisanga, the Songa [sgo] entry is spurious because Kisanga falls under the Lega-Shabunda [lea] entry.

E16/E17 Borna (Democratic Republic of Congo) [bxx], alternative name Eborna, is given as an Unclassified Southern Bantoid language of the DRC without a population estimate and without a location. The entry must emanate from Welmers (1971:786) since that is the only source mentioning such a language with exactly that much (lack of) information. Welmers’s list, in turn, derives from sheets of African languages at the US Office of Education started in 1959 (Welmers 1971:761). Since these sheets were hand-written there is considerable likelihood that Borna is a misreading of Boma or eBoma (the entry for which is Teke-Eboo [ebo]) which would have been extant in the literature at the appropriate time.

E16/E17 Similarly, Buya [byy] is given as an Unclassified Southern Bantoid language of the DRC without a location. This too, must emanate from Welmers (1971:788-789). Since that is the only source mentioning such a language with exactly that much (lack of) information. In all likelihood, Buya [byy] is a rendering of a DRC language already counted in another E16/E17 entry, presumably Buyu [byi].

E16/E17/E18 E16/E17/E18 lists a Northwestern Edoid language Aduge [adu]. But Aduge is an Okpanheri [opa]-speaking village (Jungraithmayr & Ganslmayr 1986:39), and is thus already included.
E16/E17/E18 The now extinct Baga Kaloum [bqf] attested in Koelle (1854) spoke the same language as the Baga Koga [bgo] (Dalby 1965, Voeltz 1996:5-7) so the Baga Kaloum entry is spurious.

E16/E17/E18 The Baga Binari [bcg] entry presumably\(^{12}\) refers to the indigenous language of the island of Binari. This language is the same as that in the Baga Mboteni [bgm] entry (Voeltz 1996:25-26) and is thus spurious.

E16/E17/E18 Baga Sobane [bsv] is subsumed in the Baga Sitemu [bsp] entry as the two are "sensiblement identique" and "l'intercompréhension est immédiate" (Voeltz 1996:6-7).

E16/E17/E18 E16/E17/E18 has Ndengerekö [ndg] as well as Rufiji [rui] entry, but these terms denote the same language – Ndengerekö being the name of the language and Rufiji\(^{13}\) the name of a river and the area around it (Ström 2009). KiRufiji is a natural Swahili way to denote the language spoken in the Rufiji area, and this term is used in the Tanzania census, which is a plausible origin of the E16/E17/E18 entry\(^{14}\).


E16/E17/E18 The only source on Gengle [geg] and Kugama [kow] shows them to have 100% cognacy (Meek 1931:529-531) (cf. Temple 1919:244-247) and therefore their E16/E17/E18 entries should be merged.

E16 E16 has an entry Ngong [nnx] as well as an entry Nagumi [ngv], but Ngong is a village of Nagumi (Maddieson & Williamson 1975:132-133) and should not be a separate entry.

E16/E17/E18 E16/E17/E18 has an entry Bainouk-Samik [bcb] and another Bainouk-Gunyuño [bab] but this split cannot be motivated by anything except the Senegal-Guinea Bissau national border (Lüpke 2010:159, Cobbinah 2013:31-34). Thus the entries should be merged, leaving one of them spurious.

E16/E17 E16/E17 has two entries Shubi [suj] and Subi [xsj] which are both described as being spoken in the Kagera region of Tanzania. They are the same language (Maho & Sands 2003).

E16/E17/E18 Gey [guv] (Gewe, Gueve) is listed as a separate Adamawa language in Cameroon. But this language is attested with numerals in Baudelaire (1944) which

\(^{12}\) The entry contains inconsistent information in that “Kalum” is given as an alternative name and the language is said to be similar to, and subclassified with Temne and Baga language. The indigenous language of Binari is not closely related to or saliently similar to the Temne and Baga languages.

\(^{13}\) Rufije may also be used as a name for a subfamily that includes Ndengereko and other languages (Chuo Kikuu cha Dar es Salaam 2009:xi).

\(^{14}\) I wish to thank Eva-Marie Ström for discussion on the Ndengereko [ndg] and Rufiji [rui] entries.
shows it to be the same language as Duli [duz], attested in Baudelaire (1944) and Strümpell (1922-1923). So the Gey [guy] entry duplicates Duli [duz]. A curiosity is that the numerals of this Duli-Gey language resemble numerals of the Bata group of Chadic while the rest of the vocabulary shows good resemblances to the Leko-Nimbari subgroup of Adamawa.

E16/E17/E18 Mayeka [myc], a Bantu language of the Democratic Republic of the Congo, is very likely to be the same as Ngbinda [nbd] since de Calonne-Beaufaict (1921:114, 120, 247, 251) calls Mayeka a clan of Abangbinda, though no actual linguistic data is presented. The information on Mayeka [myc] in E16/E17/E18 presumably derives from Hackett & van Bulck (1956) who took down a Mayeka wordlist but were not able to find out the location of Mayeka, and did not have access to a Ngbinda wordlist (van Bulck 1954:74-75) to check if Mayeka and Ngbinda were the same. Though, at the present time, if Hackett and van Bulck's unpublished Mayeka wordlist still exists in an archive somewhere, it should be checked whether the data is actually Ngbinda or a different, but related, language.

E16/E17/E18 Dombe [dov] is a derogatory nickname for Tonga found in Hwange district of Zimbabwe (Hachipola 1998). Their variety is closest to the Toka-Leya dialects spoken on the northern side of the Victoria Falls in Zambia, which are dialects of E16/E17/E18’s Tonga [toi].

E16/E17/E18 E16/E17/E18 has an entry for Judeo-Berber [jbe] originating in the High Atlas of Morocco. Although there were certainly Jews who spoke Berber, contra Galand-Pernet & Zafrani (1970), there is now a detailed investigation revealing that there was no normed Jewish variety that differed from the corresponding non-Jewish Berber (Chetrit 2007). The entry presumably entered on analogy with Judeo-Moroccan Arabic varieties, which are real in the sense of being normed varieties that differ from the non-Jewish ones (Lévy 1990).

E16/E17 E16/E17 has no less than three entries for Angolan varieties of Ju, i.e., !O!ung [oun], Vasekela Bushman [vaj] and Maligo [mwj]. Although the Angolan dialect situation is imperfectly known, there is no basis for dividing it into three separate languages – it should be no more than one entry (Snyman 1997, Snyman 1980, Brenzinger 2010:329-337, Heine & König 2014).

E16/E17 E16/E17 has a Kakihum [kxe] entry, describing a West Kainji language of the Kambari group, as well an entry Western Acipa [awc] in the Kamuku group of West Kainji. Western Acipa [awc] is the language spoken in Kakihum (and the variety there could be labeled the Kakihum dialect of Western Acipa or Tieuwuun locally, see McGill 2009) and there is no different language spoken there except Tsvu’â’i [vd] (another Kambari language – the Kakihum Tsvu’â’i seems to be more or less identical to the Tsvu’â’i spoken spoken by Ava’di in the Tirisino and Tikula dialect areas of Cicipu, p.c. Stuart McGill 2010). Therefore a separate Kakihum entry is spurious. The

15I wish to thank Roger Blench and Richard Gravina for discussion of this vocabulary.
inclusion of the erroneous entry was possibly catalyzed by the fact that Western Acipa [awe] is now known to belong with the Kambari group rather than the Kamuku group.

E16/E17 Uokha [uok] is a place name where the language of the clans Emai, Iuleha and Ora is spoken (Elugbe 1989:22-24, Jungraithmayr & Ganslmayr 1986:34), but this language already has an entry in Emai-Iuleha-Ora [ema].


E16 Ibilo [ibi] is a place name in the Okpamheri [opa] area (Elugbe 1989:24, Jungraithmayr & Ganslmayr 1986:38-39, Thomas 1910, Bradbury 1968), and is thus spurious.

E16/E17/E18 Kwak [kwq] appeared as Nkwak in the 1992 Index of Nigerian Languages (Crozier & Blench 1992), but was withdrawn from the 3rd (Blench no date [1999]) since it had been discovered as a place in the Hyam area\textsuperscript{16}. Despite reports to the E16/E17/E18 editor, it has not been taken out.

E16/E17 Aramanik [aam] is listed as a Southern Nilotic language of the Nandi group, presumably because the Aramanik people assimilated to the Nandi. The original Aramanik language was a Cushitic language (or a non-Nilotic language with heavy Cushitic overlay) usually called Aasáx (Fleming 1969) and is already counted in a separate Aasáx [aas] entry.

E16/E17/E18 E16/E17/E18 has three entries for “Dorobo” (hunter-gatherers in the vicinity of Maasailand) languages close to Okiek [oki] namely Mediak [mwx], Kisankasa [kqh] and Mosiro [mwy]. But our only source on Mediak [mwx] and Kisankasa [kqh] says “The three names [Mediak, Mosiro and Kisankasa – HH] given above would seem to denote one and the same type of Dorobo. I have been unable to discover any difference between them.” (Maguire 1928:134), so two of the three entries are spurious.

E16 E16 has an entry Nyeng [nfg], but Nyeng is uncontroversially a Ndun/Ahwai [nfd] variety (Blench 2006a), and thus already covered.

E16 E16 has an entry Shakara [nkf], but Shakara is uncontroversially a Ndun/Ahwai [nfd] variety (Blench 2006b), and thus already covered.

E16/E17/E18 E16/E17/E18 has an entry Yauma [yax] in the Kwango river area of Angola. But Yauma is simply the name given to Mbunda [mck] who remained on dry land (Yauma meaning “dry mound” or “dry island”, Papstein 1994:123, Barrett 2004:308-309) and does not represent a distinct language.

E16 E16 has Tingal [tie] as a separate language. This division stems from a conservative interpretation of Tucker & Bryan (1966:70)’s rendering of information provided by Stevenson. Stevenson had no first-hand data from Tingal and therefore could not

\textsuperscript{16}I wish to thank Roger Blench for clarifying this matter to me.
assert that Tingal was a either a separate language or a dialect of Tegali or of Tagoi (Stevenson 1956/1957:103). When an early version of the Ethnologue was compiled, such cases were rendered as separate languages, and while some others, such as Tukum and Turum, were weeded out later, Tingal remained (p.c. Andrew Persson 2009). In July 2009 in Kadugli, I was able to find a speaker who had grown up in Tingal and had excellent command of the language. While I only had the opportunity to spend an hour or so with him, his speech was nearly identical to the Rashad and Tegali [ras] materials in Stevenson (1956/1957:46-52) and his own listing of villages whose speech was intelligible versus unintelligible to him matched the Tagoi/Tegali speaking villages listed by Stevenson perfectly. Based on this information, Tingal should be considered a Tegali dialect [ras].

**E16/E17/E18** E16/E17/E18 has a separate entry for Odut [oda], an Upper Cross language of Nigeria. The language of this village is extinct and not attested in publication (Dimmendaal 1978:36-37). Katherine Barnwell had an unpublished 20-word list showing it to be the same as another Upper Cross language (p.c. Roger Blench 2011). The entry is thus spurious.

**E16/E17/E18** E16/E17/E18 has an entry Buso [bso], but this language is the same as Kwang [kvi] (Toureux et al. 1986:116, Jungraithmayr 1977:140).

**E16/E17/E18** E16/E17/E18 has an entry Ndonde Hamba [ndj] in the Lindi region of Tanzania. While different from its immediate neighbours, it is a dialect of the language in the Makonde [kde] entry (Kraal 2005).

**E16/E17/E18** E16/E17/E18 has two entries Teke-Eboo [ebo] and Teke-Nzikou [nzu]. Two such entities are frequently enumerated separately but the recent comparison by Raharimananitsa (2012) shows that such a distinction is untenable linguistically. Thus the two entries should be merged.


**E16/E17/E18** E16/E17/E18 has an entry Lalia [lal] in addition to Bongando [bxl]. But studies in the field emphasize that Lalia is simply a variety in the Bongando area with no special status vis-a-vis other Bongando varieties (Hulstaert 1951:22, Lingomo 1995:340, Hulstaert 1987:205-207), wherefore we merge the two should arguably be merged.

**E16/E17/E18** E16/E17/E18 has two entries Leti (Cameroon) [leo] and Mengisa [mct]. Leti [leo] is the ancestral language of the Mengisa ethnic group who are switching to Eton [eto]. The Eton that most Mengisa now speak is not linguistically remarkable (Geslin-Houdet 1984). Thus the Mengisa entry is duplicated whether or not it refers to the ancestral or the contemporary language. (In addition, the Leti [leo] language is so close to Tuki [bag] as to perhaps count as a dialect of it (Biloa 2013:37-38)).
E16/E17/E18 E16/E17/E18 has an entry Yamongeri [ymg] which does not appear in the E16/E17/E18 maps. Its location is specified only as “Equateur province: south of Congo river”, but this and the name is sufficient to locate the group (Hackett & van Bulck 1956:72). This group speaks a variety of Mongo [lol] (Hulstaert 1999) and indeed the E16/E17/E18 maps mark the territory where it is spoken as Mongo [lol].

E16/E17/E18 E16/E17/E18 differentiates between Yela [yel] and Kela [kel] corresponding to two extant geographical enclaves. Hulstaert (Hulstaert 1951:23, Hulstaert 1999:20-21) who had studied the varieties in the field, argues that this distinction is not particularly salient linguistically, wherefore they should arguably be merged.


B.2 Spurious Languages: Australia

E17/E18 E17/E18 has an entry Arakwal [rkw] for an extinct Australian language. Although Arakwal may be an (old or recent) ethnic group, linguistically it is one of many Bandjalang [bdy] varieties (Oates 1975:212, Wafer & Lissarrague 2008:354) and should not have a separate entry in a linguistic catalogue.

E16/E17/E18 E16/E17/E18 has an entry for Ngurmbur [nrx] following information given by speakers of neighbouring languages that a tribe with this name were prevalent in a certain territory (as per maps like Harvey 2001:xii). However, only second hand meta-reports are available on the supposed language, and was said to be similar to Umbugarla [umr] (Davies 1989). Consequently, it cannot be asserted that Ngurmbur [nrx] was a separate language.

E17 E17 has an entry Gugu Mini [ggm] with little information beyond this name and its classification as Rarmul Paman. The name Gugu Mini (with spelling variants) has been applied to two distinct languages, namely the Kawarrang-Ogh Undjan cluster (Sommer 1997b:1-3, Sommer 1976:133-144) and to Ikarranggal-Angkula-Alungul [ikr] (Sommer no date:1-2, Sommer 1976:133-144), but only Ikarranggal-Angkula-Alungul is Rarmul Paman. This implies that Gugu Mini [ggm] duplicates the Ikarrangal [ikr] entry and should either be removed or changed to mean Kawarrang-Ogh Undjan, which is otherwise missing from E17.

E16/E17/E18 Kuku-Mangk [xmq] is listed as an extinct separate Middle Paman language. However, out only information regarding Kuku-Mangk (von Sturmer 1978:171-172, Sutton 1978:168-171, Smith & Johnson 2000:358-361) says that some informants equate it with Kugu Iiyanh [wij] while others ascribe it a separate identity. But even if the informants giving it a separate identity are correct, they do not necessarily mean that Kuku-Mangk is not intelligible to Kugu Iiyanh or any of the other Kugu dialects. No form-meaning pairs are available to ascertain any variant of the relationship, and thus
there is insufficient basis for ascertaining that Kugu-Mankg [xmq] should be a separate entry.

**E16/E17/E18** Giyug [giy] is unattested and reports available regarding its status are all second-hand and do not necessarily indicate that the language was actually unintelligible to those of the mainland (Ford 1990). I wish to thank Ian Green and Mark Harvey for discussing this question with me.

**E16/E17/E18** The Kunggara [kvs] entry references the same language as Gurdjar [gdj] (Black 1980:187) and is thus spurious.

### B.3 Spurious Languages: Eurasia

**E16/E17/E18** E16/E17/E18 has two languages Desiya [dso] and Adivasi Oriya [ort]. These names simply divide the speakers of the same language along a state border (Gustafsson 1989, Mahapatra 2002, 1992) leaving one of the entries spurious. I wish to thank Felix Rao for clarifying the status of these two entries to me.

**E16/E17** Ethnologue has a separate entry for Xiandao [xia] noting that “Some consider it a dialect of Achang [acn]”. In fact, it is easily mutually intelligible with Achang [acn] (Dai 2005, Tsung & Dai 2004) and so has little basis for forming a separate entry.

**E16** has a separate entry Yamphe [yma] which appears to be a leftover of varieties taken up in the linguistic survey of Nepal (Hanfson 1991). However, Sibao Yamphe of Hanfson (1991) falls under Southern Lorung, i.e., the Southern Yamphu [lrr] entry, and other varieties that may be called Yamphe fall under the Yamphu [ybi] entry (Hilty & Mitchell 2012, Rutgers 1998).

**E16/E17/E18** E16/E17/E18 has two entries Lahta Karen [kvt] and Zayein Karen [kxk] with the same location, with the comment that the two be the same. The do indeed reference the same groups (Diran 1997:144).

**E16/E17/E18** E16/E17/E18 has two entries Kháng [kjm] and Khao [xao] for languages in Northwestern Vietnam associated with the Khmuic subgroup of Austroasiatic. The name Khao is an exonym used mainly in older sources (Bouchet 1935, 1936, Macey 1906) for a fraction of the Kháng in Thuan-Chau, but the language is the same (Edmondson 2010, Proshcan 1996).

**E16/E17/E18** E16/E17/E18 has an entry Thudam [thw]. But Thudam is a village where the same language as other villages to the south and east, including the Tokpe Gola and Olangchung Gola villages, is spoken (von Führer-Haimendorf 1975, Bartram 2011:27-28). This language already has an entry in Walungge [ola].

**E16/E17/E18** Ir [irr] is listed as a separate language “closest to Ong [oog]” but, in fact, the two are “parfaitement identique” (Ferlus 1974:113), i.e., completely identical.
**E16/E17/E18** Pao [ppa] is listed as a Tibeto-Burman language of the scheduled tribe Pao of Madhya Pradesh in India. However, the Pao speak a variety of Bundelkhandi to themselves and others (Singh 1994:977-979, Shukla 1986, Singh 1997:63), a language which already has an E16/E17/E18 entry in Bundeli [bns]. The misclassification of Pao as a Tibeto-Burman languages in E16/E17/E18 presumably results from confusion of Pao in Madhya Pradesh with the Tibeto-Burman language Pa'o (of the Karenic subgroup) in Myanmar, also rendered as Pao (Nishida 1966).

**E16** Dhanwar (India) [dha] is listed as the separate language of the scheduled tribe Dhanwar in Madhya Pradesh and Maharashtra (not to be confused with the Dhanwar Rai [dhw] language much further north, see Bhandari 2003). However, the Dhanwar speak Chattisgarhi [hne] in Madhya Pradesh and Marathi [mar] in Maharashtra (Singh 1997:47, Singh 1994:238-241), both of which already have entries in E16.

**E16/E17/E18** Warduji [wrd] is listed as an unclassified, unmapped language of Afghanistan. Although unmapped, the location is stated in text as “Werdoge River area west of Ishkashim”. The only (non-Persian) language spoken at the Werdoge (also rendered as Warduji) river is Sanglichki (Yusufbekov & Dodykhudoeva 2008, Payne 1989:419, Beck 2012) – a dialect of the Sanglechi-Ishkashimi-Zebaki language – which is already counted in the Sanglechi-Ishkashimi [sgl] entry.

**E16/E17/E18** E16/E17/E18 has an entry Jakati [jat] which is said to be spoken by 29,300 people in Ukraine. The alternative names, which include ‘Jat’, the classification of the language as Indo-Aryan, and a note indicating ‘nomadic’ suggests that the denotation is an itinerant population with roots on the Indian subcontinent, i.e., ‘Gypsy’ in loose terminology. 29,300 is a plausible number of Gypsies, or Roma, in Ukraine related to the Roma in countries to the west, but these Roma speak and indentify as a variety of Vlax [rmv] (Barannikov 1934:24-44, Marushiakova & Popov 2014). Thus their speech is already covered by the Vlax [rmv] entry, and, furthermore, they do not call themselves ‘Jat’ and are not called ‘Jat’ by others. Although they could not have reached Ukraine in such numbers as per the E16/E17/E18 entry, several ethnically and linguistically different communities in Afghanistan are called ‘Jat’ (Rao 1981, 1982, 1995), and the entry may be a confused attempt to render one of them. At least two mother tongues (as opposed to speech registers) are securely attested among the Afghanistan ‘Jat’s, both of which are missing from E16/E17/E18. One may be called Afghanistan Ghorbat (Rao 1995:74-81) and has a Persian grammar and a vocabulary from Persian, manipulated Persian, Indo-Aryan and other sources. The second may be called Ink [Rao 1995:82-85] and is a language of the Lahnda group of Indo-Aryan, perhaps closest to E16/E17/E18’s Siraiki [skr] (Kieffer 1983). Thus, the E16/E17/E18 Jakati [jat] entry is either spurious or erroneous (as to location and population) and ambiguous between the two Afghani Jat languages.

**E16/E17** E16/E17 lists Malakhel [mkl] as a distinct unclassified, unmapped language of Afghanistan distinct from Pashto, Dari and other languages of Afghanistan. Although unmapped, the location is given in text form as “Southwest of Kabul, Logar north of Baraki”. Thanks to the survey work of Kieffer (1989), this location (and the fact that
it is not supposed to be Pashto etc) can only match Ormuri [oru], which already has an entry. The name Malakhel ends in -khel, revealing that it is a clan name.

**E16/E17/E18** E16/E17/E18 lists Bhatola [btl] as an unclassified, unmapped language of Madhya Pradesh, India. It derives from a name in a Madhya Pradesh Census Report and is spurious as a separate language (Shukla 1986).

**E16/E17/E18** lists Lang’e [yne] as a separate language of Yongsheng county in Yunnan province, but this language is intelligible with Talu [yta] (Lama 2012:80) which already has a separate entry. There is possibly confusion with the Chinese variety of Lhao Vo [mhx] which is also called Lang’e and also in Yunnan province, but spoken closer to the border with Myanmar (Dai 1999, Mann 1998), in which case the entry is also spurious, duplicating Lhao Vo [mhx].

**E16/E17/E18** has an entry Khua [xhv] as a separate language of Laos-Vietnam. But the Khua in question speak a dialect of what E16/E17/E18 calls Eastern Bru [bru] (Vargyas 2000). The entry presumably derives from Wurm & Hattori (1981:35) where two other Bru varieties, Mangkon and Tri, are also listed separately.

**E16/E17/E18** has an entry Khlor [llo] as a separate language of Laos. But the Khlor speak a variety mutually intelligible with Ngeq [ngt] (Schliesinger 2003b:144, Bradley 2007b:316) and is thus spurious as a separate entry.

**E16/E17/E18** has an entry Tareng [tgr] as a separate Katuic language of Laos, noting that it must be different from Talieng [tdf], located to the northwest, because Talieng [tdf] is Bahnaric. The Tareng [tgr] entry presumably derives from (Wurm & Hattori 1981:35) where it is indeed listed as Katuic. But there is only one (geographically discontinuous) language in Laos matching the geographic and other data in the two Tareng [tgr] and Talieng [tdf] entries (L-Thongkum 2002:104-105, Schliesinger 2003b:47-49) and this language is Bahnaric (L-Thongkum 2002:104-105, cf. Sidwell 2015a:185), not Katuic. Presumably the existence of two entries rather than one comes from the Katuic/Bahnaric confusion. The Stieng [stg] entry of Vietnam may also represent the same language, though there is insufficient data available to me to validate one claim or the other (Sidwell 2009:83, Schliesinger 2003b:47-49).

**E16/E17/E18** E16/E17/E18 has a separate entry Tai Pao [tpo] in Laos. But this language is the same as that of the Tai Hang Tong [thc] in Vietnam, also named Tai Muong (Louppe 1934).

**E16/E17/E18** E16/E17/E18 has an entry Ahirani [ahr]. But Ahirani is simply another Khandeshi variety (Grierson 1907:203, Kulkarni 1948, Varma 1978b) already subsumed in the Khandesi [khn] entry.

**E16/E17/E18** Kang [kyp], alternative name Tai Khang, is listed as a Tai-Kadai language of Laos in the region bordering China but is not shown own the Laos map. The exhaustive survey of Tai peoples of Laos by Schliesinger (2003c) finds no Tai-Kadai language matching this name and location. The Tai-Kadai languages matching the
vaguely specified location are already counted in other entries in E16/E17/E18 and the Tay Khang [knu] language is spoken much too far south (in Bolikhamsai) and by a much smaller population (Schliesinger 2003c). Kang [kyp] is said to have have 47 600 speakers in Laos and thus could hardly have escaped survey. The entry is presumably a confusion of an existing language and the Tai-Kadai Tay Khang [tnu] language.

E16/E17/E18 Norra [nrr], is listed as a Nungish language of Northern Myanmar (no more precise location) with the dialects Nora, Byabe, Kizolo. The information almost certainly derives from Voegelin & Voegelin (1965:15), but Voegelin & Voegelin (1965:15) list Nora, Byabe, Kizolo as dialects of Nung (which has an entry Nung [nun]). Indeed, fieldwork confirms Nung in China and Nung across the border in Myanmar are mutually intelligible (Sun & Liu 2009:8).

E16/E17/E18 Mina (India) [myi] is a caste whose members speak Dhundari [dhd] (Samantha 1998) as the Mina [myi] entry itself declares.

E16/E17/E18 Pu Ko [puk] is listed as a SW Tai language in Laos without a speaker estimate or a location. No such language or ethnicity could be found in the comprehensive ethnic survey of Laos (Schliesinger 2003c:287).

E16/E17/E18 E16/E17/E18 has two entries Munda [unx] and Mundari [unr]. They are part of the same language (Cook 1965, Grierson 1906, Osada & Onishi 2012).

E16/E17/E18 E16/E17/E18 has two entries for the ethnic group Khalaj.

| code:    | klj | klf |
| name:    | Khalaj, Turkic | Khalaj |
| Region:  | Central Province, northeast of Arak. | Also in Iran. |
| Classification: | Altaic, Turkic, Southern, Azerbaijani | Indo-European, Indo-Iranian, Iranian, Western, Northwestern |
| country: | Iran | Azerbaijan, Iran |
| Language use: | All ages. Most also use Farsi [pes]. |

One entry is for the Turkic Khalaj [klj] language – well known through Doerfer (1988)’s publications – which says they also speak Farsi. The other entry Khalaj [klf] (which has the same population figure and refers to the same ethnic group) is listed as a separate Iranian language. But in addition to their Turkic language, the Khalaj speak only Farsi [pes], which already has an entry.

E16/E17/E18 E16/E17/E18 lists four entries for varieties of the dialect complex also known as Lua’/Mal/T’in/Pray (on the Border of Thailand and Laos) classified as follows:

- Mal-Phrai (3)
They are given overlapping names, speaker numbers, classification and locations on map, in a way that cannot be reconciled with detailed surveys of the Mal-Prai area (Filbeck 1978, Jordan-Diller & Diller 2004, Rischel 2007). We suggest removing Pray 3 [pry], which is the entry whose location is the least well-defined and whose existence it least motivated by detailed surveys (Filbeck 1978, Jordan-Diller & Diller 2004, Rischel 2007).

E16/E17/E18 Parsi [prp] is listed as an Iranian language of the Parsis of (predominantly) Gujarat and Maharashtra in India. These Parsis do not retain an Iranian language (Palsetia 2001)\(^ {17} \), but speak a variety of Gujarati with some Iranian lexical material not otherwise present in Gujarati (Gajendragadkar 1974, Modi 2011). Parsi is accurately listed as a dialect of Gujarati [guj], leaving the Parsi [prp] entry spurious.

E16 The Tangshewi [tnf] entry carries minimal information but is listed as an unclassified Iranian languages and is located on the map in the NE Afghani Badakhshan. The name signifies it is spoken at least in Darrah-e Tang Shew. All other sources, e.g., Dodykhudoeva (1997), Beck (2013:239) place Roshani (of the Shughni-Roshani group [sgh], see Edelman & Dodykhudoeva 2009a,b) here, and Morgenstierne (1928) has Shewa as a Shugni group village, though not impossibly a Tajik dialect could also be spoken there. In either case, the Tangshewi [tnf] entry arguably duplicates a language already included in E16, either a Tajik dialect such as Darwazi [drw] or a Shughni Pamir variety [sgh].

E16/E17/E18 E16/E17/E18 has an entry Shuadit [sdt] for an extinct allegedly Jewish variety of Provençal. But Provençal spoken by Jews has been investigated and found to be indistinguishable from non-Jewish varieties of Provençal (Banitt 1963, Pansier 1925, Guttel & Aslanov 2006:560) which already have entries in E17/E18.

E16 E16 has an entry Mahei [mja], but this language is the same as that known as Baihong in the entry for Honi [how] (Madrolle 1908, Bradley 1987:84).

E16/E17/E18 E16/E17/E18 has two entries Chetco [ctc] and Tolowa [tol] following earlier listings. Thanks to the work of Bommelyn (1997) we now know that the two are indistinguishable as languages, making one of the entries spurious.

E16/E17/E18 E16/E17/E18 lists Southern Lolopo [ysp] as a separate Central Yi language in the Jingdong, Jinggu, Lancang, Zhenyuan, Simao and Pu’er counties of Yunnan

\(^ {17} \)A distinct Iranian ancestral language does not appear to be attested from earlier times either (Gajendragadkar 1974, Grierson 1916).
province, China. The (rather vague) location and classification matches the locations (and classification) described for Lolopo [ycl] and Miqie [yiq] in other sources (Bradley 2007a, Matisoff 1996, Bradley 2002:104) and the Southern Lolopo [ysp] entry is quite possibly a confusion.

**E16** E16 lists Ayi [ayx] as an unclassified Tibeto-Burman language of in the counties Fugong and Gongshan of Northwest Yunnan province, China. Some grammatical information is also given. The location and grammatical information matches Nung [nun] exactly, and apart from Nung there is no other non-Lisu language spoken at this location (Sun & Liu 2009). Therefore the Ayi [ayx] entry is probably a duplication of the Nung [nun] entry.

**E16/E17/E18** E16/E17/E18 has a Mon-Khmer, Unclassified entry Kuanhua [xnh] with alternative name Damai, located in Southwest Yunnan Province, Xishuangbanna Dai Autonomous Prefecture, Jinghong County of China, with the comment that they are “locally considered part of Khmu” [kjg]. This entry does not seem to match any language known in the literature, although there are many languages in Yunnan and the entry has too little information to decide this properly. One possibility is that Damai is rendering of *tnɔɔi* ‘subgroup’ and that Kuan is a rendering of *kwen* (an ethronym) that passed through a Chinese transcription. If the Khmu *kwen* on the Chinese side of the border were asked are who they are, a plausible response would be *tnɔɔi kwen* ‘the kwen subgroup’. I owe this ingenious suggestion to Jan-Olof Svanesson (who himself calls it a 'wild guess', p.c. 2011). If the guess is correct, the Kuanhua [xnh] entry is just Khmu [kjg] spoken on the Chinese side of the border. If not, Kuanhua [xnh] entry should be removed on the grounds of insufficient information to decide if it is the same as an otherwise existing entry.

**E16** E16 lists Palu [pbz] as an unclassified Tibeto-Burman language of Myanmar without giving a location. Since the name is not known in the literature, the entry should be removed on the grounds of insufficient information to decide if it is the same as an otherwise existing entry.

**E16/E17/E18** E16/E17/E18 lists Lui [lba] as an unclassified Tibeto-Burman language of Myanmar (E16, 200 speakers) or of India (E17/E18, extinct) without giving a location. Lui is a derogatory term meaning ‘serf’ and has been applied to several different languages (McCulloch 1859, Luce 1985:36-46). Without a more precise denotation, such an entry is spurious.

**E16/E17/E18** E16/E17/E18 has an entry Lopi [lov] as an unclassified Loloish/Ngwi language of Myanmar (E16)/China (E17/E18). The only unique reference to such an ethnic group is the mention by Madrolle (1908:544) (as Lo-pì, No-pì or Nou-pì). It is perfectly possible that Madrolle (1908:544) means that the Lo-pì speak a variety of the language in E16/E17/E18’s Honi [how] entry, and there is nothing to assert that the Lopi speak a language mutually unintelligible to Honi [how] and other Hanic languages.

**E16/E17/E18** E16/E17/E18 has an entry Laopang [lbg] as an unclassified Loloish/Ngwi language of Myanmar (E16)/Laos (E17/E18) with the alternate name Laopa. There
appears to be only one language in the literature that can match these characteristics, namely Xa-Pho of Abadie (1924:185-186) who testifies they call themselves Lao-pa. However, this language already has an entry in E16/E17/E18 as Laghnu [lgh] (cf. Edmondson & Ziwo 1999).

**E16/E17** E16/E17 lists Adap [adp] as a separate language but the area in which it is spoken falls under Dzongkha [dzo] in van Driem (1998) and the lexicostatistical similarity is consistent with Adap being simply a dialect of Dzongkha [dzo].


**E16/E17** Yos [yos] is just another name for Zou [zom] (Matisoff 1996:121) and is thus spurious.

**E16/E17/E18** Lumba-Yakkha [hmu], Phangduwali [phw], Lambichhong [lmh] are entries from a misinterpretation of the questionnaires used in the Linguistic Survey of Nepal 1981-1984 and should simply be filed under Yakkha (van Driem 2001a:682-683).

**E16/E17/E18** E16/E17 has three entries Northwestern Tamang [tmk], Southwestern Tamang [tsf] and Western Tamang [tdg], while E18 retains two of them, removing Southwestern Tamang [tsf]. The Western Tamang entry, according to its description, subsumes both NW and SW and indicates that they are close enough to be mutually intelligible (as do all other descriptions of the Tamang dialect situation, e.g., Höfer 1969, Honda 2007, Lipp 2014, Mazaudon 2003), so the NW and SW Tamang entries are spurious. A similar conclusion is reached for the areas covered by the survey of Lipp (2014).

**E16/E17** Lingkhim [lii], is listed as a separate Western Kiranti language, but the Saam [raq] entry already represents the Western Kiranti language of Lingkhim (a place name, see van Driem 2001a:623, 704-705), so the Lingkhim (or the Saam) entry is spurious.

**E16** Yangbye [ybd], listed as a separate language of Myanmar in the Arakan region, but not further located on the map. It is simply Arakanese, i.e., Rakhine [rki] (Bradley 1995).

**E16** Chaungtha [ccq], listed as a separate language of Myanmar, whose location is not specified at all. It is simply Arakanese of Sit-Twe, i.e., Rakhine [rki] (Bradley 1995).


**E16/E17/E18** The Thu Lao [tyl] language is the same as the languages in the Dai Zhuang entry [zhd] (Kosaka 1997) but spoken in Vietnam.

**E16/E17/E18** E16/E17/E18 has two entries Gowli [gok] and Gowlan [goj], with the following information:
Classification: 

<table>
<thead>
<tr>
<th>Gowli [gok]</th>
<th>Gowlan [goj]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indo-Aryan, Central zone, Western Hindi, Unclassified</td>
<td>Indo-Aryan, Southern zone, Unclassified</td>
</tr>
</tbody>
</table>

Region

| Madhya Pradesh; Maharashtra, Amravati District | Maharashtra, Amravati District |

E16 Comment: Spearkers belong to Gowli caste. Surrounded by Korku. Belong to Gowli caste.

The information in (both) these can only match one and the same caste, i.e., the Gowli (also Gaoli/Gowli/Gowlan) caste (Russell & Hiralal 1916e), so at least one of the entries is spurious, and, on purely linguistic grounds, both are spurious, since the Gowli do not have a language that is mutually unintelligible with all surrounding languages (Grierson 1916:554-555, Russell & Hiralal 1916e).

E16/E17/E18 Bhalay [bhx] is listed as a separate language. However, Bhalay is a caste rather than an ethnic group with a particular language, and if they speak a language separate from their neighbours, it is Nimadi [noe] (Russell & Hiralal 1916c:106), which already has an E16/E17/E18 entry.

E16/E17/E18 Ethnologue lists Khuen [khf] as a separate language from Khmu [kjg], but the only difference is that the Khuen have adopted Buddhism (Izikowitz 1951:24) and language-wise they are not appreciably different from the Yuan dialect of Khmu (p.c. Jan-Olof Svan tesson 2011 whom I wish to thank for further discussion on this point).

E16/E17/E18 E16/E17/E18 has a Southwestern Nisu [nsv] as well as a Southern Nisu [nsd] said to be spoken in the same (or more) locations and with the same dialect names. The map shows S Nisu [nsd] to be spoken in the region described for both (Pu’er subprefecture), whereas SW Nisu is put much further southwest in Xishuangbanna (Dai) Prefecture. SW Nisu, as a distinct variety spoken in Pu’er or in Xishuangbanna prefecture, is not taken up in Nisu dialect surveys (Chan et al. 2008, Yang 2009). The SW Nisu entry is therefore probably spurious, or seriously confused.

E16 Chin, Khumi Awa [cka] is listed as a separate entry, but this language is the same as the language of the Chin, Mro [cmr] entry (Hartmann 2001, On Pe 1933).

E16 Pongyong [pgy] is an entry deriving from a misinterpretation of the questionnaires used in the the Linguistic Survey of Nepal 1981-1984, and is simply a clan name (van Driem 2001a:623).

E16/E17/E18 E16/E17/E18 has an entry Kayort [kyv], a language related to Bengali in Morang district of Eastern Nepal. The Rangeli-Rajbanshi [rjs] data from Toulemin (2006) covers this area (cf. van Driem 2001a:1146) and there is no no other Bengali related lect spoken there (cf. Gautam & Thapa-Magar 1994:176, Sanyal 1965). Thus the Kayort [kyv] entry is to be equated with Rajbanshi [rjs].
**E16/E17/E18** Lama (Myanmar) [lay] is a name for a section of Nung [nun] of Nungish branch of Tibeto-Burman (Voegelin & Voegelin 1965:15).

**E16/E17/E18** A superbly detailed investigation (Yang 2010) of the Lalo languages (Loloish) leaves Eastern Lalu [yit] a spurious entry.

**E16/E17/E18** The Degaru [dgu] entry (as an Indo-European-Indo-Aryan Eastern zone Unclassified language) can only match the caste Dhekaru who are not known to have a separate language (Mitra 1953).

**E16** The Sansu [sca] entry represents the same language as Hlersu [hle] in China, only that it is spoken across the border in Burma (Bradley 2007a:398).

**E16/E17** The Lamam [lmm] entry (a Bahmaric language of Cambodia) is the same language as Kaco’ [xkk] across the border in Vietnam (Bradley 2007a:413).

**E16/E17/E18** The Kiorr [xko] (with alternative name ‘Con’) and Con [cono] refer to the same language (Schliesinger 2003a).

**E16/E17/E18** The Inpui Naga [nkf] entry, also known as Kabui, is another name for a variety in the Naga-Rongmei [nbu] entry (Marrison 1967:II:361-362, Grierson 1903).

**E16/E17/E18** The entry Puimei Naga [npu], as described in E16, refers to the same language as the Naga-Rongmei [nbu] entry (Kabui 2004). There may be differences within the Naga-Rongmei [nbu] entry that justify dividing it, but the division should, in that case, not be Puimei vs Inpui vs Naga-Rongmei.

**E16/E17/E18** E16/E17 has two entries in its Northern Kuki-Chin subgroup Purum Naga [puz] (in India) and Purum [pub] (across the border in Burma). Not surprisingly, they are the same language and were recognized as such (Grierson 1904:263) before the British empire was split up.

**E16/E17/E18** The Loarki [lrk] duplicates the (more informative) Gade Lohar [gda] entry and is thus spurious (Sharma 1988).

**E16/E17/E18** The Chamari [cdg] entry describes the Chamār caste, but the Chamari do not have a language of their own (Russell & Hiralal 1916d).

**E16/E17/E18** The Savara [svr] entry listed as a Dravidian language of the Telugu group, separate from the well-known Munda Sora (also known as Savara) [srb] language. It has been checked quite carefully that there is no Dravidian language matching the name Savara or any of the other information in the entry (p.c. David Stampe 2011)\(^{18}\), nor, for that matter, the Indo-Aryan Oriya variety labeled Sahara/Soara in Mahapatra (2002:183-184), but there is a plausible way in which the erroneous entry may have entered E16/E17/E18. Dalton (1872) includes a Savara language which he (mistakenly) lists as a Dravidian language. The vocabulary included shows this Savara to be the Munda Sora [srb], not a different language. Presumably E16/E17/E18 is citing (some work that is still citing) Dalton’s listing of Savara as a Dravidian language.

---

\(^{18}\)I wish to thank David Stampe and John Peterson for discussing this entry with me.
Agaria [agi] is listed as a Munda language. First of all, the entry conflates information from different tribes with a similar sounding name (Crooke 1896). Second there is no attested Agaria language (Elwin 1942:xxvii) so any suspicions that they spoke a Gond (Dravidian!) language (Dimaggio 1991, Russell & Hiralal 1916a) is undemonstrable. E16/E17/E18 lists Agoria as a dialect name which seems to explain how the erroneous Munda affiliation came into place. It seems that E16/E17/E18 confused Agaria with “Aghoria”, an alternative name for the neighbouring group Asur who, like the Agaria, are iron-smelters (Dalton 1872, Elwin 1942:xxvii). Asuri [asr], which is an attested Munda language (Konow 1906), is already listed in E16/E17/E18. This makes Agaria [agi] a spurious entry in any corrected reading of it since Asur is already listed in the E16/E17/E18 and the assumed Dravidian-descendent Agaria language is not attested.

The entry Chuanqiandian Cluster Miao [cqd] entry subsumes varieties already included in E16/E17/E18 under different entries (Niederer 1998). As a macro-language the [cqd] makes sense, but not as a distinct language in addition to the Miao languages already having entries.

E16/E17/E18 has an entry Kannada Kurumba [kfi] which is not mapped but its name, alternative names, region and comments indicate it is meant to subsume Kurumba groups who speak varieties close to Kannada, e.g., Alu, Jennu, and Betta Kurumba as well as Palu Kurumba who speak a variety closer to Tamil (Menon 1996, Zvelebil 1981). All of these Kurumba varieties already have entries in Alu Kurumba [xua], Jennu Kurumba [xuj], Betta Kurumba [xub], and Palu Kurumba = Attapaddy Kurumba [pkr]. The large number of speakers indicated, as well as the name, is consistent with the entry referring to the large community of shepherds in Tamil Nadu (Aiyappan 1938:125-128) also called Kurumba. However, their language is a dialect of Kannada and thus covered in the Kannada [kan] entry (Varma 1978a).

## B.4 Spurious Languages: South America

E16/E17/E18 has an entry Tapeba [tbb] representing people of Ceará state in Brazil with indigenous background and a common identity. However, the common identity is recently formed and does not correspond to a past or present language, as the ancestors of the present Tapeba spoke languages such as Kariri, Tremembe and likely others (Barretto Filho 1989, Studart 1926) which either have entries of their own or are too poorly known to show if they were separate languages. Since E16/E17/E18 is a catalogue of languages, a Tapeba [tbb] entry is spurious.

E16/E17/E18 lists two languages Kanamari [knn] and Katukina [kay] as per the below.
<table>
<thead>
<tr>
<th><strong>name</strong></th>
<th>Kanamarí</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Population</strong></td>
<td>1,330 (Moore 2006), 100 Tsohom-Djapa speakers (Crevels 2007)</td>
</tr>
<tr>
<td><strong>Ethnic population</strong></td>
<td>2,750 (2006 FUNASA).</td>
</tr>
<tr>
<td><strong>Location</strong></td>
<td>Amazonas, upper regions of Jurua, Jutai, Itaquai rivers.</td>
</tr>
<tr>
<td><strong>Language Use</strong></td>
<td>Vigorous.</td>
</tr>
<tr>
<td><strong>Language Status</strong></td>
<td>6a (Vigorous).</td>
</tr>
<tr>
<td><strong>Language Maps</strong></td>
<td>Western Central Brazil</td>
</tr>
<tr>
<td><strong>ISO 639-3</strong></td>
<td>knm</td>
</tr>
<tr>
<td><strong>Dialects</strong></td>
<td>Tshom-Djapa (Txunhuá Dyapá, Txunhuá-Djapá), Tsohon-Djapa.</td>
</tr>
<tr>
<td><strong>Country</strong></td>
<td>Brazil</td>
</tr>
<tr>
<td><strong>Classification</strong></td>
<td>Maipurean, Southern, Southern Outlier, Piro</td>
</tr>
<tr>
<td><strong>Alternate Names</strong></td>
<td>Canamarí, Kanamaré</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>name</strong></th>
<th>Katukína</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Other Comments</strong></td>
<td>Different from Panoan Katukína [knt] in Amazonas and Acre.</td>
</tr>
<tr>
<td><strong>Location</strong></td>
<td>Acre.</td>
</tr>
<tr>
<td><strong>Language Status</strong></td>
<td>9 (Dormant).</td>
</tr>
<tr>
<td><strong>ISO 639-3</strong></td>
<td>kav</td>
</tr>
<tr>
<td><strong>Dialects</strong></td>
<td>Cutiadapa (Kutia-Dyapa).</td>
</tr>
<tr>
<td><strong>Country</strong></td>
<td>Brazil</td>
</tr>
<tr>
<td><strong>Classification</strong></td>
<td>Katukinan</td>
</tr>
<tr>
<td><strong>Alternate Names</strong></td>
<td>Catuquina, Katukína do Jutaí, Katukína do Rio Biá, Pidá-Djapá</td>
</tr>
</tbody>
</table>

The language matching the name, location, population and dialect information of the Kanamari [knm] entry, is well-known from the linguistic (e.g., Anonby & Holbrook 2010, Carvalho 1955, de Magalhães 2012, Groth 1985, Queixalós 2007, Queixalós & dos Anjos G.S. 2007) and ethnographic (e.g., Melatti 1981, Verneau 1921) literature. But this language is not, as the E16/E17/E18 classification has it, a Maipurean Arawakan language closest to Piro (Yine) [yib]. The idea that it is closely related to Piro ultimately stems from a name confusion with the Canamaré vocabulary in von Martius (1867b). This Canamaré vocabulary is indeed so close to Piro as to count as Piro, but it is not the same language the Kanamari indicated by the data in the Kanamari [knm] entry, as shown already by Rivet (1920). Turning now to the Katukína [kav]-entry, its location, classification, alternative names and dialect names (but not speaker number) corresponds to the Katukína known in the literature (e.g., dos Anjos 2005). However, this Katukína [kav] language is mutually intelligible with Kanamari [knm] (dos Anjos 2011:8-16) and the two should count uncontroversially as one entry. Older vocabularies are also similar enough to count as the same language (Loukotka 1963, Rivet 1920) so there is no reason to posit a separate entry for a Katukína that existed in the past. Thus, one of the Kanamari [knm] and Katukína [kav] entries is spurious. The confusion in this case may have been licensed by the existence of yet more distinct languages surfacing under the name Katukína, i.e., the Panoan Katukína (de Aguiar 1992) and the problematic Catuquinarú vocabulary (Rivet 1920).

**E16/E17/E18** E16/E17/E18 has an entry for Pokangá [pok] as an Eastern Tucanoan language of the Upper Tiquié river in northwest Brazil. The entry contains the comment
that “The people claim to speak Waimaha (Northern Barasano) [bao].” The Waimaha [bao] entry indeed has the Upper Tiquié river included in the list of locations. This area has been surveyed many times and there is no other Tucano language spoken there (Giaccone 1949, Jackson 1983, Koch-Grünberg 1912, 1913, 1914, 1914, 1915-1916, 1915-1916) except Tucano [tuo] and Waimaha [bao] (also known as Bará), and the non-Tucano languages of the hunting-gathering peoples living between the larger rivers also have entries. The Pokangá [pok] entry is spurious, and should be merged with the Waimaha [bao] entry as the information in both already indicates.

**E16/E17/E18** Iapama [iap] is listed as an unclassified language in the Pará and Amapá border region of Brazil. It is located on the map but commented as “existence uncertain”. The location on the map, near the Jari river between Apalai [apy] and Wayampi [oym], can only match one known uncontacted group, i.e., number XXXIII of Grenand & Grenand (1994:111), “Rio Ipitinga” of Brackelaire & Azanha (2006:322) or “grupo arredio do alto Ipitinga” of (Gallois & Ricardo 1983:102). The little available evidence points to this group being an offshoot Wayampi [oym] group, though not impossibly an Apalai [apy] group. In either case, no separate language can be asserted, so the entry is spurious.

**E16/E17/E18** Karahawyana [xkh] is not attested, and the only pieces of information on their language, several of which are independent, declare it a dialect of Waiwai [waw] (Frikel 1957:59, Frikel 1958:141, FUNAI 1984:59, Howard 2001a).

**E16/E17/E18** Agavotaguerra [avo] is not attested, and the only information available are reports by the Kuikuru, who say the Agavotaguerra speak the same language as the Yawalapiti [yaw] (Galvão & Simões 1965:23, Boas 1975, Villas Boas & Villas Boas 1972:45).

**E16/E17/E18** Tetete [teb] is unattested, but second hand reports testify it was intelligible with Siona [snn] (Cabodevilla 2007).

**E16/E17** Kakauhua [kbf] is listed as an extinct Alacalufan language. Caucahue is an ethnonym from colonial sources referring to non-Mapuche (Mapuche [arn], Huiliche [huh]) giant canoemen in the area of the Chonos [-]. There is no linguistic data attributable to Caucahue so the inference that Caucahue is a separate language related to Qawesqar [alc] is wholly illegitimate – if anything, Caucahue should be equated with Chono on the ethnographic and geographic evidence (Viegas Barros 2005:81).

**E16/E17/E18** E16/E17/E18 lists an unclassified language Yarí [yri] on the Yarí river. The given map location coincides with the Chiribiquete national park which is traditionally Carijona [cbd] territory (Friede 1948). Uncontacted peoples have been sighted in this area in the past (Grenand & Grenand 1994), including a Murui Huitoto [huu] group who were there until 1959. Presently, no uncontacted peoples are known to inhabit the area (Brackelaire & Azanha 2006, Franco 2002) and if there were, these would be Murui Witoto [huu] or Carijona [cbd]-speaking (Franco 2002) – languages which
already have entries in E16/E17/E1819.

**E16/E17/E18** Cauca [cca] is listed as an extinct Chocoan language of Colombia. No location is given but from the name it may be inferred it is supposed to have been spoken in the Cauca river valley. There is no attested language in the Cauca valley not already counted in other entries in E16/E17/E18 (Loukotka 1968:246-249, 255-258) except Quimbaya which is attested in about 10 words. But there are doubts as to whether 8 of these two words are really attributable to the language of the Quimbaya ethnic group (Adelaar & Muysken 2004:57). In any case, there is too little data to assert that the Quimbaya spoke a language different from all its neighbours.

**E16/E17/E18** Runa [rna] is listed as an extinct Chocoan language of Colombia without a more precise location. No separate Chocoan language with that name is attested (Landaburu 2000, Loukotka 1968, Ortiz 1965).

**E16/E17/E18** Kamba [xba] is listed as an extinct unclassified language of Mato Grosso do Sul in Brazil with the remark that it “May have been Tupí”. In fact, the Kamba identify as Chiquitano [cax] and the language is attested and is clearly Chiquitano [cax] (Paula 1979).

**E16/E17/E18** Xipináwa [xip] is unattested (Loukotka 1968:172) and cannot be asserted to have been a separate language.

**E16/E17/E18** Kapinawá [xpn] is unattested (Sampaio 1995) and cannot be asserted to have been a separate language.

**E16/E17/E18** Pisabo [pig] is unattested (Fleck 2013) and cannot be asserted to have been a separate language.

**E16/E17/E18** Shinabo [snh] (of the Mamoré river) is unattested (Fleck 2013) and cannot be asserted to have been a separate language.

**E16/E17/E18** Tukumanféd [tkf] is a presumed Kawahib language earlier spoken on the middle Machado River, but is unattested (Loukotka 1950:24, Loukotka 1968:115) and cannot be asserted to have been different from other Kawahib languages.

**E16/E17/E18** E16/E17/E18 lists a Kabixí [xbx] entry with the following information

---

19I wish to thank Frank Seifart for researching this matter and for access to Franco (2002)’s thorough historical account.
Population 100 (1986 SIL).
Region Mato Grosso, Planalto dos Parecis slopes, right bank of upper Guaporé, near Vila Bela.
Language map Brazil, reference number 116
Alternate names Cabichí, Cabishi
Dialects Related to Maxakali [mbl].
Classification Chapacura-Wanham, Guapore
Comments The name is also used for Parecis [pab]
Language map Brazil, reference number 116
Language use Both people and language may be extinct.

The name Kabixi (with variant spellings) has been applied to a range of different ethnic groups (see Price 1983 for a superb historical treatment), including Parecis (Arawakan), Huanyam (Chapacuran) and Southern Nambiquara (Nambiquaran) speakers. The location in the entry can be read to match either the Southern Nambiquara and Sararé of Price (1985), both of which correspond to another E16/E17/E18 entry, i.e., Southern Nambikuara [nab]. The classification in the entry points to it being Huanyam (Nordenskiöld 1915), which would explain why Huanyam is otherwise missing from E16/E17/E18. The location matches Southern Nambiquara better but is reconcilable with Huanyam. However, a population of 100 in 1986 is too much if it refers to Huanyam which would have been down to a few families or less in the 1980s (Becker-Donner 1974, Hanke 1975). A population of 100 is difficult to match with 20 Sararé and 350 Southern Nambiquara (Price 1985) as well, so possibly 1986 should be read as the date of the information passing to the Ethnologue rather than the date of the observation. The claim that Kabixi [xbx] is related to Maxakali [mbl] cannot be reconciled with either Southern Nambiquara or Huanyam and can only make sense if the entry refers to Rikbaktsa [rkb] (Dornstauder 1960) or Beicos-de-Pau, now integrated with the Suyá [suy] (Vicente César 1971), but these peoples live too far north of the area where the name Kabixi used to be applied. Since the information in the entry either duplicates other entries or has misleading information, it should be dropped or corrected to form a more accurate Huanyam entry.

E16/E17/E18 E16/E17/E18 has two entries

<table>
<thead>
<tr>
<th>Alternate names</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mawayana, Maipitian, Mawayana</td>
<td>50 in Brazil (1986) 50 (1986 C. Howard)</td>
</tr>
</tbody>
</table>

Not surprisingly, the two entries refer to the same group (Howard 2001b), as comparison of data from both ethnonyms confirms (Carlin 2006, Farabee 1918).

E16/E17/E18 A major confusion in E16/E17/E18 is with the languages called Karipuna. There are two entries, both classified as Tupi-Guarani, Subgroup VI:

20 At least the assertion that the entry does not represent Parecis [pab] (perhaps the language most frequently called Kabixi) is not misleading.
21 There is also Karipuna Creole French [kmv], an entry which is correct.
Karipúna [kgm] Karipuna [kuq]
Region: Amapá, on French Guiana border Rondônia, Acre, banks of Jaru, Jamery, Urupa, Cabecciras, Candéias, and Jaciparaná rivers.

Classification: Tupi, Tupi-Guarani, Subgroup VI Tupi, Tupi-Guarani, Subgroup VI
Alternate names: Karipúna do Amapá, Karipúna do Uaçá Caripuna, Jau-Navo, Juanauo, Kagwahiva, Karipuná de Rondônia, Karipuná do Guaporé

Language use: Shifted to Karipúna Creole French [kmv]. Some may also use Tenharim [pah]

Early references to the Karipuna (or similar names) of Amapá refer to a variety of Palikúr [plu] (Gallois & Ricardo 1983:62). The language of the Karipuna of Amapá in this sense is attested with a short wordlist, and is indeed a Palikúr [plu] variety (Loukotka 1963). More recent references to the Karipuna of Amapá refer to an immigrant group whose original language is not attested, but, is reported to have been Nheengatu [yrl] (Arnaud 1969, Gallois & Ricardo 1983:62). Except for the classification and the fact that Nheengatu [yrl] already has an entry, the Karipúna [kgm] entry is consistent with referring to the more recent Karipuna of Amapá. Thus, whatever it refers to, the same language already has a different E16/E17/E18 entry, either Palikur [plu] or Nheengatu [yrl].

Until the 1950s (Leonel 1995:40-45, 199-213) the only language called Karipuna in the Madeira-Guaporé region was a Panoan language also known as Jau-Navo, amply attested in wordlists (de Créqui-Montfort & Rivet 1913b, Hanke 1949, Rondon & de Faria 1948, von Martius 1867b). However, the name started to be applied to groups of hostile Indians at the Mutumparaná and Jaciparaná which were contacted in the 1950-1970s (Kästner 2005:104) and turned out to speak a Tupi-Guarani language of Subgroup VI (Betts 2012, Mello 2000, Monserrat 2000), which we may call Kawahib. FUNAI frivolously used the name Karipuna for this Kawahib group (Leonel 1995:45), encouraging the confusion we now observe. The Karipuna [kuq] entry mixes information from the Panoan Karipuna (e.g., the name Jau-Navo) and the Kawahib Karipuna (e.g., the classification). The Panoan Karipuna is otherwise missing from E16/E17/E18 whereas the Kawahib Karipuna speak the same language as the language of several other entries (see below).

E16/E17/E18 Of the languages of Tupi-Guarani subgroup VII, there is considerable overdiffererntiation. Amundava [adw], Tenharim [pah], Uru-Eu-Wau-Wau [urz], Júma or Arara-Kawahib [jua] and Karipúna (if it refers to Kawahib-Karipuna) [kuq] all speak easily mutually intelligible varieties (Kästner 2005, Leonel 1995, Lévi-Strauss 1948b, Loukotka 1950, Sampaio 1997), as also acknowledged by the E16/E17/E18 entries
themselves whose regions, dialects and speaker numbers overlap. Paranawát [paf] is a name for a people of the same complex met with (and attested by wordlists) by the Rondon commission (Rondon & de Faria 1948). Morebê [xmo] is an uncontacted group of the Rio Preto and Marmelos, who speak another variety of the this language, if they still exist (Kästner 2005). Wirafêd [wir], also known as Tupi do Machado or Bocas Pretas, is amply attested in wordlists from the past century (Hanke 1953, Koch-Grünberg 1932, Loukotka 1963, Nimuendajú 1955a, Nimuendajú & do Valle Bentes 1923) and is hardly distinguishable from Apiaká [api] (Nimuendajú & do Valle Bentes 1923:222-223, Loukotka 1950), which also is so close that it could also count as the Kawahib-language.

**E16/E17/E18** Tupinikin [tpk] is listed as a separate entry, but, in fact, only three words are attested (Garcia Rosalba 1976) which is insufficient grounds for separating it from the other entries of Tupí-Guarani, Subgroup III (Loukotka 1968:103-105).

**E16/E17/E18** Four entries are described as extinct Arawakan, Unclassified languages of Colombia with no more precise location given: Cumeral [cum], Omejes [ome], Ponares [pod], Tomedes [toe]. They are not to be found in anywhere else in the literature (except a few recent publications which are derivative of earlier Ethnologue versions), including otherwise exhaustive literature on past languages of Colombia (Landaburu 2000, Loukotka 1968, Ortiz 1965). Even if one were to find referent populations for these names, the entries should be removed on the grounds of insufficient information to decide if it is the same as an otherwise existing entry.

**E16/E17/E18** Another three entries are described as extinct Unclassified languages of Colombia with no more precise location given: Coxima [kox], Chipiajes [cbe], Cagua [cbh]. They are not to be found in anywhere else in the literature (except a few recent publication which are derivative of earlier Ethnologue version), including otherwise exhaustive literature on past languages of Colombia (Landaburu 2000, Loukotka 1968, Ortiz 1965). Even if one were to find referent populations for these names, the entries should be removed on the grounds of insufficient information to decide if it is the same as an otherwise existing entry.

**E16/E17/E18** Yet another three entries describing extinct Unclassified Colombian languages in the Tolima region are Natagaimas [nts], Pijao [pij], Coyaima [coy]. These names are known in the literature, but Natagaimas and Coyaima are villages where Pijao was spoken in its final stages (Durbin & Seijas 1973b) so the three the entries should therefore be merged into one. This language is attested with a published minuscule wordlist.

**E16/E17/E18** Tingui-Boto [tgv] is listed as an unclassified language of Brazil. There is certainly a corresponding ethnic group, but this group speaks Portuguese. They retain some of their ancestral language for ritual purposes, but this language is Dzubukuá²².

²²According to Vera Lucia Calheiros Mata [http://pib.socioambiental.org/pt/povo/tingui-boto/1050 accessed 7 February 2012.}
Dzubukuá (de Queiroz 2012) is subsumed under the (otherwise highly problematic – see below) Kariri-Xoco [kzw] entry.

E16/E17/E18 Among its many entries in the Campa group, E16/E17/E18 lists Ashéninka, South Ucayali [cpy]. However, unlike the other Campan languages, it is not placed on the map, and in the region where the entry places it (Upper Ucayali River and tributaries), the survey by Heitzman (1973) has varieties that fall under other Campan entries in E16/E17/E18.

E16/E17/E18 E16/E17/E18 has an entry for Chilean Quechua [cqu] but this variety is an extension of the variety in the South Bolivian Quechua [quh] entry, and the nation border does not correspond to any linguistic division (p.c. Paul Heggarty 2011).

B.5 Spurious Languages: Pacific

E16/E17/E18 E16/E17/E18 has an entry for Laura [lur] as a separate language. But Laura is a “wenig verschiedene dialect van Wajewa” (van der Velde 1900:57), i.e., a dialect little different from Weewa [wew]. Laura [lur] is thus spurious as a separate entry, and, in any case, prohibitively inconsistent given that there are many more more divergent lects which do not have separate entries (Asplund 2010). Since the Laura [lur] entry mentions that it is “not intelligible with Kodi [kod]” it is possible that there is confusion with the nearby Gaura which is a very divergent dialect of Kodi [kod] which is otherwise not mentioned in any entry in E16/E17/E18 (for locations, names and data, see Asplund 2010 and Putra 2007).

E16/E17/E18 E16/E17/E18 has two entries Karon Dori [kgw] and Mai Brat [ayz]. The dialects represented in the two are mutually intelligible (Dol 2007:8) and may thus be merged.

E16 Talur [ilw] of Wetar island is listed as a separate entry from Galoli [gal] in Timor. However, Talur is simply Galoli spoken on Wetar, and the two are minimally different as per any simple data inspection (Hinton 2000, Hull 2003a, Stokhof 1981). They have been known to be the same in the literature at least since (Capell 1944, 1944, 1944:40-41) and have ended up as two separate entries in E16 presumably only due to belonging to two different political spheres.

E16/E17/E18 E16/E17/E18 has two entries Tobo [tbo] and Kube [kgf]. They are mutually intelligible and 95% lexicostatistically cognate (Lee 2014 [1993]:11), superseding the lower figure of an earlier survey (McElhanon 1967:7). Since they are mutually intelligible, the two entries may be merged.

E16/E17/E18 Maramba [myd] is listed as a separate entry presumably following Wurm (1982) where it first occurs as a separate language. Earlier surveys have Maramba as an Angoram [aog] speaking village (Laycock 1973). The reason Maramba appeared as a separate language in Wurm (1982) is as follows. William Foley was told by Angoram speakers in 1977 that the language of Maramba was different from their own language and reported this to Donald Laycock who, in turn, passed it on to Stephen
Wurm. On later trips by William Foley, other Angoram speakers have claimed that the Maramba village does speak their language. Unfortunately no researcher has studied the language(s?) at Maramba so we do not know what the situation actually is, but so far, there is no insufficient information to assert that there is a separate language at Maramba.

**E16**

E16 has two entries Piru [ppr] and Luhu [lcq], but it has been known for a long time that these are two mutually intelligible varieties (Collins 1983, Stresemann 1927).

**E16**

E16 has a separate entry for Horuru [hrr], but Horuru is known to fall under the language subsumed in the Yalahatan [jal] entry (Collins 1983).

**E16/E17**

have two entries Mea [meg] and Tiri-Mea [cir] noting that there is some comprehension between the two. But this has been investigated and reliable reports state the two to be mutually intelligible (Grace 1976, Haudricourt 1971, Osumi 1995).

**E16/E17/E18**

E16/E17/E18 has a separate entry for Laba [lau] or South Loloda. But according to van Baarda (1904) Laba differs little from Loloda of Baru of the Loloda [loa] entry and other surveys see no positive reason for separating the two (Grimes & Grimes 1984, Voorhoeve 1987).

**E16/E17/E18**

Anasi [bpq] is listed as a separate language on the left-hand side of the Lower Mamberamo, but the most detailed survey of this area (Jones 1987:3, cf. Detiger 1935) finds Nisa [njs] be spoken at Anasi. The Anasi entry was, in the first place, based on the incomplete information — just the name and location — in Galis (1955).

**E16/E17/E18**

E16/E17/E18 lists Seru [szd] as an extinct separate Sarawak-Melanau-Kajang language. However, it has been known for a century that the Sru were a fraction of the Ukit [umi] who happened to end up at Sru, and indeed, “every word…” in the wordlist of Bailey (1963 [1901]) were understood by the Ukit [umi] “as belonging to their own language” (De Rozario 1963 [1901]:332).

**E16/E17/E18**

Palumata [pmc] is not attested. If anything, the available non-linguistic evidence (Grimes 2000:99) implies it should be counted as Hukumina [huw].

**E16/E17/E18**

Balau [blg] is listed as a separate language from Iban [iba], but Balau is simply one of the places where Iban is spoken (Omar 1981:2) and the vocabulary of Balau shows no significant deviation from other varieties of Iban (Ray 1913).

**E16**

Vatrata [vhr] is listed as a separate language, but this is simply a village name of the language covered in the Vera’a [vra] entry (Lynch & Crowley 2001:40).

**E16/E17/E18**

E16/E17/E18 has two entries Lun Bawang [lnd] including the alternate name “Lundaya Putuk” and Putoh [put] with the remark that it “may be the same as Lun Bawang [lnd]”. Both the name Putuk and the area marked on the map for Putuk, falls under the same language that the entry Lun Bawang [lnd] represents. Though this language has a lot of dialectal variation, the division indicated by the two Ethnologue entries is not justified from a linguistic or ethnic perspective (Kaskija 2002:10-11, Crain 1970:17-35, Hudson 1977, Omar 1983:587-588, Sellato 2009).
E16 E16 has the entry Elpaputih [elp] which is not given a more precise location than West Seram. Elpaputih is actually the name of a bay around which the languages represented in the entries Nusa Laut [nul] and Amahai [amq] are spoken (Collins 1983, Ribbe 1903, Taber 1996). The Elpaputih [elp] entry is thus spurious.

E16/E17/E18 Yarsun [yrs], a Sarmi coast Oceanic language, was not found on survey 2005 (according to E16/E17/E18 itself) and its presumed existence seems to stem from the Index of Irian Jaya languages (Silzer & Heikkinen-Clouse 1991). The first mention of it is as an island (not singled out as a separate language) between Anus and Podena is by van der Leeden (1954) who considers Anus [auq], Podena [pdn] and Yarsun [yrs] one and the same language. Thus the Yarsun [yrs] entry is either the same as Anus [auq] or Podena [pdn] or unattested.

E16/E17/E18 Wares [wai] is listed as a separate Tor language, referring to the warlike tribe on the upper Biri river (Oosterwal 1961:26-27). The Wares themselves had to flee to the coast from their original territory in the 1950s, possibly returned (Oosterwal 1961:26-27) and have been heard of since (Koentjaraningrat 1977, Koentjaraningrat 1965:135-136), but no separate Wares language has been found in old or recent surveys (Grace 1956, Smits & Voorhoeve 1994, van der Leeden 1954, Wambaliu 2006) – in particular, the only language found at the village Wares, is that covered in the entry Mawes [mgk]. Since there is no data and no intelligibility information on the Wares language, we cannot assert that they spoke language different from all others (let alone a Tor language) even if the ethnic group was (and still may be) extant.

E16/E17/E18 The data on Kazukuru, an extinct language of New Georgia attested in wordlists, shows no more variation than expected if it was only one language (rather than three). Therefore, having three entries as in E16 Dororo [drr], Guliguli [gli] and Kazukuru [kzk] overproduces two entries (Dunn & Ross 2007).

E16/E17/E18 Adabe [adb], an alleged Papuan language of Atauro island, does not exist and it has been explained how it was mistakenly thought to exist (Hull 1998:3-4) viz. a misreading of an ambiguous statement by a missionary that made it into Wurm & Hattori (1981) from where the E16/E17/E18 entry derives from (Hull 2004:41).

E16/E17/E18 Ontena [ont] is given a separate entry from Gadsup [ga] and its assigned territory on the map pertains to Oyana [auy] on extant survey maps (Capell 1948-1949, McKaughan 1964). In fact, Ontena is place where a Gadsup offshoot lives (Radford 1987:9) and they speak a variety of Gadsup [ga] intelligible to other Gadsup dialects (Frantz 1994, Frantz 1976:75)\(^23\).

E16/E17/E18 The languages covered in the E16/E17/E18 entries Minanibai [mcv] and Foia Foia [ffi] are the same language (Usher & Suter 2015:16-17).

\(^23\)I wish to thank Tim Usher for discussion on this matter.
B.6 Spurious Languages: North America

**E16** E16 has an entry Chumash [chs] with information that could match several individual Chumashan languages. E16 also has entries for the individual Chumashan languages Obispeño [obi], Cruzeño [crz], Ventureño [veo], Purisimeño [puy], Ineseño [inz] and Barbareño [boi], which exhaust the attested Chumashan languages (Golla 2011). This leaves the Chumash [chs]-entry spurious.

**E16/E17/E18** Quetzaltepec Mixe [pxm] is given its own entry, but no positive evidence exists that this is not intelligible with other Midland Mixe varieties. While some distinctiveness as to this variety may be expected, it is arbitrary to pick this town over others in the area with the same expectation. In this sense the entry is spurious (p.c. Søren Wichmann 2011).

**E16/E17/E18** E16/E17/E18 has three entries Lushootseed [lut], Skagit [ska] and Snohomish [sno]. The Lushootseed name (and entry) is an umbrella concept, encompassing both Skagit [ska] and Snohomish [sno] which are mutually intelligible (Thompson & Kinkade 1990:38). Thus, Skagit [ska] and Snohomish [sno] have spurious entries given a Lushootseed [lut]-entry. If one were to separate Skagit [ska] and Snohomish [sno] despite the intelligibility, it would render the Lushootseed [lut]-entry spurious.

**E16** E16 has two entries Lumbee [lmz] and Carolina Algonquian [crr], both classified as Algonquian. However, there is only one securely attested language in that region which is Algonquian, and that is the one represented by the vocabularies in Geary (1955), Hariot et al. (1999 [1590]). E16 is either wrong in a) having two entries for the same language or b) wrong in calling Lumbee an Algonquian language (Lumbee, in modern times, can be understood to refer to a variety of English, see Wolfram & Dannenberg 1999). I wish to thank Ives Goddard for discussing this entry with me.

C Examples of Bookkeeping Errors in the E16 classification

There are two kinds of incongruencies between the E16/E17/E18 classification of languages into families and that of specialists on the respective languages/families. One category are errors of an elementary kind that are not controversial: bookkeeping, name confusion, misunderstanding of linguistic vs non-linguistic classification, not checking relevant research and not keeping up with relevant research. The second category are cases where expert opinions differ and a non-trivial amount of research would be required to assess the merits of each proposal. Below we list a number of specific cases of the first kind, taken from E16 (for E17/E18 it would not be appreciably different). In the interest of space, this is not (in fact, far from) an exhaustive list.

- All so-called Tati languages, i.e., both the Caucasian Tat comprising Muslim Tat [ttt] and Judeo-Tat [jdt] as well Tatic dialects subsumed under the entries Takestani
[tks], Eshtehardi [esh], Alviri-Vidari [avd], Vafsi [vaf], Ashtiani [atn], Razajerdi [rat], Maraghei [vmh] and Gozarkhani [goz], are classified as Southwestern Iranian. In actuality, only the Caucasian Tat are Southwestern Iranian, while the Tatic group is Northwestern Iranian (Schmitt 2000).

- The Mek languages are subdivided into a Western group comprising only Ketengban [xte] and an Eastern group comprising the rest. No such division is in evidence and the label Western for Ketengban is infelicitous as Ketengban lies east of the others. If not a smooth dialect chain, one may posit a Western group comprising Korupum-Sela [kpq], Nakca [nlc], Nipsan [nps], Kosarek Yale [kkl] and an Eastern group comprising Eipomek [eip], Ketengban [xte], Una [mtg] (Heeschen 1978, 1992).

- The classification of the languages in E16s Yi and Ngwi groups are mixed up and mis-subclassified in too many ways to go through here, see, e.g. Pelkey (2011).

- Hamba [hba] is given as a Lega-Kalanga (D.20) Bantu language, but it belongs to the (Greater) Tetela group (Jacobs 1962, 1964).

- Majhwar [mmj] is listed as an unclassified language. It is a Kherwari Munda language (Grierson 1906:135-146).

- Adangbe [adq] is classified as Niger-Congo, Atlantic-Congo, Volta-Congo, Kwa, Left Bank, Kpöso-Ahlo-Bowili. But the language of the Adangbe [adq]24 entry (not to be confused with the language that the Adangme [ada] entry refers to) is a variety of Ewe that the Adangbe shifted to, and should thus be classified as a member of the Gbe group (Harley 2005:31).

- Kaingang [kgp] of Paraná state is classified outside the subgroup that bears its name:

  - Kaingang (2)
    Northern (2)
    Kaingáng, São Paulo [zkp] (Brazil)
    Xokleng [xok] (Brazil)
  - Kaingang [kgp] (Brazil)

A linguistically justified classification is the following (Jolkesky 2010):

  - Kaingang-Xokleng (3)
    Kaingangic (2)
    Kaingáng, São Paulo [zkp]
    Kaingang [kgp]
  - Xokleng [xok]

24I wish to thank Mark Dingemanse for discussion of this entry.
• Tchumbuli [bqa] is given as an Akanic language closely related to Akan [aka] but it is River Oti North Guang (Bertho 1951).

• Isanzu [isn] is listed as an Unclassified Narrow Bantu language, but it is uncontroversially a member of the Nyaturu-Nilamba group (Masele 2001).

• Luo (aka Kasabe) [luw] and Yeni [yei], if at all different from Njerep (Connell 1998:214-217), are close relatives of Njerep rather than unclassifieds (Connell & Zeitlyn 2000).

• Bung [bqd] is listed as Unclassified but it is certainly a Benue-Congo language, most probably in the Konja-Mambila-Vute subgroup (Connell 1998:216).

• Irántxe-Münkii [irn] is classified as Arawakan – an idea that is original only to Rondon (1922:6, 12), but has not been supported by evidence and has not won support in the intervening century.

• Marriammu [xru] is placed in its own subgroup within Daly. In fact it is mutually intelligible with either Maridjabin [zm] or Marithiel [mfr] and should be in the Bringin subgroup of Western Daly (Tryon 1974:94-103).

• Bharia [bha] is classified as an unclassified Dravidian language, but data shows it to be an Indo-Aryan language (close to Bundeli [bns], Grierson 1906).

• Bendi [bct] is listed as a language in the Lendu group (of Central Sudanic) but it has long been known that this is a Mangbutu-Efe group language, though also Central Sudanic (Liesenborghs 1934, Kutsch Lojenga 1994:451, Bokula & Irumu 1994:226).

• Kamar [keq] is listed as a Dravidian language, presumably deriving indirectly from Russell & Hiralal (1916f)’s racially based classification. The language, however, is Indo-European, in the Halbi group (Grierson 1905:386).

• Lodhi [lbm] is listed as a Munda language but with the highly suspect information that it is lexically close to a number of Indo-Aryan languages: “Related to Sora [srh]. Lexical similarity: with Oriya [ori] 59%-67%, with Bangla (Bengali [ben]) 56%-72%, Santali [sat] 20%, Mundari [unr], Munda [unx] 10%; 66%-85% between varieties of Lodhi; 57% to 75% with Kharia Thar [ksy].”. The language is in fact a variety close to Bengali (Dasgupta 1978:153), and there is no evidence that it is or ever was Munda (an erroneous piece of information emanating from the Indian Language Census of 1961, Dasgupta 1978:152).

• Mukha-Dora [mmk] is classified as Dravidian, South-Central, Gondi-Kuí, Konda-Kuí, Konda. In fact they speak only Telugu [mmk] (von Fürer-Haimendorf 1945) and, if given an distinct entry at all, the language should be classified in the same subgroup as Telugu.

• In E16, Sambe [xab] is misclassified as the sole language of the subgroup Niger-Congo, Atlantic-Congo, Volta-Congo, Benue-Congo, Plateoid, Plateau, Western, Southwestern, B. There is another path under which the remaining several dozen Plateau languages
(including Sambe [xab]'s closest relatives, see Blench 2010a) are found: Niger-Congo, Atlantic-Congo, Volta-Congo, Benue-Congo, Plateau. Platoid is the name in the literature for the hypothesized node Plateau+Jukumoid+Kainji which is not adopted by E16, except in the confused path for Sambe [xab].

- E16 has four languages classified as Arai-Kwomtari, Arai (Left May) and another two language classified as (a stand-alone family) Left May. The languages in the two respective branches belong to the same subfamily Arai (the indigenous name for the Left May river), regardless of the validity of the potential wider affiliations of the Arai languages (Conrad & Dye 1975, Laycock 1975).

- Ketangalan [kae] is given as an Austronesian, Unclassified language. It is uncontroversially an Austronesian, East Formosan, Northern East Formosan language, i.e., with Kavalan [ckv] and Basay [byq] (Blust 2009).

- Arára, Pará [aap] is given a profoundly confused classification as the sole member of a spurious “North Brazil” family:
  North Brazil (1)
  - Carib (1)
    * Northern (1)
      · Arára, Pará [aap]

  Arára of Pará [aap] is, in fact, uncontroversially a Cariban language of the Xinguan subgroup (Gildea 2012, Meira 2006).

- Furu [fuu], also known as Bagiro, is classified in the Kara subgroup of Central Sudanic (Central Sudanic, West, Bongo-Bagirmi, Kara: with Gula [kcm] and Yulu [yul]). This is quite far from its actual classification as a member of the Sara subgroup of Western Sara-Bongo-Bagirmi (Boyeldieu 2000a,b).

- Possibly related to the above confusion over the place of Furu [fuu], the Kara [kah] language (also known as Fer) is given as Unclassified. It is uncontroversially a member of the Fer-Gula-Koto subgroup of Sara-Bongo-Bagirmi, i.e. with Gula [kcm] (Boyeldieu 1987, 2000a).

- Shanenawa [swo] is listed as Unclassified. It is uncontroversially a Pano language in the Yaminawa subgroup (Vieira Cândido 2004).

- The Pashai subgroup is not a member of the Kunar group (of Indo-Aryan), though this appeared in linguistic publications as a printing error (Bashir 2003:826).

- An unfortunate innovation in E16 (not in previous editions) is the merging of the Mabaan-Burun subgroup of Western Nilotic (in Eastern North/South Sudan, near the Ethiopian border, see Andersen 2006) and the very different Maban family (in Western Sudan and across the borders in Chad and Central African Republic, see Edgar 1991). This appears to have happened solely on basis of the coincidence of names
(Mabaan [ethnonym] versus Maban [ethnonym Maba plus English -n]). While there are hypotheses worth discussing that unite the Mabaan-Burun and Maban languages under a much higher node (Bender 2000), there is no justification neither for the immediate merging of the two subgroups nor the equally confused subclassification within the merging.

- E16 has two different subgroup of Tupi-Guarani, as follows:
  
  - Guarayu-Siriono-Jora II (2)
    
    Jorá [jor]
    
    Yuqui [yuq]
  
  - Subgroup II (2)
    
    Guarayu [gyr]
    
    Sirionó [srq]

As the name suggests, the four languages uncontroversially belong to the same subgroup (see, e.g., Dietrich 2008, Rodrigues & Cabral 2001).

- Hunsriker [hrx] is listed as a separate branch of Germanic. This is uninformed – it should be Westphalic Low Saxon (Harbert 2007).

- The Tangkic subgroup is given with its outdated affiliation in Pama-Nyungan. It has subsequently been taken out of Pama-Nyungan (Evans 2005:24-27).

- Megam [mef] is listed as a Tibeto-Burman language. This is a confusion of ethnic versus linguistic heritage, as the language is Austroasiatic of the Khasian subgroup (Burling 2004).

- Pataxó-Hâhaáí [pth] is listed as unclassified, but it is long known to be a Maxakalian language (Loukotka 1968, Urban 1985).

- Andh [anr] is listed as unclassified but it closely related to Marathi if at all different from it (Russell & Hiralal 1916b).

- Andaqui [ana] is listed as a branch of Barbacoan presumably as a leftover of a now outdated attempt at a Greater Chibchan family. The attachment of Andaqui [ana] to Barbacoan is equally unfounded (Adelaar & Muysken 2004).

- Cofan [con] is listed as a branch of Chibchan presumably as a leftover of a now outdated attempt at a Greater Chibchan family. It cannot be sustained (Constenla-Umaña 2012).

- E16 has one branch Trade Malay, East Indonesian and another Trade, Malay East-Indonesian which should be the same branch.

- Sartang [onp], a close relative of Sherdukpen (Blench & Post 2011) in the Kamengic subgroup, is wrongly listed as a subgroup of its own under Tibeto-Burman.
- Of Tibeto-Burman, there is one Jinghpō-Konyak-Bodo, Jinghpō-Luish, Luish branch and one Jingpho-Konyak-Bodo, Jingpho-Luish, Luish branch – an obvious typo (Jinghpō for Jingpho).

- Gongduk [goe] is listed as a Tibetan language, but Gongduk is not genealogically a Tibetan language. In fact, it is not closely related to any other Tibeto-Burman language and therefore represents an independent branch (van Driem 1995, 2001a).

There are languages for which no linguistic data (form-meaning pairs) exists, yet they can be argued to exist or have existed as separate languages since there is sociolinguistic information to the effect that the language was not intelligible to any of its neighbours. We may call such languages “unattested”. Unattested languages obviously cannot be classified on linguistic grounds. E16, nevertheless, classifies a number of them. For example:

- Sentinel [std] is unattested (Sarkar & Pandit 1994).
- Arma [aoh], a presumed Chocoan language, is unattested (Uricoechea 1854).
- Villa Viciosa Agta [dyg] is unattested (Headland 2009).
- Badeshi [bdz] is unattested (Khan & Sagar 2004).
- Nagarchal [nbg] is unattested (Valte et al. 2009).

E16 has the category of Unclassified languages in addition to Language Isolates. Many of the languages in the Unclassified category are, in fact, perfectly classifiable, while other are attested and have been compared (without success) to other languages, and are therefore on par with Language Isolates. In order to clear up the confusion, unless mentioned elsewhere in this review, we go through the status of all E16s Unclassified languages here:

- Kehu [khh] is an isolate on the evidence available so far (Kamholz 2012a).
- Turumsa [tqm] and Doso [dol] have 61% lexicostatistical similarity which indicates that they form a small family. They may further belong to the same family as Dibiyasu [dby] since Turumsa [tqm] and Dibiyaso [dby] have 19% similarity, pace the known caveats of lexicostatistical comparison (Tupper 2007).
- Kembra [xkw] is attested, but only in 50 or so words (Doriot 1991) which shows that it is a different language from any of its neighbours and, probably even an isolated language.
- Lepki [lpe] shows some vocabulary resemblances to the neighbouring Murkim [rmh] perhaps enough to establish the two as a family. In any case, the language is attested, so if it is not related to Murkim [rmh], it is an isolate (Hammarström 2010).
• Murkim [rmh] shows some vocabulary resemblances to the neighbouring Lepki [lpe] perhaps enough to establish the two as a family. In any case, the language is attested, so if it is not related to Lepki [lpe], it is an isolate (Hammarström 2010).

• Wakaná [waf] is unattested (Meader 1978).

• Arára do Mato Grosso [axg] is attested (Hargreaves 2007, Souza 2008). The majority of items on the wordlist do not have cognates in neighbouring families, in particular the language does not come out as a Tupí (suggested by its location) or Arawak (suggested by one pronominal resemblance) language. There were four rememberers left in 2001. In 2008, there were only two (Souza 2008), neither of which are fluent speakers.

• Tremembé [tme] is unattested (Métraux 1946).

• Kambiwá [xbw] has so little data attested that it can indeed be called unclassified (Meader 1978).

• Kaimbé [xai] has so little data attested that it can indeed be called unclassified (Meader 1978).

• Aguano [aga] is unattested (Loukotka 1968, Tessmann 1930).

• Wasu [wsu] is unattested (Fabre 2005).

• Beothuk [bue] is an extinct language isolate attested in 400 or so words (Hewson 1978).

• Pumé [yae] is well-attested language isolate (Almenar & Jackson 1988).

• Abishira [ash] is near extinct but attested with wordlists and appears to be an isolate (Michael & Beier 2012).

• Yuwana [yau] is well-attested (Quatra 2008, Vilera 1987, Vilera Díaz 1985) and is either a language isolate or related to Sáliba-Piaroa (Jolkesky 2009).

• Puquina [puq] is an extinct language isolate (Adelaar & van de Kerke 2009, Torero Fernández de Cordoba 2002).


• Rer Bare [rer] is unattested (Bender 1975).

• Mawa [wma] is unattested (Temple 1919:271,430).

• Lufu [ldq] is attested and clearly a language of the Yukub en subgroup of Benue-Congo (Prischneg 2008).

• Kujarge [vkj] is an isolate or an East Chadic language (Hammarström 2010, Lovestrand 2012).
• Laal [gdm] is an isolate (Lionnet 2010).

• Carabaylo [cby] is very poorly attested and may rightly be called an unclassified language (Hammarström 2010).

• Monimbo [mom] is attested (Lehmann 1920a:814-816) with some family names and proves to be the Nicaraguan variant of the language best called Mangue, which is an Otomanguean language (Quirós Rodríguez 2002).

• Truká [tka] is unattested (Fabre 2005).

• Uamúe [uam] poorly attested, perhaps some 100 words (see Meader 1978). Borderline between an isolate or an unclassified language.

• Papavô [ppv] is a name given to uncontacted peoples of the Xinane, Envira, Humaitá, Tarahuacá, Jordao and Yaco rivers. Their language (or languages) is unattested but believed to be Panoan speaking (Huertas Castillo 2004:143-156, Nogueira 1991).

• Pankararé [pax] is unattested (Hohenthal 1960).

• Gail [gic] is an English-Afrikaans-based argot and is thus a Germanic language (Cage 2003).

• Himarimã [hir] is an uncontacted group whose language is unattested (Kroemer 1990, Nogueira 1991, Pohl 2000), but believed to be Arawán as per testimonies from the Suruwahá and Banawá who say they have met them (Everett 1996:62).

• Miarrã [xmi] is unattested (Boas 1975, Galvão & Simões 1965:24).
D Language/Dialect Divisions on a Sample of 100 Languages

To address the question of whether there is overcounting in general in E16, and, to obtain a sharper estimate of the number of mutually intelligible languages (henceforth MI-languages) in the world, 100 entries have been sampled at random, and checked with respect to mutual intelligibility.

-1 represents varieties all intelligible to those of some other entry

OK represents varieties all intelligible to themselves but not to those of some other entry, or

+1 represents varieties not intelligible to themselves and neither to those of some other entry.

The languages sampled and the individual assessment plus source and comments are given below. In all cases, the information in the cited sources is preferable to E16 since the sources explain how and where they obtained the information presented.

<table>
<thead>
<tr>
<th>Language</th>
<th>Code</th>
<th>Macro-Area</th>
<th>Motivation</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>//Gana</td>
<td>gnk</td>
<td>Africa</td>
<td>Intelligible to /Gwi [gwj] (Nakagawa 2006:14)</td>
<td>-1</td>
</tr>
<tr>
<td>A-Pucikwar</td>
<td>apq</td>
<td>Eurasia</td>
<td>Close to but not assertably intelligible with Kede [akx], Kol [aky] or Juwoi [okj] (Portman 1898)</td>
<td>OK</td>
</tr>
<tr>
<td>Akawaio</td>
<td>ake</td>
<td>South America</td>
<td>Intelligible to Patamona [pbc] (Sousa Cruz 2005:36-40)</td>
<td>-1</td>
</tr>
<tr>
<td>Amatlán Zapotec</td>
<td>zpo</td>
<td>North America</td>
<td>The detailed survey by Egland et al. (1983:81) does indicate low levels of intelligibility to the neighbouring San Juan Mixtepec [zpm] and Cuixtla Zapotec [zam]</td>
<td>OK</td>
</tr>
<tr>
<td>Andaman Creole Hindi</td>
<td>hca</td>
<td>Eurasia</td>
<td>Intelligibility status unclear but the examples in Subbakrishna (1994) strongly suggest intelligibility to Urdu [urd]</td>
<td>-1</td>
</tr>
<tr>
<td>Angami Naga</td>
<td>njm</td>
<td>Eurasia</td>
<td>Not mutually intelligible to its closest relative Chokri Naga [nri] (Bielenberg &amp; Zhalie 2001:86)</td>
<td>OK</td>
</tr>
</tbody>
</table>

25 This indicates that the entry, based on unintelligibility, should be split. In cases encountered in the sample, the entry should be split in two, rather than some higher number.

26 I wish to thank Raoul Zamponi for discussion on this point.
Avau avb Pacific Arguably unintelligible to all other Arawe languages (MacKenzie et al. 2011) OK

Awak awo Africa Arguably unintelligible to all other Tulaic languages (Kleinewillinghöfer 1996:98-100) OK

Babatana baa Pacific Groups together a number of dialects (Tryon & Hackman 1983) of which at least Sisiqa is different enough to be a different language (Ross 2002) +1

Bari bfa Africa Although Bari’s closest relatives Kakwa [k eo] and Mandari [mqu] are often called “dialects” (e.g., Spagnolo 1933) the lexicostatistical percentages calculated by Vossen (1982:113-115) at 71% and 73% respectively, appear to be more consistent with Kakwa and Mandari being inherently unintelligible to Bari OK

Barok bjk Pacific The lexicostatistical study of Lithgow & Claassen (1968:388) shows Barok’s closest lexicostatistical relationship to be only 52% (to Patpatar [gfk]) clearly indicating that Barok is not intelligible to any of its relatives OK

Bishnupriya bpy Eurasia Not intelligible to Bengali [ben] or other surrounding languages (Kim & Kim 2008, Sinha 1981) OK

Bidyara bym Australia Intelligible to Gungabula [gyf] -1 (Breen 1973:4)

Big Nam- nmb Pacific Information implicit in Fox (1979:9) and Lynch & Crowley (2001) suggests Big Nambas is not intelligible to any of its neighbours OK
<table>
<thead>
<tr>
<th>Language</th>
<th>Code</th>
<th>Region</th>
<th>Relationship Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boghom</td>
<td>bux</td>
<td>Africa</td>
<td>Boghom is not intelligible to its closest relatives Kir-Balar [kkr] and Mangas [zns] (Shimizu 1978)</td>
</tr>
<tr>
<td>Bon Gula</td>
<td>glc</td>
<td>Africa</td>
<td>Faris &amp; Meundeung (1993:7) reports only 46% lexicostatistical similarity with its closest relative Zan Gula [zna]</td>
</tr>
<tr>
<td>Border</td>
<td>kvn</td>
<td>South America</td>
<td>Intelligible to San Blas Kuna [cuk] (Llerena Villalobos 2000)</td>
</tr>
<tr>
<td>Kuna</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northern</td>
<td>tsp</td>
<td>Africa</td>
<td>Not intelligible to Southern Toussian [wib] – when speakers meet they use Dioula (Zaugg-Coretti 2003:4)</td>
</tr>
<tr>
<td>Toussian</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breri</td>
<td>brq</td>
<td>Pacific</td>
<td>Arguably a separate language as Z’graggen (1969:151-155) reports 73% (with Romkun [rmk]) as the highest lexicostatistical percentage with any other Tamolan language</td>
</tr>
<tr>
<td>Budukh</td>
<td>bdk</td>
<td>Eurasia</td>
<td>The closest relative Kryts [kry] has dialects which are intelligible to Budukh [bdk] but the most distant ones are not intelligible (as evidenced by the use of Azeri to communicate) (Authier 2009:9)</td>
</tr>
<tr>
<td>Burmese</td>
<td>mya</td>
<td>Eurasia</td>
<td>Intelligible to Intha [int], Rakhine [rki] and Tavoyan [tvn] after a week or two (Okell 1995:1)</td>
</tr>
<tr>
<td>Butbut</td>
<td>kyb</td>
<td>Pacific</td>
<td>Intelligible to Southern Kalinga [ksc] (Takaki 1977:vi)</td>
</tr>
<tr>
<td>Kalinga</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kulina</td>
<td>xpk</td>
<td>South America</td>
<td>Kulina Pano of the Curuça is partially intelligible to other languages in the Mayo group, but scores no more than 72% lexical similarity with any of them (Fleck 2007a) leaving the separate entry marginally justifiable</td>
</tr>
<tr>
<td>Pano</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language</td>
<td>Code</td>
<td>Region</td>
<td>Status</td>
</tr>
<tr>
<td>----------</td>
<td>------</td>
<td>--------------</td>
<td>------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Canichana</td>
<td>caz</td>
<td>South America</td>
<td>Certainly unintelligible to all surrounding languages (Crevels 2012)</td>
</tr>
<tr>
<td>Central Okinawan</td>
<td>ryu</td>
<td>Eurasia</td>
<td>Not entirely clear but arguably not intelligible to the other Okinawan languages in E16 (Loveless 1963:1-3)</td>
</tr>
<tr>
<td>Mixtec, Jamiltepec</td>
<td>mxt</td>
<td>North America</td>
<td>The Jamiltepec Mixtec [mxt] entry represents the intelligible Jamiltepec, Mechoacán San Cristobal varieties but is not (sufficiently) intelligible (Egland et al. 1983:34) to the other East Coast Mixtec languages Ixtayutla Mixtec [vmj], Tututepec Mixtec [mtu], San Juan Colorado Mixtec [mjc] and Chayuco Mixtec [mih] (Josserand 1983)</td>
</tr>
<tr>
<td>Dibiyaso</td>
<td>dby</td>
<td>Pacific</td>
<td>Lexicostatistically not near intelligibility with any of its neighbours (Reesink 1976, Shaw 1986, Tupper 2007)</td>
</tr>
<tr>
<td>Dongxiang</td>
<td>sce</td>
<td>Eurasia</td>
<td>Not intelligible with any of the other Qinghai-Gansu Mongolic languages (Field 1997:4)</td>
</tr>
<tr>
<td>Emerillon</td>
<td>eme</td>
<td>South America</td>
<td>Not intelligible to Wayampi of Trois Sauts [oym] without exposure or considerable talent (p.c. François Copin 2012 via Françoise Rose). Intelligibility to Zo’è [pto] appears not to be known, although some varieties of Wayampi [oym] are intelligible to Zo’è (Kästner 2007)</td>
</tr>
<tr>
<td>Language</td>
<td>Code</td>
<td>Region</td>
<td>Notes</td>
</tr>
<tr>
<td>----------</td>
<td>------</td>
<td>--------</td>
<td>-------</td>
</tr>
<tr>
<td>Gadang</td>
<td>gdk</td>
<td>Africa</td>
<td>Information implicit in Arditi (1969), Caprile &amp; Jungraithmayr (1973), Jungraithmayr (1977) suggests that Gadang is unintelligible to its closest relatives Sarwa [swy] and Miltu [mlj], and presumably Boor [bvf] as well, further to the south</td>
</tr>
<tr>
<td>Gera</td>
<td>gew</td>
<td>Africa</td>
<td>Gera is said to be mutually intelligible “to some extent” with Geruma [gea] (Schuh 1978a:91)</td>
</tr>
<tr>
<td>Kuman</td>
<td>kue</td>
<td>Pacific</td>
<td>The closest varieties, subsumed under the Dom [doa] entry are not all inherently intelligible to Kuman (Tida 2006:1-8)</td>
</tr>
<tr>
<td>Cinda-Regi-Tiyal</td>
<td>cdr</td>
<td>Africa</td>
<td>Lexicostatistical percentages in Dettweiler &amp; Dettweiler (2003) indicate that not all three Cinda-Regi-Tiyal lects are intelligible to any other the other Kamuku languages</td>
</tr>
<tr>
<td>Guyani</td>
<td>gvy</td>
<td>Australia</td>
<td>Intelligible with Adnymathanha [adt] (Hercus 1979:626-627)</td>
</tr>
<tr>
<td>Shekhawati</td>
<td>swv</td>
<td>Eurasia</td>
<td>Lexicostatistical figures indicate that no other Rajasthani variety in E16 is intelligible to all varieties subsumed under Shekhawati (Abraham et al. 2012:26)</td>
</tr>
<tr>
<td>Hungarian</td>
<td>hun</td>
<td>Eurasia</td>
<td>Not intelligible with any other Ugric language (Kenesei et al. 1998)</td>
</tr>
<tr>
<td>Ibuoro</td>
<td>ibr</td>
<td>Africa</td>
<td>Mutually intelligible with Itu Mbon Uzo [itm] (Connell 1991:10)</td>
</tr>
<tr>
<td>Idon</td>
<td>idc</td>
<td>Africa</td>
<td>Implicit in the comparative discussion in Blench (2009b) is that Idon is not close enough to any of the other Northern Plateau to be intelligible</td>
</tr>
<tr>
<td>Language</td>
<td>Code</td>
<td>Region</td>
<td>Notes</td>
</tr>
<tr>
<td>----------</td>
<td>------</td>
<td>--------</td>
<td>-------</td>
</tr>
<tr>
<td>Iresim</td>
<td>ire</td>
<td>Pacific</td>
<td>Not intelligible to any of its neighbours or relatives (Kamholz 2012b,c)</td>
</tr>
<tr>
<td>Chikuli</td>
<td>cik</td>
<td>Eurasia</td>
<td>Not intelligible to (Sangla) Kinnauri [kfk] or to other languages of the Kinnauri group as evidenced by lexical similarity (Saxena 2011:23)</td>
</tr>
<tr>
<td>Janji</td>
<td>jni</td>
<td>Africa</td>
<td>Janji [jni] is implicitly described as unintelligible to its neighbours in Shimizu (1975)</td>
</tr>
<tr>
<td>Gugu Badhun</td>
<td>gdc</td>
<td>Australia</td>
<td>Not intelligible to its closest relative Warungu [wrg] (Sutton 1976b, Tsunoda 2011:225)</td>
</tr>
<tr>
<td>Lele</td>
<td>lel</td>
<td>Africa</td>
<td>Lexicostatistical figures barely over 70% to varieties of Bushoong [buf] and Wongo [won] suggests unintelligibility (Bastin et al. 1999:209)</td>
</tr>
<tr>
<td>Kadiwéu</td>
<td>kbc</td>
<td>South America</td>
<td>Kadiwéu is the only Guaicuran language and not intelligible to any other language (Sándalo 1997, Viegas Barros 2011)</td>
</tr>
<tr>
<td>Amdang</td>
<td>amj</td>
<td>Africa</td>
<td>Not at all intelligible to its closest relative Fur [fvr] (Wolf 2010)</td>
</tr>
<tr>
<td>Kalapuya</td>
<td>kyl</td>
<td>North America</td>
<td>The Kalapuya entry is described as subsuming Lukamiute, Santiam and Wapatu, but Wapatu (also known as Tualatin) was not intelligible to Santiam-Lukamiute or the other Central dialects (Mithun 1999:431-433). Furthermore, Yoncalla, also unintelligible to the Central and Northern (including Wapatu) dialect clusters, being extinct before 1950, is missing altogether.</td>
</tr>
<tr>
<td>Language</td>
<td>Code</td>
<td>Region</td>
<td>Description</td>
</tr>
<tr>
<td>---------------</td>
<td>------</td>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Wanggom</td>
<td>wng</td>
<td>Pacific</td>
<td>Only 61% cognate with a variety of its closest relative Kombai [tyn] suggesting lack of inherent intelligibility (Versteeg 1983:22)</td>
</tr>
<tr>
<td>Kinaray-a</td>
<td>ktjr</td>
<td>Pacific</td>
<td>Intelligible to (at least) Sulod [srg] (Zorc 1972)</td>
</tr>
<tr>
<td>Towei</td>
<td>ttn</td>
<td>Pacific</td>
<td>Not intelligible to Dubu [dmu] or any other neighbour (Wambaliau 2004)</td>
</tr>
<tr>
<td>Krongo</td>
<td>kgo</td>
<td>Africa</td>
<td>Krongo is a little over 70% lexicostatistically similar to Tumtum [tbr] (Thelwall &amp; Schadeberg 1983:225) and seems thus borderline intelligible only (Stevenson 1956/1957:106-107)</td>
</tr>
<tr>
<td>Lafofa</td>
<td>laf</td>
<td>Africa</td>
<td>Certainly not intelligible to any surrounding languages (Stevenson 1956/1957:102). The El Amira dialect looks fairly different and may constitute a separate language, though there is too little data to tell (MacDiarmid &amp; MacDiarmid 1931:154-155)</td>
</tr>
<tr>
<td>Lahu Shi</td>
<td>lhi</td>
<td>Eurasia</td>
<td>Not mutually intelligible to the varieties under the Lahu [lhu] entry (Matisoff 1973:xlii-xliii, Cooper 2002)</td>
</tr>
<tr>
<td>Lefa</td>
<td>lfa</td>
<td>Africa</td>
<td>Although they might insist on a separate identity, Lefa is intelligible to and lexicostatistically at least 78% with all varieties of Bafia [ksf] surveyed (Guarisma &amp; Paulian 1986, Mbongue 1999)</td>
</tr>
<tr>
<td>Language</td>
<td>Code</td>
<td>Region</td>
<td>Intelligibility</td>
</tr>
<tr>
<td>-----------</td>
<td>-------</td>
<td>----------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Lakota Dida</td>
<td>dic</td>
<td>Africa</td>
<td>OK</td>
</tr>
<tr>
<td>Lokoya</td>
<td>lky</td>
<td>Africa</td>
<td>OK</td>
</tr>
<tr>
<td>Machame</td>
<td>jmc</td>
<td>Africa</td>
<td>-1</td>
</tr>
<tr>
<td>Mambae</td>
<td>mgm</td>
<td>Pacific</td>
<td>+1</td>
</tr>
<tr>
<td>Mander</td>
<td>mqr</td>
<td>Pacific</td>
<td>OK</td>
</tr>
<tr>
<td>Mehri</td>
<td>gdq</td>
<td>Eurasia</td>
<td>OK</td>
</tr>
<tr>
<td>Language</td>
<td>Region</td>
<td>Type</td>
<td>Notes</td>
</tr>
<tr>
<td>--------------</td>
<td>--------</td>
<td>------</td>
<td>-------</td>
</tr>
<tr>
<td>Mescalero-Chiricahua Apache</td>
<td>North America</td>
<td>apm</td>
<td>Despite high lexicostatistical figures in Hoijer (1971) the Western Southwestern Apachean languages Mescalero-Chiricahua Apache [apm], Navajo [nav] and Western Apache [apw] are not fully intelligible (p.c. Willem de Reuse 2012)</td>
</tr>
<tr>
<td>Tibea</td>
<td>Africa</td>
<td>ngy</td>
<td>Not intelligible to other languages of the Bafia (A.50) group (Bradley 1992)</td>
</tr>
<tr>
<td>Lauje</td>
<td>Pacific</td>
<td>law</td>
<td>Likely two languages based on intelligibility as “.. both Lauje and Ampibabo-Lauje speakers do not consider their speech varieties mutually intelligible” (Himmelmann 2001:21)</td>
</tr>
<tr>
<td>Inebu One</td>
<td>Pacific</td>
<td>oin</td>
<td>Not fully intelligible to any other Central-Northern One variety subsumed under the Kabore One [onk], Molmo One [aun] and Northern One [onr] entries (Crowther 2001:76)</td>
</tr>
<tr>
<td>Ouma</td>
<td>Pacific</td>
<td>oum</td>
<td>Not intelligible to the Magoric or any other closely related languages (Dutton 1976:615)</td>
</tr>
<tr>
<td>Namonuito</td>
<td>Pacific</td>
<td>nmt</td>
<td>From the perspective of intelligibility, nothing hinder the Namonuito [nmt] to be merged with Murilo [pfa] since all varieties in both entries are intelligible (Quackenbush 1968:107)</td>
</tr>
<tr>
<td>Language</td>
<td>Code</td>
<td>Region</td>
<td>Note</td>
</tr>
<tr>
<td>-----------</td>
<td>------</td>
<td>--------</td>
<td>------</td>
</tr>
<tr>
<td>Nkukoli</td>
<td>nbo</td>
<td>Africa</td>
<td>A statement in Talbot (1912:245) suggests that there is not full intelligibility to its closest relative (Dimmendaal 1978) Lubila [kcc]</td>
</tr>
<tr>
<td>North Watut</td>
<td>una</td>
<td>Pacific</td>
<td>Not inherently intelligible to Middle Watut [mpl] or South Watut [mcy] (Holzknecht 1989:33-35)</td>
</tr>
<tr>
<td>Northern Nuni</td>
<td>nuv</td>
<td>Africa</td>
<td>Not all varieties are intelligible to all varieties under the Southern Nuni [nnw] entry (Malgoubri 2011)</td>
</tr>
<tr>
<td>Northwest esk</td>
<td>North America</td>
<td>Intelligible to North Alaskan Inupiatun [esi] (Lanz 2010)</td>
<td></td>
</tr>
<tr>
<td>Alaska Imupiatun</td>
<td>Andio</td>
<td>Pacific</td>
<td>Maximally 66% lexicostatistically similar to its nearest relative in the Saluan-Banggai subgroup, suggesting non-intelligibility (Mead 2003:68)</td>
</tr>
<tr>
<td>Penadu ums</td>
<td>Pacific</td>
<td>Not intelligible to other Southern Tomini languages (Quick 2007:1-21)</td>
<td></td>
</tr>
<tr>
<td>Pwo Western pwo</td>
<td>Eurasia</td>
<td>Not intelligible to Pwo Eastern Karen [kjp] (Kato 1995:36)</td>
<td></td>
</tr>
<tr>
<td>Labuk-Kinabatangan dtb</td>
<td>Pacific</td>
<td>Unintelligible to other Dusunic languages, and insufficient intelligibility between the two clusters (Spitzack 1984)</td>
<td></td>
</tr>
<tr>
<td>Kadazan</td>
<td>Safeyoka apz</td>
<td>Pacific</td>
<td>Only 51% lexicostatistical similarity to the closest relative Kawacha [kcb] (Lloyd 1973:36)</td>
</tr>
<tr>
<td>Bench bcq</td>
<td>Africa</td>
<td>Not intelligible to any other Ta-Ne-Omotic language (Rapold 2006:1-36)</td>
<td></td>
</tr>
<tr>
<td>Language</td>
<td>Code</td>
<td>Region</td>
<td>Intelligibility Notes</td>
</tr>
<tr>
<td>----------</td>
<td>------</td>
<td>-----------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Woria</td>
<td>wor</td>
<td>Pacific</td>
<td>At most 69% lexicostatically similar to any other Geelvink Bay language, suggesting unintelligibility (Jones 1987:5)</td>
</tr>
<tr>
<td>Upper Chehalis</td>
<td>cjh</td>
<td>North America</td>
<td>Only partially intelligible to its closest relative Cowlitz [cow] (Kinkade 1963:1)</td>
</tr>
<tr>
<td>Sila</td>
<td>slt</td>
<td>Eurasia</td>
<td>Implicit that there is no other Bisoid-Hanic intelligible language (Bradley 2007a, Kingsada &amp; Shin-tani 1999, Schliesinger 1998)</td>
</tr>
<tr>
<td>Siwi</td>
<td>siz</td>
<td>Africa</td>
<td>Not intelligible to its closest relative Soknah [swn] (Sarnelli 1924-1925, Vycichl 2005)</td>
</tr>
<tr>
<td>Somray</td>
<td>smu</td>
<td>Eurasia</td>
<td>Close enough to be intelligible to Samre [sxm] (Headley 1977, 1978)</td>
</tr>
<tr>
<td>Suena</td>
<td>sue</td>
<td>Pacific</td>
<td>Although intelligible to some extent to Zia [zia], lexicostatistical values drop to 59% on the Mawae dialect of Zia (Wilson 1969:76)</td>
</tr>
<tr>
<td>Tenis</td>
<td>tns</td>
<td>Pacific</td>
<td>Only 66% lexicostatically cognate, suggesting unintelligibility, with the only other St. Matthias Oceanic language Mussau-Emira [emi] (Blust 1984)</td>
</tr>
<tr>
<td>Tarok</td>
<td>yer</td>
<td>Africa</td>
<td>The presentation in Blench (2004) implicitly entails that Tarok is not intelligible to its closest relative Pe [pai]</td>
</tr>
<tr>
<td>Language</td>
<td>Code</td>
<td>Region</td>
<td>Notes</td>
</tr>
<tr>
<td>----------</td>
<td>------</td>
<td>----------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Mbedam</td>
<td>xmd</td>
<td>Africa</td>
<td>Close to but not intelligible with any of the other Sukur-Matakam-Daba group languages (Lamberty 2003)</td>
</tr>
<tr>
<td>Tol</td>
<td>jic</td>
<td>North America</td>
<td>Not intelligible to any other language (Campbell &amp; Oltrogge 1980). Western Jicaque, a separate non-intelligible language, possibly went extinct after 1950 and is missing (Holt 1999:5)</td>
</tr>
<tr>
<td>Ngbundu</td>
<td>nuu</td>
<td>Africa</td>
<td>Although close, lexicostatistically and phonologically not explainable as an intelligible dialect of any other E16 Banda entry (Cloarec-Heiss 1978)</td>
</tr>
<tr>
<td>Tonsawang</td>
<td>tnw</td>
<td>Pacific</td>
<td>Although partially intelligible lexicostatistical figures are maximally 65% to any other Minahasan language (Merrifield &amp; Salea 1996:11-14)</td>
</tr>
<tr>
<td>Tuwali Ifugao</td>
<td>ifk</td>
<td>Pacific</td>
<td>Not intelligible to all varieties subsumed in Amganan Ifugao [ifa] – the closest Ifugaw language (Walrod 1978)</td>
</tr>
<tr>
<td>Urapmin</td>
<td>urm</td>
<td>Pacific</td>
<td>Argued by Robbins (1998:25-29) to be intelligible neither to the neighbouring Telefol [tlf] nor to Tifal [tif] (though most Urapmin also speak Telefol)</td>
</tr>
<tr>
<td>Tajuasohn</td>
<td>tja</td>
<td>Africa</td>
<td>Not intelligible to its closest relative Klao [klu] (Duitsman et al. 1975)</td>
</tr>
</tbody>
</table>
Aranadan  aaf  Eurasia  Intelligible to Malayalam 1
[Mal] (Menon 1996:12-16)
Morouas  mrp  Pacific  Argued by Lynch & Crowley (2001:56) to be a separate language on the basis of lexico-statistical figures in the 70%
Western Maninkakan  mlq  Africa  Intelligible with Jahanka [jad] (Long 1971)
Yecuatla Totonac  tlc  North America  From the discussion in Brown et al. (2011) it is clear that it is not intelligible to any of the Central Totonacan (or Tepehuan languages)
Yei  yey  Africa  Not intelligible to any other languages (Seidel 2005, 2008)

References


Abraham, Binny; Kara Sako; Elina Kinny; and Isapdaile Zeliang. 2005. A Sociolinguistic Research Among Selected Groups in Western Arunachal Pradesh Highlighting MONPA. Ms.


Bai, Bibo, and Xianming Xu. 2015. Documentation of endangered Cosao "green codes" in Yunnan, China. Paper presented at the ICLDC 4, Universiti of Hawai'i at Mānoa.


Barannikov, Aleksej P. 1934. The Ukrainian and South-Russian Gypsy Dialects. Leningrad: NAUKA.


Bowern, Claire. 2010. Two missing pieces in a Nyulnyulan jigsaw. LSA Annual Winter Meeting.


Boyeldieu, Pascal. 2000b. La langue Bagiró (République Centrafricaine), *Schriften zur Afrikanistik*, vol. 4. Frankfurt am Main: Peter Lang.


Bueno, Cosme. 1770. *Descripción del Obispado de Santa Cruz, El conocimiento de los tiempos, ephemeride del año de 1771*. El conocimiento de los tiempos, ephemeride del año de 1771. Lima: En la Imprenta Real; Calle de Palacio. I think that the title "Descripción de las misiones del Alto-Perú" is nothing more than an ad hoc title for the original title "Descripción del Obispado de Santa Cruz" which is a kind of appendicle in the original book.


Bunce, Daniel. 1851. *Language of the Aborigines of the Colony of Victoria and other Australian districts, with parallel translations and familiar specimens in dialogue as a guide to Aboriginal Protectors and others engaged in ameliorating their condition*. Melbourne: Daniel Harrison.


Cardús, José. 1886. *Las Misiones Franciscanas entre los Infiernos de Bolivia: descripción del estado de ellas en 1883 y 1884*. Barcelona: Barcelona: Librería de la inmaculada concepción.


CARTER, JOHN; KATIE CARTER; JOHN GRUMMITT; BONNIE MACKENZIE; and JANELL MASTERS. 2012. A Sociolinguistic Survey of the Mur Village Vernaculars, SIL Electronic Survey Reports 2012-043. SIL Electronic Survey Reports 2012-043. SIL International.


COOK, WALTER ANTHONY. 1965. A Descriptive Analysis of Mundari: A Study of the Structure of the Mundari Language According To the Methods of Linguistic Science, With Particular Attention To the Units of Sound, the Units of Meaning, the Units of Grammar, and Their Mutually Contrastive Arrangement Patterns. Georgetown University dissertation.


DE LUZENA, GERÓNIMO JOSEF, no date. Traducion de la lengua española á la otomaca, taparita y yarura. Manuscrito II/2927 Biblioteca del Palacio Real de Madrid.


DE VERGARA, GABRIEL. 1965 [1732]. El cuadernillo de la lengua de los indios pajalates (1732) y El confesonario de índios en lengua coahuilteca. Instituto Tecnológico y de Estudios Superiores.


Egland, Steven; Doris Bartholomew; and Saúl Cruz Ramos. 1983. La inteligibilidad interdialectal en México: resultados de algunos sondeos. México: ILV.


Fitz-Roy, Robert. 1839. Fragment of a vocabulary of the Alikhoolip and the Tekeenica languages, also some words of those spoken by the Patagonian (Tehuel-Het) and Chonos Indians. *Appendix to the second volume, Narrative of the surveying voyages of His Majesty's Ships Adventure and Beagle between the years 1826 and 1836, describing their examination of the southern shores of South America, and the Beagle’s circumnavigation of the globe*, vol. 2, 135-141. London: Henry Colburn.


HAJEK, JOHN. 1996. The Mystery of the Kenaboi: A First Report. Royal Institute of Linguistics and Anthropology, International workshop on South-East Asian studies ; No. 11. (The study of) endangered languages and literatures of South-East Asia.


HOLTON, GARY; MARIAN KLAME; FRANTIŠEK KRATOCHVÍL; LAURA ROBINSON; and ANTOINETTE SCHAPPER. 2012. The historical relation of the Papuan languages of Alor Oceanic Linguistics. *Oceanic Linguistics* 51.86–122.


JIJÓN Y CAAMAÑO, JACINTO. 1919. Contribución al conocimiento de las lenguas indígenas que se hablaron en el Ecuador interandino y occidental con anterioridad a la Conquista española. *Boletín de la Sociedad Ecuatoriana de Estudios Históricos* 2.340-413.


KÄSTNER, KLAUS-PETER. 2005. Beiträge zur Geschichte und Kultur der Urueuwauwau-Stämme (Rondonia, Brasilien). Abhandlungen und Berichte der Staatlichen Ethnographischen Sammlungen Sachsen 52.91–133.


Lafone Quevedo, Samuel A. 1897. Los Índios Chanases y su Lengua, con Apuntes sobre los Querandies, Timbíes, Yaros, Boanes, Guenoas o Minuanes, y un mapa étnico. *Boletín del Instituto Geográfico Argentino* XVIII.115–154.


Loukotka, Čestmír. 1932. La Familia Lingüística Kamakan del Brasil (Segunda Parte del Estudio de los Že). Revista del Instituto de etnología de la Universidad nacional de Tucumán 2.493–524.


Masele, Balla F. Y. P. 2001. The linguistic history of SiSuumbwa, KISukuma and KINyanweezi in Bantu Zone F. Memorial University of Newfoundland dissertation.


McCulloch, W. 1859. *Account of the Valley of Munnipore and of the hill tribes, with a comparative vocabulary of the Munnipore and other languages, Selections from the records of the Gov’t of India (Foreign Department)*, vol. 27. Calcutta: Bengal Printing Co.


PIERSON, ABRAHAM. 1658. *A further accoempt of the progresse of the gospel amongst the Indians in New-England, and of the means used effectually to advance the same.*: Set forth in certaine letters sent from thenc e declaring a purpose of printing the Scriptures in the Indian tongue into which they are already translated. With which letters are likewise sent an epitome of some exhortations delivered by the Indians at a fast, as testimonies of their obedience to the gospell. As also Some helps directing the Indians how to improve naturall reason unto the knowledge of the true God. Cambridge, England: Samuel Green.


Quesada Pacheco, Miguel Ángel. 1990. La lengua huetar. Estudios de Lingüística Chibcha 9.7–64.


Ramos Boerr, Nancy. [ca. 2007]. *Lenguaje charrúa, chaná, guenoa y algunos términos de la toponimia, fauna y flora en guaraní, Testimonios (lenguas indígenas). Testimonios (lenguas indígenas)*. Montevideo: [s.l.].


Reid, Lawrence A. no date(a). Malaweg Lexicon. Ms.

Reid, Lawrence A. no date(b). Malaweg Syntax. Ms.


Robbins, Joel Lee. 1998. Becoming sinners: Christian transformations of morality and

Robson, Laura. 2011. The Documentation of the Language Ecology of Njanga, a Moribund

Rocca, Manuel, and Juan José Rossi. 2004. Los Chané-Chiriguano: (Arawak y
Guaraní), Indígenas de la Argentina. Indígenas de la Argentina, 1 edn. Buenos Aires:
Galerna. Desde Sudamérica.


Cambridge University Press.

Revendo a classificação interna da família Tupi-Guarani. Línguas indígenas brasileiras.
fonologia, gramática e história. atas do i encontro internacional do gtl, ed. by Ana Suelly
Editora Universitaria da UFPA.

Rona, José Pedro. 1964. Nuevos elementos acerca de la lengua charrúa, Publicación del

Rondon, Cândido M. S. 1922. Conferências realizadas em 1910 no Rio de Janeiro e em

Rondon, Cândido M. S., and João Barbosa de Faria. 1948. Glossário geral das tribos
silvícolas de Mato Grosso e outras da Amazônia e do Norte do Brasil: Tomo I, Publicação
76 do Conselho nacional de proteção dos Indios, Anexo, vol. 5. Rio de Janeiro: Comissão
Rondon.

Rosenblat, Angel. 1936. Los Otomacos y Taparitas de los llanos de Venezuela. Estudio

Ross, Malcolm. 2002. Sisiqa. The oceanic languages, ed. by John Lynch, Malcolm Ross,
and Terry Crowley, Curzon Language Family Series, 456–466. Richmond: Richmond: Cur-
zon.


Rudes, Blair A. 1997. Resurrecting Wampano (Quiripi) from the Dead: Phonological

Rudes, Blair A. 2000. Resurrecting coastal Catawban: The reconstituted phonology and
Rueck, Michael J. no date. Unubahe: A nearly extinct language. Ms.


Sommer, Bruce A. no date. Ikarranggal. Ms.


TYLOR, E. B.; W. BLOXAM; DANIEL WILSON; G. M. DAWSON; H. LEFROY; and R. G. HALIBUTON. 1891. Sixth report of the committee appointed to investigate the physical characters, languages, and industrial and social conditions of the North-Western Tribes of the Dominion of Canada. *Report of the meeting of the British Association for the Advancement of Science* 60.553-715.


