The status of the least documented language families in the world

Harald Hammarström

Radboud Universiteit, Nijmegen and
Max Planck Institute for Evolutionary Anthropology, Leipzig

This paper aims to list all known language families that are not yet extinct and all of whose member languages are very poorly documented, i.e., less than a sketch grammar's worth of data has been collected. It explains what constitutes a valid family, what amount and kinds of documentary data are sufficient, when a language is considered extinct, and more. It is hoped that the survey will be useful in setting priorities for documentation fieldwork, in particular for those documentation efforts whose underlying goal is to understand linguistic diversity.

1. INTRODUCTION. There are several legitimate reasons for pursuing language documentation (cf. Krauss 2007 for a fuller discussion).[1] Perhaps the most important reason is for the benefit of the speaker community itself (see Voort 2007 for some clear examples). Another reason is that it contributes to linguistic theory: if we understand the limits and distribution of diversity of the world's languages, we can formulate and provide evidence for statements about the nature of language (Brenzinger 2007; Hyman 2003; Evans 2009; Harrison 2007). From the latter perspective, it is especially interesting to document languages that are the most divergent from ones that are well-documented—in other words, those that belong to unrelated families. I have conducted a survey of the documentation of the language families of the world, and in this paper, I will list the least-documented ones. The aim is to inform readers of their existence and documentation status, with the ultimate hope of increasing their chances of being documented.

All and only the language families that satisfy all of the four following criteria are included in the present list:

1. The language family is known through at least a wordlist (i.e., this excludes languages known to exist, but for which there are no data, such as the languages of 'isolados'[2]).
2. The language family, at the present state of knowledge, is not demonstrably related to any other known family.
3. There are no viable grounds for concluding that the language is extinct, i.e., that it does not have fluent speakers.
4. All languages of the family are poorly documented, in the sense that there is less documentation than a rudimentary grammar sketch, and there is no ongoing documentation effort for any of them.

The listing is summarized in Table 1. In section 2 I provide full bibliographic data, references, potential links to other families, endangerment status, and a history of knowledge for each case. In section 3 I clarify why a number of less obvious cases do not satisfy (at least one of) the above criteria, and are thus not included in the list in section 2 and Table 1.

I wish to stress that the listing does not take the "density" of documentation of a language family into account. Even for a very large family, if only one language has a grammar sketch, it is not the case that all of its languages are poorly documented. Such a family does not satisfy the last of the four criteria and therefore will not appear in the present listing. The density of documentation for a language family is a legitimate concern, but is, unfortunately, beyond the scope of the present paper.

All the judgments as to family membership, i.e., what counts as demonstrably related and so on, were made by the author, based on inspection of all relevant sources. These judgments are potentially more superficial than those of a specialist, but since there are no active specialists for the bulk of the languages in question, this is unavoidable. In addition, because there are frequent disagreements among scholars on such matters, the entry for each individual case[3] includes an explanation of the comparative situation and the choices made.

The present listing is based on information (classification, extinct status, descriptive work, etc.) that is subject to change in the future and may additionally contain errors and omissions on my part. Therefore, I maintain a website at http://haraldhammarstrom.ruhosting.nl/least_documented/ for changes and updates to the present listing. I encourage input and feedback from other scholars. Since there is no space limitation on a website, it also contains a full listing of families that clearly do have a grammar sketch for at least one of its languages.

---

[2] See Brackelaire & Azanha 2006; Kästner 2005:99–103; Eberhard 2009:31; Kästner 2007:160–162; Queixalós & Anjos G.S. 2007:29; Forline 1997:1; Silva Magalhães 2002:5; Eyzaguirre Beltroy et al. 2008; Crevels & Voort 2008:162–163; Survival International 2009b; Cabodevila 2007; Rodríguez Bazán 2000; Mora B. 2007; Valenzuela 2006; Meira 2000:19–20; Survival International 2009a; Kumar 2000; Kaplan & Hill 1984; Huertas Castillo 2004, Michael & Beier 2003; Keifenheim 1997; Wallace 2003; Hughes 2009; Clouse 1993:2; and Vries & Enk 1997:5 for pieces of information on groups so far uncontacted.

[3] I do this also for the unclear cases wherever relevant.

Every language and language family considered is cited with its iso-639-3 three-letter unique identifier for ease of comparison with standard reference works, e.g., *Ethnologue* (Lewis 2009). The information in the present paper differs from that in standard reference works mainly in that it systematically investigates the status of documentation and uses it as the criterion for inclusion. For example, while *Ethnologue* lists a certain number of references to descriptive data, it does not aim to systematically list all (or the most extensive) references, so the status of a language's documentation is not directly deducible from these listings. Other recent reference works, e.g., Brown & Ogilvie 2009 and Brown 2006, do have rigid bibliographic listings but do not aim to be complete in their coverage of the language families of the world. The present listing also differs somewhat from general reference works on genetic classification and speaker numbers. As explained above, I state the reason for every innovative choice taken as to genetic classification and give individual references to the relevant comparative-historical literature. For speaker numbers, I have consulted specialist literature wherever possible and used figures from *Ethnologue* (Lewis 2009) in the rest of the cases or when there is reason to believe that it holds the most recent information. A final point of difference is that I mention the first appearance of each language in the linguistic literature in order to give a general sense of how accessible the language is and how long it has gone without documentation.

Whalen and Simons (2009) present a survey of endangered language families, i.e., families whose member languages are at risk of extinction. It differs from the present study in that it uses information (classification as well as speaker numbers) solely from Lewis (2009). More importantly, it does not discuss the status of documentation of different families, which is the focus of the present survey.

In total, twenty-seven families are listed, all but four of them one-member families, i.e., isolates. Since most of the world's language families are very small (Hammarström 2007)—roughly half of them even isolates—it is not surprising that the set of the least documented ones is dominated by isolates.

The lack of geographical diversity in the resulting list is striking. The bulk of the language families discussed are situated in the lowland rainforest area between the Sobger (Papua, Indonesia) and the Upper Sepik (Papua New Guinea) Rivers of the island of New Guinea. A few more on the list are in other remote areas of Papua, Indonesia, namely, Sause between the Tor and the Lakes Plain, Bayono-Awbono just south of the highlands, Dem in a pocket of the highlands, Mawes on the northern coastline, as well as Mor and Tanahmerah in the northeast of the Bomberai peninsula. In contrast, the listed cases from South America come from much more explored areas and no longer maintain their original territory. Kujarge, the only African case, was reported to be on the Chad side in the very remote area near the border of Chad, Sudan and the Central African Republic. In India, Nihali is in an accessible location in central India, while Shom Pen requires travel to the Nicobar Islands.

Name	Location	iso	Size	Endanger-ment	Documentation
Awaké	Venezuela	atx	Isolate	HE	Wordlists, some phrases
Sapé	Venezuela	spc	Isolate	HE	Wordlists, some phrases
Tekiráka	Peru	ash	Isolate	HE	Wordlists
Yurí	Colombia	cby?	Isolate	Unknown	Wordlists
Kujarge	Chad	vkj	Isolate	Unknown	200-item wordlist
Nihali	India	nll	Isolate	E	Long wordlists, a little text, possibly more
Shom Pen	Nicobar	sii	Isolate?	Unknown	Wordlists and phrases, but quality unclear
Afra	Papua, Indonesia	ulf	Isolate	E	ca 250 words and 15 sentences
Bayono-Awbono_		2-family			
-Bayono	Papua, Indonesia	byl		NE	ca 200 words and a few sentences
-Awbono	Papua, Indonesia	awh		NE	ca 200 words and a few sentences
Lepki	Papua, Indonesia	lpe	Isolate	Unknown	ca 200 words
Asaba	PNG	seo	Isolate?	NE	Short wordlist
Baiyamo	PNG	ppe	Isolate?	NE	Short wordlist
Busa	PNG	bhf	Isolate	Unknown	Short wordlist
Dem	Papua, Indonesia	dem	Isolate	NE	Wordlist and some sentences
Guriaso	PNG	grx	Isolate	HE	Short wordlist and short grammar notes
Inland Gulf_		7-family			
-Ipiko	PNG	ipo		Unknown	Wordlists
-Foia Foia	PNG	ffi		Unknown	Wordlists

-Hoia Hoia	PNG	hhi		Unknown	Wordlists
-Hoyahoya	PNG	hhy		Unknown	Wordlists
-Karami	PNG	xar		D	Short wordlist
-Minanibai	PNG	mcv		Unknown	Wordlists
-Mubami	PNG	tsx		NE	Wordlists
Kapauri	Papua, Indonesia	khp	Isolate	NE	ca 250 words and 15 sentences
Kimki	Papua, Indonesia	sbt	Isolate	NE	ca 250 words and 15 sentences
Mawes	Papua, Indonesia	mgk	Isolate	E	ca 250 words and 15 sentences
Mor	Papua, Indonesia	moq	Isolate?	HE	Short wordlist, possibly sentences
Murkim	Papua, Indonesia	rmh	Isolate	NE	ca 250 words and 15 sentences
Namla-Tofanma_			2-family		
-Namla	Papua, Indonesia	naa		HE	ca 250 words and 6 sentences
-Tofanma	Papua, Indonesia	tlg		NE	ca 250 words and 15 sentences
Pyu	PNG	pby	Isolate	HE	Short wordlist
Sause	Papua, Indonesia	sao	Isolate	Unknown	ca 200 words
Tanahmerah	Papua, Indonesia	tcm	Isolate?	Unknown	Short wordlist
Walio_			4-family		
-Pei	PNG	ppq		Unknown	Short wordlist
-Tuwari	PNG	tww		Unknown	Short wordlist
-Walio	PNG	wla		Unknown	Short wordlist
-Yawiyo	PNG	ybx		Unknown	Short wordlist
Yetfa/Biksi	Papua, Indonesia	yet	Isolate	NE	ca 250 words and 15 sentences

TABLE 1. D = Extinct, HE = Highly Endangered, E = Endangered,
NE = Not Presently Endangered, Unknown = No recent data (though in all cases, a good guess
is that they are endangered).

2. LISTINGS
2.1. SOUTH AMERICA

2.1.1. AWAKÉ [atx]. A survey of the Rio Branco area from 1787 (d'Almada 1787:677) mentions an ethnic group "Aoaquis," whose name and location match with Awaké of later reports (Koch-Grünberg 1922:226). The first report of Awaké, together with a vocabulary, was by Koch-Grünberg (1913:455–457).

The collected vocabularies bear no significant relationship to neighboring languages (Loukotka 1968; Migliazza 1980, 1983, 1985). In particular, there is no evidence (Migliazza & Campbell 1988) for an Arutani-Sape family consisting of Awaké [atx] and Sape [spc]. The listings of such a family (Kaufman 1990:50; Lewis 2009) is merely the result of grouping "leftover" languages.

There are published short vocabularies and some phrases (Koch-Grünberg 1928; Matallana & Armellada 1943; Migliazza 1978, Perozo et al. 1978), a short unpublished vocabulary (Coppens 1976), and a minuscule amount of analyzed grammar (Migliazza 1980, 1983, 1985). Ernest Migliazza is currently re-working field notes from the 1960s for publication (p.c. Ernest Migliazza 2010). Further, there is a fairly extensive ethnographic study in Coppens 1983b.

We know that the Awaké have been a small nation for at least a century, since the first expedition in 1882 to visit the Awaké in their home territory counted a mere eighteen people (Koch-Grünberg 1922:226). However, it is not clear how many competent speakers, if any, remain at present. Migliazza, with excellent knowledge of the region, counted only five speakers among fifteen ethnic Awaké in 1964 (Migliazza 1978:135–136, cf. 1972:20). The 2001 Venezuela census counts twenty-nine members of the ethnic group (Mattei-Müller 2006; Amodio 2007), but it is not recorded how many of them speak Awaké (rather than Yanam [shb]). The situation on the Brazilian side appears to be similar (Lewis 2009; Fabre 2005).

2.1.2. SAPÉ [spc]. Sapé was first reported by Koch-Grünberg (1928) who also published a vocabulary.

Like those for Awaké, the collected Sapé vocabularies bear no significant relationship to neighboring languages (Loukotka 1968; Migliazza 1980, 1983, 1985). Additionally, there is no evidence (Migliazza & Campbell 1988) for an Arutani-Sapé family consisting of Awaké [atx] and Sapé [spc]. The listing of such a family (Kaufman 1990:50; Lewis 2009) is merely the result of grouping "leftover" languages.

Short vocabularies and some phrases (Koch-Grünberg 1928; Matallana & Armellada 1943; Migliazza 1978), a short unpublished vocabulary (Coppens 1976), and a minuscule amount of analyzed grammar (Migliazza 1980, 1983, 1985) exist for Sapé. Migliazza is currently re-working field notes from the 1960s for publication (p.c. 2010). Further, a fairly extensive ethnographic study can be found in Coppens 1983a.

The number of competent speakers is presently unclear. Migliazza counted only five speakers among twenty-five ethnic Sapé in 1977 (Migliazza 1978:135–136), though census data has listed a number around one hundred for the ethnic group (Fuchs 1967:78–79). The 2001 Venezuela census counts six members of the ethnic group (Mattei-Müller 2006; Amodio 2007), which conflicts with a slightly higher figure in Mosonyi 2003 (109–110). A recent anthropological report confirms that there are still elderly speakers of the language (Perozo et al. 2008).

2.1.3. TEKIRÁKA [ash]. The Tekiráka people may have been mentioned as early as 1621 (Espinosa Pérez 1955:62) but this, and subsequent early mentions without accompanying linguistic data, may refer to other ethnic groups also labelled "Avijiras". The first unequivocal reference to Tekiráka, together with a wordlist, is Tessmann 1930:475–485.

The collected vocabularies bear no significant relations to neighboring languages (Loukotka 1968:156; Adelaar & Muysken 2004:456).

There is a wordlist in Tessmann 1930:475–485, and Espinosa Pérez 1955:66–68 includes an extract of another wordlist collected by Ismael Barrio in 1936. Lev Michael has recently collected another wordlist from one of the last remaining speakers (p.c. 2010).

The language has been presumed extinct (Lewis 2009), but recently Lev Michael has located a few speakers or semi-speakers (p.c. 2010).

2.1.4. YURÍ [cby?]. The first mention of an ethnic group Yuri between the Putumayo and the Marañon Rivers is probably the conversation record of two Yuri in Ruiz de Quijano et al. (1765:228). The first report with linguistic data is Wallace 1853:510, which locates the Yuri at the Solimões River between the Içá (Putumayo) and Japurá (Caquetá) Rivers.

The vocabularies bear no significant resemblance to any other language in the region (Ortiz 1965; Loukotka 1968), and the possibility that Yurí and Ticuna are related (Nimuendajú 1977:62) still awaits explicit comparisons.

The only certain Yurí vocabularies are those collected by Wallace (1853:fold-out appendix) and Martius (1867:268–272), reproduced in Ortiz 1965:232–244. An additional wordlist which may or may not be Yurí (see below) is mentioned in Vidal y Pinell 1970:108, of which three words are reproduced there.

The language has not been sighted since the nineteenth century and therefore was suspected to be extinct (cf. Ortiz 1965). However, there are uncontacted peoples at the Rio Puré in Colombia, touching the historical territory of the Yurí. Lewis 2009 has the name Carabayo [cby] for an uncontacted group at the Rio Puré. Given the geographical proximity, the Rio Puré uncontacted groups (or one the groups, if there are several), are often suspected to be the descendants of the century-old Yurí (Trupp 1974; Patiño Rosselli 2000; Fabre 2005; Landaburu 2000:30). Vidal y Pinell (1970) makes the strongest version of this case and is the only author in a position to adduce linguistic evidence. In 1969, a brief episode of contact allowed the collection of a wordlist[4] (which presumably contains a fair number of misunderstandings) of an uncontacted Rio Puré group (Font 1969). Vidal y Pinell (1970:108), finds 21% cognation between the 1969 Rio Puré list and Yurí of Wallace (1853:fold-out appendix), as opposed to 0–8% to various surrounding Tucano and Arawak languages. This is taken to be decisive evidence ("parece acertada") that the 1969 Rio Puré list is Yurí. However, the three example comparisons given contain semantic and formal discrepancies, and even if the 21% figure is correct, it seems too large a divergence from the Yurí of a century and a half earlier. If the 1969 Rio Puré language nevertheless is Yurí, or a different language forming a small family with Yurí, then the Yurí (family) is still alive. Otherwise, the 1969 Rio Puré language is a non-extinct language attested in only a wordlist with no known relatives (and the Yurí language of Wallace 1853 may be presumed extinct).

If the entry for Carabayo [cby] of the Rio Puré turns out to be Yurí, then the number of speakers is estimated, from aerial observations, at 150 (Lewis 2009).

---

[4] Not seen by the present author since the publication in which it appears is very difficult to access—the relevant issue is missing from archives both in Leticia and Bogotá (p.c. Frank Seifart 2010).

2.2. AFRICA
2.2.1. KUJARGE [vkj]

Kujarge was first reported by Doornbos and Bender (1983:59–60) with a 100-word list, and this remains the only known sighting of the language.[5] It is based on reports by Doornbos (1981), who met speakers on two different occasions in 1981 near Foro Boranga, on the Sudan side of the Chad-Sudan border. The informant reported that the language is spoken in seven villages in Chad, near Jebel Mirra (11°45' N, 22°15' E)[6] and scattered among the Fur and Sinyar in the lower Wadi Azum valley.

The language was classified as Chadic, in the Mubi group, following 1979 personal communication between Marvin Lionel Bender and Paul Newman (Doornbos & Bender 1983:76).[7] However, Paul Newman does not remember the precise details of this classification, and in particular, whether it was based on a 100-word or 200-word list (p.c. Sept. 2006). When shown the 200-word list, Paul Newman sees Chadic elements in it, but does not want to commit to Kujarge being a Chadic language (p.c. Sept. 2006). Low numerals and pronouns, among other things, look very un-Chadic, so it is likely that the words in Kujarge shared with Chadic are loans from a Mubi group language (see Blench 2008 for a different interpretation of the same evidence).

The only published data are in a 100-word list in Doornbos and Bender 1983, which was taken from a 200-word list obtained by Paul Doornbos. The full 200-word list has been typed up by Paul Whitehouse and is available to interested linguists (Doornbos 1981).

The number of speakers was estimated at 1,000 in Doornbos and Bender 1983:59–60. Nothing further is known about its endangerment status.

2.3. EURASIA
2.3.1. NIHALI [nll]

The Nihali ethnic group was first reported in 1868 (Driem 2001:243–244), but the first language data are from Konow (1906).

It is clear that Nihali (also known as Nahali) has a heavy overlay from neighboring Munda, Dravidian, and Indo-Aryan languages, but some core vocabulary remains distinct. There have been many attempts, old and new, to relate this chunk of vocabulary to other families (cf. a special issue of *Mother Tongue* in 1996), but so far there is no convincing case for a relationship (Shafer 1940; Kuiper 1962, 1966; Driem 2001:242–253).

Published data on Nihali include long wordlists, a little textual material, and tiny amounts of grammar (Mundlay 1996; Bhattacharya 1957; Konow 1906).

---

[5] Names similar to Kujarge occur in Lebeuf 1959:116 and MacMichael 1918:45 but without any language data, and seemingly designating a different group or caste than Doornbos's Kujarge. (The name Kujarge is held to mean 'sorcerers.').

[6] Not to be confused with the more famous Jebel Marra, on the Sudan side.

[7] Also, the remark "I am assuming that […] Doornbos's Kujarke is Newman's Birgit, 1977:6" (Doornbos & Bender 1983:76) suggests that the classification is (partly) the result of a confused language identity. Comparing the actual language data shows that Newman's Birgit is not the same language as Doornbos' Kujarge, and Newman's Birgit is definitely an East Chadic language (Jungraithmayr 2004).

Speaker numbers generally range between 1,000 and 2,000 (Lewis 2009), mostly bilinguals, though it is possible that there are also a few monolinguals (Driem 2001:252). I know no further information about the language's endangerment status.

2.3.2. SHOM PEN [sii]. The first linguistic data were collected in 1876 (Man 1886:432). Shom Pen has been assumed[8] to be a Mon-Khmer language in the Nicobar group (Lewis 2009), but recently Blench (2007a) has argued convincingly that this affiliation is unjustified.

There is an old, short wordlist (Man 1886), along with some data on counting (Temple 1907), as well as a more recent, larger collection of words and short phrases (Chattopadhyay and Mukhopadhyay 2003), though there are question marks for the quality (Blench 2007a). Finally, a little more data appeared in Driem 2008. There are also some very brief ethnographic studies (Raja Ram 1960; Lal 1969; Haider 1994; Rizvi 1990).

Lewis (2009) gives a speaker number of 400 (from 2004), while Chattopadhyay and Mukhopadhyay (2003:3) estimate some 300 speakers and the general conditions indicate that the language is being transmitted to children.

2.4. PAPUA

2.4.1. AFRA [ulf]. First reported (with a wordlist) as Oeskoe by Galis (1956), whose information remained the only source for almost half a century.

Voorhoeve (1971) has Afra (under the name Usku) as "unclassified," by which he means that no significant lexical relations are found with its neighbors, or, in other words, it is a language isolate. In Voorhoeve 1975a, however, it is classified as Trans New Guinea, but no evidence or arguments were ever adduced. Hammarström (2008b) finds that any link to Trans New Guinea is premature.

Wordlists are published in Smits & Voorhoeve 1994 in addition to an SIL Indonesia survey report to appear, which contains 250 words and fifteen sentences (Im & Lebold 2006). There is also a brief anthropological report by Dumatubun & Wanane (1989).

At present, there are about 115 speakers, but the language is not immediately in danger. However, the younger generation is just as strong in Indonesian as it is in Afra (Im & Lebold 2006), which points to a weakening of the vernacular.

2.4.2. ASABA [seo]. The Asaba language was probably first reported (under the name Suarmin) by Healey (1964:108).

---

[8] I am not sure who the first author was to assume this. Man (1886:436) was fully aware of the distinctness of Shom Pen, as he writes "of words in ordinary use there are very few in the Shom Pen dialect which bear any resemblance to the equivalents in the language of the coast people." Nevertheless, Man, in his dictionary of Central Nicobarese (1889), includes an appendix with comparative wordlists of the six "dialects" of Nicobarese, where Shom Pen is not given any special position next to the five other Nicobarese varieties which are much more alike. Because of the particular use of the word 'dialect,' subsequent authors appear to assume that all six varieties are, if not even mutually intelligible, closely related, and file Shom Pen under "Nicobarese" (e.g., Anonymous 1908:66–68; Przyluski 1924:391). Even such a careful observer as Schmidt (1906:23) also calls Shom Pen a dialect without paying attention to its divergence from the other Nicobarese lects.

Typological arguments are not sufficient to conclude a Leonard Schultze family with Walio (Laycock & Z'Graggen 1975). The lexical evidence does not show any conclusive genetic relationship either, be it inside or outside Leonard Schultze (Conrad & Dye 1975), with Sepik-Hill (as suggested in Lewis 2009), or with Baiyamo (Papi) (Conrad & Lewis 1988). However, a higher figure (29%) of Baiyamo-Asabo (Papi-Duranmin) lexicostatistical relations was quoted by Laycock & Z'Graggen (1975:753) before the later revised lower figure (10%) posed by Conrad & Lewis (1988:259), and some lexical data recently collected by anthropologists do contain matches between the two. It remains to be determined whether these are loans or indicators of a genetic relationship.

There are some very brief notes on grammar in Laycock & Z'Graggen 1975. There are extensive anthropological studies of the people (Lohmann 2000; Little 2008).

There are about 180 speakers (Little 2008:2), and the language is still being transmitted to children (p.c. Roger Lohmann 2009).

2.4.3. BAIYAMO [ppe]. Baiyamo was first reported by Laycock (1973) as Papi (a village name).

As with Asaba, typological arguments are not sufficient to conclude a Leonard Schultze family with Walio (Laycock & Z'Graggen 1975). In addition, the lexical evidence does not show any conclusive genetic relationship either, be it inside or outside Leonard Schultze (Conrad & Dye 1975), with Sepik-Hill (as suggested in Lewis 2009), or with Asaba (as Duranmin) (Conrad & Lewis 1988). However, Laycock & Z'Graggen (1975:753) quoted a higher figure (29%) of Baiyamo-Asabo (as Papi-Duranmin) lexicostatistical relations, before the later superseding lower figure (10%) quoted by Conrad & Lewis (1988:259). Some lexical data collected recently by anthropologists does contain matches between the two. It remains to be determined whether these are loans or indicative of a genetic relationship.

There is a wordlist in Conrad & Dye 1975 and some very brief notes in Laycock & Z'Graggen 1975 (752–753).

Lewis (2009) cites seventy speakers. However, the language is still being transmitted to children (p.c. Jack Kennedy 2009).

2.4.4. BAYONO-AWBONO [byl, awh]. The missionary linguists Vries and Enk (1997:10) reported that northwest of the Korowai territory "… Ulakhin is spoken, a totally unknown language," but were unable to obtain any data on Ulakhin. Later, in 1998, Mark Donohue collected wordlists for Bayono and Awbono during a helicopter visit to the area northwest of the Korowai territory. Finally, an amateur wordlist of Urajin[9] was compiled by Hischier in 2006 from the same general area. The Urajin wordlist matches the Bayono and Awbono data, so that Urajin must be a language closely related to Bayono and Awbono, if not identical to one of them. Since the name and location of Urajin match Ulakhin, we assume that Ulakhin (Vries & Enk 1997:10) is indeed the first mention of a people of this language family.

Bayono and Awbono are not obviously related to any of their neighbors; however, very few linguists have actually seen the wordlists and made the comparisons, so there is an unusual amount of uncertainty surrounding the languages' classification. Nevertheless,

---

[9] I wish to thank Nicholas Ossart for bringing this wordlist to my attention.

Timothy Usher (p.c. 2010) has found a sizeable number of parallels between Bayono and Awbono and the Awyu-Dumut family (neighboring to the south) that may or may not prove to be due to inheritance from a common language.

The only data are the unpublished wordlists, including some sentences, of Mark Donohue and Phyllis Hischier (Hischier 2006), all of which were collected monolingually under challenging circumstances.

Speaker estimates are one hundred each for Bayono and Awbono (Lewis 2009). The languages are certainly being transmitted to children.

2.4.5. BUSA [bhf]. Busa was first reported by Loving & Bass (1964).

Busa lexicon bears no significant relation to any other language in the region (Laycock 1975a; Conrad & Dye 1975).

There is a wordlist in Conrad & Dye 1975 and some very brief notes on grammar in Laycock 1975a.

Lewis (2009) cites 240 (2000 census) speakers. In 1980, Busa was spoken by 238 people, and, though Tok Pisin usage was growing, Busa was not endangered (Graham 1981).

2.4.6. GURIASO [grx]. Guriaso was first reported in 1983 as a new language and named after a central village (Baron 1983:27). Previously, Guriaso and a few smaller villages were thought to speak (a dialect of) Kwomtari. It was subsequently grouped with Kwomtari on very low cognate counts (3%–13%) and shared typological features (Baron 1983:27–29). In my judgment of the same data, these resemblances can just as well be explained by chance.

The only data (basic lexical and grammatical data) appear to be the 1983 unpublished SIL Survey (Baron 1983) and five numerals in Lean 1986.

The most recent data put the speaker number at 160 (2003 SIL) as of Lewis 2009, and the language is endangered (p.c. Murray Honsberger, SIL-PNG via Ian Tupper, SIL-PNG 2007).

2.4.7. DEM [dem]. The first data from Dem appear in Le Roux 1950 (776–835), which remains the main source of information about the language.

Dem was grouped with neighboring highland languages based on lexicostatistical counts (Larson 1977), but these cognation judgments relied on the matching of only one segment, obviously yielding inconsistent sound correspondences. The lexicostatistic argument for relatedness is the only one offered so far, and apart from probable borrowings, I cannot find any cognates in vocabulary or morphology.

A rather extensive wordlist as well as sample sentences can be found in Le Roux 1950 (776–835). There is also a wordlist in Stokhof 1983 (219–221).

Lewis (2009) cites 1,000 speakers (1987 SIL), and the language is still strong in the community (p.c. Donohue 2008).

2.4.8. INLAND GULF [ipo, ffi, hhi, hhy, xar, mcv, tsx]. The first vocabulary of an Inland Gulf language to appear in print is the Karami wordlist of Flint (1918), and other government patrol wordlists appeared shortly thereafter (references in Franklin 1973:269–270).

Internally, the membership of the geographically non-adjacent Ipikoi in the family was realized only in the early 1970s (Franklin 1973:267–273). Evidence for a Trans New Guinea membership are the singular pronouns in the Minanibai branch and a few lexical items (Wurm 1975:509–510), and Ross (1995:152, 157) takes the pronoun evidence to be probative. However, the pronouns that look most like Trans New Guinea have not yet been shown to go back to proto-Inland Gulf, and even if we assume they are characteristic, the total evidence for a Trans New Guinea affiliation is very slight. Therefore, it would be premature to call Inland Gulf a branch of the Trans New Guinea family. No stronger cases for Inland Gulf affiliations to other (sub)families have been put forward.

There are wordlists and very scant grammar notes on Inland Gulf in Reesink 1976 and Franklin 1973 (269–273, 577–578, and references therein). Further, SIL survey data from 2006, including wordlists and some sentences, have not yet been published.

For all Inland Gulf languages except Karami and Mubami, Lewis (2009) cites speaker numbers of 80–200, emanating from the 2000 census or surveys two decades earlier. I have not had access to more recent (2006) survey data, but given their size and location, one can expect speakers to be under pressure. Mubami, on the other hand, has 1,730 speakers (SIL 2002) and is used in all age groups and represented in all domains, including education (Lewis 2009). The villages in which the Karami wordlist was collected have not been located again, and the language is presumed extinct (Franklin 1973:269).

2.4.9. KAPAURI [khp]. Probably the first mention of Kapauri as a separate language, along with language data, appears in Voorhoeve 1975b (45) and is based on a wordlist furnished by Myron Bromley. Voorhoeve (1975b) grouped Kapauri with the Kaure languages based on some lexical correspondences. However, a newer evaluation of the lexical relationships reveals figures that are only in the range of 5–6%, shedding considerable doubt on a genetic relation between the Kaure languages and Kapauri (Rumaropen 2006:13).

A short wordlist (forty words) appears in Voorhoeve 1975b. Two hundred fifty words and fifteen sentences are to appear in an SIL Indonesia survey report (Rumaropen 2006), which also mentions translated biblical texts. Unpublished survey reports referred to by Silzer & Heikkinen (1984:31) may contain further wordlists.

The number of native speakers is approximately 200, and transmission of the language from older to younger generations is regular. At present, Kapauri is not an endangered language (Rumaropen 2006).

2.4.10. KIMKI [sbt]. A tiny eleven-word list of what are probably Kimki lexical items (Hammarström 2008c) was collected as early as 1914 (Langeler 1915). Otherwise, references to Kimki go back no earlier than 1978 in unpublished SIL Indonesia surveys (Silzer & Heikkinen 1984; Silzer & Heikkinen-Clouse 1991).

The language was listed as "unclassified" (Silzer & Heikkinen 1984; Silzer & Heikkinen-Clouse 1991) until between 1996 and 2000, when Grimes (2000) grouped it with neighboring Yetfa-Biksi. However, the lexical evidence is not sufficient for concluding a genetic relation between the two (Hammarström 2008b). The only substantial body of data

is an unpublished 250-word list and fifteen sentences in an SIL survey report to appear (Rumaropen 2004).[10]

At present, Kimki is being transmitted to children and thus is not an endangered language (Rumaropen 2004).

2.4.11. LEPKI [lpe]. As with Kimki, a fifteen-word list of what are probably Lepki lexical items (Hammarström 2008c) was collected as early as 1914 (Langeler 1915). Otherwise, Lepki is first listed in Silzer & Heikkinen-Clouse 1991, presumably deriving from Doriot 1991.

Wherever it appears, the language is listed as "unclassified" (Lewis 2009; Silzer and Heikkinen-Clouse 1991). The label "isolate" is more appropriate, as the wordlist shows no significant relationship to any of the neighboring languages.

For Lepki, there exists an unpublished, hastily collected wordlist by Donohue (no date) and a few songs plus a short wordlist contained in an unpublished anthropological report (Andersen 2007). Doriot (1991) also refers to another unpublished wordlist.

Lewis (2009) cites 530 speakers (1991 SIL). Andersen (2007) counts exactly 328 Lepki speakers and records the clan membership of each one. Though there has been no investigation of language shift, one may suspect that Lepki is under pressure from Ketengban and Indonesian in recently founded villages, which have attracted many Lepki (Andersen 2007).

2.4.12. MAWES [mgk]. Mawes was first reported as a separate language in Robidé van der Aa 1879 (112) but without accompanying data. Likewise, Leeden (1954) noted the separate identity of the language, but no actual language data surfaced until the twenty words in Galis 1955 (118). Voorhoeve (1975b:40, 60) classified Mawes as a family-level isolate within his Tor-Lakes-Plain stock using unpublished lexical data (of which 40 words were published). This classification was retained in all later listings (e.g., Lewis 2009), but the Lakes Plain languages were later excised (Clouse 1997), pushing Mawes into a subfamily with Tor and Orya. To be a family-level isolate (Voorhoeve 1975b:16) within the Tor-Lakes-Plain stock means that the language "shares 12%–27% cognates on a 100-word list" with at least one other Tor-Lakes-Plain language. However, the cognate identifications supporting this classification were never published, and modern lexical data fail as evidence for it (Hammarström 2010a). Indeed, another independent count (Wambaliau 2006) shows Mawes cognate percentages never exceeding 6% with any Tor language (nor with any other language in the immediate region). Therefore, it seems best to consider Mawes an isolate (Hammarström 2010a).

A substantial wordlist was finally published in Smits and Voorhoeve 1994, of which twenty words had appeared in Galis 1955 and 40 words in Voorhoeve 1975b. Two hundred fifty words and fifteen sentences will appear in an SIL Indonesia survey report (Wambaliau 2006).

Though the speaker number is not low (ca 850), Mawes is under pressure from Indonesian and can be considered an endangered language (Wambaliau 2006).

---

[10] Doriot (1991) refers to an unpublished wordlist of Kimki from Mot, but Mot is listed in survey maps as Murkim-speaking (Wambaliau 2004)..

2.4.13. MOR (OF BOMBERAI) [moq]. Mor was first reported by Anceaux (1958) with a short wordlist.

Lexical and pronominal evidence for inclusion in Trans New Guinea is weak (Voorhoeve 1975a:431).

A wordlist can be found in Smits and Voorhoeve 1998, and it appears that Anceaux (1958) collected grammatical data. I searched the Anceaux Nachlass in June 2008 for these grammatical data but could locate only wordlists for Bomberai Mor.[11] Additionally, there are unpublished wordlists and some grammatical data in an SIL Indonesia survey (Walker 1983).[12] Finally, an unpublished 200-word list was compiled by Mark Donohue in February 2010.

Lewis (2009) cites twenty-five speakers, giving Stephen Wurm 2000 as the source. Since there is no record of Wurm having collected independent data for this region, the figure was presumably taken over from an earlier source, possibly an earlier *Ethnologue* edition. After a very brief trip to the edge of the traditional Mor area in February 2010, Mark Donohue established that there are at least a few Mor speakers in Goras and that an unknown but non-zero number of speakers are still to be found in the traditional Mor area (p.c. Donohue June 2010).

2.4.14. MURKIM [rmh]. Murkim was first reported in Silzer and Heikkinen-Clouse 1991 and is usually listed as an unclassified language (Silzer and Heikkinen-Clouse 1991; Lewis 2009). It has some lexical matches with neighboring languages, but they look more like loans than the outcome of genetic inheritance (Hammarström 2008b).

The only data are 250 words and fifteen sentences from an SIL survey not yet published (Wambaliau 2004).

At this time, Murkim is being transmitted to children and is thus not an endangered language (Wambaliau 2004).

2.4.15. NAMLA-TOFANMA [naa, tlg]. Tofanma was first reported (with wordlist) by Galis (1956) whose information remained the only source for almost half a century. Namla was considered by Galis (1956), Voorhoeve (1971), and Anceaux (no date) to be a separate language. Since there was no published or unpublished wordlists or other language data to be found from these authors, the language failed to appear in subsequent listings. Much later, Doriot (1991) "re-discovered" the language and compiled a wordlist, but since neither the wordlist nor the survey was published the language still failed to appear in listings. Namla was "discovered" a third time in an SIL Indonesia survey more than a decade later (Lee 2005).

The only known data indicate that Namla and Tofanma are genetically related, because there are good matches in the basic lexicon, which are arguably not loans (Hammarström 2008b). Voorhoeve (1971) lists Tofanma as "unclassified," meaning that no significant lexical relations are found with its neighbors. It is, in other words, a language isolate.

---

[11] The Anceaux Nachlass is held at KITLV manuscripts, Leiden, with the call number Or 615 (especially anvulling 4–23 is relevant for Mor).

[12] I wish to thank Mark Donohue for bringing my attention to this survey.

In Voorhoeve 1975a, however, it is classified as Trans New Guinea, but no evidence or arguments were ever adduced. Hammarström (2008b) finds any link to Trans New Guinea premature.

Published wordlists for Tofanma are collected in Smits and Voorhoeve 1994 and there are also a 250-word list and fifteen sentences in a survey report to appear (Wambaliau 2005). The only data for Namla are 250 words and eight sentences from an SIL survey (Lee 2005) and a wordlist by Doriot (1991), both unpublished.

At present, Tofanma (251 speakers) is still being learned by children and so is not an endangered language (Wambaliau 2005). Namla speaker numbers are twenty-five speakers in Galis 1956, twenty-seven in Anceaux's notebooks (no date), and thirty-six in Lee 2005. Early researchers found the Namla village deserted and suspected that an epidemic could explain the low number of speakers (Beer 1959). At present, all Namla speakers are bilingual in Tofanma. However, the youngest generations are not learning the Namla language fluently but, rather, are using Indonesian instead. Therefore, Namla is a highly endangered language (Lee 2005).

2.4.16. PYU [pby]. Pyu was first reported in the linguistic literature by Laycock (1972). Pyu was grouped in the Kwomtari-Baibai-Pyu phylum, but no real evidence for this classification was ever presented (Laycock 1975b). There are no significant lexical links with neighboring languages (Conrad & Dye 1975).

Two short wordlists were compiled by Conrad & Dye (1975) and Laycock (1972), and the latter included a sentence or two on grammar in a later work (1975b:854). Further, an unpublished 200-word list was collected by Arjen Lock in 1992.

Lewis (2009) cites a speaker number of 100 (2000 census), and more recent information suggests that the language is highly endangered. According to a 1992 report by Arjen Lock, the language is spoken only in the village of Biake 2 within its hamlets (north of the Sepik River and just east of the PNG-Indonesia border), and in an unlocated village on the bend of the Sepik within Indonesian territory. According to Lock's consultant (who came from Biake 2), "people who are over thirty years and older are bilingual in Abau and Pyu. The children are claimed to lack fluency in both Abau and Pyu. They prefer to communicate in Tok Pisin." Although Lock's data did not come from observations in the language area, it seems very plausible that the language is highly endangered (p.c. Tupper SIL-PNG Sept 2008).

2.4.17. SAUSE [sao]. Sause is mentioned as an ethnic group in early patrol reports, based on secondhand information (Hoogland 1939:7). Probably the first mention of Sause as a separate language, along with supporting language data (forty words), is in Voorhoeve 1975b (45), based on Anceaux's collection of wordlists.

Based on some lexical correspondences, Voorhoeve (1975b) grouped Sause with the Kapauri and Kaure languages. At some point, presumably on geographical grounds, the language started to be listed in the Tor-Lakes Plain stock (Silzer & Heikkinen-Clouse 1991:28–29). When the Lakes Plain languages were excised (Clouse 1997), Sause remained a Tor-related language (Lewis 2009), but the lexical data available do not support this classification.

The only published source of data on Sause is a wordlist in Smits and Voorhoeve 1994, of which forty words appear in Voorhoeve 1975b. Unpublished data include a minuscule wordlist collected by Donohue from a transient speaker (p.c. Aug 2008), as well as wordlists presumably contained in survey reports referred to in Silzer & Heikkinen-Clouse 1991:74.

Lewis (2009) cites 250 speakers. I know nothing further about the endangerment status of Sause.

2.4.18. TANAHMERAH (OF BOMBERAI) [tcm]. As far as I can tell, Bomberai Tanahmerah was first reported by Galis (1955). Suggested links with Mairasi based on lexical and pronominal comparisons are unconvincing (Voorhoeve 1975a:424–431).

There are some very brief notes on Bomberai Tanahmerah grammar in Voorhoeve 1975a:424–431 and numerals in Galis 1960. Very short wordlists were published in Galis 1955 and Anceaux 1958, though a slightly longer one appears in Smits and Voorhoeve 1998. A note on p. 18 of the latter makes it clear that additional grammatical data were collected by Anceaux. I searched the Anceaux Nachlass for these grammatical data in June 2008 but could not locate anything beyond wordlists for Bomberai Tanahmerah.[13] It is likely that Lloyd Peckham, an SIL member working with the nearby Mairasi languages, has newer data for Tanahmerah (p.c. Mark Donohue 2008). In addition, unpublished wordlists and some grammatical data can be found in an SIL Indonesia survey conducted by Walker in 1983.[14]

Lewis (2009) cites 500 (SIL 1978) speakers. I have not had access to any other information, such as the 1983 survey data.[15]

2.4.19. WALIO [ppq,tww,wla,ybx]. The Walio languages were probably first reported by Healey (1964:108).

Typological arguments are not sufficient to conclude a Leonard Schultze family with Baiyamo (Papi) (Laycock & Z'Graggen 1975:752–753; Laycock 1973:32–33). The lexical evidence does not show any conclusive genetic relationship either, be it inside or outside Leonard Schultze (Conrad & Dye 1975; Conrad & Lewis 1988).

There are published wordlists in Conrad & Dye 1975 and in Conrad & Lewis 1988, numerals in Lean 1986, and some very minimal notes on grammar in Laycock & Z'Graggen 1975 (752–753) and Laycock 1973 (32–33). Saito 1998 contains ethnographic data on Yabio.

Lewis (2009) cites speaker numbers of 50–360 emanating from the 2000 census, and Saito (1998:95) cites 122 Yawiyo speakers in 1988. I know of no further information on endangerment status.

---

[13] The Anceaux Nachlass is held at KITLV manuscripts, Leiden, with the call number Or 615 (especially anvulling 4–23 is relevant for Tanahmerah).

[14] I wish to thank Mark Donohue for bringing my attention to this survey.

[15] However, the unpublished survey report *Roland Walker and Michael Werner 1978 Bomberai Survey Report MS, SIL Indonesia* contains no original information, as the survey team did not travel to the Tanahmerah area (p.c. Walker Aug 2008).

2.4.20. YETFA/BIKSI [yet]. Biksi was first reported in the linguistic literature by Laycock (1972), who had met with transients from Papua, Indonesia (then West Irian) while doing fieldwork on the Papuan (then Australian) side in 1970. Yetfa is mentioned for the first time in the 2nd edition of the *Index of Irian Jaya Languages* (Silzer & Heikkinen-Clouse 1991) and, without any references to data, is labeled as an unclassified language. This information presumably derives from Doriot (1991) who trekked in parts of the Yetfa-speaking area from April–May 1991. Some time between the 14th (Grimes 2000) and 15th (Gordon 2005) editions of *Ethnologue*, it was realized that Yetfa and Biksi are similar enough to be regarded as one language.

Biksi (Biksi-Yetfa) was placed in the Sepik language group by Laycock & Z'Graggen (1975:740–741), and this has often been repeated since (Lewis 2009). Due to lack of data, Biksi-Yetfa was not considered by Foley in his re-assessment of the Sepik family (Foley 2005:126–127). The lexical matches adduced by Laycock for various Sepik languages are sporadic and look more like loans or chance resemblances than the results of genetic inheritance (Hammarström 2008b). In addition, the lexical relations were investigated independently by Conrad & Dye (1975:19), who found that Biksi shared no more than 4% probable cognates with any of the surrounding languages to the east, including Abelam.[16] (This lexical comparison includes numerals, but no demonstratives or pronouns.) For the alleged connection with Kimki, e.g., Lewis 2009, see the section on Kimki above.

Scanty notes on Biksi grammar can be found in Laycock & Z'Graggen 1975:740–741, and short wordlists are published in Laycock 1972 and Conrad & Dye 1975. An unpublished SIL survey of Indonesia contains 250-word lists for Yetfa from five locations and 15 sentences in Yetfa (Kim 2006). There are additional unpublished wordlists from several locations collected by Doriot (1991), and missionaries (SIL or UFM) in the area may have even more unpublished materials from the past decade, but I do not know their extent.

At this time, Yetfa is still being transmitted to children and so is not an endangered language (Kim 2006).

3. UNCLEAR CASES
3.1. SOUTH AMERICA. In his 2005 work, Moore mentioned in passing an unpublished wordlist for Arara do Rio Branco [axg?] collected by Inês Hargreaves. This wordlist is now posted on the internet (Hargreaves 2007), and a rechecked version appears in Souza 2008. The majority of items on the wordlist do not have cognates in neighboring families. Notably, Arara do Rio Branco is prima facie neither a Tupí language (as suggested by its location and some matches–presumably loans–with the Mondê group) nor an Arawak language (as suggested by one pronominal resemblance). There were four rememberers left in 2001. In 2008, there were only two (Souza 2008), neither of which is a fluent speaker.

Quite a lot of data, perhaps enough for a basic grammar sketch, were collected by Tastevin on Katawixi [xat] (Anjos 2005; Adelaar 2007), and these data are enough to support the claim that Katawixi is related to Harakmbut, Katukina, or both (Adelaar 2007).

---

[16] The exact languages in question are Yerakai (0%), Chenapian (0%), Bahinemo (1%), Washkuk (1%), Yessan-Mayo (4%), Abelam (1%), Namie (0%), Abau (0%), May River Iwam (1%), Musan (0%), Amto (1%), Rocky Peak (0%), Ama (0%), Nimo (1%), Bo (0%), Iteri (0%), Owiniga (2%), Woswari (0%), Walio (0%), Paupe (0%), South Mianmin (0%), Nagatman (0%), Busan (1%), Pyu (1%).

Except for a possibly Katawixi-speaking isolated group, the language appears to be extinct (Queixalós & Anjos G.S. 2007:29).

The Maco [wpc] language of the Lower Ventuari, Venezuela,[17] is known only from some thirty-eight words collected a century or longer ago. It is occasionally listed as an independent "not definitively classified" language (Mosonyi 2003:109) or as an "unclassified" language (Mattei-Müller 2006:295). However, further analysis (Hammarström 2010b) of the thirty-eight words does allow the conclusion that it is a dialect of Piaroa or a closely related language (as had been assumed already, on more meager evidence, by many authors, following Koch-Grünberg (1913:468–469)).

The Máku [iso-639-3 code lacking] language isolate (Migliazza 1985) is reasonably well-documented in accessible documents (see Maciel 1991, Migliazza 1966, 1965 and references therein), and enough data for a 300-page grammar have been collected (p.c. Zamponi 2006). Also, the language is extinct, as the last speaker died sometime between 2000 and 2002 (p.c. Zamponi 2006).

The isolate Taruma [iso-639-3 code lacking] was believed to be extinct, but the last speakers have been located by Eithne Carlin, who is actively working with them to document it (p.c. Adelaar 2009).

3.2. AFRICA. Oropom [iso-639-3 code lacking] of Uganda is mentioned together with a wordlist in Wilson 1970, but the language (and ethnic group) is otherwise previously wholly unattested and could not be located in the area in the 1970s (Souag 2004:1). This has led to the suspicion that the 97-item long wordlist may be spuriously elicited forms rather than the last pieces of a lost language (Fleming 1987:203). However, even if so, it is both extinct and has too few non-Nilotic elements to be considered an isolate (Souag 2004).

The Mpra (= Mpre) [iso-639-3 code lacking] language in Ghana has lexical items with cognates in Atlantic-Congo, as well as lexical items without plausible cognates (Goody 1963). Regardless of the genetic status of the language, the language is practically extinct (Blench 2007b).

A number of language families often subsumed under "Nilo-Saharan" have documentation statuses bordering that of a grammar sketch. I list the most unclear cases (only the most extensive data collections are mentioned):

Birri [bvq]: The Birri language, even though clearly reported as a distinct language in Santandrea 1950 and 1964, was not listed in overview works such as Tucker & Bryan 1956 and 1966 and Greenberg 1966 and 1971. Wherever it appears later it is subsumed under Central Sudanic (Lewis 2009). With a more modern understanding of Central Sudanic, it is not clear whether Birri is a bona fide Central Sudanic language or not (Boyeldieu 2010). In any case, a short grammar sketch of Birri exists (Santandrea 1966).

Daju [byg, djc, daj, dau, njl]: A grammar and dictionary of Daju-Eref by Pierre Palayer is forthcoming (p.c. Pascal Boyeldieu 2007).

---

[17] There are a multitude of languages/ethnic groups referred to with a name similar to Maco (Hammarström 2010b). The Maco language discussed here is the one defined by the vocabularies furnished by Vráz, Koch-Grünberg (Loukotka 1949:56–57), and Humboldt (1825:V7:155–157).

Eastern Jebel [soh, xel, zmo, tbi]: Enough data for a sketch have been collected (Stirtz 2006).

Kresh group [krs, aja]: The Kresh group is often subsumed under Central Sudanic, but with a more modern understanding of Central Sudanic, it is not clear whether the Kresh group is a bona fide Central Sudanic subfamily or not (Boyeldieu 2010; Boyeldieu & Nougayrol 2008). In any case, there are enough data for a grammar sketch of at least Kresh-Gbaya [krs] (Struck 1930; Santandrea 1976; Brown 1994).

Shabo [sbf]: There is a grammar sketch, though admittedly brief (Teferra 1991), and the language is being further documented by Tyler Schnoebelen (Stanford University).

Tama [mgb, sjg, tma]: Three (admittedly brief) sketches are available (Lukas 1938, 1933; Kellermann 2000), and Edgar 1989 contains some very brief grammar notes on Miisiirii. Gerrit Dimmendaal (Cologne University) has field data on Tama from the 2000s that are sufficient for a grammar sketch (Dimmendaal 2009).

Temeinan [teq, keg]: A phonological description (Yip 2004) as well as collections by Stevenson (Blench 2006) give the impression that enough data for a sketch have been collected.

The Nuba mountains outlier language Warnang [wrn] is counted here as a divergent Heiban language (Schadeberg 1981), since this affiliation has been cleared by Nicolas Quint (p.c. 2010).

Robin Thelwall collected a fair amount of grammatical data on Tegem/Lafofa [laf] (p.c. May 2008).

There are good reasons to believe that substantial amounts of data have been—and are being—collected for the Mao [myf, gza, hoz, sze] languages (Yimam 2007, 2006; Dumessa 2007; Ahland 2009), and there is an unpublished, rudimentary Ganza grammar sketch (Hammarström 2008a).

Stefan Elders collected enough data on Bangeri Me [dba] for at least a grammar sketch (cf. Elders 2006) before he died, and this is posted online at http://www.dogonlanguages.org/ (accessed 29 Sept 2008). In addition, Hantgan (2010) is now conducting further documentation of the language.

Jalaa [cet] is now presumably extinct (Kleinewillinghöfer 2001).

Smith (1897) first mentions Dūmē [iso-639-3 code lacking], including a tiny vocabulary, which is not obviously related to any neighboring language. Later surveys of the region have failed to find any trace of Dūmē, and scholars have therefore regarded the original information as suspect (Jensen 1952:57–58; Haberland & Jensen 1959; Haberland 1962:142). In any case, if it is genuine, Dūmē must be presumed extinct.

3.3. EURASIA. It has recently been suggested that Jiamao [jio] is a language isolate with heavy Hlai overlay (Norquest 2007). However, if the theory of massive borrowing from Hlai into Jiamao is in fact correct, so little residual vocabulary remains for Jiamao that it appears insufficient for positing a new language isolate.

3.4. PAPUA. There is too little data on Kehu to say whether it is to be considered an isolate or related to some other language(s), since the only known wordlists contain, for example, neither numerals nor pronouns (Whitehouse: no date; Moxness 1998).

I have had no access to data on Kembra [xkw] (Doriot 1991), Yerakai [yra] (Laycock 1973), or the Mongol-Langam [lnm, mgt, yla] languages (Laycock 1973) and insufficent access to data on Dibiyaso [dby], Doso [dol], and Turumsa [tqm] (Tupper 2007)[18] and therefore cannot vouch for their statuses either as isolates or relatives of some other language(s).

Foau [flh] and Diebroud [tbp] are related to the Lakes Plain languages (Clouse 1997), though not in an obvious way (p.c. Mark Donohue 2008).

Karkar-Yuri [yuj] was, for a long time, believed to be a language isolate by researchers working on the Papua New Guina side of the border (Laycock 1975a:882; Loving and Bass 1964). Similarly, the Eastern Pauwasi languages Zorop [wfg] and Emem [enr], on the Indonesian side, were not thought to be related to any of the neighboring languages in the Upper Sepik region across the border (Voorhoeve 1975a:418–419, 1971). However, Timothy Usher, after inspecting the wordlists (Whitehouse 2006), spotted the close relationship that in fact exists between Karkar-Yuri and the Eastern Pauwasi languages—Emem and Karkar-Yuri may even constitute a dialect continuum.[19] Therefore, since there are extensive materials on Karkar-Yuri (Price 1987; Price et al. 1994; Rigden: no date), the family counts as reasonably well-documented. (Several lexical cognates suggest that the Western Pauwasi languages Tebi [dmu] and Towei [ttn] form a bona fide family with the Eastern Pauwasi languages and Karkar-Yuri, but it is not impossible that the links are loans and that we are dealing with two (or three) small families with a lot of interaction (Hammarström 2008b)).

Mark Donohue has collected enough data for sketches of Abinomn [bsa], Powle-Ma [msl], Tanglapui [swt] (of the Kolana-Tanglapui family/subfamily), Masep [mvs], Elseng [mrf], Moraori [mok], Yoke [yki][20], and Damal [uhn] (p.c. 2008). In addition, Arka (2010) is currently documenting Moraori [mok].

SIL members and other researchers have compiled grammar sketches, or at least enough materials to produce them, for the following language families (only the most extensive data collections are mentioned):

> **Amto-Musan [amt, mmp]:** Linda Krieg et al. of the New Tribes Mission is in
>   the process of translating the Bible into Siawi, also known as Musan, and

---

[18] Doso and Turumsa have 61% lexicostatistical similarity and Turumsa and Dibiyaso have 19%, which, pace known caveats, indicates that the three form a small (sub-)family (Tupper 2007).

[19] It may be that this piece of information was known earlier to non-linguists. A handwritten map accompanying Hoogland 1940 shows Enam (= Emem) territory, which extends across the border into what is now labelled as Karkar-Yuri territory (Lewis 2009). Similarly, the anthropologist Hanns Peter seems to have been well-aware that a dialect continuum streches from Karkar-Yuri across the border into Papua, Indonesia, into what is now known to be Emem-speaking territory (Peter 1974:28–31), but he appears not to have made an explicit case that Yuri and Eastern Pauwasi are related.

[20] In any case, it is expected that this data will show that Yoke is Austronesian.

so far, has produced unpublished phonemic and grammar sketch write-ups (p.c. Linda Krieg 2007).

Arafundi [afd, afk, afp]: Notes on Arafundi are included in the writings of William Foley (1986, 2006), who presumably has extensive fieldnotes. Darja Hoenigman (Australian National University) is writing an anthropological-linguistic dissertation on the Awiakay variety. An SIL survey report was also produced in 2005 (p.c. Ian Tupper 2005).

Awin-Pa [awi, ppt]: There are some very brief grammar notes in Voorhoeve 1975a and a full New Testament (Stewart 1987).

East Strickland [agl, goi, kxw, jko, kkc, smq]: Britten Årsjö of the SIL has prepared an extensive unpublished grammar sketch of Konai (p.c. 2009).

Gogodala [aac, ggw, wrv]: There is a New Testament translation (Partridge 1981).

Kamula [xla]: An extensive grammar sketch has now been posted on the internet (Routamaa 1994).

Kayagar [aqm, kyt, tcg]: Unpublished, short grammar sketches are referenced in Silzer & Heikkinen-Clouse 1991.

Kol [kol]: Stellan and Eivor Lindrud produced unpublished manscripts (Akerson & Moeckel 1992), and the publication of a New Testament translation is due soon (p.c. Lindrud 2006). Furthermore, Ger Reesink has collected a few weeks-worth of grammatical data (p.c. 2010).

Konda-Yahadian [knd, ner]: Enough data for a rudimentary grammar sketch has been collected by SIL Indonesia members (Berry & Berry 1987).

Pahoturi [kit, idi]: An unpublished rudimentary (twenty-page) grammar sketch of Idi is stored in the SIL archives (No Author Stated: no date), and raw data collected by Wurm (1971) in 1966 and 1970 may be enough for a grammar sketch.

Pele-Ata [ata]: A dictionary (Hashimoto 2008) as well as the unpublished "Ata grammar essentials" are stored in the SIL (Ukarumpa) archives. Tatsuya Yanagida (Australian National University) is currently writing a PhD dissertation on Pele-Ata (Yanagida 2004).

Piawi [tmd, pnn]: There is an unpublished grammar sketch of Pinai (Melliger 2000), as well as published works on the grammar of Haruai (e.g., Comrie 1991).

Porome [prm]: Martin Steer (Australian National University) is currently writing a PhD dissertation on the language.

Somahai [mmb, mqf]: Martha Reimer (1986) has collected a large amount of data for Momuna, of which little has been published.

Suki [sui]: There is a New Testament translation (Bidri et al. 1981).

Tirio [aob, bmz, mcc, aup, wei]: An unpublished SIL survey from the 2000s contains collective sociolinguistic, lexical, and grammatical data (p.c. Tupper 2007). The raw data collected by Wurm in 1966 and 1970 may be enough for a grammar sketch (Wurm 1971). Ray (1923:360) mentions a

Tirio grammar manuscript by the Reverend Riley of unknown size and location.[21]

Yalë [nce]: An unpublished grammar sketch was compiled by SIL missionaries (Campbell & Campbell 1987).

Yuat [bwm, buv, cga, kql, myd, mvk]: Unpublished notes are contained in the Mead/Fortune fieldnotes (McDowell 1991:23). James McElvenny (Sydney University) did two months of fieldwork on Mudukumo and has written a draft grammar sketch (p.c. 2008).

Waia [knv]: There is an unpublished grammar (2004) in the SIL archives (p.c. Schlatter 2006). Translations of the New Testament have appeared in both the Aramia River (No Author Stated 2006a) and Fly River dialects (No Author Stated 2006b).

4. CONCLUSION. In this paper, I have presented a survey of the documentation of the language families of the world. I have listed all the known least documented language families of the world that are not yet known to be extinct. I have also included borderline cases, unclear cases, and cases for which there exist little-known data (which are frequently unpublished). I hope that this compilation will be useful for setting priorities in documentation fieldwork, especially for those documentation efforts whose underlying goal is a better understanding of linguistic diversity.

REFERENCES

Adelaar, Willem F. H. 2007. Ensayo de clasificación del Katawixí dentro del conjunto Harakmbut-Katukina. In Romero-Figueroa, Andres, Ana Fernández-Garay & Ángel Corbera Mori (eds.), *Lenguas indígenas de América del Sur: Estudios descriptivo-tipológicos y sus contribuciones para la lingüística teórica*, 159–169. Caracas: Universidad Católica Andres Bello.

Adelaar, Willem F. H. & Pieter C. Muysken. 2004. *The languages of the Andes*. Cambridge Language Surveys. Cambridge University Press.

Ahland, Michael. 2009. Aspects of Northern Mao phonology. *Linguistic Discovery* 7(1). 1–40.

Akerson, Paula & Bonita E. R. Moeckel. 1992. *Bibliography of the Summer Institute of Linguistics Papua New Guinea Branch 1956–1990*. Ukarumpa, Eastern Highlands Province, Papua New Guinea: Summer Institute of Linguistics.

Amodio, Emanuele. 2007. La república indígena. Pueblos indígenas y perspectivas políticas en Venezuela. *Revista Venezolana de Economía y Ciencias Sociales* 13(3). 175–188.

Anceaux, Johannes Cornelis. 1958. Languages of the Bomberai Peninsula. *Nieuw-Guinea Studiën* 2. 109–121.

---

[21] I could not find any further clues about the Tirio grammar manuscript in the Nachlass of Sidney H. Ray at SOAS Library (Aug 2008).

Anceaux, Johannes Cornelis. (no date). District Jafi-Jamas. KITLV Manuscripts and Archives, Leiden [Or 615 Anvulling 12].

Andersen, Øystein Lund. 2007. The Lepki People of Sogber [sic!] River, New Guinea. Unpublished.

Anjos, Zoraide dos. 2005. Fonologia Katukina (Dialeto Katukina do Biá). Brasília: Universidade de Brasília master's thesis.

Anonymous. 1908. Nicobars. In William Stevenson Meyer, Richard Burn, James Sutherland Cotton & Herbert Hope Risley (eds.), *Nāyakanhatti to Parbhani*. The Imperial Gazetteer of India XIX, 59–84. 3rd edn. Oxford: Clarendon Press.

Arka, I Wayan. 2010. Projecting Morphology and Agreement in Morori. Paper presented at the Workshop on the Languages of Papua 2. Manokwari, Indonesia, 8–12 February 2010.

Baron, Wietze. 1983. Kwomtari Survey. Unpublished manuscript, SIL Survey office, Ukarumpa. http://www.kwomtari.net/kwomtari_survey.pdf (15 Dec, 2008.)

Beer, H. de. 1959. Patrouilleverslag naar de Meervlakte. Nationaal Archief, Den Haag, Ministerie van Koloniën: Kantoor Bevolkingszaken Nieuw-Guinea te Hollandia: Rapportenarchief, 1950–1962, nummer toegang 2.10.25, inventarisnummer 12.

Berry, Keith & Christine Berry. 1987. A survey of the South Bird's Head Stock. *Workpapers in Indonesian Languages and Cultures* 4. 81–117.

Bhattacharya, Sudhibhushan. 1957. Field-Notes on Nahāli. *Indian Linguistics* 17. 245–258.

Bidri, Midim, Ivy Lindsay & Grahame Martin. 1981. *Godte gi amkari titrum ine [Suki New Testament]*. Port Moresby: Bible Society Papua New Guinea.

Blench, Roger. 2007a. The language of the Shom Pen: a language isolate in the Nicobar islands. *Mother Tongue* XII. 179–202.

Blench, Roger. 2007b. Recovering data on Mpra [=Mpre] a possible language isolate in North-Central Ghana. Draft Manuscript March 10, 2007.

Blench, Roger M. 2006. Temein, Tese and Keiga Jirru wordlists from R. C. Stevenson's Nachlass. Ms.

Blench, Roger M. 2008. Links between Cushitic, Omotic, Chadic and the position of Kujarge. Paper presented at the 5th International Conference of Cushitic and Omotic languages. Paris.

Boyeldieu, Pascal. 2010. Evaluating the genetic unity of Central Sudanic: Lexical and morphological evidence. Paper presented at the Workshop on Genealogical Classification of African Languages Beyond Greenberg. Berlin, 21–22 February 2010.

Boyeldieu, Pascal & Pierre Nougayrol. 2008. Les langues soudaniques centrales: Essai d'évaluation. In Bechhaus-Gerst, Marianne, Dymitr Ibriszimow & Rainer Vossen (eds.), *Problems of linguistic-historical reconstruction in Africa* (Sprache und Geschichte in Afrika: Beiheft 19). Köln: Rüdiger Köppe.

Brackelaire, Vincent & Gilberto Azanha. 2006. Últimos pueblos indígenas aislados en América Latina: Reto a la supervivencia. In *Lenguas y tradiciones orales de la Amazonía. ¿diversidad en peligro?*, 313–367. La Habana: Casa de las Américas.

Brenzinger, Matthias. 2007. Language endangerment throughout the world. In Brenzinger, Matthias (ed.). *Language diversity endangered*. Trends in Linguistics: Studies and Monographs, ix–xviii. Berlin: Mouton de Gruyter.

Brown, D. Richard. 1994. Kresh. In Peter Kahrel & René van den Berg (eds.), *Typological studies in negation*. Typological studies in language 29, 163–189. Amsterdam: John Benjamins.

Brown, Keith (ed.). 2006. *Encyclopedia of language and linguistics*. 2nd edn. Amsterdam: Elsevier.

Brown, Keith & Sarah Ogilvie (eds.). 2009. *Concise encyclopedia of languages of the world*. Amsterdam: Elsevier.

Cabodevila, Miguel Ángel. 2007. *El exterminio de los pueblos ocultos*. 2nd edn. Ecuador: CICAME.

Campbell, Carl & Jody Campbell. 1987. Yade grammar essentials. Ukarumpa: Summer Institute of Linguistics. Unpublished ms.

Chattopadhyay, Subhash Chandra & Asok Kumar Mukhopadhyay. 2003. *The Language of the Shompen of Great Nicobar: A preliminary appraisal*. Kolkata: Anthropological Survey of India.

Clouse, Duane. 1993. Languages of the Western Lakes Plains. *Irian* XXI. 1–31.

Clouse, Duane A. 1997. Toward a reconstruction and reclassification of the Lakes Plain languages of Irian Jaya. In Karl J. Franklin (ed.), *Papers in Papuan Linguistics*, No. 2, Pacific Linguistics: Series A 85, 133–236. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Comrie, Bernard. 1991. How much pragmatics and how much grammar: The case of Haruai. In Jef Verschueren (ed.), *Pragmatics at Issue*, 81–92. Amsterdam: John Benjamins.

Conrad, Robert J. & T. Wayne Dye. 1975. Some language relationships in the Upper Sepik region of Papua New Guinea. *Papers in New Guinea Linguistics* 18. Pacific Linguistics: Series A 40, 1–35. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Conrad, Robert J. & Ronald K. Lewis. 1988. Some language and sociolinguistic relationships in the Upper Sepik region of Papua New Guinea. *Papers in New Guinea Linguistics* 26. Pacific Linguistics: Series A 76, 243–273. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Coppens, Walter. 1976. Breve vocabulario: Sape y Arutani. Fundación La Salle, Caracas. Ms.

Coppens, Walter. 1983a. Los Sapé. In Walter Coppens (ed.), *Los Aborigenes de Venezuela,* Vol II (Monografía / Fundación la Salle 29), 381–406. Caracas: Fundación la Salle.

Coppens, Walter. 1983b. Los Uruak (Arutani). In Walter Coppens (ed.), *Los Aborigenes de Venezuela,* Vol II (Monografía / Fundación la Salle 29), 407–426. Caracas: Fundación la Salle.

Crevels, Mily & Hein van der Voort. 2008. The Guaporé-Mamoré region as a linguistic area. In Pieter Muysken (ed.), *From linguistic areas to areal linguistics*. Studies in Language Companion Series 90, 151–179. Amsterdam: John Benjamins.

d'Almada, M. de Gama Lobo. 1861 [1787]. Descrição Relativa ao Rio Branco e seu Território (1787). *Revista Trimensal do Instituto Historico, Geographico e Etnographico do Brasil* XXIV. 617–683.

Dimmendaal, Gerrit J. 2009. Tama. In Gerrit J. Dimmendaal. (ed.), *Coding Participant Marking: Construction Types in Twelve African Languages*. Studies in Language Companion Series 110, 305–330. Amsterdam: John Benjamins.

Donohue, Mark. (no date). Lepki. Hurriedly filled-in SIL-Indonesia 1998 wordlist form.

Doornbos, Paul. 1981. Kujarge. Field notes on Kujarge, language metadata, 200-word list plus numerals and pronouns.

Doornbos, Paul & Lionel M. Bender. 1983. Languages of Wadai-Darfur. In Marvin Lionel Bender (ed.), *Nilo-Saharan language studies* (Monograph / Committee on Northeast African studies 13), 43–79. East Lansing: African Studies Center, Michigan State University.

Doriot, Roger E. 1991. 6-2-3-4 Trek, April-May, 1991. Ms.

Driem, George van. 2001. *Languages of the Himalayas*. Handbuch der Orientalistik: Section Two: India 10. E. J. Brill. 2 Vols.

Driem, George van. 2008. The Shompen of Great Nicobar Island: new linguistic and genetic data. *Mother Tongue* XIII. 227–247.

Dumatubun, A. E. & Teddy K. Wanane. ca 1989. *Orang Usku di daerah batas Timur, Senggi, Irian Jaya*. Jayapura: s.n.

Dumessa, Alemayehu. 2007. Word Formation in Diddessa Mao. Addis Ababa: Addis Ababa University master's thesis.

Eberhard, David M. 2009. Mamaindê Grammar: A Northern Nambikwara language and its cultural context. Amsterdam: Vrije Universiteit doctoral dissertation.

Edgar, John. 1989. *A Masalit grammar: With notes on other languages of Darfur and Wadai*. Sprache und Oralität in Afrika: Frankfurter Studien zur Afrikanistik 3. Berlin: Dietrich Reimer.

Elders, Stefan. 2006. Présentation du bangeri me. Atélier sur le projet dogon, vendredi 8 décembre 2006, Bamako.

Espinosa Pérez, Lucas. 1955. *Contribuciones lingüísticas y etnograficas sobre algunos pueblos índigenas del amazonas peruano*. Madrid: Bernardino de Sahagun.

Evans, Nicholas. 2009. *Dying words: Endangered languages and what they have to tell us*. John Wiley & Sons.

Eyzaguirre Beltroy, Carlos Francisco, Neptalí Cueva Maza, Roberto Basilio Quispe Vilca & Graciela Sánchez Navarro. 2008. *Norma y Guías Técnicas en Salud: Indígenas en aislamiento y contacto inicial*. Lima: Ministerio de Salud.

Fabre, Alain. 2005. Diccionario Etnolingüístico y guía Bibliográfica de los Pueblos Indigenas Sudamericanos. Book in Progress at http://butler.cc.tut.fi/~fabre/BookInternetVersio/Alkusivu.html (May 2005)

Fleming, Harold C. 1987. Review article: Towards a definitive classification of the world's languages: Review of *A guide to the world's languages* by Merritt Ruhlen. *Diachronica* 4. 159–223.

Flint, L. A. 1917–1918. Vocabulary: Name of tribe, Karami. People. Name of village, Kikimairi and Aduahai. *Commonwealth of Australia*. *Papua: Annual Report for the Year* 1917–1918. 96–96.

Foley, William. 2006. Universal constraints and local conditions in Pidginization: Case studies from New Guinea. *Journal of Pidgin and Creole Languages* 21(1). 1–44.

Foley, William A. 1986. *The Papuan languages of New Guinea*. Cambridge Language Surveys. Cambridge: Cambridge University Press.

Foley, William A. 2005. Linguistic prehistory in the Sepik-Ramu Basin. In Andrew Pawley, Robert Attenborough, Jack Golson & Robin Hide (eds.), *Papuan pasts: Studies in the cultural, linguistic and biological history of the Papuan-speaking peoples*. Pacific Linguistics 572, 109–144. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Font, Antonio. 1969. Breve Vocabulario. *Amanecer Amazónico* XVI(745). 16–17. 31 Mayo.

Forline, Louis Carlos. 1997. The persistence and cultural transformation of the Guajá indians: Foragers of Maranhão State, Brazil. Gainesville: University of Florida doctoral dissertation.

Franklin, Karl J. 1973. Other language groups in the Gulf District and adjacent areas. In Karl J. Franklin (ed.), *The linguistic situation in the Gulf District and adjacent areas, Papua New Guinea*. Pacific Linguistics: Series C 26, 263–277. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Fuchs, Helmuth. 1967. Urgent Tasks in Eastern Venezuela. *Bulletin of the International Committee on Urgent Anthropological Ethnological Research* 9. 69–103.

Galis, Klaas Wilhelm. 1955. Talen en dialecten van Nederlands Nieuw-Guinea. *Tijdschrift Nieuw-Guinea* 16. 109–118, 134–145, 161–178.

Galis, Klaas Wilhelm. 1956. *Ethnologische Survey van het Jafi-district (Onderafdeling Hollandia)*, vol. 102. Hollandia [Jayapura]: Gouvernement van Nederlands Nieuw-Guinea, Kantoor voor Bevolkingszaken.

Galis, Klaas Wilhelm. 1960. Telsystemen in Nederlands-Nieuw-Guinea. *Nieuw Guinea Studien* 4(2). 131–150.

Goody, Jack R. 1963. Ethnological notes on the distribution of the Guang languages. *Journal of African Languages* 2(3). 173–189.

Gordon, Raymond G. Jr. (ed.). 2005. *Ethnologue: Languages of the world*. 15th edn. Dallas: SIL International.

Graham, Glenn H. 1981. A sociolinguistic survey of Busa and Nagatman. In Richard Loving (ed.), *Sociolinguistic surveys of Sepik languages*. Workpapers in Papua New Guinea Languages 29, 177–192. Ukarumpa: Summer Institute of Linguistics.

Greenberg, Joseph H. 1966. *The Languages of Africa*. 2nd edn. The Hague: Mouton de Gruyter.

Greenberg, Joseph H. 1971. Nilo-Saharan and Meroitic. In Thomas A. Sebeok (ed.), *Linguistics in Sub-Saharan Africa*. Current Trends in Linguistics 7, 421–442. Mouton de Gruyter.

Grimes, Barbara F. (ed.). 2000. *Ethnologue: Languages of the world*. 14th edn. Dallas: SIL International.

Haberland, Eike. 1962. Zum Problem der Jäger und besonderen Kasten in Nordost- und Ost-Afrika. *Paideuma* 8. 136–155.

Haberland, Eike & Adolf E. Jensen. 1959. *Altvölker Süd-Äthiopiens* (Völker Süd-Äthiopiens: Ergebnisse der Frobenius-Expeditionen 1950-52 und 1954-56 I). Stuttgart: Frobenius-Institut an der Johann Wolfgang Goethe-Universität (Frankfurt am Main).

Haider, R. 1994. Shompen. In T. N. Pandit & B. N. Sarkar (eds.), *Andaman and Nicobar Islands*. People of India XII, Anthropological Survey of India, 188–195.

Hammarström, Harald. 2007. *Handbook of descriptive language knowledge: A full-Scale reference guide for typologists*. LINCOM Handbooks in Linguistics 22. München: Lincom.

Hammarström, Harald. 2008a. Notes on the Ganza Language from unpublished notes by Reidhead and Disney. Ms in preparation.

Hammarström, Harald. 2008b. A reclassification of some West Papua languages. Paper presented at the International Workshop on Minority Languages in the Malay/Indonesian Speaking World. Universiteit Leiden, Leiden, The Netherlands, 28 June 2008.

Hammarström, Harald. 2008c. Two hitherto unnoticed languages from Sobger River, West Papua, Indonesia. Ms.

Hammarström, Harald. 2010a. The genetic position of the Mawes language. Paper presented at the Workshop on the Languages of Papua 2. Manokwari, Indonesia, 8–12 February 2010.

Hammarström, Harald. 2010b. A note on the Maco [wpc] (Piaroan) language of the lower Ventuari, Venezuela. *Cadernos de Etnolingüistica*. To appear.

Hantgan, Abbie. 2010. Does tone polarity exist? Evidence from plural formation among Bangime nouns. In Jonathan C. Anderson, Christopher R. Green & Samuel G. Obeng (eds.), *African linguistics across the discipline*. Indiana University Working Papers in Linguistics 8. Bloomington: Indiana University Linguistics Club.

Hargreaves, Inês. 2007. Lista de palavras transcritas por Inês Hargreaves, de dois grupos ao norte do Parque Aripuanã, RO. Manuscript made available with the help of Denny Moore.

Harrison, David K. 2007. *When languages die*. Oxford Studies in Sociolinguistics. Oxford: Oxford University Press.

Hashimoto, Kazuo. 2008. *Ata-English dictionary with English-Ata finder list*. Ukarumpa: Summer Institute of Linguistics.

Healey, Alan. 1964. The Ok language family in New Guinea. Canberra: Australian National University doctoral dissertation. (Sometimes cited as *A Survey of the Ok Family of Languages* presumably because part of the thesis II-IV, which contains all linguistic data, carries this title.)

Hischier, Phyllis. 2006. Exploration of the remote Kopayap and Urajin areas in West Papua, Indonesia: A first contact in Kopayap and Urajin. Ms.

Hoogland, J. 1939. Uittreksel uit het Dagboek van Hollandia over het tijdvak 16 Juni to en met 13 Juli 1939. Nationaal Archief, Den Haag, Ministerie van Koloniën: Kantoor Bevolkingszaken Nieuw-Guinea te Hollandia: Rapportenarchief, 1950–1962, nummer toegang 2.10.25, inventarisnummer 24A.

Hoogland, J. 1940. Gezagsinvloed. Nationaal Archief, Den Haag, Ministerie van Koloniën: Kantoor Bevolkingszaken Nieuw-Guinea te Hollandia: Rapportenarchief, 1950–1962, nummer toegang 2.10.25, inventarisnummer 25.

Huertas Castillo, Beatriz. 2004. *Indigenous peoples in isolation in the Peruvian Amazon: their struggle for survival and freedom* (IWGIA document 100). Copenhagen: IWGIA.

Hughes, Jock. 2009. Upper Digul survey. SIL International. SIL Electronic Survey Reports 2009–003.

Humboldt, Alexandre de. 1815, 1815, 1816, 1817, 1820, 1820, 1822, 1822, 1825. *Voyage aux régions équinoxiales du Noveau Continent*. Paris: N. Maze. 9 vols.

Hyman, Larry M. 2003. Why describe African languages? Keynote address at the World Congress of African Linguistics & Annual Conference of African Linguistics. New Jersey, Rutgers University, 18 June 2003.

Im, Youn-Shim & Randy Lebold. 2006. Draft Survey Report on the Usku Language of Papua. To appear in the SIL Electronic Survey Reports.

Jensen, Adolf E. 1952. Forschungsreise nach Süd-Abessinien. *Zeitschrift für Ethnologie* 77. 57–61.

Jungraithmayr, Herrmann. 2004. Das Birgit, eine Osttschadische Sprache–Vokabular und Grammatische Notizen. In Takács, Gábor (ed.), *Egyptian and Semito-Hamitic (Afro-Asiatic) studies in memoriam W. Vycichl*. Studies in Semitic Languages and Linguistics 39, 342–371. E. J. Brill.

Kaplan, Hilliard & Kim Hill. 1984. The Mashco-Piro Nomads of Peru. *Anthroquest* 29. 1, 13–17.

Kästner, Klaus-Peter. 2005. Beiträge zur Geschichte und Kultur der Urueuwauwau-Stämme (Rondônia, Brasilien). *Abhandlungen und Berichte der Staatlichen Ethnographischen Sammlungen Sachsen* 52. 91–133.

Kästner, Klaus-Peter. 2007. *Zoé: Materielle Kultur, Brauchtum und kulturgeschichtliche Stellung eines Tupí-Stammes im Norden Brasiliens*. Abhandlungen und Berichte des Staatlichen Museums für Völkerkunde Dresden 53. VWB–Verlag für Wissenschaft und Bildung.

Kaufman, Terrence. 1990. Language history in South America: What we know and how to know more. In Doris L. Payne (ed.), *Amazonian linguistics: Studies in lowland South American languages*. Texas Linguistics Series, 13–73. Austin: University of Texas Press.

Keifenheim, Barbara. 1997. Futurs Beaux-Frères ou Esclaves? Les Kashinawa découvrent des Indiens non contactés. *Journal de la Société des Américanistes* 83. 141–158.

Kellermann, Petra. 2000. Eine grammatische Skizze des Tama auf der Basis der Daten von R.C. Stevenson. Mainz: Johannes Gutenberg-Universität master's thesis.

Kim, So Hyun. 2006. Draft Survey Report on the Yetfa Language of Papua, Indonesia. To appear in the SIL Electronic Survey Reports.

Kleinewillinghöfer, Ulrich. 2001. Jalaa—An almost forgotten language of northeastern Nigeria: A language isolate. In Derek Nurse (ed.), *Historical Language Contact in Africa* (Sprache und Geschichte in Afrika 16/17), 239–271. Köln: Rüdiger Köppe.

Koch-Grünberg, Theodor. 1913. Abschluß meiner Reise durch Nordbrasilien zum Orinoco, mit besonderer Berücksichtigung der von mir besuchten Indianerstämme. *Zeitschrift für Ethnologie* 45. 448–474.

Koch-Grünberg, Theodor. 1922. Die Völkergruppierung zwischen Rio Branco, Orinoco, Rio Negro und Yapurá. In W. Lehmann (ed.), *Festschrift Eduard Seler dargebracht zum 70: Geburtstag von Freunden, Schülern und Verehrern*, 205–266. Stuttgart: Stecker und Schröder.

Koch-Grünberg, Theodor. 1928. *Sprachen* (Von Roroima zum Orinoco: Ergebnisse einer Reise in Nordbrasilien und Venezuela in den Jahren 1911-13 4). Stuttgart: Strecker und Schröder.

Konow, Sten. 1906. Nahālī. In G. A. Grierson (ed.), *Mundā and Dravidian Languages*. Linguistic Survey of India IV, 185–189. Calcutta: Office of the Superintendent of Goverment Printing.

Krauss, Michael E. 2007. Mass Language Extinction and Documentation: The Race against Time. In O. Miyaoka, O. Sakiyama & M. Krauss (eds.), *Vanishing languages of the Pacific Rim*, 3–24. Oxford: Oxford University Press.

Kuiper, F. B. J. 1962. *Nahali: A comparative study* (Mededelingen der Koninklijke Akademie van Wetenschappen, afdeling letterkunde, nieuwe reeks, deel 25 nr 5). Amsterdam: Noord Hollandsche Uitgeversmatschappij.

Kuiper, F. B. J. 1966. The sources of the Nahali Vocabulary. In Norman H. Zide (ed.), *Studies in comparative Austroasiatic linguistics*, 57–81. Berlin: Mouton de Gruyter.

Kumar, Satinder. 2000. The Sentinelese. In *Encyclopaedia of South-Asian tribes*, 3163–3165. New Delhi: Anmol Publications.

Lal, Parmanand. 1969. The Shompen of Great Nicobar. *Bulletin of the Anthropological Survey of India* 9. 247–254.

Landaburu, Jon. 2000. Clasificación de la lenguas indígenas de Colombia. In María Stella González de Pérez & María Luisa Rodríguez de Montes (eds.), *Lenguas indígenas de Colombia: Una visión descriptiva*, 25–50. Santafé de Bogotá: Instituto Caro y Cuervo.

Langeler, J. W. 1915. Sobger-Rivier: 'Journaal loopende van 22 October 1914 t/m 7 Januari 1915. Betreffende eene patrouille te water tot exploratie van de twee groote linker zijrivieren der Idenburg-rivier op ongeveer 140 10/4 O.L. en tot aanpeiling van het hooggebergte'. KITLV Manuscripts and Archives, Leiden [D H 1163].

Larson, Gordon F. 1977. Reclassification of some Irian Jaya highlands language families: A lexicostatical cross-family subclassification with historical implications. *Irian* VI(2). 3–40.

Laycock, Donald C. 1972. Looking westward: Work of the Australian National University on languages of West Irian. *Irian* 1(2). 68–77.

Laycock, Donald C. 1973. *Sepik Languages: Checklist and Preliminary Classification* (Pacific Linguistics: Series B 25). Canberra: Research School of Pacific and Asian Studies, Australian National University.

Laycock, Donald C. 1975a. Isolates: Sepik Region. In Stephen A. Wurm (ed.), *New Guinea area languages and language study vol. 1: Papuan languages and the New Guinea linguistic scene*. Pacific Linguistics: Series C 38, 879–886. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Laycock, Donald C. 1975b. Sko, Kwomtari and Left May (Arai) Phyla. In Stephen A. Wurm (ed.), *New Guinea Area Languages and Language Study vol. 1: Papuan Languages and the New Guinea linguistic scene* (Pacific Linguistics: Series C 38, 849–858. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Laycock, Donald C. & John A. Z'Graggen. 1975. The Sepik-Ramu Phylum. In Stephen A. Wurm (ed.), *New Guinea Area Languages and Language Study vol. 1: Papuan Languages and the New Guinea linguistic scene*. Pacific Linguistics: Series C 38, 731–764. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Le Roux, C. C. F. M. 1950. 25: Taalkundige Gegevens. In *De Bergpapoea's van Nieuw-Guinea en hun Woongebied*, vol. II, 776–913. Leiden: E. J. Brill.

Lean, Glendon A. 1986. *Sandaun Province*. Counting Systems of Papua New Guinea 7. Port Moresby: Papua New Guinea University of Technology. Draft Edition.

Lebeuf, Annie M.-D. 1959. *Les populations du Tchad (Nord du 10e parallèle)*. Ethnographic survey of Africa, Western Africa, French series. part 8. Paris: Presses Universitaires de France for the International African Institute (IAI).

Lee, Myung Young. 2005. Draft Survey Report on the Namla Language of Papua. To appear in the SIL Electronic Survey Reports.

Leeden, Alexander Cornelis van der. 1954. *Verslag over taalgebieden in het Sarmische van de Ambtenaar van het Kantoor voor Bevolkingszaken* volume 35. Hollandia: Gouvernement van Nederlands-Nieuw-Guinea, Dienst van Binnenlandse Zaken, Kantoor voor Bevolkingszaken.

Lewis, Paul M. (ed.). 2009. *Ethnologue: Languages of the world*. 16th edn. Dallas: SIL International.

Little, Christopher A. J. L. 2008. Becoming an Asabano: The socialization of Asabano children, Duranmin, West Sepik Province, Papua New Guinea. Canada: Trent University master's thesis.

Lohmann, Roger Ivar. 2000. Cultural reception in the contact and conversion history of the Asabano of Papua New Guinea. Madison: University of Wisconsin-Madison doctoral dissertation.

Loukotka, Čestmír. 1949. Sur Quelques Langues Inconnues de l'Amerique du Sud. *Lingua Posnaniensis* I. 53–82.

Loukotka, Čestmír. 1968. *Classification of the South American Indian Languages*. Reference Series 7. Los Angeles: Latin American Center, University of California.

Loving, Richard & Jack Bass. 1964. *Languages of the Amanab sub-district*. Port Moresby: Department of Information and Extension Services.

Lukas, Johannes. 1933. Beiträge zur Kenntnis der Sprachen von Wadái (Mararet, Mába). *Journal de la Société des Africanistes* 3(1). 25–55.

Lukas, Johannes. 1938. Die Sprache der Sungor in Wadai. *Mitteilungen der Ausland-Hochschule an der Universität Berlin* XLI (III). 171–246.

Maciel, Iraguacema. 1991. Alguns aspectos fonológicos e morfológicos da língua Máku. Brasília: Universidade de Brasil master's thesis.

MacMichael, H. A. 1918. Nubian Elements in Darfur. *Sudan Notes and Records* I. 30–48.

Man, Edward H. 1886. A Brief Account of the Nicobar Islanders. With Special Reference to the Inland Tribe of Great Nicobar. *Journal of the Royal Anthropological Institute of Great Britain and Ireland* 15. 428–451.

Man, Edward H. 1889. *A dictionary of Central Nicobarese*. London: W. H. Allen.

Martius, Carl Friedrich Philip von. 1867. *Zur Sprachenkunde* (Beiträge zur Ethnographie und Sprachenkunde Amerikas zumal Brasiliens II). Leipzig: Friedrich Fleischer.

Matallana, B. de & Cesareo de Armellada. 1943. Exploración del Paragua. *Boletín de la Sociedad Venezolana de ciencias naturales* VIII (53). 61–110.

Mattei-Müller, Marie-Claude. 2006. Lenguas Indígenas de Venezuela en peligro de extinción. In *Lenguas y tradiciones orales de la Amazonía. ¿diversidad en peligro?*, 281–312. La Habana: Casa de las Américas.

McDowell, Nancy. 1991. *The Mundugumor: From the Field Notes of Margaret Mead and Reo Fortune* (Smithsonian Series in Ethnographic Inquiry). Washington, D.C.: Smithsonian Institution Press.

Meira, Sérgio. 2000. *A reconstruction of Proto-Taranoan: phonology and morphology*. LINCOM Studies in Native American Linguistics 30. München: Lincom.

Melliger, Markus. 2000. Pinai-Hagahai. In John Brownie (ed.), *Sociolinguistic and literacy studies: Highlands and islands*. Data papers on Papua New Guinea languages 45, 64–122. Ukarumpa: Summer Institute of Linguistics.

Michael, Lev & Christine Beier. 2003. Poblaciones indígenas en aislamiento voluntario en la región del Alto Purús. In Renata Leite Pitman, Nigel Pitman & Patricia Álvarez (eds.), *Alto Purús: Biodiversidad, conservación y manejo*, 149–164. Durham: Center for Tropical Conservation, Duke University.

Migliazza, Ernest. 1972. Yanomama grammar and intelligibility. Bloomington: Indiana University doctoral dissertation.

Migliazza, Ernest & Lyle Campbell. 1988. *Panorama General de las Lenguas Indígenas*. Historia General de América: Periodo Indígena 10. Caracas: Academiá Nacional de la Historia de Venezuela.

Migliazza, Ernesto C. 1965. Fonología Makú. *Boletim do Museu Paraense Emílio Goeldi, Série Antropologia* 25. 1–17.

Migliazza, Ernesto C. 1966. Esbôço sintático de um corpus da língua Makú. *Boletim do Museu Paraense Emílio Goeldi, Série Antropologia* 32. 1–38.

Migliazza, Ernesto C. 1978. Maku, Sape and Uruak Languages: Current Status and Basic Lexicon. *Anthropological Linguistics* XX(3). 133–140.

Migliazza, Ernesto C. 1980. Languages of the Orinoco-Amazon Basin: Current Status. *Antropológica* 53. 95–162.

Migliazza, Ernesto C. 1983. Lenguas de la Región Orinoco Amazonas: Estado Actual. *América Indígena* 43. 703–784.

Migliazza, Ernesto C. 1985. Languages of the Orinoco-Amazon region: Current status. In Harriet E. Manelis Klein & Louisa Stark (eds.), *South American Indian languages: Retrospect and prospect*, 17–139. Austin: Texas University Press.

Moore, Denny. 2005. Classificação interna da família lingüística Mondé. *Estudos Lingüísticos* XXXIV. 515–520.

Mora B., Carlos. 2007. Propuesta de creación de reserva indígena para nativos en aislamiento: Un análisis de caso. Paper presented in Iquitos, 26 de Octubre 2007.

Mosonyi, Esteban Emilio. 2003. Situación actual de las lenguas indígenas de Venezuela. In Esteban Emilio Mosonyi, Arelis Barbella & Silvana Caula (eds.), *Situación de las lenguas indígenas en Venezuela*, 86–116. Caracas: Casa de Las Letras-Casa de Bello.

Moxness, Mike. 1998. A Brief, Second-hand Report on the Kehu (Keu?). Ms. 24 words from the memory of a non-native speaker.

Mundlay, Asha. 1996. Nihali Lexicon. *Mother Tongue* II. 17–48.

Nimuendajú, Curt. 1977. Os indios Tucuna. *Boletim do Museu do Indio. Antropologia* 7. 1–69.

No Author Stated. 2006a. *Godokono Hido Tabo: Aramia River Tabo Testament*. Port Moresby: Bible Society of Papua New Guinea.

No Author Stated. 2006b. *Godokono Wade Tabo: Fly River Tabo New Testament*. Port Moresby: Bible Society of Papua New Guinea.

No Author Stated. (no date). The Dibla:g Language. SIL, Ukarumpa. Ms.

Norquest, Peter K. 2007. A phonological reconstruction of Proto-Hlai. Tucson: University of Arizona doctoral dissertation.

Ortiz, Sergio Elías. 1965. *Prehistoria Tomo 3: Lenguas y dialectos indígenas de Colombia*. Historia Extensa de Colombia I. Bogotá: Ediciones Lerner.

Partridge, Edna. 1981. *Sa:lenapa wala gilala dote bata ete miyana gi kanika:*. Port Moresby: Bible Society Papua New Guinea.

Patiño Rosselli, Carlos. 2000. Lenguas aborígenes de la Amazonia meridional de Colombia. In María Stella González de Pérez & María Luisa Rodríguez de Montes (eds.), *Lenguas indígenas de Colombia: Una visión descriptiva*, 16–170. Santafé de Bogotá: Instituto Caro y Cuervo.

Perozo, Laura, Ana Liz Flores, Abel Perozo & Mercedes Aguinagalde. 2008. Escenario histórico y sociocultural del alto Paragua, Estado Bolívar, Venezuela. In Josefa Celsa Señaris, Carlos A. Lasso & Ana Liz Flores (eds.), *Evaluación rápida de la biodiversidad de los ecosistemas acuáticos de la cuenca alta del río Paragua, estado Bolívar*, Boletín RAP de Evaluación 49, 169–180. Arlington, VA: Conservation International.

Peter, Hanns. 1973–1974. Vorstellungen über Krankheiten und Krankenbehandlungen bei den Gargar im West-Sepik-District. *Wiener Völkerkundliche Mitteilungen, N.S.* 15–16. 27–62.

Price, Dorothy. 1987. Some Karkar-Yuri orthography and spelling decisions. In John M. Clifton (ed.), *Studies in Melanesian orthographies*, Data Papers on Papua New Guinea Languages 33, 57–76. Ukarumpa: Summer Institute of Linguistics.

Price, Dorothy, Veda Rigden & Maramia Nkonifa. 1994. *Kwaromp kwapwe kare kar (God's truly good talk) [New Testament]*. USA: The Bible League, South Holland, Illinois.

Przyluski, Jean. 1924. Les langues Austroasiatiques. In Antoine Meillet & Marcel Cohen (eds.), *Les langues du monde*, Collection linguistique 16, 385–403. Paris: Librarie Ancienne Édouard Champion.

Queixalós, Francesc & Zoraide dos Anjos G.S. 2007. A língua Katukína-Kanamarí. *LIAMES* 6. 29–60.

Raja Ram, M. G. 1960. My contact with the Shom-Pens of Dogmar river-Great Nicobar (February 1960). *Bulletin of the Anthropological Survey of India* 9. 74–80.

Ray, Sidney H. 1923. The languages of the Western Division of Papua. *Journal of the Royal Anthropological Institute of Great Britain and Ireland* 53. 332–360.

Reesink, Ger P. 1976. Languages of the Aramia River Area. In *Papers in New Guinea Linguistics* 19, Pacific Linguistics: Series A 45, 1–37. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Reimer, Martha. 1986. *The notion of topic in Momuna narrative discourse*. Pacific Linguistics: Series A 74. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Rigden, Veda. (no date). Karkar Grammar Essentials. Ukarumpa: SIL, Unpublished ms.

Rizvi, S. N. H. 1990. *The Shompen: A Vanishing Tribe of the Great Nicobar Island*. Calcutta: Seagull Books.

Robidé van der Aa, Pieter Jan Baptist Carel. 1879. *Reizen naar Nederlandsch Nieuw-Guinea ondernomen op last der Regeering van Nederlandsche Indie in de jaren 1871, 1872, 1875-1876 door de Heeren P. van Crab en J.E. Teysmann, J.G. Coornengel, A.J. Langeveldt van Hemert en P. Swaan*. The Hague: Martinus Nijhoff.

Rodríguez Bazán, Luis Antonio. 2000. Estado de las lenguas indígenas del Oriente, Chaco y Amazonia Bolivianos. In Francesc Queixalós & Odile Renault-Lescure (eds.), *As Línguas Amazônicas Hoje/Las Lenguas Amazonicas Hoy/Les Langues d'Amazonie aujourd'hui/The Amazonian Languages Today*, 129–150. São Paulo: Museo Paraense Emilio Goeldi.

Ross, Malcolm. 1995. The Great Papuan Pronoun Hunt: Recalibrating Our Sights. In Connie Baak, Mary Bakker & Dick van der Meij (eds.), *Tales from a concave world: Liber amicorum Bert Voorhoeve*, 139–168. Leiden: Department of Languages and Cultures of Southeast Asia and Oceania, Leiden University.

Routamaa, Judy. 1994. Kamula grammar essentials. Ms. http://www.sil.org/pacific/png/abstract.asp?id=50209 (1 August, 2008.)

Ruiz de Quijano, Tomas, Fabricio de Yanguas & Manuel del Sorillo Luz. 1894 [1765]. Informan difusamente de cuanto les consta y experimentan, de tocante á los adelantamientos y nuevos descubrimientos de derrotas y naciones gentiles de yindios que han convertido los padres misioneros en diferentes regiones incognitas hasta la hora presente. In Antonio B. Cuervo (ed.), *Sección segunda: Geografía, Viajes, Misiones y Límites*, Colección de documentos inéditos sobre la historia y geografía de Colombia IV, 226–232. Bogotá: Zalamea Hermanos.

Rumaropen, Benny. 2004. Draft Survei Sosiolinguistik pada ragam Bahasa Kimki di Bagian Tenggara Gunung Ji, Papua, Indonesia. To appear in the SIL Electronic Survey Reports.

Rumaropen, Benny. 2006. Draft Survey Report on the Kapauri Language of Papua. To appear in the SIL Electronic Survey Reports.

Saito, Hisafumi. 1998. We are one flesh: Unity and migration of the Yabio. In Shuji Yoshida & Yukio Toyoda (eds.), *Fringe Area of Highlands in Papua New Guinea*, Senri Ethnological Studies 47, 93–112. Osaka: National Museum of Ethnology.

Santandrea, Stefano. 1950. Gleanings in the Western Bahr el Ghazal. *Sudan Notes and Records* XXXI. 54–64.

Santandrea, Stefano. 1964. *A tribal history of the western Bahr El Ghazal*. Museum Combonianum 17. Bologna: Editrice Nigrizia.

Santandrea, Stefano. 1966. *The Birri language: Brief elementary notes*. Museum Combonianum 20. Bologna: Editrice Nigrizia. Extract from Afrika und Übersee 49(2): 81–105, 196–234, 1965–1966.

Santandrea, Stefano. 1976. *The Kresh Group, Aja and Baka Languages (Sudan): A Linguistic Contribution*. Napoli: Istituto Universitario Orientale.

Schadeberg, Thilo C. 1981. *A Survey of Kordofanian vol. 1: The Heiban Group* (Sprache und Geschichte in Afrika: Beiheft 1). Hamburg: Helmut Buske.

Schmidt, Wilhelm. 1906. Die Mon-Khmer-Völker, ein Bindeglied zwischen Völkern Zentralasiens und Austronesiens. *Archiv für Anthropologie* 5. 59–109.

Shafer, Robert. 1940. Nahālī: A singuistic study in paleoethnography. *Harvard Journal of Asiatic Studies* 4. 346–371.

Silva Magalhães, Marina Maria. 2002. Aspectos fonológicos e morfossintáticos da língua Guajá. Brasília: Universidade de Brasil master's thesis.

Silzer, Peter J. & Heljä Heikkinen. 1984. Index of Irian Jaya languages. *Irian* XII. 1–124.

Silzer, Peter J. & Heljä Heikkinen-Clouse. 1991. Index of Irian Jaya languages. *Special Issue of Irian: Bulletin of Irian Jaya*. 2nd edn. Jayapura: Program Kerjasama Universitas Cenderawasih and SIL.

Smith, Donaldson A. 1897. *Through unknown African countries: The first expedition from Somaliland to Lake Lamu*. New York: Edward Arnold.

Smits, Leo & C. L. Voorhoeve. 1994. *The J. C. Anceaux collection of wordlists of Irian Jaya languages B: Non-Austronesian (Papuan) languages*. part I. Irian Jaya Source Material No. 9 Series B 3. Leiden-Jakarta: DSALCUL/IRIS.

Smits, Leo & C. L. Voorhoeve. 1998. *The J. C. Anceaux collection of wordlists of Irian Jaya languages B: Non-Austronesian (Papuan) languages*. part II. Irian Jaya Source Material No. 10 Series B 4. Leiden-Jakarta: DSALCUL/IRIS.

Souag, Lameen M. 2004. Oropom etymological lexicon: Exploring an extinct, unclassified Ugandan language. Paper in progress, http://lameen.googlepages.com/oropomed.pdf (20 September, 2008.)

Souza, Larissa da Silva Lisboa. 2008. O processo de revitalização de uma língua: Mecanismos para documentação e clasificaçao da língua dos Arara do Rio Branco. *Língua, Literatura e Ensino* 3. 555–561.

Stewart, Jean. 1987. *God ya tyo kimina, God ya swagumin nin [New Testament in Aekyom]*. Port Moresby: Bible Society Papua New Guinea.

Stirtz, Timothy M. 2006. Possession of Alienable and Inalienable Nouns in Gaahmg. In Al-Amin Abu-Manga, Leoma Gilley & Anne Storch (eds.), *Insights into Nilo-Saharan Language, History and Culture: Proceedings of the 9th Nilo-Saharan Linguistic Colloquium, Institute of African and Asian Studies, University of Khartoum, 16-19 February 2004*, Nilo-Saharan 23, 377–392. Köln: Rüdiger Köppe.

Stokhof, W. A. L. (ed.). 1983. *Holle lists: Vocabularies in languages of Indonesia, Vol.5/2: Irian Jaya: Papuan Languages, Northern Languages, Central Highlands Languages*. Pacific Linguistics: Series D 53. Research School of Pacific and Asian Studies, Australian National University.

Struck, Bernhard. 1930. Die Gbaya-Sprache (Dar-Fertit). *Mittheilungen des Seminars für Orientalische Sprachen* XXIII. 53–82.

Survival International. 2009a. Chronology of evidence of uncontacted Indians fleeing from Peru to Brazil. London: Survival International.

Survival International. 2009b. One year on: Uncontacted tribes face extinction. London: Survival International.

Teferra, Anbessa. 1991. A Sketch of Shabo Grammar. In M. Lionel Bender (ed.), *Proceedings of the Fourth Nilo-Saharan Linguistics Colloquium,* Nilo-Saharan: Linguistics Analyses and Documentation 7, 371–387. Hamburg: Helmut Buske.

Temple, Richard Carnac. 1907. A Plan for Uniform Scientific Record of the Languages of Savages Applied to the Languages of the Andamanese and Nicobarese. *Indian Antiquary: A Journal of Oriental Research* XXXVI. 181–203, 217–251, 317–347, 353–369.

Tessmann, Günter. 1930. *Die Indianer Nordost-Perus: grundlegende Forschungen für eine systematische Kulturkunde* (Veröffentlichung der Harvey-Bassler-Stiftung 2). Hamburg.

Trupp, Fritz. 1974. Una Tribu Desconocida en la Amazonia Colombiana. *Bulletin of the International Committee on Urgent Anthropological Ethnological Research* 16. 109–110.

Tucker, Archibald N. & Margaret A. Bryan. 1956. *The Non-Bantu languages of North-Eastern Africa* (Handbook of African Languages III). Oxford: Oxford University Press.

Tucker, Archibald N. & Margaret A. Bryan. 1966. *Linguistic analyses: The Non-Bantu Languages of North-Eastern Africa* (Handbook of African Languages). 2nd edn. Oxford: Oxford University Press.

Tupper, Ian. 2007. Endangered Languages Listing: TURUMSA [tqm]. http://www.pnglanguages.org/pacific/png/show_lang_entry.asp?id=tqm (1 May, 2007.)

Valenzuela, Pilar M. 2006. Peru: Language Situation. In Keith Brown (ed.), *Encyclopedia of Language and Linguistics*, vol. 9, 316–319. 2nd edn. Amsterdam: Elsevier.

Vidal y Pinell, Ramón. 1969–1970. Identificación de la tribu de los Yuríes en el Amazonas de Colombia. *Amazonía Colombiana Americanista* 7. 95–109.

Voorhoeve, C. L. 1971. Miscellaneous Notes on Languages in West Irian, New Guinea. In *Papers in New Guinea Linguistics 14* Pacific Linguistics: Series A 28, 47–114. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Voorhoeve, C. L. 1975a. The Central and Western Areas of the Trans-New Guinea Phylum: Central and Western Trans-New Guinea Phylum Languages. In Stephen A. Wurm (ed.), *New Guinea Area Languages and Language Study Vol 1: Papuan Languages and the New Guinea linguistic scene*, Pacific Linguistics: Series C 38, 345–460. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Voorhoeve, C. L. 1975b. *Languages of Irian Jaya, Checklist: Preliminary classification, language maps, wordlists*. Pacific Linguistics: Series B 31. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Voort, Hein van der. 2007. Theoretical and social implications of language documentation and description on the eve of destruction in Rondônia. In Peter K. Austin, Oliver Bond & David Nathan (eds.), *Proceedings of Conference on Language Documentation and Linguistic Theory*, 251–259. London: SOAS.

Vries, Lourens de & Gerrit J. van Enk. 1997. *The Korowai of Irian Jaya: Their language and its cultural context*. Oxford Studies in Anthropological Linguistics 9. Oxford: Oxford University Press.

Walker, Roland. 1983. Fakfak Survey Report. Abepura. Unpublished report.

Wallace, Alfred R. 1853. *Travels on the Amazon and Rio Negro, with an account of the Native Tribes and Observations on the climate, geology, and natural history of the Amazon Valley*. London: Reeve & Co.

Wallace, Scott. 2003. Into the Amazon. *National Geographic* 196(Aug). 2–23.

Wambaliau, Theresia. 2004. Draft Laporan Survei pada Bahasa Murkim di Papua, Indonesia. To appear in the SIL Electronic Survey Reports.

Wambaliau, Theresia. 2005. Draft Laporan Survei pada Bahasa Tofanma di Papua, Indonesia. To appear in the SIL Electronic Survey Reports.

Wambaliau, Theresia. 2006. Draft Laporan Survei pada Bahasa Mawes di Papua, Indonesia. To appear in the SIL Electronic Survey Reports.

Whalen, D. H. & Gary F. Simons. 2009. Endangered Language Families. Paper presented at the 1st International Conference on Language Documentation and Conservation. University of Hawaiʻi at Mānoa, Honolulu, Hawaiʻi, 12–14 March 2009.

Whitehouse, Paul. 2006. The "Lost" paper: A belated conference postscript. *Mother Tongue* XI. 262–274.

Whitehouse, Paul. (no date). Type-up of anonymous Kehu wordlist from SIL Indonesia. Ms. (The wordlist presumably comes from Ron Baird in the 1980s.)

Wilson, J. G. 1970. Preliminary observations on the Oropom people of Karamoja, their ethnic status, culture, and postulated relation to the peoples of the late Stone Age. *The Uganda Journal* 34(2). 125–145.

Wurm, Stephen A. 1971. Notes on the Linguistic Situation of the Trans-Fly Area. In *Papers in New Guinea Linguistics 14*, Pacific Linguistics: Series A 28, 115–172. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Wurm, Stephen A. 1975. Eastern Central Trans-New Guinea Phylum Languages. In Stephen A. Wurm (ed.), *New Guinea Area Languages and Language Study Vol 1: Papuan Languages and the New Guinea linguistic scene*, Pacific Linguistics: Series C 38, 461–526. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Yanagida, Tatsuya. 2004. Socio-historic overview of the Ata language, an endangered Papuan language in New Britain, Papua New Guinea. In Shibata Norio & Toru Shionoya (eds.), *Kan minami Taiheiyoo no gengo 3 [Languages of the South Pacific Rim 3]*, ELPR Publications Series A1-008, 61–94. Suita: Faculty of Informatics, Osaka Gakuin University.

Yimam, Baye. 2006. A general description of Mao. *Ethiopian Languages Research Centre (ELRC) Working Papers* 2(2). 166–225.

Yimam, Baye. 2007. Mao of Bambasi. In Siegbert Uhlig (ed.), *Encyclopaedia Aethiopica* vol. III, 760–761. Wiesbaden: Otto Harrassowitz.

Yip, May. 2004. Phonology of the These language. *Occasional Papers in the Study of Sudanese Languages* 9. 93–117.

Harald Hammarström
h.hammarstrom@let.ru.nl