Sampling and genealogical coverage in WALS

HARALD HAMMARSTRÖM

Abstract

WALS was designed with the goal of providing a “systematic answer” to questions about the geographical distribution of language features. In order to achieve this goal, there must be an adequate sample of the world’s languages included in WALS. In this article we investigate to what extent WALS fulfils its aim of maximizing the genealogical diversity of the samples of languages included. For this we look at the core-200 sample (included on almost all maps) as well as the 1,370 sample for the feature OV/VO word order (the sample with the largest number of languages). The genealogical diversity in these samples is compared against a database of “what could have been done”, i.e., a database of which language families have adequate descriptive resources for the task at hand. In the 200 sample, we find a highly significant overinclusion of Eurasian languages at the expense of South American and Papuan languages. In the 1,370 sample, we find a highly significant overinclusion of North American languages at the expense of South American and Papuan languages. It follows that statistics based on these WALS samples cannot be used straightforwardly for sound inferences about the distribution of the features in question.

Keywords: genealogical classification, linguistic atlas, sampling, word order

1. Introduction

WALS was designed with the goal of providing a “systematic answer” to questions about the geographical distribution of language features (Comrie et al. 2005: 1). As stressed in the introduction (Comrie et al. 2005: 1, 4), in order to achieve this goal, there must be an adequate sample of the world’s languages included in WALS.
In this article we will investigate to what extent WALS\textsuperscript{1} fulfils its goals and its claims, especially as it pertains to the desideratum of maximizing the genealogical diversity of the samples of languages included. For this we chose to investigate the core-200 sample and the 1,370 sample for the feature OV/VO word order. We chose the core-200 sample because this set was specially designed with the guiding principle to maximize genealogical diversity and the sampled languages are included on almost all maps. We chose the 1,370 sample of the OV/VO word order feature because it had the largest number of languages\textsuperscript{2} included. The genealogical diversity in these samples will be compared against a database of “what could have been done”, i.e., a database of which language families have adequate descriptive resources for the task at hand. In other words, we contrast the “breadth” and “depth” of the genealogical diversity of WALS with the state-of-the-art possibility.

2. Preliminaries and data matters

2.1 WALS languages

WALS contains approximately 2,560 languages\textsuperscript{3} that appear on at least one map (Comrie et al. 2005). Sign languages, pidgins, and creoles have origins of a different kind, and fall outside the scope of this study (of genealogical coverage). There is a certain amount of language/dialect inconsistency, and a certain amount of coding inconsistency in WALS. However, in no case does this affect the genealogical classification of any WALS entry in this study, so such matters can be safely disregarded here.

2.2 Genealogical classification

Without a full genealogical classification from the start, it is impossible to assess how well a certain sample covers the languages of the world. WALS provides a classification of the languages included in WALS, which is only a minority of languages in the world (Dryer 2005a). Thus, to assess the coverage it is necessary to use information from outside WALS.

WALS works with two different levels of genealogical classification (Dryer 2005a: 584): families (“the highest level accepted by specialists”) and genera (“fairly obvious without systematic comparative analysis, and which even the most conservative ‘splitter’ would accept”). We have chosen not to work with

---

1. For this article, we used a hard copy with the accompanying CD-ROM purchased in August 2005 (Haspelmath et al. (eds.) 2005). If subsequent paper or web editions contain updates, they are not taken into account here.

2. It is also the feature in WALS with the most genealogical diversity by any measure.

3. The exact number is given as 2,559 twice on page 3, but 2,560 on pages 4 and 584. From the data tables, 2,560 appears to be the correct number of entries.
Sampling and genealogical coverage in WALS

exactly these concepts, both for theoretical and practical reasons. In short, the family concept as of WALS is theoretically weak because it is not clear who is a specialist, that specialists agree, that all specialists know all the data, that specialists across areas have the same traditions on what to “accept”, and so on. In practice, there is no tangible evidence from specialists adduced to support the actual WALS-families as given – that is, there is no declaration of who the specialists are and where they argue their expert opinion, for each language family. Indeed, the outcome list is honestly described as an “educated guess”. However, the outcome is quite different from, e.g., our educated guesses even about some undeniable trends in the opinion of the specialists. Similarly, the genus concept lacks a threshold for “obviousness” and there are practical problems here too – there is no evidence adduced to for the obviousness/non-obviousness of the various listed genera, and inconsistencies are easily spotted.

For reasons just explained, it was unfeasible for us to compile a list of genera or families in the WALS sense. Instead, we used the related concept of a “D-family” (for demonstrated family), defined as

(i) a set of languages (possibly a one-member set)
(ii) with at least one sufficiently attested member language
(iii) that has been demonstrated in publication
(iv) to stem from a common ancestor
(v) by orthodox comparative methodology (Campbell & Poser 2008)
(vi) for which there are no convincing published attempts to demonstrate a wider affiliation.

To support the actual choices, in each case, we give a reference to a publication pointing to the evidence necessary to establish the above, possibly adducing comments, in an appendix of supplementary online material to this article.

While this concept is not free from theoretical or practical problems, we feel that it is preferable in terms of tangibility, as all choices are made more explicit. A fuller discussion is beyond the scope of this study. We feel that this practice

4. To take just one example, Khoesan specialists agree that a “Khoisan” family as listed in WALS, is not in evidence (Güldemann & Vossen 2000, Traill 1995, Westphal 1979). As the references in this article show, there are very many more examples.

5. Matthew Dryer (personal communication, May 2008) admits that the “obviousness” criterion may have to be adjusted to exclude the use of numerals as a criterion. The relatedness of most of the Indo-European branches can arguably be said to be obvious since several early amateurs independently saw the relatedness (using numerals as one of the arguments). If we want to call the various branches separate genera, then this fact must be dealt with – one way out is to rephrase “obvious” as “obvious considering everything except the numerals”.

6. To take just one example here too, Central Solomons is listed as a genus, but these languages cannot even be shown to be related (Terrill 2006), while Germanic and Romance are listed as separate genera, despite the clear relatedness, as evidenced, for instance, in the numeral series.
Harald Hammarström

is justified because, on the whole, it does not really matter for the results what level of genealogical relatedness one counts, be it WALS-genera or D-families, as long as one is consistent across the world. As noted, we do not believe this property holds for specialists across different areas of the world, so it is not clear that WALS-families have this property.

2.3. Database of descriptive resources

According to WALS (Comrie et al. 2005: 3), only 10 to 15 percent of the world’s languages are comprehensively described. However, no evidence, no distribution, and no discussion surrounds this figure.

We have compiled a database of available descriptive resources for each D-family, similar to Hammarström 2007b which is already becoming outdated. For each language family, one of the best descriptive resources for the best described language is listed, and categorized as “(Full) Grammar”, “Grammar sketch”, or “Less than grammar sketch”. Ideally, the “(Full) Grammar” category would correspond to “sufficiently described for inclusion in the WALS core-200 sample”, and “Grammar sketch” to “sufficiently described to decide the OV/VO feature”, but the matter is not so simple. Thus, we took care to check that our “(Full) Grammar” category included only languages for which the size of the description as a whole was similar to or larger than the least well-described languages actually included in the core-200 sample, and, for the question of the OV/VO feature, we took care to double-check also exactly which grammar sketches give sufficient information to decide that feature, and which languages with poorer descriptive resources do in fact give sufficient information for OV/VO. The sources are listed explicitly in an Appendix of supplementary online material to this article (http://dx.doi.org/10.1515/lity.2009.006_supp-1).

There is a little discrepancy in that a few sources in our database (from 2008) became available too late for WALS. There are references in WALS from no later than 2004, so we take anything from 2005 (inclusive) and on to have

7. In this study, we used the break of 2004 to 2005 as the limit, for the reason just explained. Matthew Dryer (personal communication, 2009) has since informed me that the core-200 sample was prepared in 1999, and therefore does not include languages for which a description appeared after 1999. This information was hardly deducible from WALS as it is not mentioned in the section about the core-200 sample, nor can it be inferred from the source lists for the core-200 languages, where, in several cases the best (or one of the best) source mentioned post-dates 1999, e.g., Ket (2000), Lepcha (2003), Shipibo-Condorí (2003), Avariná (2000), Aymara (2001). The matter is not insignificant, because a lot of good descriptions for underdescribed families appeared in 2000–2004. Should the genealogical coverage of the core-200 sample be evaluated on the existence of grammatical descriptions as of the end of 1999 (rather than 2004, as in this article), it seems that the Papuan and South American underinclusions would disappear, though we lack the database annotation needed to compute this exactly.
been “too late” for WALS. When we evaluate the WALS samples, we take this into account, by “excusing” WALS in the posterior discussion if a source from after 2004 makes a difference. We keep good track of all such cases, and in no case do we fail to state them if they bear on the conclusions we draw.

Note also that some of the yet “poorly described” languages are extinct while others are not, which means that the status is subject to change in the future.

3. Evaluations of coverage

3.1. The core-200 sample

The description of the construction of the core-200 sample is vague, but it is clear that it is not the result of a formal procedure. Rather it is an ad hoc procedure guided by the following principles: maximizing genealogical diversity and areal diversity, existence of a grammatical description, inclusion of major languages, inclusion of geographically disparate languages, hampered by availability of grammatical descriptions. Existence of a grammatical description must be understood as a mandatory criterion. The desiderata of maximizing genealogical and areal diversity are described with the word “major”, and the discourse indicates the remaining criteria to be minor (Comrie et al. 2005: 4–6).

In spite of vagueness, we (and as we expect most WALS readers to do as well) find the following as the only consistent reading of the objective. A set of languages is selected, the members of which should be maximally genealogically and geographically diverse, and, in addition, a sprinkle of further languages are added (by majorness and geographical disparateness) which do not increase to the genealogical and geographical diverseness of the whole set. The geographical diverseness maxim requires (at least) that large reasonably disjoint geographical regions are equally considered. The genealogical diverseness maxim requires that only languages from different families are selected. The goal of WALS, i.e., to provide a “systematic answer” to questions about the geographical distribution of language features, implies that no region should be overfocussed or underfocussed (that would hardly be systematic). The demands on systematicity and maximization leave no room for deviations. Thus, the objective of the core-200 sample must entail that languages included are from families evenly represented in large reasonably disjoint geographical regions. We now discuss whether this is indeed the case.

We focus on coverage. The 200 sample contains languages from 110 D-families. We recognize a total of 394 D-families in the world. However, only 212 of them contain a language for which there is a “(Full) Grammar”, i.e., is described comprehensively enough to be included in WALS on most features. In other words, WALS could be expanded to include most features for no less than $212 - 110 = 112$ further language families.
However, it is not necessary to cover all possible families to achieve the goals of WALS, i.e., to provide a “systematic answer” to questions about the geographical distribution of language features, as long as the families included are evenly sampled. We will now go on to discuss whether this is true for the WALS core-200 sample.

Table 1 shows the continental break-up of the language family coverage of the WALS core-200 sample. Again, the figures refer to sufficiently well-described D-families versus D-families included in the WALS 200 sample. Well-described D-families that are not included in the WALS 200 sample are shown in italics.

As can be seen from Table 1, the sampling is not even across continents, ranging from 33.3% (Papua 12/36) to 84.6% (Eurasia 22/26) of the total number of families. If the inclusion of families were even across continents then all would have a coverage of around 51.9% (110/212). Are the differences we see statistically significant? We estimate the $p$-values by simulation as follows: (i) generate 1000 random 110-member subsets of the 212 well-described families; (ii) for each family, ask how many of the subsets have more/less than the number found in the WALS 200 sample. For instance, if, say, 100 of those 1000 random subsets contain more Eurasian languages than in the WALS 200 sample, then there is no statistically significant oversampling from Eurasia in the WALS 200 sample, because higher inclusion of Eurasian families are included too often just by random. The results are shown in Table 2.

In plain words, we find that overinclusion of Eurasian D-families is highly significant in the WALS 200 sample, and that this happens at the expense of South American and Papuan D-families (but the significance levels in of these underinclusions are much lower). Overinclusion of African D-families is not significant with conventional levels of significance. The underinclusion of Papuan languages is somehow excusable because of a number of borderline choices. Firstly, based mostly on pronominal morphemes, Abun, Hatam, and North Halmahera are widely held to be genetically related, in spite of lacking lexical correspondences (Klamer et al. 2008). Secondly, extensive descriptions of Inanwatan and Sko member languages did not appear until too late for WALS, and for Awin-Pa languages there were text collections out (Stewart 1987), but lengthy written-up grammar papers were not easily accessible until recently (2008), when Routamaa 1994 was posted online. Thirdly, a few more Papuan D-families which are here listed as being sufficiently well-described, are actually debatable (though still comparable in descriptive status to the least well-described Papuan languages in the WALS 200 sample). If these borderline cases are turned in favour of WALS, the Papuan coverage is well within non-significance limits.

The vast overinclusion of Eurasian D-families is disturbing, especially since the pitfalls of Eurasian oversampling is precisely what is highlighted in the
### Table 1: Continental break-up of the language family coverage of the WALS core-200 sample

<table>
<thead>
<tr>
<th>Area</th>
<th>Coverage</th>
<th>D-families</th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa</td>
<td>13/18</td>
<td>72.2%</td>
</tr>
<tr>
<td>America</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>11/21</td>
<td>52.3%</td>
</tr>
<tr>
<td>Eurasia</td>
<td>23/26</td>
<td>84.6%</td>
</tr>
<tr>
<td>North America</td>
<td>26/49</td>
<td>53.0%</td>
</tr>
<tr>
<td>Papua</td>
<td>12/36</td>
<td>33.3%</td>
</tr>
</tbody>
</table>

D-families versus D-families included in the WALS 200 sample. Well-described D-families that are not included in the WALS 200 sample are shown in italics.
Table 1. (continued)

<table>
<thead>
<tr>
<th>Area</th>
<th>Coverage</th>
<th>D-families</th>
</tr>
</thead>
</table>
Table 2. Tests for statistical significance of over/under-inclusion in the WALS 200 sample. All i’s range up to 1000. $S_i[C] = \{ x \in S_i \text{ such that } x \text{ is from continent } C \}$.

<table>
<thead>
<tr>
<th>Continent</th>
<th>W200 Question</th>
<th>Test</th>
<th>Outcome</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eurasia</td>
<td>22/26 Overinclusion</td>
<td>$</td>
<td>{ i</td>
<td>S_i[Eurasia] \geq 22 }</td>
</tr>
<tr>
<td>Africa</td>
<td>13/18 Overinclusion</td>
<td>$</td>
<td>{ i</td>
<td>S_i[Africa] \geq 13 }</td>
</tr>
<tr>
<td>South America</td>
<td>26/62 Underinclusion</td>
<td>$</td>
<td>{ i</td>
<td>S_i[South America] \leq 26 }</td>
</tr>
<tr>
<td>Papua</td>
<td>12/36 Underinclusion</td>
<td>$</td>
<td>{ i</td>
<td>S_i[Papua] \leq 12 }</td>
</tr>
</tbody>
</table>

WALS sampling section (Comrie et al. 2005: 3)! Note that this fact has nothing to do with the desideratum of including extra “major” languages (Comrie et al. 2005: 3), because, e.g., the Eurasian families Abkhaz-Adyge, Ainu, Burushaski, Basque, Chukotko-Kamchatkan, Miao-Yao, Nivkh, Tungusic, Yeniseic, and Yukaghir contain no major language in terms of speaker numbers (Gordon (ed.) 2005). Furthermore, we may look at the Eurasian families which were not included; they are the ancient families, all of them long extinct – ancient families were consciously excluded from WALS on other considerations – and the isolate Kusunda, for which a full-ish description only became available too late for WALS (Watters 2005). Thus, considering that the WALS core-200 sampling was designed never to catch these families, the Eurasian bias in the selected set is even stronger than the presented figures (every available Eurasian family was caught!).

3.2. The OV/VO-1,370 sample

As mentioned already, the OV/VO feature is the feature in WALS with the largest number of languages included (and also the feature which includes languages for the largest number of D-families). It is not clear how the data points/languages were selected, but it may be guessed that it is some kind of convenience sample (Dryer 2005b).

Again, it may be interesting to see to what extent the languages included are evenly sampled. The 1,370 sample contains languages from 244 D-families. We recognize a total of 394 D-families in the world. However, only 339 of them contain a language for which there is a publication with information to decide the OV/VO feature. In other words, WALS could be expanded to include the OV/VO feature for no less than 339 – 244 = 95 further language families.

However, it is not necessary to cover all possible families to achieve the goals of WALS, i.e., to provide a “systematic answer” to questions about the

---

8. One may also perhaps exclude Hurro-Urartian and Elamite, as we do with Etruscan, on the grounds that they are not sufficiently well-known.
geographical distribution of language features, as long as the families included are evenly sampled. We will now go on to discuss whether this is true for the WALS OV/VO-1,370 sample.

Table 3 shows the continental break-up of the language family coverage of the WALS OV/VO-1,370 sample. The figures refer to the number of D-families for which there is a publication with information to decide the OV/VO feature, versus D-families included in the WALS OV/VO-1,370 sample. The D-families with a description that are not included in the WALS OV/VO-1,370 sample are shown in italics.

As can be seen from Table 3, the sampling is not even across continents, ranging from 59.7% (Papua 55/92) to 86.1% (North America 56/65). If the inclusion of families were even across continents then all would have a coverage of around 72.0% (244/339). Are the differences we see statistically significant? Again, we estimate the $p$-values by simulation as follows: (i) generate 1000 random 244-member subsets of the 339 families for which sufficient information exists; (ii) for each family, ask how many of the subsets have more/less than the number found in the WALS OV/VO-1,370 sample. The results are shown in Table 4.

In plain words, we find that overinclusion of North American D-families is highly significant in the WALS OV/VO-1,370 sample, and that this happens at the expense of South American and Papuan D-families (but the significance levels of these underinclusions are lower). Overinclusion of Australian D-families is not significant with conventional levels of significance (Eurasian even less so).

Again, in fact, the underinclusion of Papuan D-families is excusable for the following reasons. The data for the Kwomtari, Baibai, Guriaso, and Yuat-Muramba languages is difficult to access (or too recent). The published basis available for assignment in the Pahoturi, Bayono-Awbono, Elseng, Abinomn, and Lepki is just a plain statement of the basic word order, without any explicit source or examples – a prudent researcher would perhaps require at least a sentence example to be convinced. For Amto-Musan only unanalysed text material is available (Krieg 1992) and in the case of Waia and Inanwatan, material is too recent for WALS. If we remove these 12 families from the list of sufficiently described families, 327 families remain in total, 80 (rather than 92) being Papuan. The inclusion proportion for Papuan languages rises from 59.7% (55/92) to 68.7% (55/80), and most importantly, underexclusion for Papuan families is no longer statistically significant ($p \approx 0.118$).

There are no similar considerations which would explain the dearth of South American families. It is clear that the neglect of South American languages lamented two decades ago by Derbyshire & Pullum (1986) is not quite over yet.
Sampling and genealogical coverage in WALS 115

Table 3. The continental break-up of the language family coverage of the WALS OV/VO sample. The figures refer to the number of D-families for which there is a publication with information to decide the OV/VO feature, versus D-families included in the WALS OV/VO sample. The D-families with a description that are not included in the WALS OV/VO sample are shown in italics.

<table>
<thead>
<tr>
<th>Area</th>
<th>Coverage</th>
<th>D-families</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>22/26</td>
<td>Anindilyakwa, Anson Bay, Bunaban, Eastern Daly, Garrwan, Gunwinyguan, Iwaidjan Proper, Larrakinyan, Limilngan, Mangarrayn-Maran, Manangrida, Minkin-Tangkil, Mirndi, Northern Daly, Nuynulunan, Pama-Nyungan, Southern Daly, Tiwi, Wagiman, Western Daly, Wonorrn, Yangmanic, Gaagudju, Giimbiyu, Jarrakan, Umbugarla</td>
</tr>
</tbody>
</table>
Table 3. (continued)

<table>
<thead>
<tr>
<th>Area</th>
<th>Coverage</th>
<th>D-families</th>
</tr>
</thead>
</table>
Sampling and genealogical coverage in WALS 117

Table 4. Tests for statistical significance of over/under-inclusion in the WALS OV/OV-1,370 sample. All $i$'s range up to 1000. $S_i[C] = |\{x | x \in S_i \text{ such that } x \text{ is from continent } C\}|$.

<table>
<thead>
<tr>
<th>Continent</th>
<th>W1,370</th>
<th>Question</th>
<th>Test</th>
<th>Outcome</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>North America</td>
<td>56/65</td>
<td>Overinclusion</td>
<td>$</td>
<td>{i</td>
<td>S_i[\text{North America}] \geq 56}</td>
</tr>
<tr>
<td>Australia</td>
<td>22/26</td>
<td>Overinclusion</td>
<td>$</td>
<td>{i</td>
<td>S_i[\text{Eurasia}] \geq 22}</td>
</tr>
<tr>
<td>South America</td>
<td>58/90</td>
<td>Underinclusion</td>
<td>$</td>
<td>{i</td>
<td>S_i[\text{South America}] \leq 58}</td>
</tr>
<tr>
<td>Papua</td>
<td>55/92</td>
<td>Underinclusion</td>
<td>$</td>
<td>{i</td>
<td>S_i[\text{Papua}] \leq 55}</td>
</tr>
</tbody>
</table>

4. Discussion

WALS is aimed at giving researchers a tool to investigate frequencies and correlations of language features (Comrie et al. 2005: 1). For this to be meaningful, the WALS data, at least if further refined, should be sufficient for approaching some level of inferential validity, without recourse to collecting further data.

We have shown that there are significant gaps in the breadth and depth coverage of the WALS data. Whether these gaps alter conclusions drawn in studies based on the WALS data depends on the specifics of the individual studies. What we indicate here is that the WALS data cannot be used blindly to draw statistically valid inferences about the state-of-the-art knowledge of the world’s languages.

There is a further point not to be forgotten as to the WALS data and statistical inferences, which has traditionally been overlooked in typology (but cf. Hammarström 2007a and Janssen et al. 2006). Statistically valid conclusions about a population can only be drawn if data is sampled at random from the population. Whatever the method(s) to used to include languages in WALS, it was not random, in the sense of picking from a plain list of languages, from lists of languages stratified into genera, or from a list of languages stratified into families. It is discernable that WALS data points were included based on convenient availability of data, and this may or may not turn out to be functionally equivalent (in terms of feature distributions) to a random selection. To test whether this is the case is beyond the scope of this study, but we may nevertheless speculate on how a skewing may come about; languages which are deep in the lowlands of New Guinea and deep in the Amazon forest may be less influenced by SVO contact languages, Malay/Indonesian and Spanish/Portuguese respectively, than their more accessible better documented counterparts. Now, of course, typologists must use data from documented languages rather than undocumented ones: we are certainly not attempting to imply any fault here;
what we do wish to emphasize is that blind statistical inferences about the world’s languages based on such data are not necessarily sound.

5. Conclusion

We have shown that even in the cases of the broadest and deepest data in WALS, they are not a perfect mirror of the genealogical diversity of the languages of the world as of today’s knowledge. In the 200 sample, there is a strong bias favouring Eurasian families, and in the 1,370 sample there is a strong bias favouring North American families. In both cases, there is also a weak but statistically significant bias disfavouring Papuan and South American families. In both cases, the WALS may be “excused” for the underinclusion of the Papuan families, but in neither case is the underexclusion of South American families or the overinclusion of Eurasian/North American families justifiable. The study depends a little on the classification of data size and availability, the demarcation of which is not watertight, and as more data becomes available in the future, the situation will change further. One may also discuss the level and detail of genealogical classification used here, but even so, the conclusions above are likely to remain. Caution is due when using the WALS data to draw statistical inferences.

Appendix: Descriptive resources for the languages of the world

http://dx.doi.org/10.1515/lity.2009.006_supp-1

Received: 8 August 2008
Revised: 6 January 2009

Correspondence address: Datavetenskap, Chalmers tekniska högskola, 412 96 Göteborg, Sweden; e-mail: harald2@chalmers.se

Acknowledgements: I wish to thank Matthew Dryer for a number of objections raised to a near-final draft of this article, some of which led to improvements in the explanation of the objective of this article, and some of which merit a separate response from him. I wish to thank the following libraries for granting access and services: Centralbiblioteket (Göteborg), Institutionen för orientaliska och afrikanska språk (Göteborg), Etnografiska Museet (Göteborg), LAL (Göteborg), Carolina Rediviva (Uppsala), NAI (Uppsala), Karin Boye (Uppsala), KB (Stockholm), SUB (Stockholm), LAL (Stockholm), Universiteitsbibliotheek (Leiden), KITLV (Leiden), Universiteitsbibliotheek (Amsterdam), Institute for Asian and African Studies (Helsinki), MPI-EVA (Leipzig), Universitätsbibliothek (Leipzig), Butler/Columbia University (New York City), IEA (Köln), BNF (Paris), SOAS (London), ILPGA (Paris), ZAS (Zürich). I am also indebted to (in no particular order) Hein van der Voort, Lincoln Almir Amarante Ribeiro, Eduardo Rivail Ribeiro, Michael Cysson, Nathan Hill, Jesús Mario Girón, Karsten Legère, Helene Fatima Idris, Bernard Comrie, Lionel M. Bender, John Kalespi, Hilaria de Sousa, Frank Seidel, Tor Guldemann, Lourens de Vries, Ian Tupper, Johanna Fenton, Randy Lebold, Willem Adelaar, Lyle Campbell, Norbert Cyffer, Maarten Mous, Thilo Schadeberg, Raoul Zamponi, Paul Whitehouse, Swintha Dambisa,
Lauren Campbell, Dmitry Idiatov, Nick Evans, Matthew Dryer, Mark Donohue, and Peter Bakker for help with access to data. The bibliographies by Alain Fabre (for South America) and Jouni Filip Maho (for Africa) have been very helpful in bibliographical searching leading up to this study.

References


What does this document list?

1. Each language family (or isolate) in the world which is:
   - “sufficiently” well-attested,¹
   - shown in publication to be internally coherent by orthodox comparative methodology (i.e., all members should be related to each other),
   - not demonstrably related (in publication) to any other family by orthodox comparative methodology.

   This delineation is meant to be authoritative (wherefore a “canonical source” is given, possibly with comments), and any errors/omissions/inconsistences pointed out will be greatly appreciated.

2. For each language family, the “most extensive” descriptive source for the most well-described language of the group. Most extensive descriptive source is defined according to the following hierarchy:

   (a) a published (full-length) descriptive grammar,
   (b) a published grammar sketch,
   (c) a published description of some element of grammar (i.e., noun class system, verb morphology, etc.),
   (d) a published phonological description,
   (e) a published dictionary,
   (f) a published text (collection),

¹The borderline cases and the insufficiently well-attested languages are NOT listed in this document, as they have no relevance to the points made in the main body of the article.
(g) a published wordlist,
(h) a published document with meta-information about the language
   (i.e., where spoken, non-intelligibility to other languages etc.),
(i) notes on unpublished manuscripts or people engaged in studying
   the language.

3. If it is not immediately clear that the given reference for a family
   contains enough data to set the OV/VO-parameter (see the main body
   of the paper for the relevance of this), an additional reference (perhaps
   with page numbers) is listed as well where OV/VO-information can
   be found. This reference is not always to the same language as the
   language with the (in general) most extensive description.

Correspondence address: Datavetenskap, Chalmers tekniska högskola, 412
96 Göteborg, Sweden; e-mail: harald2@chalmers.se

Note: This document reflects my point of view at the time of submission
with the original article on 1 August 2008. Following new data and inter-
pretations that have appeared since has caused me to change opinions (in
about ten cases) in the language family classification, but these have not
been incorporated in the present document (as that would be different from
what I submitted for review).
# Contents

1 Africa (37)

1.1 Afro-Asiatic (346) ............................................. 1
1.2 Atlantic-Congo (1400) ........................................ 2
1.3 Bangi Me (1) ..................................................... 4
1.4 Berta (1) ......................................................... 4
1.5 Central Sudanic (66) ........................................... 4
1.6 Dogon (9) ......................................................... 5
1.7 East Sudanic (92) ................................................ 6
1.8 Furan (3) .......................................................... 7
1.9 Gumuz (1) .......................................................... 7
1.10 Hadza (1) .......................................................... 8
1.11 Heiban (10) ....................................................... 8
1.12 *Hôã (1) ........................................................... 9
1.13 Ijoid (10) .......................................................... 10
1.14 Jalaa [probably extinct] (1) ................................ 10
1.15 Ju (6) ............................................................... 10
1.16 Kadugli-Krongo (6) ............................................. 11
1.17 Katla-Tima (2) ................................................... 11
1.18 Khoe-Kwadi (13) ................................................. 12
1.19 Koman (5) .......................................................... 12
1.20 Kujargé (1) ......................................................... 13
1.21 Kuliak (3) ......................................................... 13
1.22 Kunama (1) ......................................................... 14
1.23 Laal (1) ............................................................. 14
1.24 Maban (9) .......................................................... 14
1.25 Mande (71) ......................................................... 15
1.26 Mao (4) ............................................................. 16
1.27 Meroitic [extinct] (1) ............................................ 16
1.28 Narrow Talodi (8) ............................................... 17
1.29 Omotic (24) ....................................................... 18
1.30 Ongota (1) ........................................................ 19
1.31 Rashad (3) ........................................................ 19
1.32 Saharan (9) ....................................................... 20
## CONTENTS

1.33 Sandawe (1) .............................................. 21  
1.34 Shabo (1) .............................................. 21  
1.35 Songhay (8) ............................................ 22  
1.36 Tegem (1) .............................................. 22  
1.37 Tuu (5) ................................................ 23

2 Australia (30) ............................................. 25  
2.1 Anindilyakwa (1) .................................... 25  
2.2 Anson Bay (2) ........................................ 25  
2.3 Bunaban (2) .......................................... 26  
2.4 Eastern Daly (2) ....................................... 26  
2.5 Gaagudju (1) .......................................... 27  
2.6 Garrwan (1) ........................................... 28  
2.7 Giimbuyu [extinct] (3) .......................... 28  
2.8 Gunwinyguan (9) .................................... 29  
2.9 Iwaidjan Proper (5) .............................. 29  
2.10 Jarrakan (3) .......................................... 30  
2.11 Kungarakany [extinct] (1) ................. 31  
2.12 Larrakiyian (1) ..................................... 31  
2.13 Limilngan (2) ........................................ 32  
2.14 Maran (4) ............................................ 32  
2.15 Maningrida (4) ..................................... 33  
2.16 Marrku-Wurrugu [extinct] (2) ........ 34  
2.17 Minkin-Tangkic (4) .............................. 34  
2.18 Mirndi (5) ........................................... 35  
2.19 Northern Daly (2) ............................... 35  
2.20 Nyulnyulan (9) ..................................... 36  
2.21 Pama-Nyungan (175) .......................... 36  
2.22 Southern Daly (2) ............................... 37  
2.23 Oyster Bay, Big River, Little Swanport (Tasmanian) [extinct] (3) .......................... 37  
2.24 Piper River, Cape Portland, Ben Lomond (Tasmanian) [extinct] (3) .......................... 37  
2.25 Tiwi (1) .............................................. 38  
2.26 Umbugarla-Ngurmbur (2) .................. 38  
2.27 Wagiman (1) ........................................ 39  
2.28 Western Daly (10) .............................. 39  
2.29 Worroran (7) ...................................... 40  
2.30 Yangmanic (3) ..................................... 40

3 Eurasia (33) ................................................. 43  
3.1 Abkhaz-Adyge (5) .................................. 43  
3.2 Ainu (1) ............................................... 43  
3.3 Austroasiatic (168) ............................. 43
3.4 Basque (3) ............................................ 44
3.5 Burushaski (1) ........................................ 44
3.6 Chukotko-Kamchatkan (5) .......................... 44
3.7 Dravidian (73) ......................................... 45
3.8 Elamite [extinct] (1) ................................. 45
3.9 Etruscan [extinct] (1) ............................... 46
3.10 Great Andamanese (10) ............................ 46
3.11 Hattic [extinct] (1) ................................. 46
3.12 Hurro-Urartian [extinct] (2) ....................... 47
3.13 Iberian [extinct] (1) ............................... 47
3.14 Indo-European (449) .................. 48
3.15 Japanese (12) ......................................... 48
3.16 Jarawa-Onge (2) ...................................... 48
3.17 Kartvelian (5) ......................................... 49
3.18 Korean (1) ........................................... 49
3.19 Kusunda (1) ........................................... 50
3.20 Miao-Yao (35) ........................................ 50
3.21 Mongolian (14) ....................................... 50
3.22 Nakh-Dagestanian (29) ............................ 51
3.23 Nihali (1) ............................................. 51
3.24 Nivkh (1) ............................................. 51
3.25 Shom Pen (1) ......................................... 52
3.26 Sino-Tibetan (402) ................................. 52
3.27 Sumerian [extinct] (1) ............................ 52
3.28 Tai-Kadai (76) ....................................... 53
3.29 Tungusic (12) ........................................ 53
3.30 Turkic (40) ............................................ 53
3.31 Uralic (39) ........................................... 54
3.32 Yeniseian (2) ........................................ 54
3.33 Yukaghir (2) .......................................... 54

4 North America (72) ..................................... 55
4.1 Adai [extinct] (1) ....................................... 55
4.2 Algin (44) ............................................. 55
4.3 Alsea [extinct] (1) ..................................... 56
4.4 Atakapan [extinct] (1) .............................. 56
4.5 Beothuk [extinct] (1) ............................... 56
4.6 Caddoan (5) ........................................... 57
4.7 Cayuse [extinct] (1) ................................... 57
4.8 Chimakuan [extinct] (2) ............................ 57
4.9 Chimariko [extinct] (1) ............................ 58
4.10 Chinook (2) ........................................... 58
4.11 Chitimacha [extinct] (1) .......................... 58
4.12 Chumashan [extinct] (7) ......................... 59
<table>
<thead>
<tr>
<th>Section</th>
<th>Language</th>
<th>Codes</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.13</td>
<td>Coahuilteco [extinct] (1)</td>
<td></td>
<td>59</td>
</tr>
<tr>
<td>4.14</td>
<td>Cochimi-Yuman (9)</td>
<td></td>
<td>60</td>
</tr>
<tr>
<td>4.15</td>
<td>Comecrudan [extinct] (1)</td>
<td></td>
<td>60</td>
</tr>
<tr>
<td>4.16</td>
<td>Coosan (1)</td>
<td></td>
<td>60</td>
</tr>
<tr>
<td>4.17</td>
<td>Cotoname [extinct] (1)</td>
<td></td>
<td>61</td>
</tr>
<tr>
<td>4.18</td>
<td>Cuitlatec [extinct] (1)</td>
<td></td>
<td>61</td>
</tr>
<tr>
<td>4.19</td>
<td>Eskimo-Aleut (11)</td>
<td></td>
<td>61</td>
</tr>
<tr>
<td>4.20</td>
<td>Esselen [extinct] (1)</td>
<td></td>
<td>62</td>
</tr>
<tr>
<td>4.21</td>
<td>Eyak-Athapaskan-Tlingit (45)</td>
<td></td>
<td>62</td>
</tr>
<tr>
<td>4.22</td>
<td>Guaicurian [extinct] (1)</td>
<td></td>
<td>63</td>
</tr>
<tr>
<td>4.23</td>
<td>Haida (2)</td>
<td></td>
<td>63</td>
</tr>
<tr>
<td>4.24</td>
<td>Huavean (4)</td>
<td></td>
<td>63</td>
</tr>
<tr>
<td>4.25</td>
<td>Iroquoian (11)</td>
<td></td>
<td>64</td>
</tr>
<tr>
<td>4.26</td>
<td>Jicaquean (1)</td>
<td></td>
<td>64</td>
</tr>
<tr>
<td>4.27</td>
<td>Kalapuyan (1)</td>
<td></td>
<td>65</td>
</tr>
<tr>
<td>4.28</td>
<td>Karankawa [extinct] (1)</td>
<td></td>
<td>65</td>
</tr>
<tr>
<td>4.29</td>
<td>Karuk (1)</td>
<td></td>
<td>65</td>
</tr>
<tr>
<td>4.30</td>
<td>Keresan (2)</td>
<td></td>
<td>66</td>
</tr>
<tr>
<td>4.31</td>
<td>Kiowa-Tanoan (6)</td>
<td></td>
<td>66</td>
</tr>
<tr>
<td>4.32</td>
<td>Klamath-Modoc (1)</td>
<td></td>
<td>66</td>
</tr>
<tr>
<td>4.33</td>
<td>Kutenai (1)</td>
<td></td>
<td>67</td>
</tr>
<tr>
<td>4.34</td>
<td>Lencan [extinct] (1)</td>
<td></td>
<td>67</td>
</tr>
<tr>
<td>4.35</td>
<td>Maiduan (4)</td>
<td></td>
<td>68</td>
</tr>
<tr>
<td>4.36</td>
<td>Maratino [extinct] (1)</td>
<td></td>
<td>68</td>
</tr>
<tr>
<td>4.37</td>
<td>Mayan (69)</td>
<td></td>
<td>68</td>
</tr>
<tr>
<td>4.38</td>
<td>Misumalpan (4)</td>
<td></td>
<td>69</td>
</tr>
<tr>
<td>4.39</td>
<td>Miwok-Costanoan (10)</td>
<td></td>
<td>69</td>
</tr>
<tr>
<td>4.40</td>
<td>Mixe-Zoque (17)</td>
<td></td>
<td>70</td>
</tr>
<tr>
<td>4.41</td>
<td>Molala [extinct] (1)</td>
<td></td>
<td>70</td>
</tr>
<tr>
<td>4.42</td>
<td>Muskogean (6)</td>
<td></td>
<td>70</td>
</tr>
<tr>
<td>4.43</td>
<td>Natchez [extinct] (1)</td>
<td></td>
<td>71</td>
</tr>
<tr>
<td>4.44</td>
<td>Otomanguean (179)</td>
<td></td>
<td>71</td>
</tr>
<tr>
<td>4.45</td>
<td>Palaihnihan (2)</td>
<td></td>
<td>72</td>
</tr>
<tr>
<td>4.46</td>
<td>Pomoan (7)</td>
<td></td>
<td>72</td>
</tr>
<tr>
<td>4.47</td>
<td>Sahaptian (5)</td>
<td></td>
<td>73</td>
</tr>
<tr>
<td>4.48</td>
<td>Salinan [extinct] (1)</td>
<td></td>
<td>73</td>
</tr>
<tr>
<td>4.49</td>
<td>Salishan (27)</td>
<td></td>
<td>73</td>
</tr>
<tr>
<td>4.50</td>
<td>Seri (1)</td>
<td></td>
<td>74</td>
</tr>
<tr>
<td>4.51</td>
<td>Shasta [extinct] (1)</td>
<td></td>
<td>74</td>
</tr>
<tr>
<td>4.52</td>
<td>Siouan (17)</td>
<td></td>
<td>74</td>
</tr>
<tr>
<td>4.53</td>
<td>Sinlaw [extinct] (1)</td>
<td></td>
<td>75</td>
</tr>
<tr>
<td>4.54</td>
<td>Takelma (1)</td>
<td></td>
<td>75</td>
</tr>
<tr>
<td>4.55</td>
<td>Tarascan (2)</td>
<td></td>
<td>76</td>
</tr>
<tr>
<td>4.56</td>
<td>Tequistlatecan (2)</td>
<td></td>
<td>76</td>
</tr>
</tbody>
</table>
CONTENTS

4.57 Timucua [extinct] (1) ........................................ 76
4.58 Tonkawa [extinct] (1) ........................................ 77
4.59 Totonacan (11) ................................................ 77
4.60 Tsimshian (3) .................................................. 77
4.61 Tunica [extinct] (1) ............................................ 78
4.62 Uto-Aztecan (61) .............................................. 78
4.63 Wakashan (5) ................................................... 78
4.64 Wappo (1) ....................................................... 79
4.65 Washo (1) ....................................................... 79
4.66 Wintuan (1) ..................................................... 80
4.67 Xincan (1) ....................................................... 80
4.68 Yana (1) ........................................................ 81
4.69 Yokutsan (1) .................................................... 81
4.70 Yuchi (1) ........................................................ 81
4.71 Yuki (1) ........................................................ 82
4.72 Zuni (1) ........................................................ 82

5 Papua (111) ....................................................... 83
5.1 Abinomn (1) ..................................................... 83
5.2 Abun (1) ........................................................ 83
5.3 Amto-Musan (2) ............................................... 84
5.4 Aném (1) ........................................................ 84
5.5 Angan (13) ....................................................... 85
5.6 Arafundi (1) ..................................................... 86
5.7 Ata (1) ........................................................... 86
5.8 Austronesian (1275) .......................................... 87
5.9 Awin-Pa (3) ..................................................... 88
5.10 Baibai (2) ....................................................... 88
5.11 Baining (6) ..................................................... 89
5.12 Bayono-Awbono (2) ......................................... 89
5.13 Biksi (2) ........................................................ 89
5.14 Bilua (1) ........................................................ 90
5.15 Binanderean (11) ............................................ 91
5.16 Border (15) ..................................................... 91
5.17 Bosavi (10) ..................................................... 91
5.18 Bulaka River (2) ............................................. 92
5.19 Burmese (1) .................................................... 92
5.20 Busa (Odiai) (1) ............................................. 93
5.21 Cenderawasih Bay (11) .................................... 93
5.22 Dem (1) ........................................................ 94
5.23 Duna-Bogaya (2) ............................................. 94
5.24 Duramin (1) ................................................... 95
5.25 East Bird’s Head (3) ....................................... 95
5.26 East Kutubu (2) ............................................. 96
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.27</td>
<td>East Timor (3)</td>
<td>96</td>
</tr>
<tr>
<td>5.28</td>
<td>Eastern Trans-Fly (4)</td>
<td>97</td>
</tr>
<tr>
<td>5.29</td>
<td>Eleman (6)</td>
<td>98</td>
</tr>
<tr>
<td>5.30</td>
<td>Elseng (1)</td>
<td>99</td>
</tr>
<tr>
<td>5.31</td>
<td>Fasu (1)</td>
<td>99</td>
</tr>
<tr>
<td>5.32</td>
<td>Goilalan (5)</td>
<td>100</td>
</tr>
<tr>
<td>5.33</td>
<td>Guriaso (1)</td>
<td>100</td>
</tr>
<tr>
<td>5.34</td>
<td>Hatam (1)</td>
<td>101</td>
</tr>
<tr>
<td>5.35</td>
<td>Inanvatan (2)</td>
<td>101</td>
</tr>
<tr>
<td>5.36</td>
<td>Kaki Ae (1)</td>
<td>101</td>
</tr>
<tr>
<td>5.37</td>
<td>Karkar (Yuri) (1)</td>
<td>102</td>
</tr>
<tr>
<td>5.38</td>
<td>Kaure-Kapori (4)</td>
<td>102</td>
</tr>
<tr>
<td>5.39</td>
<td>Kayagar (3)</td>
<td>103</td>
</tr>
<tr>
<td>5.40</td>
<td>Kehu (1)</td>
<td>103</td>
</tr>
<tr>
<td>5.41</td>
<td>Kembra (1)</td>
<td>103</td>
</tr>
<tr>
<td>5.42</td>
<td>Kiwaian (6)</td>
<td>104</td>
</tr>
<tr>
<td>5.43</td>
<td>Koianian (7)</td>
<td>104</td>
</tr>
<tr>
<td>5.44</td>
<td>Kol (1)</td>
<td>104</td>
</tr>
<tr>
<td>5.45</td>
<td>Kolana-Tinglapui (3)</td>
<td>105</td>
</tr>
<tr>
<td>5.46</td>
<td>Kolopom (3)</td>
<td>105</td>
</tr>
<tr>
<td>5.47</td>
<td>Konda-Yahadian (2)</td>
<td>106</td>
</tr>
<tr>
<td>5.48</td>
<td>Koia (1)</td>
<td>106</td>
</tr>
<tr>
<td>5.49</td>
<td>Kwerba (8)</td>
<td>107</td>
</tr>
<tr>
<td>5.50</td>
<td>Kwomtari (2)</td>
<td>107</td>
</tr>
<tr>
<td>5.51</td>
<td>Lakes Plain (20)</td>
<td>108</td>
</tr>
<tr>
<td>5.52</td>
<td>Lavukaleve (1)</td>
<td>108</td>
</tr>
<tr>
<td>5.53</td>
<td>Left May (6)</td>
<td>109</td>
</tr>
<tr>
<td>5.54</td>
<td>Leplki (1)</td>
<td>109</td>
</tr>
<tr>
<td>5.55</td>
<td>Lower Mamberamo (2)</td>
<td>110</td>
</tr>
<tr>
<td>5.56</td>
<td>Lower Sepik-Ramu (Extended)</td>
<td>110</td>
</tr>
<tr>
<td>5.57</td>
<td>Mairasi (3)</td>
<td>110</td>
</tr>
<tr>
<td>5.58</td>
<td>Marind (6)</td>
<td>111</td>
</tr>
<tr>
<td>5.59</td>
<td>Masep (1)</td>
<td>112</td>
</tr>
<tr>
<td>5.60</td>
<td>Maybrat (2)</td>
<td>112</td>
</tr>
<tr>
<td>5.61</td>
<td>Molof (1)</td>
<td>113</td>
</tr>
<tr>
<td>5.62</td>
<td>Mombum (2)</td>
<td>113</td>
</tr>
<tr>
<td>5.63</td>
<td>Mongol-Langam (3)</td>
<td>114</td>
</tr>
<tr>
<td>5.64</td>
<td>Mor (1)</td>
<td>114</td>
</tr>
<tr>
<td>5.65</td>
<td>Moraori (1)</td>
<td>115</td>
</tr>
<tr>
<td>5.66</td>
<td>Morehead and Upper Maro Rivers (17)</td>
<td>115</td>
</tr>
<tr>
<td>5.67</td>
<td>Mpur (1)</td>
<td>116</td>
</tr>
<tr>
<td>5.68</td>
<td>Murkim (1)</td>
<td>117</td>
</tr>
<tr>
<td>5.69</td>
<td>Nimboran (5)</td>
<td>117</td>
</tr>
<tr>
<td>5.70</td>
<td>North Bougainville (4)</td>
<td>117</td>
</tr>
<tr>
<td>5.71</td>
<td>North Halmahera (16)</td>
<td>118</td>
</tr>
<tr>
<td>5.72</td>
<td>Oksapmin (1)</td>
<td>118</td>
</tr>
<tr>
<td>5.73</td>
<td>Paloturi (2)</td>
<td>119</td>
</tr>
<tr>
<td>5.74</td>
<td>Papi (1)</td>
<td>119</td>
</tr>
<tr>
<td>5.75</td>
<td>Pauwasi (4)</td>
<td>120</td>
</tr>
<tr>
<td>5.76</td>
<td>Paweia (1)</td>
<td>121</td>
</tr>
<tr>
<td>5.77</td>
<td>Piawi (2)</td>
<td>121</td>
</tr>
<tr>
<td>5.78</td>
<td>Porome (Kibiri) (1)</td>
<td>122</td>
</tr>
<tr>
<td>5.79</td>
<td>Pyu (1)</td>
<td>122</td>
</tr>
<tr>
<td>5.80</td>
<td>Savosavo (1)</td>
<td>123</td>
</tr>
<tr>
<td>5.81</td>
<td>Senagi (2)</td>
<td>123</td>
</tr>
<tr>
<td>5.82</td>
<td>Sentani (4)</td>
<td>124</td>
</tr>
<tr>
<td>5.83</td>
<td>Sepik (48)</td>
<td>124</td>
</tr>
<tr>
<td>5.84</td>
<td>Sko (7)</td>
<td>125</td>
</tr>
<tr>
<td>5.85</td>
<td>South Bird’s Head Proper (6)</td>
<td>126</td>
</tr>
<tr>
<td>5.86</td>
<td>South Bougainville (9)</td>
<td>126</td>
</tr>
<tr>
<td>5.87</td>
<td>Suki-Gogodala (4)</td>
<td>127</td>
</tr>
<tr>
<td>5.88</td>
<td>Sulka (1)</td>
<td>127</td>
</tr>
<tr>
<td>5.89</td>
<td>Taiap (1)</td>
<td>127</td>
</tr>
<tr>
<td>5.90</td>
<td>Tambora [extinct] (1)</td>
<td>128</td>
</tr>
<tr>
<td>5.91</td>
<td>Tanahmerah (1)</td>
<td>128</td>
</tr>
<tr>
<td>5.92</td>
<td>Taulil-Butam (1)</td>
<td>129</td>
</tr>
<tr>
<td>5.93</td>
<td>Teberan (2)</td>
<td>129</td>
</tr>
<tr>
<td>5.94</td>
<td>Tirio (5)</td>
<td>130</td>
</tr>
<tr>
<td>5.95</td>
<td>Tofanma (1)</td>
<td>131</td>
</tr>
<tr>
<td>5.96</td>
<td>Tor-Orya (13)</td>
<td>131</td>
</tr>
<tr>
<td>5.97</td>
<td>Torricelli (53)</td>
<td>132</td>
</tr>
<tr>
<td>5.98</td>
<td>Touo (1)</td>
<td>132</td>
</tr>
<tr>
<td>5.99</td>
<td>Trans New Guinea (338)</td>
<td>133</td>
</tr>
<tr>
<td>5.100</td>
<td>Turama-Kikori (3)</td>
<td>134</td>
</tr>
<tr>
<td>5.101</td>
<td>Uhunduni (1)</td>
<td>134</td>
</tr>
<tr>
<td>5.102</td>
<td>Usku (1)</td>
<td>135</td>
</tr>
<tr>
<td>5.103</td>
<td>Waia (1)</td>
<td>135</td>
</tr>
<tr>
<td>5.104</td>
<td>Walio (4)</td>
<td>136</td>
</tr>
<tr>
<td>5.105</td>
<td>West Bird’s Head (5)</td>
<td>137</td>
</tr>
<tr>
<td>5.106</td>
<td>West Bomberai (3)</td>
<td>138</td>
</tr>
<tr>
<td>5.107</td>
<td>West Timor-Alor-Pantar (16)</td>
<td>139</td>
</tr>
<tr>
<td>5.108</td>
<td>Yalé (Nagatman) (1)</td>
<td>140</td>
</tr>
<tr>
<td>5.109</td>
<td>Yawa (2)</td>
<td>141</td>
</tr>
<tr>
<td>5.110</td>
<td>Yéli Dnye (1)</td>
<td>141</td>
</tr>
<tr>
<td>5.111</td>
<td>Yuat-Maramba (6)</td>
<td>141</td>
</tr>
<tr>
<td>Section</td>
<td>Language</td>
<td>Code</td>
</tr>
<tr>
<td>--------------</td>
<td>---------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>6.1</td>
<td>Aikaná (1)</td>
<td></td>
</tr>
<tr>
<td>6.2</td>
<td>Aimoré (1)</td>
<td></td>
</tr>
<tr>
<td>6.3</td>
<td>Andaqui [extinct] (1)</td>
<td></td>
</tr>
<tr>
<td>6.4</td>
<td>Andoque (1)</td>
<td></td>
</tr>
<tr>
<td>6.5</td>
<td>Arara do Rio Branco (1)</td>
<td></td>
</tr>
<tr>
<td>6.6</td>
<td>Araucanian (2)</td>
<td></td>
</tr>
<tr>
<td>6.7</td>
<td>Arawá (8)</td>
<td></td>
</tr>
<tr>
<td>6.8</td>
<td>Arawak (62)</td>
<td></td>
</tr>
<tr>
<td>6.9</td>
<td>Atacame (Esmeraldeño) [extinct] (1)</td>
<td></td>
</tr>
<tr>
<td>6.10</td>
<td>Atacameño (Kunza) [extinct] (1)</td>
<td></td>
</tr>
<tr>
<td>6.11</td>
<td>Awaké (Arutani, Uruak) (1)</td>
<td></td>
</tr>
<tr>
<td>6.12</td>
<td>Awshiri (Tekiráka) [presumed extinct] (1)</td>
<td></td>
</tr>
<tr>
<td>6.13</td>
<td>Aymara (3)</td>
<td></td>
</tr>
<tr>
<td>6.14</td>
<td>Barbacoan (6)</td>
<td></td>
</tr>
<tr>
<td>6.15</td>
<td>Betoi-Jirara [extinct] (1)</td>
<td></td>
</tr>
<tr>
<td>6.16</td>
<td>Bora-Huitoto (6)</td>
<td></td>
</tr>
<tr>
<td>6.17</td>
<td>Bororo (3)</td>
<td></td>
</tr>
<tr>
<td>6.18</td>
<td>Cahuapanan (2)</td>
<td></td>
</tr>
<tr>
<td>6.19</td>
<td>Candoshi-Shapra (1)</td>
<td></td>
</tr>
<tr>
<td>6.20</td>
<td>Canichana [extinct] (1)</td>
<td></td>
</tr>
<tr>
<td>6.21</td>
<td>Carib (32)</td>
<td></td>
</tr>
<tr>
<td>6.22</td>
<td>Cayuvava (1)</td>
<td></td>
</tr>
<tr>
<td>6.23</td>
<td>Chapacura-Wanham (5)</td>
<td></td>
</tr>
<tr>
<td>6.24</td>
<td>Charrúa [extinct] (3)</td>
<td></td>
</tr>
<tr>
<td>6.25</td>
<td>Chibchan (21)</td>
<td></td>
</tr>
<tr>
<td>6.26</td>
<td>Chiquitano (1)</td>
<td></td>
</tr>
<tr>
<td>6.27</td>
<td>Chocoan (12)</td>
<td></td>
</tr>
<tr>
<td>6.28</td>
<td>Chonan (2)</td>
<td></td>
</tr>
<tr>
<td>6.29</td>
<td>Chono [extinct] (1)</td>
<td></td>
</tr>
<tr>
<td>6.30</td>
<td>Cofán (1)</td>
<td></td>
</tr>
<tr>
<td>6.31</td>
<td>Culli [probably extinct] (1)</td>
<td></td>
</tr>
<tr>
<td>6.32</td>
<td>Fulniö (1)</td>
<td></td>
</tr>
<tr>
<td>6.33</td>
<td>Guahibo (5)</td>
<td></td>
</tr>
<tr>
<td>6.34</td>
<td>Guacuruan (5)</td>
<td></td>
</tr>
<tr>
<td>6.35</td>
<td>Guamo [extinct] (1)</td>
<td></td>
</tr>
<tr>
<td>6.36</td>
<td>Guató (1)</td>
<td></td>
</tr>
<tr>
<td>6.37</td>
<td>Harakmbut (2)</td>
<td></td>
</tr>
<tr>
<td>6.38</td>
<td>Hibito-Cholon [extinct] (2)</td>
<td></td>
</tr>
<tr>
<td>6.39</td>
<td>Huarpean [extinct] (2)</td>
<td></td>
</tr>
<tr>
<td>6.40</td>
<td>Iranxe (Münkii) (1)</td>
<td></td>
</tr>
<tr>
<td>6.41</td>
<td>Itonama (1)</td>
<td></td>
</tr>
<tr>
<td>6.42</td>
<td>Jabutí (2)</td>
<td></td>
</tr>
<tr>
<td>6.43</td>
<td>Jê (16)</td>
<td></td>
</tr>
</tbody>
</table>
CONTENTS

6.44 Jirajaran [extinct] (1) ........................................... 160
6.45 Jivaro (4) ........................................................... 160
6.46 Jodi (1) ............................................................. 161
6.47 Kakua-Nukak (2) .................................................. 161
6.48 Kamakã [extinct] (1) .............................................. 162
6.49 Kamsá (1) ........................................................... 162
6.50 Kanoê (Kapixaná) (1) ............................................. 163
6.51 Karajá (1) ........................................................... 163
6.52 Kariri/Kiriri [extinct] (1) ........................................ 163
6.53 Katukina (3) ......................................................... 164
6.54 Kawesqar (2) ........................................................ 165
6.55 Kwazá (1) ........................................................... 165
6.56 Leko (1) ............................................................. 165
6.57 Lengu-Masoy (Enlhet) (5) ...................................... 166
6.58 Lule [extinct] (1) .................................................... 166
6.59 Máku [extinct] (1) ................................................... 166
6.60 Matacoan (Mataguayo) (7) ..................................... 167
6.61 Matanawí [extinct] (1) ............................................ 167
6.62 Maxakalí (1) ........................................................ 168
6.63 Mochica [probably extinct] (1) ................................ 168
6.64 Mosetén-Chimane (1) ............................................ 169
6.65 Movima (1) ........................................................ 169
6.66 Muniche (1) ........................................................ 169
6.67 Mura-Pirahã (1) .................................................... 169
6.68 Nadahup (4) ......................................................... 170
6.69 Nambiquaran (3) .................................................. 170
6.70 Ofaíé (1) ........................................................... 170
6.71 Omurano [extinct] (1) ............................................. 170
6.72 Oti [extinct] (1) ....................................................... 171
6.73 Otomaco [extinct] (1) ............................................. 171
6.74 Páez (1) ............................................................. 172
6.75 Pankararu (Panakararu) [extinct] (1) ..................... 172
6.76 Panoan (28) ........................................................ 172
6.77 Peba-Yagua (2) ..................................................... 173
6.78 Puelche (1) .......................................................... 173
6.79 Puinave (1) .......................................................... 174
6.80 Puquina [extinct] (1) .............................................. 174
6.81 Puri [extinct] (1) .................................................... 174
6.82 Quechuan (46) ....................................................... 175
6.83 Rikbaktsá (1) ......................................................... 175
6.84 Saliban (3) .......................................................... 175
6.85 Sapé (Kaliana) (1) ................................................ 176
6.86 Sechuran [extinct] (1) ............................................ 176
6.87 Tacanan (6) ......................................................... 177
6.88 Tallán [extinct] (1) .......... 177
6.89 Taruma [extinct] (1) .......... 178
6.90 Taushiro (1) ................. 178
6.91 Ticuna (Tucuna) (1) .......... 178
6.92 Timote-Cuica [probably extinct] (1) .......... 179
6.93 Tinigua (1) .................. 179
6.94 Trumai (1) .................. 180
6.95 Tucanoan (25) ............... 180
6.96 Tupí (76) .................. 180
6.97 Tuxá [extinct] (1) .......... 181
6.98 Urarina (1) ................ 181
6.99 Uru-Chipaya (2) ............ 182
6.100 Vilela [extinct] (1) ....... 182
6.101 Waorani (1) ............... 182
6.102 Warao (1) ................ 183
6.103 Xukuru [extinct] (1) ....... 183
6.104 Yámana (Yahgan) (1) .... 183
6.105 Yanomam (4) ............... 184
6.106 Yaruro (Pumé) (1) ......... 184
6.107 Yurakare (1) ............... 184
6.108 Yurí (1) .................. 185
6.109 Yurumanguí [extinct] (1) ... 185
6.110 Zamucoan (2) .............. 185
6.111 Zaparoan (6) .............. 186
Chapter 1

Africa (37)

1.1 Afro-Asiatic (346)

Core area: N Africa

Canonical source: [1], [2], [3]

Classification comment: Omotic does not show diagnostic Afro-Asiatic features ([4], [5], [6], [7]), and is therefore excluded until such features can be shown.

Documentation: (Full) Grammar ([8])


CHAPTER 1. AFRICA (37)


1.2 Atlantic-Congo (1400)

Core area: W and E Africa

Canonical source: [9]

Classification comment: The core is held together by regular sound correspondences in lexical items between subfamily languages, less systematic verbal extensions, and noun class systems ([10], [11], [12], [13], [14], [15], [16], [17]). Mande, Kordofanian, Ijoid, and Dogon have not yet been shown to stand up to these criteria wherefore they are excluded. The Pre or Mbre language in Côte d’Ivoire, known from wordlists only, is Atlantic-Congo but cannot be assigned to a specific branch at this time ([18]). The Mpra [= Mpre] language in Ghana has cognates with Atlantic-Congo especially Guang ([19]), including numerals 2 to 5, but the bulk of the little vocabulary that is known is not Atlantic-Congo ([20]). On the grounds that the numerals are less likely to be borrowed, and that lexical innovation may produce vocabulary that looks unrelated to anything else, we count Mpra as an Atlantic-Congo language.

Documentation: (Full) Grammar ([21])


1.2. ATLANTIC-CONGO (1400)


1.3 Bangi Me (1)

Core area: Mali

Canonical source: [22]

Documentation: Grammar sketch ([22], [23])


1.4 Berta (1)

Core area: Sudan, Ethiopia

Canonical source: [24], [25]

Documentation: Grammar sketch ([25], [26])


1.5 Central Sudanic (66)

Core area: Chad, Sudan, CAR

Canonical source: [27], [28]
1.6. DOGON (9)

Classification comment: No conclusive, methodologically sound basis for assigning Central Sudanic to an alleged full or partial Nilo-Saharan has been presented ([29], [30]).

Documentation: (Full) Grammar ([31])


1.6  Dogon (9)

Core area: W Africa

Canonical source: [32], [33]

Classification comment: Dogon has not yet been shown to contain systematic sound correspondences, noun class systems, or verbal extensions to Niger-Congo in some form. Bangi Me is taken out of Dogon in view of [34].

Documentation: (Full) Grammar ([35])

CHAPTER 1. AFRICA (37)


1.7 East Sudanic (92)

Core area: E Africa

Canonical source: [36]

Classification comment: No conclusive, methodologically sound basis for assigning Eastern Sudanic to an alleged full or partial Nilo-Saharan has been presented ([37], [38]). The latest attempt at demonstrating the unity of East Sudanic ([36]) has not yet been discredited wherefore it is instated tentatively.

Documentation: (Full) Grammar ([39])


1.8 Furan (3)

Core area: W Sudan

Canonical source: [40]

Classification comment: No conclusive, methodologically sound basis for assigning Furan to an alleged full or partial Nilo-Saharan has been presented ([41], [42]).

Documentation: (Full) Grammar ([40])


1.9 Gumuz (1)

Core area: Sudan, Ethiopia

Canonical source: [43], [44]

Classification comment: Gumuz doesn’t not show much in the way of alleged Nilo-Saharan features ([43], [44]).

Documentation: Grammar sketch ([43: 50–51])


1.10 Hadza (1)

Core area: Tanzania

Canonical source: [45]

Documentation: Grammar sketch ([46]; in particular, for setting the OV/VO-feature: [47: 120])


1.11 Heiban (10)

Core area: Sudan

Canonical source: [48]

Classification comment: Heiban is held together by pronouns, lexical cognates, and form-meaning resemblances in the noun class systems, though with a small question mark for Warnang. The alleged noun class form/meaning correspondences to other Kordofanian groups, namely Rashad and Talodi (with Tegem), in [49] and [50] are not convincing as conclusively genetic; as to form, there is much eclectic selection as well as irregular correspondences, and as to meaning, the methodology of finding non-cognate shared class items is highly dubious. Chance resemblance has not been excluded. The lexical evidence for a Kordofanian unity is weak in the extreme ([51]). It follows that the arguments presented for a Niger-Congo affiliation are not compelling either.

Documentation: Grammar sketch ([52]; in particular, for setting the OV/VO-feature: [53: 282–287])

1.12  *Hõã* (1)


[52] Black, K. & B. Black. 1971. *The Moro language: Grammar and dictionary* (Linguistics Monograph Series 6). Khartoum: Sudan Research Unit, University of Khartoum. [States authors as "Mr. and Mrs. K. Black". However, the authors’ full first names are Keith and Betty.]


1.12  *Hõã* (1)

Core area: Botswana

Canonical source: [54]

Documentation: Grammar sketch ([55]; in particular, for setting the OV/VO-feature: [56])


1.13  Ijoid (10)

Core area: Nigeria

Canonical source: [57]

Classification comment: Ijoid has not yet been shown to contain systematic sound correspondences, noun class systems, or verbal extensions with Niger-Congo in some form.

Documentation: (Full) Grammar ([58])


1.14  Jalaa [probably extinct] (1)

Core area: NE Nigeria

Canonical source: [59]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([59])


1.15  Ju (6)

Core area: N Kalahari

Canonical source: [60]

Documentation: (Full) Grammar ([61])


1.16 Kadugli-Krongo (6)

Core area: Nuba Mountains

Canonical source: [62]

Documentation: (Full) Grammar ([63])


1.17 Katla-Tima (2)

Core area: Sudan

Canonical source: [64]

Classification comment: Katla and Tima are related ([65: 190–196]). There is no reason to adhere to the suggestion of a relationship to Heiban, Narrow Talodi, Tegem, or Rashad ([64], [66]), since there are no noun classes and the lexical evidence is weak in the extreme ([67]).

Documentation: Grammar sketch ([68]; in particular, for setting the OV/VO-feature: [65: 1195], [69: 705])


1.18 Khoe-Kwadi (13)

Core area: SW Africa

Canonical source: /70/

Documentation: (Full) Grammar (/71/)


1.19 Koman (5)

Core area: Sudan, Ethiopia

Canonical source: /72/

Classification comment: No conclusive, methodologically sound basis for assigning Koman to an alleged full or partial Nilo-Saharan has been presented (/73/, /74/).

Documentation: Grammar sketch (/75; in particular, for setting the OV/VO-feature: /76/)
1.20 Kujargé (1)

Core area: Wadai, Darfur

Canonical source: [77]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([77])


1.21 Kuliak (3)

Core area: N Uganda

Canonical source: [78]

Documentation: (Full) Grammar ([79])


1.22 Kunama (1)

Core area: Sudan, Eritrea

Canonical source: [80]

Classification comment: No conclusive, methodologically sound basis for assigning Kunama to an alleged full or partial Nilo-Saharan has been presented ([81], [82]).

Documentation: (Full) Grammar ([80])


1.23 Laal (1)

Core area: Chad

Canonical source: [83]

Documentation: (Full) Grammar ([83])


1.24 Maban (9)

Core area: W Sudan

Canonical source: [84]

Classification comment: No conclusive, methodologically sound basis for assigning Maban to an alleged full or partial Nilo-Saharan has been presented ([85], [86]).
1.25. MANDE (71)

Core area: W Africa

Canonical source: [88], [89], [90]

Classification comment: Mande has not yet been shown to contain systematic sound correspondences, noun class systems, or verbal extensions with Niger-Congo in some form.

Documentation: (Full) Grammar ([91])


1.26 Mao (4)

Core area: Ethiopia

Canonical source: [92], [93], [94]

Classification comment: The Mao languages (Hozo, Seze, and Bambeshi) have not yet been shown to be Omotic in morphology and/or core vocabulary, though there are some resemblances with North Omotic.

Documentation: Less than sketch but sufficient data to decide OV/VO-feature. ([95])


1.27 Meroitic [extinct] (1)

Core area: NW Sudan, S Egypt

Canonical source: [96], [97]

Documentation: Grammar sketch ([98]; in particular, for setting the OV/VO-feature: [99: 500])


1.28 Narrow Talodi (8)

Core area: Sudan

Canonical source: [100]

Classification comment: Narrow Talodi (excluding Tegem) is held together by pronouns, lexical cognates, and form-meaning resemblances in the noun class systems. Tegem systematically stands out lexically, pronominally, in the noun class system, and alleged sound correspondences are often irregular. The alleged noun class form/meaning correspondences to other Kordofanian groups, namely Rashad and Heiban, in [101] and [102] are not convincing as conclusively genetic; as to form, there is much eclectic selection as well as irregular correspondences, and as to meaning, the methodology of finding non-cognate shared class items is highly dubious. Chance resemblance has not been excluded. The lexical evidence for a Kordofanian unity is weak in the extreme ([103]). It follows that the arguments presented for a Niger-Congo affiliation are not compelling either.

Documentation: Grammar sketch ([104]; in particular, for setting the OV/VO-feature: [105: 38–39], [106: 284–287])


1.29 Omotic (24)

Core area: Ethiopia

Canonical source: [107]

Classification comment: Omotic does not show diagnostic Afro-Asiatic features ([108], [107], [109], [110]), and is therefore excluded until such features can be shown. See also [111] for the coherence of North and South Omotic. The Mao languages (Hozo, Seze and Bambeshi) have not yet been shown to be Omotic in morphology and/or core vocabulary, though there are some resemblances with North Omotic.

Documentation: (Full) Grammar ([112])


1.30  Ongota (1)

Core area: Ethiopia

Canonical source: [113], [114], [115]

Documentation: Grammar sketch ([114: 101])


1.31  Rashad (3)

Core area: Sudan

Canonical source: [116]

Classification comment: Rashad is held together lexically and pronominally ([117: 46–52]). The noun class system cannot be shown conclusively to be original as there are no secure traces of a loss. The alleged noun class form/meaning correspondences to other Kordofanian groups, namely Talodi (with Tegem) and Heiban, in [116] and [118] are not convincing as conclusively genetic; as to form, there is much eclectic selection as well as irregular correspondences, and as to meaning, the methodology of finding non-cognate shared class items is highly dubious. Chance resemblance has not been excluded. The lexical evidence for a Kordofanian unity is weak in the extreme ([119]). It follows that the arguments presented for a Niger-Congo affiliation are not compelling either.

Documentation: Grammar sketch ([120: 59], [121: 174], [117: 50])
1.32 Saharan (9)

Core area: Chad

Canonical source: [122], [123]

Classification comment: No conclusive, methodologically sound basis for assigning Saharan to an alleged full or partial Nilo-Saharan has been presented ([124], [125]).

Documentation: (Full) Grammar ([126])

1.33 Sandawe (1)

Core area: Tanzania

Canonical source: [127]

Documentation: Grammar sketch ([128])


1.34 Shabo (1)

Core area: Ethiopia

Canonical source: [129]

Documentation: Grammar sketch ([129: 384])

1.35 Songhay (8)

Core area: Arid W Africa

Canonical source: [130]

Classification comment: The Nilo-Saharan affiliation à la Ehret or Bender is thoroughly dismantled by Nicolaï, and the Berber-like lexical stratum remains inconclusive.

Documentation: (Full) Grammar ([131])


1.36 Tegem (1)

Core area: Sudan

Canonical source: [132]

Classification comment: Tegem systematically stands out from Narrow Talodi lexically, pronominally, in the noun class system, and alleged sound correspondences are often irregular. The alleged noun class form/meaning correspondences to other Kordofanian groups, namely Rashad and Heiban, in [132] and [133] are not convincing as conclusively genetic; as to form, there is much eclectic selection as well as irregular correspondences, and as to meaning, the methodology of finding non-cognate shared class items is highly dubious. Chance resemblance has not been excluded. The lexical evidence for a Kordofanian unity is weak in the extreme ([134]). It follows that the arguments presented for a Niger-Congo affiliation are not compelling either.

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([135: 45], [136: 284–287])

Tuu (5)  

Core area: S Africa

Canonical source: Güldemann

Documentation: Grammar sketch


Chapter 2

Australia (30)

2.1 Anindilyakwa (1)

Core area: Australia

Canonical source: [139], [140]

Documentation: (Full) Grammar ([140])


2.2 Anson Bay (2)

Core area: Australia

Canonical source: [141: 13], [142]

Documentation: Grammar sketch ([143: 203–222], [144])

2.3 Bunaban (2)

Core area: Australia

Canonical source: [145]

Documentation: (Full) Grammar ([146])


2.4 Eastern Daly (2)

Core area: Australia

Canonical source: [147: 13], [148], [149]

Documentation: (Full) Grammar ([150])

2.5  GAAGUDJU (1)

Core area: Australia

Canonical source: [151], [152]

Classification comment: Evidence for an Arnhem family or a Gunwinyguan affiliation is not sufficient.

Documentation: (Full) Grammar ([153])


2.6 Garrwan (1)

Core area: Australia

Canonical source: [154], [155]

Documentation: (Full) Grammar ([156])


2.7 Giimbiyu [extinct] (3)

Core area: Australia

Canonical source: [157]

Classification comment: Mangeri and Uningangk are dialects of the same language ([158: 372]).

Documentation: Grammar sketch ([159: 96–97])

2.8. **Gunwinyguan (9)**

Core area: Australia

 Canonical source: [160], [161]

Classification comment: Including Ngandi-Nunggubuyu, Rembargic, Gunwinygic, Ngalkbun, and Western Gunwinyguan (Jawoyn and Warray) but excluding Maningrida, Kunkarakany, Gaagudju, Mangarrayi, Wagiman, and Yangmanic. The evidence for an Arnhem family is not sufficient.

Documentation: (Full) Grammar ([162])


2.9 **Iwaidjan Proper (5)**

Core area: Australia
Classification comment: The evidence for including Marrku-Wurrugu in Iwaidjan is not sufficient as the similarities may well be due to borrowing in the case of Marrku ([165]) and the very poorly attested Wurrugu was included because of its lexical correspondences with Marrku rather than the rest of Iwaidjan ([166]).

Documentation: (Full) Grammar ([167])


2.10 Jarrakan (3)

Core area: Australia

Canonical source: [168], [169]

Documentation: (Full) Grammar ([170])

2.11 Kungarakany [extinct] (1)

Core area: Australia

Canonical source: \[171\], \[172\]

Classification comment: Evidence for an Arnhem family or a Gunwinyguan affiliation is not sufficient.

Documentation: Grammar sketch (\[173\])


2.12 Larrakiyan (1)

Core area: Australia

Canonical source: \[174\], \[175\]

Documentation: Grammar sketch (\[176\])


2.13 Limilngan (2)

Core area: Australia

Canonical source: [177]

Documentation: (Full) Grammar ([177])


2.14 Maran (4)

Core area: Australia

Canonical source: [178], [179]

Classification comment: Verb morphology evidence for a Gunwinyguan affiliation ([180]) or an Arnhem family ([181]) is still inconclusive.

Documentation: (Full) Grammar ([182])


2.15 Maningrida (4)

Core area: Australia

Canonical source: [183], [184]

Classification comment: Membership is Burarra, Gurr-Goni, Na-kara, and Ndjébbana. Evidence for inclusion in an Arnhem family is still inconclusive.

Documentation: (Full) Grammar ([185])


2.16 Marrku-Wurrugu [extinct] (2)

Core area: Australia

Canonical source: [186], [187]

Classification comment: The evidence for including Marrku in Iwaidjan is not sufficient as the similarities may well be due to borrowing ([186]). The very poorly attested Wurrugu has lexical correspondences with Marrku rather than the rest of Iwaidjan ([187]).

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([186])


2.17 Minkin-Tangkic (4)

Core area: Australia

Canonical source: [188], [189]

Documentation: (Full) Grammar ([190])


2.18  Mirndi (5)

Core area: Australia

Canonical source: [191], [192]

Documentation: (Full) Grammar ([193])


2.19  Northern Daly (2)

Core area: Australia

Canonical source: [194: 13]

Documentation: Grammar sketch ([195]; in particular, for setting the OV/VO-feature: [196: 39])


2.20 Nyulnyulan (9)

Core area: Australia

Canonical source: [197], [198]

Documentation: (Full) Grammar ([199])


2.21 Pama-Nyungan (175)

Core area: Australia

Canonical source: [200], [201]

Classification comment: Including the Western Torres Strait language.

Documentation: (Full) Grammar ([202])


2.22  SOUTHERN DALY (2)


2.22  Southern Daly (2)

Core area: Australia

Canonical source: [203]

Documentation: (Full) Grammar ([204])


2.23  Oyster Bay, Big River, Little Swanport (Tasmanian) [extinct] (3)

Core area: Tasmania

Canonical source: [205]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([205: 419])


2.24  Piper River, Cape Portland, Ben Lomond (Tasmanian) [extinct] (3)

Core area: Tasmania

Canonical source: [206]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([206: 419])
2.25 Tiwi (1)

Core area: Australia

Canonical source: [207]

Documentation: (Full) Grammar ([207])


2.26 Umbugarla-Ngurmbur (2)

Core area: Australia

Canonical source: [208]

Classification comment: Dixon’s suggestion ([209]) received a shattering review by Evans ([208]). For the internal coherence, it seems that Ngurmbur is so close to Umbugarla that they could be considered one language ([210]), though this is not clear.

Documentation: Grammar sketch ([211: 8])


2.27 Wagiman (1)

Core area: Australia

Canonical source: [212]

Classification comment: So far there has been no published case for a relation with Yangmanic beyond typological similarities and very low cognate percentages ([213: 5–6]), and there is no obvious relation between the two ([214]).

Documentation: (Full) Grammar ([214])


2.28 Western Daly (10)

Core area: Australia

Canonical source: [215: 13, [216]

Documentation: (Full) Grammar ([217])


2.29 Worrorran (7)

Core area: Australia

Canonical source: [218], [219]

Documentation: (Full) Grammar ([220])


2.30 Yangmanic (3)

Core area: Australia

Canonical source: [221]

Classification comment: So far there has been no published case for a relation with Wagiman beyond typological similarities and very low cognate percentages ([222: 5–6]), and there is no obvious relation between the two ([223]).

Documentation: (Full) Grammar ([224])

2.30. YANGMANIC (3)


Chapter 3

Eurasia (33)

3.1 Abkhaz-Adyge (5)

Core area: NW Caucasus

Canonical source: [225]

Documentation: (Full) Grammar ([226])


3.2 Ainu (1)

Core area: N Japan

Canonical source: [227]

Documentation: (Full) Grammar ([227])


3.3 Austroasiatic (168)

Core area: India, SE Asia

Canonical source: [228], [229]

Documentation: (Full) Grammar ([230])

43
3.4 Basque (3)

Core area: SW France, NE Spain

Canonical source: [231]

Documentation: (Full) Grammar ([232])


3.5 Burushaski (1)

Core area: N Pakistan

Canonical source: [233]

Documentation: (Full) Grammar ([233])


3.6 Chukotko-Kamchatkan (5)

Core area: Siberia

Canonical source: [234], [235]

Classification comment: Allegations of areal rather than genetic relationship have yet to reply to [234].
3.7 Dravidian (73)

Core area: S India

Canonical source: [237]

Documentation: (Full) Grammar ([238])


3.8 Elamite [extinct] (1)

Core area: W Iran

Canonical source: [239]

Documentation: (Full) Grammar ([239])

3.9 Etruscan [extinct] (1)

Core area: Italy

Canonical source: [240], [241]

Documentation: Grammar sketch ([240: 961])


3.10 Great Andamanese (10)

Core area: Andamans

Canonical source: [242]

Documentation: Grammar sketch (see reference to older references in [243] which make up a sketch; in particular, for setting the OV/VO-feature: [242].)


3.11 Hattic [extinct] (1)

Core area: Anatolia

Canonical source: [244]

Documentation: Grammar sketch ([245], [246])


3.12  **Hurro-Urartian [extinct] (2)**

Core area: S Caucasus

Canonical source: [247], [248]

Documentation: (Full) Grammar ([249])


3.13  **Iberian [extinct] (1)**

Core area: E Spain

Canonical source: [250]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([251])


3.14 Indo-European (449)

Core area: Europe-, ndia

Canonical source: [252]

Documentation: (Full) Grammar ([253])


3.15 Japanese (12)

Core area: Japan

Canonical source: [254]

Classification comment: See [255] for the latest critique of Altaic.

Documentation: (Full) Grammar ([256])


3.16 Jarawa-Onge (2)

Core area: Andamans

Canonical source: [257]

Classification comment: Tantalizing, but insufficient, parallels with Austronesian can be found in [258].
3.17 Kartvelian (5)

Core area: S Caucasus

Canonical source: [259]

Documentation: (Full) Grammar ([260])

*Lingua* 115. 5–89.


3.18 Korean (1)

Core area: Korea

Canonical source: [261]

Classification comment: See [262] for the latest critique of Altaic.

Documentation: (Full) Grammar ([261])


3.19 Kusunda (1)

Core area: Nepal

Canonical source: [263]

Documentation: (Full) Grammar ([263])


3.20 Miao-Yao (35)

Core area: S China

Canonical source: [264]

Documentation: (Full) Grammar ([265])


3.21 Mongolian (14)

Core area: Mongolia

Canonical source: [266]

Documentation: (Full) Grammar ([267])


3.22  Nakh-Dagestanian (29)

Core area: NE Caucasus

Canonical source: [268]

Documentation: (Full) Grammar ([269])


3.23  Nihali (1)

Core area: India

Canonical source: [270: 242–253]

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([271])


3.24  Nivkh (1)

Core area: Siberia

Canonical source: [272]

Documentation: (Full) Grammar ([272])

3.25 Shom Pen (1)

Core area: Nicobar

Canonical source: [273]

Classification comment: The language described in [273] has a vocabulary that is clearly not cognate with Austroasiatic, though there are question marks for some of the data.

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([273])


3.26 Sino-Tibetan (402)

Core area: China, Nepal

Canonical source: [274]

Documentation: (Full) Grammar ([275])


3.27 Sumerian [extinct] (1)

Core area: Mesopotamia

Canonical source: [276]

Documentation: (Full) Grammar ([276])

3.28 Tai-Kadai (76)

Core area: SE Asia

Canonical source: [277], [278]

Documentation: (Full) Grammar ([279])


3.29 Tungusic (12)

Core area: Siberia

Canonical source: [280]

Classification comment: See [280] for the latest critique of Altaic.

Documentation: (Full) Grammar ([281])


3.30 Turkic (40)

Core area: Central Asia

Canonical source: [282]

Documentation: (Full) Grammar ([283])
3.31 Uralic (39)

Core area: E Europe
Canonical source: [284]
Documentation: (Full) Grammar ([285])


3.32 Yeniseian (2)

Core area: Siberia
Canonical source: [286]
Documentation: (Full) Grammar ([287])


3.33 Yukaghir (2)

Core area: Siberia
Canonical source: [288]
Documentation: (Full) Grammar ([289])


Chapter 4

North America (72)

4.1 Adai [extinct] (1)

Core area: USA

 Canonical source: [290]

 Documentation: Less than sketch and insufficient to decide OV/VO-feature ([291: 326])


4.2 Algic (44)

Core area: USA

 Canonical source: [292]

 Documentation: (Full) Grammar ([293])


4.3 Alsea [extinct] (1)

Core area: USA

Canonical source: [294]

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([295])


4.4 Atakapan [extinct] (1)

Core area: USA

Canonical source: [296]

Documentation: Grammar sketch ([297])


4.5 Beothuk [extinct] (1)

Core area: USA

Canonical source: [298]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([299])


4.6 Caddoan (5)

Core area: USA

Canonical source: [300]

Documentation: (Full) Grammar ([301])


4.7 Cayuse [extinct] (1)

Core area: USA

Canonical source: [302]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([302])


4.8 Chimakuan [extinct] (2)

Core area: USA

Canonical source: [304]

Classification comment: Including Quileute

Documentation: (Full) Grammar ([305])


4.9 Chimariko [extinct] (1)

Core area: USA

Canonical source: [306]

Documentation: (Full) Grammar ([306])


4.10 Chinook (2)

Core area: USA

Canonical source: [307]

Documentation: (Full) Grammar ([308])


4.11 Chitimacha [extinct] (1)

Core area: USA

Canonical source: [309]

Documentation: Grammar sketch ([310: 332], [311])


4.12  Chumashan [extinct] (7)

Core area: USA

Canonical source: [312], [313]

Documentation: (Full) Grammar ([314])


4.13  Coahuilteco [extinct] (1)

Core area: USA

Canonical source: [315]

Documentation: Grammar sketch ([316]; in particular, for setting the OV/VO-feature: [317])


4.14 Cochimi-Yuman (9)

Core area: USA

Canonical source: [318]

Documentation: (Full) Grammar ([319])


4.15 Comecrudan [extinct] (1)

Core area: USA

Canonical source: [320]

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([321])


4.16 Coosan (1)

Core area: USA

Canonical source: [322]

Documentation: (Full) Grammar ([323])


4.17  Cotoname [extinct] (1)

Core area: USA

Canonical source: [324]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([325: 395])


4.18  Cuitlatec [extinct] (1)

Core area: Mexico

Canonical source: [326]

Documentation: Grammar sketch ([327])


4.19  Eskimo-Aleut (11)

Core area: Arctic N America

Canonical source: [328], [329]

Documentation: (Full) Grammar ([330])


CHAPTER 4. NORTH AMERICA (72)


4.20 Esselen [extinct] (1)

Core area: USA

Canonical source: /331/

Documentation: Grammar sketch (/332: 218–219/)


4.21 Eyak-Athapaskan-Tlingit (45)

Core area: NW North American Coast

Canonical source: /333/

Classification comment: Interesting well-analysed parallels between Haida, Eyak-Athapaskan, and Tlingit have surfaced recently (/334/) but is not enough to conclude a genetic relationship.

Documentation: (Full) Grammar (/335/)


4.22 Guaicurian [extinct] (1)

Core area: Mexico

Canonical source: [336]

Documentation: Grammar sketch ([336: 172])


4.23 Haida (2)

Core area: NW North American Coast

Canonical source: [337]

Classification comment: Interesting well-analysed parallels between Haida, Eyak-Athapaskan, and Tlingit have surfaced recently ([338]) but is not enough to conclude a genetic relationship.

Documentation: (Full) Grammar ([339])


4.24 Huavean (4)

Core area: Mexico

Canonical source: [340]

Documentation: Grammar sketch ([341]; in particular, for setting the OV/VO-feature: [342])

4.25 **Iroquoian (11)**

Core area: USA

Canonical source: [343]

Documentation: (Full) Grammar ([344])


4.26 **Jicaquean (1)**

Core area: Mexico

Canonical source: [345]

Documentation: Grammar sketch ([346: 49])


4.27 Kalapuyan (1)

Core area: USA

Canonical source: [347]

Documentation: Grammar sketch ([348]; in particular, for setting the OV/VO-feature: [347: 432].)


4.28 Karankawa [extinct] (1)

Core area: USA

Canonical source: [349]

Documentation source: Grammar sketch ([350: 6])


4.29 Karuk (1)

Core area: USA

Canonical source: [351]

Documentation: (Full) Grammar ([352])


4.30 Keresan (2)

Core area: USA

Canonical source: [353]

Documentation: (Full) Grammar ([354])


4.31 Kiowa-Tanoan (6)

Core area: USA

Canonical source: [355]

Documentation: (Full) Grammar ([356])


4.32 Klamath-Modoc (1)

Core area: USA

Canonical source: [357]

Classification comment: There are interesting parallels to Wintuan and the rest of Plateau Penutian (= Sahaptian and Molala) ([358]).

Documentation: (Full) Grammar ([359])

4.33 KUTENAI (1)

Core area: USA

Canonical source: [360]

Documentation: (Full) Grammar ([361])


4.33 Kutenai (1)

Core area: USA

Canonical source: [360]

Documentation: (Full) Grammar ([361])


4.34 Lencan [extinct] (1)

Core area: Honduras, Nicaragua

Canonical source: [362]

Classification comment: Interesting links have been noted with Misumalpan ([363]).

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([364: 319], [365: 19])


4.35 Maiduan (4)

Core area: USA

Canonical source: [366]

Documentation: (Full) Grammar ([367])


4.36 Maratino [extinct] (1)

Core area: Mexico

Canonical source: [368]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([368])


4.37 Mayan (69)

Core area: Mexico

Canonical source: [369]

Documentation: (Full) Grammar ([370])
4.38 Misumalpan (4)

Core area: Honduras, Nicaragua

Canonical source: [371]

Classification comment: Interesting links have been noted with Lencan ([372]).

Documentation: (Full) Grammar ([373])


4.39 Miwok-Costanoan (10)

Core area: USA

Canonical source: [374]

Documentation: (Full) Grammar ([375])


4.40 Mixe-Zoque (17)

Core area: Mexico

Canonical source: [376], [377]

Documentation: (Full) Grammar ([378])


4.41 Molala [extinct] (1)

Core area: USA

Canonical source: [379]

Classification comment: There are interesting parallels to Wintuan and the rest of Plateau Penutian (= Sahaptian and Klamath-Modoc) ([380]).

Documentation: (Full) Grammar ([379])


4.42 Muskogean (6)

Core area: USA

Canonical source: [381], [382]

Documentation: (Full) Grammar ([383])
4.43  Natchez [extinct] (1)

Core area: USA

Canonical source: 

Documentation: Grammar sketch


4.44  Otomanguean (179)

Core area: Mexico

Canonical source: 

Classification comment: Including Subtiaba-Tlapanec

Documentation: (Full) Grammar

4.45 Palaihnihan (2)

Core area: USA

Canonical source: 4/388/

Documentation: Grammar sketch (4/389, 4/390)


4.46 Pomoan (7)

Core area: USA

Canonical source: 4/391/

Documentation: (Full) Grammar (4/392)


4.47 Sahaptian (5)

Core area: USA

Canonical source: [393]


Classification comment: There are interesting parallels to Wintuan and the rest of Plateau Penutian (= Molala and Klamath-Modoc) ([394]).

Documentation: (Full) Grammar ([395])


4.48 Salinan [extinct] (1)

Core area: USA

Canonical source: [396]

Documentation: (Full) Grammar ([397])


4.49 Salishan (27)

Core area: USA

Canonical source: [398]

Documentation: (Full) Grammar ([399])
4.50  Seri (1)

Core area: Mexico

Canonical source: [400]

Documentation: (Full) Grammar ([401])


4.51  Shasta [extinct] (1)

Core area: USA

Canonical source: [402]

Documentation: (Full) Grammar ([403])


4.52  Siouan (17)

Core area: USA

Canonical source: [404]

Documentation: (Full) Grammar ([405])

[405]
4.53  **Siuslaw [extinct] (1)**

Core area: USA

Canonical source: [406]

Documentation: (Full) Grammar ([407])


4.54  **Takelma (1)**

Core area: USA

Canonical source: [408]

Documentation: (Full) Grammar ([409])


4.55 Tarascan (2)

Core area: Mexico

Canonical source: [410]

Documentation: (Full) Grammar ([411])


4.56 Tequistlatecan (2)

Core area: Mexico

Canonical source: [412]

Documentation: (Full) Grammar ([413])


4.57 Timucua [extinct] (1)

Core area: USA

Canonical source: [414]

Documentation: (Full) Grammar ([415])


4.58  **Tonkawa [extinct] (1)**

Core area: USA

Canonical source: [416]

Documentation: (Full) Grammar ([417])


4.59  **Totonacan (11)**

Core area: Mexico

Canonical source: [418], [419]

Documentation: (Full) Grammar ([420])


4.60  **Tsimshian (3)**

Core area: Canada

Canonical source: [421]

Documentation: (Full) Grammar ([422])
4.61 Tunica [extinct] (1)

Core area: USA

Canonical source: [423]

Documentation: (Full) Grammar ([424])


4.62 Uto-Aztecan (61)

Core area: USA

Canonical source: [425]

Documentation: (Full) Grammar ([426])


4.63 Wakashan (5)

Core area: Canada

Canonical source: [427]

Documentation: (Full) Grammar ([428])
4.64 WAPP (1)


4.64 Wappo (1)

Core area: USA

Canonical source: [429]

Documentation: (Full) Grammar ([430])


4.65 Washo (1)

Core area: USA

Canonical source: [431]

Documentation: (Full) Grammar ([432])


CHAPTER 4. NORTH AMERICA (72)

4.66 Wintuan (1)

Core area: USA

Canonical source: [433]

Classification comment: There are interesting parallels to Plateau Penutian (= Klamath-Modoc, Sahaptian, and Molala) ([434]).

Documentation: (Full) Grammar ([435])


4.67 Xincan (1)

Core area: Guatemala

Canonical source: [436]

Documentation: Grammar sketch ([437]; in particular, for setting the OV/VO-feature: [438: 547].)


4.68  Yana (1)

Core area: USA

Canonical source: [439]

Documentation: Grammar sketch ([440: 9])


4.69  Yokutsan (1)

Core area: USA

Canonical source: [441]

Documentation: (Full) Grammar ([442])


4.70  Yuchi (1)

Core area: USA

Canonical source: [443]

Documentation: (Full) Grammar ([444])


4.71 Yuki (1)

Core area: USA

Canonical source: [445]

Documentation: Grammar sketch ([446: 372], [447: 76–77])


4.72 Zuni (1)

Core area: USA

Canonical source: [448]

Documentation: (Full) Grammar ([449])


Chapter 5

Papua (111)

5.1 Abinomn (1)

Core area: W Irian

Canonical source: [450]

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([451])


5.2 Abun (1)

Core area: NW Irian

Canonical source: [452], [453], [454]

Documentation: (Full) Grammar ([455])

CHAPTER 5. PAPUA (111)


5.3 Amto-Musan (2)

Core area: NW Papua

Canonical source: 


5.4 Anêm (1)

Core area: New Britain

Canonical source: 

Classification comment: Pronoun resemblances (Wurm 1996) are not enough for concluding a Yele-West New Britain Family.

Documentation: Grammar sketch (Wurm 1996: 30–31); in particular, for setting the OV/VO-feature: 

5.5. **ANGAN (13)**

Core area: Gulf, Morobe

Canonical source: [463]

Classification comment: As has been clear at least since [464] there are insufficient lexical links to posit a relationship with Trans New Guinea.

Documentation: (Full) Grammar ([465])


5.6 Arafundi (1)

Core area: East Sepik

Canonical source: [466]

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([467])


5.7 Ata (1)

Core area: New Britain

Canonical source: [468]

Classification comment: Pronoun resemblances [469] are not enough for concluding a Yele-West New Britain Family.

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([468: 84], [470])


5.8  Austronesian (1275)

Core area: Insular SE Asia

Canonical source: [471], [472]

Classification comment: Tantalizing, but insufficient, parallels with Jarawa-Onge can be found in [473]. Makuva [lva] of East Timor is also Austronesian ([474]). Utupua, Vanikoro, and Reef-Santa Cruz have been shown to be bona fide Oceanic ([475]). Also, the best take on the poorly attested Kazukuru language(s) is that they were New Georgia Austronesian ([476]).

Documentation: (Full) Grammar ([477])


5.9 Awin-Pa (3)

Core area: Western Province

Canonical source: [478]

Documentation: (Full) Grammar ([479])


5.10 Baibai (2)

Core area: NE Irian

Canonical source: [480]

Classification comment: Laycock never presented real evidence for a Kwomtari-Baibai-Pyu family ([481]). The membership is Baibai [bbf] and Fas [fqs] and not Biaka/Nai [bio] as many sources have erroneously repeated.

Documentation: Grammar sketch ([482: 11])


5.11 Baining (6)

Core area: E New Britain

Canonical source: [483: 250], [484: 311]

Documentation: Grammar sketch ([485]; in particular, for setting the OV/VO-feature: [486])


5.12 Bayono-Awbono (2)

Core area: Irian

Canonical source: [487]

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([487])


5.13 Biksi (2)

Core area: N Mid Papua

Canonical source: [488], [489]
Classification comment: Evidence for a Sepik affiliation is too scant, though data is very scant too.

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([490: 740–741])


5.14 Bilua (1)

Core area: Solomon Islands

Canonical source: [491]

Documentation: (Full) Grammar ([492])


5.15  Binanderean (11)

Core area: SE Papua

Canonical source: [493]

Documentation: (Full) Grammar ([494])


5.16  Border (15)

Core area: W Papua

Canonical source: [495]

Classification comment: Waris, Taikat, Bewani

Documentation: (Full) Grammar ([496])


5.17  Bosavi (10)

Core area: Western Province

Canonical source: [497], [498]

Classification comment: We also include Doso and Turumsa as Bosavi languages as per a SIL PNG document on endangered languages ([499]).

Documentation: Grammar sketch ([498])
5.18  Bulaka River (2)

Core area: S Irian

Canonical source: [500]

Classification comment: Wurm’s arguments ([500: 324]) for a Trans-Fly assignment were based on low (about 9%) lexicostatistical figures and typological characteristics.

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([501])


5.19  Burmeso (1)

Core area: Mid-Mamberamo

Canonical source: [502]

Documentation: Grammar sketch ([502: 98])
5.20  Busa (Odiai) (1)

Core area: Sandaun

Canonical source: [503]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([504])


5.21  Cenderawasih Bay (11)

Core area: NW Irian

Canonical source: [505]

Documentation: (Full) Grammar ([506])


5.22  Dem (1)

Core area: Irian

Canonical source: [507]

Classification comment: The cognition judgments involving Dem are warped in that a match is judged if at least one segment matches. Needless to say, this given 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.

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([508: 892–895])


5.23  Duna-Bogaya (2)

Core area: Western Province

Canonical source: [509: 395–396]

Classification comment: Arguments for the relatedness for Duna and Bogaya are given in [509: 395–396] but pronouns do not match sufficiently well for an immediate Trans New Guinea affiliation, and apart from this, there are only capricious lexical similarities to other families ([510]).

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([509: 396])

5.24. **DURANMIN (1)**

Core area: E Sepik

Canonical source: [511], [512], [513]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([511], [512])


5.25 **East Bird’s Head (3)**

Core area: E Vogelkop

Canonical source: [514], [515]

Documentation: (Full) Grammar ([516])

CHAPTER 5. PAPUA (111)


5.26 East Kutubu (2)

Core area: S Highlands

Canonical source: [517]

Classification comment: The link to Fasu is premature because counting framework and kinship terms are precisely the kind of argument that is not conclusive of a genetic relationship ([517: 311]).

Documentation: Grammar sketch ([518], [519])


5.27 East Timor (3)

Core area: Timor-Alor-Pantar

Canonical source: [520], [521]

Classification comment: The group is clearly internally coherent. I have not been able to replicate the lexicostatistic argument for a relation between all Timor-Alor-Pantar languages, i.e., with West Timor-Alor-Pantar and Kolana-Tanglapui ([522]), and suggested correspondences do not show much systematicity. Likewise, the Bomberai/Alor comparisons in [520] are flimsy.
5.28. EASTERN TRANS-FLY (4)

Documentation: Grammar sketch ([523], [524])


5.28 Eastern Trans-Fly (4)

Core area: Fly River

Canonical source: [525]

Classification comment: Wurm’s arguments ([525: 327–335]) for a Trans New Guinea affiliation appear to be unreliable lexicostatistics and typological features. Likewise, the lexical and pronominal evidence for a Trans New Guinea affiliation is weak. See [526] for additional lexical data on the internal coherence of the group.

Documentation: Grammar sketch ([527: 546–547], [528])

5.29 Eleman (6)

Core area: Gulf

Canonical source: [529]

Documentation: Grammar sketch (see the extensive materials by Brown ([530], [531]), by far enough to make up a sketch; in particular, for setting the OV/VO-feature: [532: 67–68], [531: 311–316].)


5.30 Elseng (1)

Core area: NE Irian

Canonical source: [533]

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([534: 423])


5.31 Fasu (1)

Core area: S Highlands

Canonical source: [535]

Classification comment: The link to East Kutubuan is premature because counting system and kinship terms are precisely the kind of argument that is not conclusive of a genetic relationship ([535: 311]).

Documentation: (Full) Grammar ([536])


5.32 Goilalan (5)

Core area: SE Papua

Canonical source: [537]

Documentation: (Full) Grammar ([538])


5.33 Guriaso (1)

Core area: NE Irian

Canonical source: [539]

Classification comment: Laycock never presented real evidence for a Kwomtari-Baiabi-Pyu family ([540]). It is clear from the data collected so far ([539]) that Guriaso [grx] shares no more lexical cognates with Kwomtari and Biaka than expected at random, and that’s not even when borrowing is discounted (Kwomtari neighbours Guriaso). Further correspondences presented are merely typological or random enough to make Japanese a Kwomtari language ([539: 29]).

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([539: 15])


5.34  Hatam (1)

Core area: Vogelkop

Canonical source: [541]

Documentation: (Full) Grammar ([542])


5.35  Inanwatan (2)

Core area: NW Irian

Canonical source: [543], [544]

Documentation: (Full) Grammar ([545])


5.36  Kaki Ae (1)

Core area: Gulf

Canonical source: [546]

Documentation: Grammar sketch ([546]; in particular, for setting the OV/VO-feature: [547])
5.37  **Karkar (Yuri) (1)**

Core area: Sandaun

Canonical source: [548]

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([549: 92–95], [550])


5.38  **Kaure-Kapori (4)**

Core area: NE Irian

Canonical source: [551]

Documentation: Grammar sketch ([552: 101–113])

5.39  **KAYAGAR (3)**

Core area: Irian

Canonical source: [553: 366–369]

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([553: 369])


5.40  **Kehu (1)**

Core area: N Irian

Canonical source: Mark Donohue, personal communication, e-mail 17 May 2007

Classification comment: There are some parallels drawn up in [554].

Documentation: Less than sketch and insufficient to decide OV/VO-feature (Mark Donohue, personal communication, e-mail 17 May 2007)


5.41  **Kembra (1)**

Core area: N Irian

Canonical source: Mark Donohue, personal communication, e-mail 17 May 2007

Documentation: Less than sketch and insufficient to decide OV/VO-feature (Mark Donohue, personal communication, e-mail 17 May 2007)
5.42  Kiwaian (6)

Core area: Gulf

Canonical source: [555]

Documentation: (Full) Grammar ([556])


5.43  Koiarian (7)

Core area: SE Papua

Canonical source: [557]

Documentation: (Full) Grammar ([557])


5.44  Kol (1)

Core area: New Britain

Canonical source: [558], [559]

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([558])


5.45 Kolana-Tanglapui (3)

Core area: Timor-Alor-Pantar

Canonical source: [560], [561]

Classification comment: I have not been able to replicate the lexicostatistic argument for a relation between all Timor-Alor-Pantar languages, i.e., with West Timor-Alor-Pantar, and East Timor ([560]), and suggested correspondences do not show much systematicity. Lexical evidence uniting Kolana and Tanglapui is much stronger, especially the numerals. Lexical or pronominal evidence for a Trans New Guinea affiliation has not been put forward independently for Kolana-Tanglapui.

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([560: 61], [562: 108])


5.46 Kolopom (3)

Core area: Frederik Hendrik Island

Canonical source: [563]

Classification comment: I am unable to find arguments for Trans New Guinea affiliation in [563] and there is no obvious relation.

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([564])

Guinea linguistic scene (Pacific Linguistics C-38), 345–460. Canberra: Research School of Pacific and Asian Studies, Australian National University.


### 5.47 Konda-Yahadian (2)

Core area: NW Irian

Canonical source: [565], [566: 437–446]

Classification comment: Evidence for inclusion in Trans New Guinea is weak ([566: 437–446]), especially lexically. The same can be said for a relation with South Bird’s Head, Konda-Yahadian, and any West Papuan affiliation ([565]).

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([565: 89])


### 5.48 Kuot (1)

Core area: New Britain

Canonical source: [567]

Documentation: (Full) Grammar ([567])

5.49 **Kwerba (8)**

Core area: N Irian

Canonical source: [568]

Classification comment: Including Isirawa and Samarokena.

Documentation: Grammar sketch ([569: 4])


5.50 **Kwomtari (2)**

Core area: NE Irian

Canonical source: [570]

Classification comment: Laycock never presented real evidence for a Kwomtari-Baibai-Pyu family ([571]). The membership is Kwomtari [kwo], Biaka/Nai [bio], and not Fas [fps] as many sources have erroneously repeated. It is clear from the data collected so far ([570]) that Guriaso [grx] shares no more lexical cognates with Kwomtari and Biaka than expected at random, and that’s not even when borrowing is discounted (Kwomtari neighbours Guriaso). Further correspondences presented are merely typological or random enough to make Japanese a Kwomtari language ([570: 29]).

Documentation: Grammar sketch ([572], [570])


linguistic scene (Pacific Linguistics C-38), 849–858. Canberra: Research School of Pacific and Asian Studies, Australian National University.


5.51 Lakes Plain (20)

Core area: Irian

Canonical source: [573]

Documentation: Grammar sketch ([574]; in particular, for setting the OV/VO-feature: [575], [576: 19].)


5.52 Lavukaleve (1)

Core area: Solomon Islands

Canonical source: [577]

Documentation: (Full) Grammar ([578])


5.53 Left May (6)

Core area: NW Papua

Canonical source: [579], [580]

Classification comment: From [580] we know that the family is internally coherent (with sound correspondences) and that there are no convincing external relations revealed in the lexicon.

Documentation: Grammar sketch ([579], [581])


5.54 Lepki (1)

Core area: N Irian

Canonical source: [582]

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([583])


CHAPTER 5. PAPUA (111)

5.55 Lower Mamberamo (2)

Core area: N Papua

Canonical source: [584]

Documentation: (Full) Grammar ([584])


5.56 Lower Sepik-Ramu (Extended) (33)

Core area: N Papua

Canonical source: [585]

Classification comment: Lower Sepik, Lower Ramu, Tamolan, Rao, Aian, Grass (including Banaro but excluding Kambot), Kambot.

Documentation: (Full) Grammar ([586])


5.57 Mairasi (3)

Core area: Irian

Canonical source: [587]

Classification comment: Links with Tanahmanah are unconvincing lexically and pronominally ([588: 424–431], [589]).

Documentation: Grammar sketch ([590: 75])

5.58  MARIND (6)

Core area: SW Irian

Canonical source: [591]

Classification comment: Not including Inanwatan, though typological affinities have been noted [592]

Documentation: (Full) Grammar ([593])


5.59 Masep (1)

Core area: N Irian

Canonical source: [594]

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([594: 6])


5.60 Maybrat (2)

Core area: NW Irian

Canonical source: [595], [596], [597]

Documentation: (Full) Grammar ([598])


5.61  Molof (1)

Core area: S Jayapura

Canonical source: [599]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([600])


5.62  Mombum (2)

Core area: Komolom Island

Canonical source: [601: 396–398]

Classification comment: Pronouns do not match sufficiently well for an immediate Trans New Guinea affiliation, and apart from this, there are only capricious lexical similarities to other families ([601: 396–398]). Internally, Koneraw and Mombum (aka Komelom) can be seen to be related from the basic vocabulary correspondences in [602].

Documentation: Grammar sketch ([603: 563])


5.63 Mongol-Langam (3)

Core area: E Sepik

Canonical source: [604], [605]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([605])


5.64 Mor (1)

Core area: Bomberai

Canonical source: [606]

Classification comment: Evidence for inclusion in Trans New Guinea is weak ([606: 431]), both lexically and pronominally.

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([606], [607])


5.65 Moraori (1)

Core area: S Border

Canonical source: [608]

Classification comment: Wurm’s arguments ([608: 327–335]) for a Trans-Fly assignment are not convincing as the only argument appears to be unreliable lexicostatistical calculations.

Documentation: Grammar sketch (see [609: 15–26] or [610] based on Drabbe’s work; in particular, for setting the OV/VO-feature: [610])


5.66 Morehead and Upper Maro Rivers (17)

Core area: S Border

Canonical source: [611]

Classification comment: Wurm’s arguments ([611: 327–335]) appear to be unreliable lexicostatics and typological features.

Documentation: Grammar sketch (see [612: 26–35] on Jei, [612: 36–52] on Kanum) or [613] based on Drabbe’s work; there is also some oft-forgotten linguistic data in [614] and [615]; in particular, for setting the OV/VO-feature: [616], [617: 202–205].)


**5.67 Mpur (1)**

Core area: NW Irian

Canonical source: [618], [619]

Documentation: Grammar sketch ([620])


5.68 Murkim (1)

Core area: N Irian

Canonical source: [621]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([621])


5.69 Nimboran (5)

Core area: NE Irian

Canonical source: [622]

Documentation: (Full) Grammar ([623])


5.70 North Bougainville (4)

Core area: Bougainville

Canonical source: [624], [625]

Documentation: (Full) Grammar ([626])


5.71 North Halmahera (16)

Core area: N Halmahera

Canonical source: [627], [628], [629], [630]

Documentation: (Full) Grammar ([631])


5.72 Oksapmin (1)

Core area: Sandaun

Canonical source: [632]

Documentation: Grammar sketch (There is a dictionary ([633]) and some published aspects of grammar by the same author. I guess this can make up a sketch. Robyn Loughnane (University of Melbourne) is writing a full grammar. In particular, for setting the OV/VO-feature: [634].)
5.73 Pahoturi (2)

Core area: Western Province

Canonical source: [635]

Classification comment: Wurm’s arguments ([635: 327–335]) appear to be unreliable lexicostatics and typological features.

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([636])


5.74 Papi (1)

Core area: E Sepik

Canonical source: [637], [638], [639]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([637])
5.75 Pauwasi (4)

Core area: NE Irian

Canonical source: [640]

Classification comment: It appears that Pauwasi pronouns do not show Trans New Guinea lookalike-ness, and other arguments for a Trans New Guinea affiliation are never mentioned ([641: 418–419]).

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([642])


Guinea linguistic scene (Pacific Linguistics C-38), 345–460. Canberra: Research School of Pacific and Asian Studies, Australian National University.


5.76 Pawaia (1)

Core area: Simbu

Canonical source: [643]

Classification comment: Despite vocabulary cognacy of 5% or so, Pawaia was included in Trans-New-Guinea because of pronoun resemblances to Kuman and on typological similarities. The typological similarities involve function only ([643]), and thus count for nothing. The pronoun resemblances do not generalize to the Chimbu family ([644: 69–71]) and match only an n anyway, so they are better accounted for as accidental similarities than deep relationship.

Documentation: Grammar sketch ([643: 20–22])


5.77 Piawi (2)

Core area: Upper Yuat

Canonical source: [645], [646]

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([647: 88])

CHAPTER 5. PAPUA (111)


5.78 Porome (Kibiri) (1)

Core area: Gulf Province

Canonical source: [648]

Classification comment: The suggestion of a Kiwai affiliation is based on pronouns only ([649]).

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([648])


5.79 Pyu (1)

Core area: October River

Canonical source: [650]

Classification comment: Laycock never presented evidence for a Kwomtari-Baibai-Pyu family ([651]).

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([651])


---

**5.80 Savosavo (1)**

Core area: Solomon Islands

Canonical source: [652]

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([653], [654])


---

**5.81 Senagi (2)**

Core area: NW Papua

Canonical source: [655]

Documentation: (Full) Grammar ([655])

5.82 Sentani (4)

Core area: N Irian

Canonical source: [656], [657]

Classification comment: The relation of Sentani-Nafri-Tabla (SNT) to Demta is best argued in [657: 161–163], see also [658], and can be verified with the subsequent SNT phonological reconstruction in [656] and the longer wordlists in [659].

Documentation: (Full) Grammar ([660])


5.83 Sepik (48)

Core area: N Papua

Canonical source: [661]

Classification comment: Abau, Yellow River, Iwam, Ram (Pouye, Karawa, Awtuw), Wogumusin-Chenapian, Tama, Kwoma-Kwanga (Kwoma, Kwanga, Mende), Ndu, Sepik Hill, and Yerakai. For Yerakai, there is some scant lexical data available that points to a Sepik affiliation ([662: 14–15]), but no actual argument are offered in [663: 738] and Yerakai is not mentioned in Foley’s demonstration of the Sepik family ([661]).
Documentation: (Full) Grammar ([664])


5.84 Sko (7)

Core area: N Papua

Canonical source: [665]

Documentation: (Full) Grammar ([666])


5.85 South Bird’s Head Proper (6)

Core area: NW Irian

Canonical source: [667], [668: 437-446]

Classification comment: Evidence for inclusion in Trans New Guinea is weak ([668: 437–446]), especially lexically. The same can be said for a relation with Inanwatan, Konda-Yahadian, and the older West Papuan affiliation ([667]).

Documentation: Grammar sketch ([667: 89], [669: 4])


5.86 South Bougainville (9)

Core area: Bougainville

Canonical source: [670], [671]

Documentation: (Full) Grammar ([672])


5.87  Suki-Gogodala (4)

Core area: W Fly

Canonical source: [673]

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([674], [675: 70–72])


5.88  Sulka (1)

Core area: New Britain

Canonical source: [676]

Documentation: Grammar sketch ([676])


5.89  Taiap (1)

Core area: E Sepik

Canonical source: [677: 61ff.]

Classification comment: Laycock’s assignment to Sepik-Ramu was for mainly typological reasons ([678: 757]).

Documentation: Grammar sketch ([679])
5.90 Tambora [extinct] (1)

Core area: C Sumbawa

Canonical source: [680]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([680])


5.91 Tanahmerah (1)

Core area: Irian

Canonical source: [681], [682: 424–431]

Classification comment: Links with Mairasi are unconvincing lexically and pronominally ([682: 424–431]).

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([682: 424–431]).

5.92  Taulil-Butam (1)

Core area: E New Britain

Canonical source: [683: 250], [684: 311]

Documentation: (Full) Grammar ([685])


5.93  Teberan (2)

Core area: S Highlands

Canonical source: [686]
Classification comment: The suggested Pawaian relation is based on lexicostatistics and typological features ([687]), while, e.g., the pronouns do not match systematically ([688: 501–504]).

Documentation: (Full) Grammar ([689])


5.94 *Tirio (5)*

Core area: Western Province

Canonical source: [690]

Classification comment: Wurm’s arguments ([690: 327–335]) appear to be unreliable lexicostatics and typological features.

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([690])

5.95  **Tofanma (1)**

Core area: NE Irian

Canonical source: [691]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([692])


5.96  **Tor-Orya (13)**

Core area: N Irian

Canonical source: [693], [694]

Classification comment: Betaf included on the basis of an unpublished wordlist ([695]). The pronouns for Tor are not Trans New Guinea and other arguments have not been offered ([693: 413–414]), nor are there any apparent relations in newer data published after Voorhoeve.

Documentation: Grammar sketch ([696])


5.97 Torricelli (53)

Core area: N Papua

Canonical source: [697]

Documentation: (Full) Grammar ([698])


5.98 Touo (1)

Core area: Solomon Islands

Canonical source: [699]

Documentation: Grammar sketch ([700]; in particular, for setting the OV/VO-feature: [701])


5.99 Trans New Guinea (338)

Core area: PNG

Canonical source: [702]

Classification comment: Includes Finisterre-Huon, Eastern Highlands ([703]), Irian Highlands (Dani and Paniai Lakes subgroups), Madang, Ok, Awyu-Dumut, Asmat-Kamoro ([704]), Mek ([705]), East Strickland ([706]), Inland Gulf ([707: 509–510]) + probable members Engan, Chimbu. See also [708: 97] and [709: 146] for re-subgrouping of a few languages.

Documentation: (Full) Grammar ([710])


5.100 Turama-Kikori (3)

Core area: Gulf

Canonical source: [711]

Documentation: Grammar sketch ([712: 16])


5.101 Uhunduni (1)

Core area: Irian

Canonical source: [713]

Classification comment: The cognition judgments involving Damal are warped in that a match is judged if at least one segment matches. Needless to say, this given 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.

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([714: 410–411])
5.102  **USKU (1)**

Core area: NE Irian

Canonical source: [715]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([716])


5.103  **Waia (1)**

Core area: Western Province

Canonical source: [717]

Classification comment: Pronouns do not match sufficiently well for a Pahoturi affiliation, contra [718].

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([719])

CHAPTER 5. PAPUA (111)


5.104 Walio (4)

Core area: E Sepik

Canonical source: [720], [721], [722]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([720])


5.105 West Bird’s Head (5)

Core area: NW Irian

Canonical source: [723], [724], [725], [726]

Documentation: Grammar sketch ([727]; in particular, for setting the OV/VO-feature: [728], [729: 67])


5.106 West Bomberai (3)

Core area: Bomberai

Canonical source: [730: 432–437]

Classification comment: The inclusion of the poorly known Karas is best argued in [731: 33–36], with system correspondences in pronominals and a few items of basic vocabulary. Evidence for inclusion in Trans New Guinea is weak ([730: 432–437]), both lexically and pronominally, cf. [732: 94–95]. Likewise, the East Timor/Alor comparisons in [733] are flimsy.

Documentation: Grammar sketch ([734]; in particular, for setting the OV/VO-feature: [735])


5.107. WEST TIMOR-ALOR-PANTAR (16)


5.107 West Timor-Alor-Pantar (16)

Core area: Timor-Alor-Pantar

Canonical source: [736], [737]

Classification comment: The group is held together on pronominal and lexical grounds. The lexical and pronominal evidence for a Trans New Guinea affiliation is still weak, especially lexically ([738: 683]), cf. [739: 94–95]. The newest comparison of cognates ([740: 6–11]) cannot muster a strong case (correspondences are few, weak, and not systematic enough). I have not been able to replicate the lexicostatistic argument for a relation between all Timor-Alor-Pantar languages, i.e., with Kolana-Tanglapui and East Timor ([741]), and suggested correspondences do not show much systematicity.

Documentation: (Full) Grammar ([742])


5.108  Yalë (Nagatman) (1)

Core area: Sandaun

Canonical source: [743]

Documentation: (Full) Grammar ([744])


5.109  Yawa (2)

Core area: Serui Island

Canonical source: [745], [746]

Documentation: Grammar sketch ([747])


5.110  Yéli Dnye (1)

Core area: Rossel Island

Canonical source: [748]

Documentation: (Full) Grammar ([749])


5.111  Yuat-Maramba (6)

Core area: E Sepik

Canonical source: [750], [751], [752]

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([753])


Chapter 6

South America (111)

6.1 Aikanã (1)

Core area: Rondônia

Canonical source: [754]

Documentation: Grammar sketch ([755])


6.2 Aimoré (1)

Core area: Amazon

Canonical source: [756], [757]

Documentation: Grammar sketch ([758: 188], [756], [759: 125], [760: 136–140])


6.3 Andaquí [extinct] (1)

Core area: Colombia

Canonical source: [761] Documentation: Less than sketch and insufficient to decide OV/VO-feature ([761: 138–141])


6.4 Andoque (1)

Core area: W Amazon

Canonical source: [762] Documentation: (Full) Grammar ([763])


6.5 Arara do Rio Branco (1)

Core area: Rondônia, Mato Grosso

Canonical source: [764]
Classification comment: There is only a short wordlist for which the majority of entries are not Tupí cognates. Cognates for these words have so far not been uncovered in other families (but the search has been limited since the wordlist is not yet published).

Documentation: Less than sketch and insufficient to decide OV/VO-feature


6.6 Araucanian (2)

Core area: S Andes

Canonical source: [765]

Documentation: (Full) Grammar ([766])


6.7 Arawá (8)

Core area: W Amazon

Canonical source: [767]

Documentation: (Full) Grammar ([768])


6.8 Arawak (62)

Core area: Amazon

Canonical source: [769]
Documentation: (Full) Grammar ([770])


6.9 Atacame (Esmeraldeño) [extinct] (1)

Core area: Ecuador

Canonical source: [771]

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([772]: [773: 86])


6.10 Atacameño (Kunza) [extinct] (1)

Core area: Andes

Canonical source: [773]

Documentation: Grammar sketch ([774]; in particular, for setting the OV/VO-feature: [775: 499], [773: 380].)


6.11  **Awaké (Arutani, Uruak) (1)**

Core area: Orinoco, Amazon

Canonical source: [776]

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([776: 50])


6.12  **Awshiri (Tekiráka) [presumed extinct] (1)**

Core area: Peruvian Amazon

Canonical source: [777]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([778])


6.13  **Aymara (3)**

Core area: Andes

Canonical source: [779]

Documentation: (Full) Grammar ([780])


CHAPTER 6. SOUTH AMERICA (111)

6.14 Barbacoan (6)

Core area: NW Sphere

Canonical source: [781]

Documentation: (Full) Grammar ([782])


6.15 Betoi-Jirara [extinct] (1)

Core area: NW Sphere

Canonical source: [783]

Documentation: Grammar sketch ([784]; in particular, for setting the OV/VO-feature: [785: 218–219])


6.16 Bora-Huitoto (6)

Core area: Peru, Colombia

Canonical source: [786]

Classification comment: The attempt [787] to demonstrate the unity of the family, on the grounds of shared basic lexicon (including pronouns), is current as [788] does not comment on it. Neither is it contradicted by a recent grammar ([789: 1.3.3]). However, a much-desired demonstration of relatedness through shared morphology is lacking.
6.17. **BORORO (3)**

Documentation: (Full) Grammar ([790])


Core area: Amazon

Canonical source: [791]

Documentation: (Full) Grammar ([792])

6.18 Cahuapanan (2)

Core area: Peruvian Amazon

Canonical source: [793]

Documentation: (Full) Grammar ([794])


6.19 Candoshi-Shapra (1)

Core area: Peruvian Amazon

Canonical source: [795]

Documentation: Grammar sketch ([796], [797])


6.20 Canichana [extinct] (1)

Core area: Bolivia

Canonical source: [798]
Documentation: Less than sketch and insufficient to decide OV / VO-feature ([799], [800: 165]. There are the prayers Pater Noster, Ave Maria, and Credo, as well as some phrases ([801]), but sparseness of transitives in them and the poor knowledge of the verbal system prevents a credible basic word order analysis.)


[801] Cardús, José. 1886. Las misiones franciscanas entre los infieles de Bolivia: Descripción del estado de ellas en 1883 y 1884. Barcelona: Librería de la inmaculada concepción.

6.21 Carib (32)

Core area: N Amazon

Canonical source: [802], [803], [804]

Documentation: (Full) Grammar ([805])


6.22 Cayuvava (1)

Core area: Bolivia

Canonical source: [806]

Documentation: (Full) Grammar ([807])


6.23 Chapacura-Wanham (5)

Core area: Rondônia

Canonical source: [808]

Documentation: (Full) Grammar ([809])


6.24 Charrúa [extinct] (3)

Core area: Andes

Canonical source: [810]

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([810])

6.25 Chibchan (21)

Core area: NW Sphere

Canonical source: [811]

Classification comment: Including Cuna, Chimila, Languages of Sierra Nevada de Santa María, Muisca, Tunebo.

Documentation: (Full) Grammar ([812])


6.26 Chiquitano (1)

Core area: Bolivia

Canonical source: [813]

Documentation: (Full) Grammar ([814])


6.27 Chocoan (12)

Core area: NW Sphere

Canonical source: [815]

Documentation: (Full) Grammar ([816])


6.28 Chonan (2)

Core area: Tierra del Fuego

Canonical source: [817], [818: 556–558]

Classification comment: Includes Querandi. The parallels with Günüña Küne are interesting but not conclusive.

Documentation: (Full) Grammar ([819])


6.29 Chono [extinct] (1)

Core area: Tierra del Fuego

Canonical source: [820: 552–558]

Classification comment: There are lexical parallels with Mapuche as well as Qawesqar ([821: 82]) but the core is clearly unrelated.

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([822])


6.30  Cofán (1)

Core area: Ecuador

Canonical source: [823]

Documentation: Grammar sketch ([824], [825], [826])


6.31  Culli [probably extinct] (1)

Core area: Andes

Canonical source: [827]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([827: 401–405])


6.32  Fulniô (1)

Core area: Amazon

Canonical source: [828]
6.33 Guahibo (5)

Core area: NW Sphere

Canonical source: [830]

Documentation: (Full) Grammar ([831])


6.34 Guaicuruan (5)

Core area: Andes

Canonical source: [832]

Documentation: (Full) Grammar ([833])


6.35  **Guamo [extinct] (1)**

Core area: NW Sphere

Canonical source: [834]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([835])


6.36  **Guató (1)**

Core area: Amazon

Canonical source: [836]

Documentation: (Full) Grammar ([837])


6.37  **Harakmbut (2)**

Core area: Peruvian Jungle

Canonical source: [838]

Classification comment: There are promising lexical links with Katukina ([839]), with a fair amount of near-identical forms, but the systems of pronouns, numerals, or bound morphology show no cognition.

Documentation: (Full) Grammar ([840])


### 6.38 Hibito-Cholon [extinct] (2)

Core area: Peruvian Amazon

Canonical source: [841: 461–463]

Documentation: (Full) Grammar ([842])


### 6.39 Huarpean [extinct] (2)

Core area: S Andes

Canonical source: [843]

Documentation: (Full) Grammar ([844])


De Valdivia, Luis. 1894 [1607]. *Doctrina christiana y catecismo con un confesionario, arte y vocabulario breves en lengua allentiac*. Sevilla: Rasco.
6.40  Iranxe (Münkü) (1)

Core area: Mato Grosso

Canonical source: [845]

Documentation: (Full) Grammar ([846])


6.41  Itonama (1)

Core area: Bolivia

Canonical source: [847]

Documentation: Grammar sketch ([847: 5])


6.42  Jabutí (2)

Core area: Rondônia

Canonical source: [848]

Documentation: (Full) Grammar ([849])


6.43  Jê (16)

Core area: Amazon

Canonical source: [850], [851]

Classification comment: The Macro-Jê family has yet to be proven beyond superficial similarities.

Documentation: (Full) Grammar ([852])


6.44  Jirajaran [extinct] (1)

Core area: NW Sphere

Canonical source: [853]

Documentation: Less than sketch but sufficient data to decide OV/VO-feature ([853: 129–130])


6.45  Jivaro (4)

Core area: Peruvian Amazon

Canonical source: [854]

Documentation: (Full) Grammar ([855])


6.46 Jodi (1)

Core area: Orinoco, Amazon

Canonical source: [856]

Classification comment: There are some miscellaneous resemblances to language in the Nadahup family [857].

Documentation: Grammar sketch ([858])


6.47 Kakua-Nukak (2)

Core area: SE Colombia

Canonical source: [859]

Documentation: Grammar sketch ([860]; in particular, for setting the OV/VO-feature: [861])


6.48 Kamakã [extinct] (1)

Core area: Amazon

Canonical source: [862]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([863: 188], [862])


6.49 Kamsá (1)

Core area: NW Sphere

Canonical source: [864]

Documentation: Grammar sketch ([865], [864], [866])


6.50 Kanoê (Kapixaná) (1)

Core area: Rondônia

Canonical source: [867]

Documentation: (Full) Grammar ([868])


Nijmegen: Katholieke Universiteit Nijmegen doctoral dissertation.

6.51 Karajá (1)

Core area: Amazon

Canonical source: [869]

Documentation: Grammar sketch ([870]; in particular, for setting the 
OV/VO-feature: [871])

http://butler.cc.tut.fi/~fabre/BookInternetVersio/Alkusivu.html 
(May 2005).


6.52 Kariri/Kiriri [extinct] (1)

Core area: Amazon

Canonical source: [872]

Documentation: Grammar sketch ([873]; in particular, for setting the 
OV/VO-feature: [874], [875])

http://butler.cc.tut.fi/~fabre/BookInternetVersio/Alkusivu.html 
(May 2005).
6.53 Katukina (3)

Core area: Amazon

Canonical source: [876]

Classification comment: There are promising lexical links with Katukina ([877]), with a fair amount of near-identical forms, but the systems of pronouns, numerals, or bound morphology show no cognition.

Documentation: Grammar sketch ([878], [879], [880])


6.54 Kawesqar (2)

Core area: Tierra del Fuego

Canonical source: [881]

Documentation: (Full) Grammar ([882])


6.55 Kwazá (1)

Core area: Rondônia

Canonical source: [883]

Documentation: (Full) Grammar ([884])


6.56 Leko (1)

Core area: Bolivia

Canonical source: [885]

Documentation: Grammar sketch ([885])

6.57 **Lengua-Mascoy (Enlhet)** (5)

Core area: S Andes

Canonical source: [886]

Documentation: (Full) Grammar ([887])


6.58 **Lule** [extinct] (1)

Core area: Andes

Canonical source: [888]

Classification comment: Lule shares some cultural vocabulary with Vilela, but otherwise the two are clearly distinct.

Documentation: (Full) Grammar ([889])


6.59 **Máku** [extinct] (1)

Core area: Orinoco, Amazon

Canonical source: [890]

Documentation: Grammar sketch ([891]; in particular, for setting the OV/VO-feature: [892])

6.60 Matacoan (Mataguayo) (7)

Core area: S Andes

Canonical source: [893], [894]

Documentation: (Full) Grammar ([895])


6.61 Matanawí [extinct] (1)

Core area: Madeira River

Canonical source: [896]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([896])

6.62 Maxakalí (1)

Core area: Amazon

Canonical source: [897]

Documentation: Grammar sketch ([898], [899], [900])


6.63 Mochica [probably extinct] (1)

Core area: Andes

Canonical source: [901]

Documentation: (Full) Grammar ([902])


6.64  Mosetén-Chimane (1)

Core area: Andes

Canonical source: [903]

Documentation: (Full) Grammar ([903])


6.65  Movima (1)

Core area: Bolivia

Canonical source: [904]

Documentation: (Full) Grammar ([904])


6.66  Muniche (1)

Core area: Peruvian Amazon

Canonical source: [905]

Documentation: Grammar sketch ([906: 26])


6.67  Mura-Pirahã (1)

Core area: Maici River

Canonical source: [907]

Documentation: (Full) Grammar ([907])

6.68 Nadahup (4)
Core area: Vaupés
Canonical source: [908], [909]
Documentation: (Full) Grammar ([908])


6.69 Nambiquaran (3)
Core area: Rondônia
Canonical source: [910]
Documentation: (Full) Grammar ([910])


6.70 Ofaié (1)
Core area: Amazon
Canonical source: [911]
Documentation: (Full) Grammar ([911])


6.71 Omurano [extinct] (1)
Core area: Peruvian Amazon
Canonical source: [912]
6.72  **Oti [extinct] (1)**

Core area: Peixe/Pando Rivers

Canonical source: [914]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([915: 85])


6.73  **Otomaco [extinct] (1)**

Core area: NW Sphere

Canonical source: [916]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([917])


6.74 Páez (1)

Core area: Colombia

Canonical source: [918]

Documentation: (Full) Grammar ([919])


6.75 Pankararu (Panakararé) [extinct] (1)

Core area: Amazon

Canonical source: [920]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([920])


6.76 Panoan (28)

Core area: W Amazon

Canonical source: [921: 11]

Classification comment: Pano-Tacanan family is considered premature (probable only) in the two most recent grammars of Panoan ([921: 11]) and Tacanan ([922: 9]) languages, therefore we go against Adelaar’s ([923: 419]) following of earlier works by Girard and Key.

Documentation: (Full) Grammar ([924])

6.77  Peba-Yagua (2)

Core area: Peruvian Amazon

Canonical source: [925]

Documentation: (Full) Grammar ([926])


6.78  Puelche (1)

Core area: Patagonia

Canonical source: [927]

Classification comment: The parallels with Chon or Chon-Querandí are interesting but not conclusive.

Documentation: (Full) Grammar ([928])


6.79 Puinavé (1)

Core area: SE Colombia

Canonical source: [929: 419–439], [930]

Classification comment: There are some lexical parallels with Nadahup and Kakua-Nukak ([929: 419–439]) but these are hardly conclusive of a genetic relation. The pronouns, morphology, and bulk of the vocabulary have no cognates.

Documentation: (Full) Grammar ([929])


6.80 Puquina [extinct] (1)

Core area: Andes

Canonical source: [931]

Documentation: Grammar sketch ([932]; in particular, for setting the OV/VO-feature: [931: 438] and [933])


6.81 Puri [extinct] (1)

Core area: Amazon

Canonical source: [934]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([935: 66–68])
6.82 Quechuan (46)

Core area: Andes

Canonical source: [936]

Documentation: (Full) Grammar ([937])


6.83 Rikbaktsá (1)

Core area: Amazon

Canonical source: [938]

Documentation: Grammar sketch ([939])


6.84 Saliban (3)

Core area: NW Sphere

Canonical source: [940]

Documentation: (Full) Grammar ([941])
6.85 Sapé (Kaliana) (1)

Core area: Orinoco, Amazon

Canonical source: [942]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([943])


6.86 Sechuran [extinct] (1)

Core area: Andes

Canonical source: [944]

Classification comment: There are occasional lexical links to the Tallán language(s) ([945: 398–400]), but the very limited data available is not compelling towards a genetic relationship.

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([945])


6.87 Tacanan (6)

Core area: Bolivia

Canonical source: [946: 9]

Classification comment: Pano-Tacanan family is considered premature (probable only) in the two most recent grammars of Panoan ([947: 11]) and Tacanan ([946: 9]) languages, therefore we go against Adelaar’s ([948: 419]) following of earlier works by Girard and Key.

Documentation: (Full) Grammar ([946])


6.88 Tallán [extinct] (1)

Core area: Andes

Canonical source: [949]

Classification comment: There are occasional lexical links to the Sechuran language ([950: 398–400]), but the very limited data available is not compelling towards a genetic relationship.

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([950])


6.89 **Taruma [extinct] (1)**

Core area: Guianas/Brazil

Canonical source: [951: 150]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([951: 150])


6.90 **Taushiro (1)**

Core area: W Amazon

Canonical source: [952]

Documentation: (Full) Grammar ([953])


6.91 **Ticuna (Tucuna) (1)**

Core area: W Amazon

 Canonical source: [954]

Documentation: (Full) Grammar ([955])


6.92  Timote-Cuica [probably extinct] (1)

Core area: NW Sphere

Canonical source: [956]

Documentation: Grammar sketch ([957]; in particular, for setting the OV/VO-feature: [958: 55], [959: 101–102])


6.93  Tinigua (1)

Core area: NW Sphere

Canonical source: [960]

Documentation: Grammar sketch ([961])


6.94 Trumai (1)

Core area: Upper Xingu
Canonical source: [962]
Documentation: (Full) Grammar ([962])


6.95 Tucanoan (25)

Core area: W Amazon
Canonical source: [963], [964]
Documentation: (Full) Grammar ([965])


6.96 Tupí (76)

Core area: Amazon
Canonical source: [966], [967], [968]
Documentation: (Full) Grammar ([969])


6.97  **Tuxá [extinct] (1)**

Core area: Pernambuco

Canonical source: [970]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([970], [971])


6.98  **Urarina (1)**

Core area: Peruvian Amazon

Canonical source: [972]

Documentation: (Full) Grammar ([972])

6.99 Uru-Chipaya (2)

Core area: Andes

Canonical source: [973]

Documentation: (Full) Grammar ([974])


6.100 Vilela [extinct] (1)

Core area: Andes

Canonical source: [975]

Classification comment: Vilela shares some cultural vocabulary with Lule, but otherwise the two are clearly distinct.

Documentation: Grammar sketch ([976])


6.101 Waorani (1)

Core area: Peru, Ecuador

Canonical source: [977]

Documentation: (Full) Grammar ([978])
6.102. **Warao (1)**

Core area: Orinoco

Canonical source: [979]

Documentation: (Full) Grammar ([979])


6.103. **Xukuru [extinct] (1)**

Core area: Pernambuco State

Canonical source: [980]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([981])


6.104. **Yámana (Yahgan) (1)**

Core area: Tierra del Fuego

Canonical source: [982]

Documentation: (Full) Grammar ([983])
6.105 Yanomam (4)

Core area: Orinoco, Amazon

Canonical source: [984]

Documentation: (Full) Grammar ([985])


6.106 Yaruro (Pumé) (1)

Core area: Venezuela

Canonical source: [986]

Documentation: Grammar sketch (987: 556)


6.107 Yurakare (1)

Core area: Bolivia

Canonical source: [988]

Documentation: (Full) Grammar ([988])

6.108  Yurí (1)

Core area: W Amazon

Canonical source: [989]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([990: 232–244])


6.109  Yurumanguí [extinct] (1)

Core area: NW Sphere

Canonical source: [991]

Documentation: Less than sketch and insufficient to decide OV/VO-feature ([991: 60–61])


6.110  Zamucoan (2)

Core area: Andes

Canonical source: [992]

Documentation: (Full) Grammar ([993])


6.111 Zaparoan (6)

Core area: Peruvian Amazon

Canonical source: [994]

Documentation: Grammar sketch ([995]; in particular, for setting the OV/VO-feature: [996: 21])

