Achievements and Challenges in the Description of the Languages of Melanesia

Harald Hammarström and Sebastian Nordhoff

Abstract

The present paper assesses the state of grammatical description of the languages of the Melanesian region based on database of semi-automatically annotated aggregated database of bibliographical references. 150 years of language description in Melanesia has produced at least some grammatical information for almost half of the languages of Melanesia, almost evenly spread among coastal/non-coastal, Austronesian/non-Austronesian and isolates/large families. Nevertheless, only 15.4% of these languages have a grammar and another 18.7% have a grammar sketch. Compared to Eurasia, Africa and the Americas, the Papua-Austronesian region is the region with the largest number of poorly documented languages and the largest proportion of poorly documented languages. We conclude with some dicussion and remarks on the documentational challenge and its future prospects.

1 Introduction

Melanesia is the sub-region of Oceania extending from the western end of the Pacific Ocean to the Arafura Sea and eastward to Fiji – see the map in Figure 1. This region is home to no less than 1 347 (1315 living + 32 recently extinct) attested indigenous languages[1] as per the language/dialect divisions of Lewis (2009), with small adjustments and adding attested extinct languages given in Table 1.

The present paper seeks to describe the current state of description of the languages of Melanesia in detail (in the online appendix) and in general (in the body of the paper) based on a database of annotated bibliographical

[1] Excluded from this count are a) non-Melanesian languages that immigrated in historical times, e.g., New Caledonian Javanese, b) languages lexified from immigrant non-Melanesian languages, e.g., Tok Pisin, and c) all pidgin languages since we are unable to count them properly given their state of documentation and fluctuating nature (Bakker and Parkvall 2010).

Action	Language	Location	Living/Extinct	Brief Rationale
Added	Bai of Miklucho-Maclay	PNG, Madang	Presumed Extinct	Not the same as Dumun (Z'graggen 1975:13-14)
Added	Nori	PNG, Sandaun	Extinct	Not the same as Warapu (Corris 2005, Donohue and Crowther 2005, Wilkes 1926)
Added	Kaniet of Dempwolff	PNG, Manus	Presumed extinct	Not the same as Kaniet of Thilenius (Blust 1996)
Added	O'oku	PNG, Northern Province	Presumed Extinct	Seemingly a Yareban language (Ray 1938a)
Added	Butam	PNG, New Britain	Extinct	Laufer 1959
Added	Pauwi of Moszkowski	Indonesia, Papua	Presumed Extinct	May be a confused elicitation session (Moszkowski) 1913), but in any case not the same as Robidé van der Aa's Pauwi (Robidé van der Aa 1885) which we count as Warembori [wsa]
Added	Batanta	Indonesia, Raja Ampat	Presumed Extinct	Remijsen (2002:42) cites reports of unintelligibility with neighbouring languages and data appears in Cowan (1953)
Added	Mansim	Indonesia, Bird's Head	Rumours of ca 50 speakers in the Manokwari area	Reesink 2002
Added	Binahari-Ma	PNG, Northern Province	Alive	Arguably a different language from Binahari-Neme (Dutton 1999)
Added	Nese	Vanuatu	Alive	Crowley 2006a
Added	Womo-Sumararu	PNG, Sandaun	Alive	Donohue and Crowther 2005
Removed	Dororo [drr]	Solomon Islands	Extinct	Not different from Kazukuru (Dunn and Ross 2007)
Removed	Guliguli [gli]	Solomon Islands	Extinct	Not different from Kazukuru (Dunn and Ross 2007)
Removed	Makolkol [zmh]	PNG, New Britain	Possibly Extinct	Unattested (Stebbins 2010:226)
Removed	Wares [wai]	-	Indonesia, Papua	Unattested or same as Mawes [mgk] (Wambaliau 2006)
Removed	Yarsun [yrs]	-	Indonesia, Papua	Unattested or same as Anus [auq] or Podena [pdn] (van der Leeden 1954)

Table 1: Adjustments concerning the languages of Melanesia to the language catalogue of Lewis (2009). We have not added totally unattested, very poorly attested languages (e.g., Ambermo, attested in two numerals, Fabritius 1855), or once attested languages whose attestation has disappeared (e.g., Rutan, only 3 words now remaining, Crowley 2006b:3).

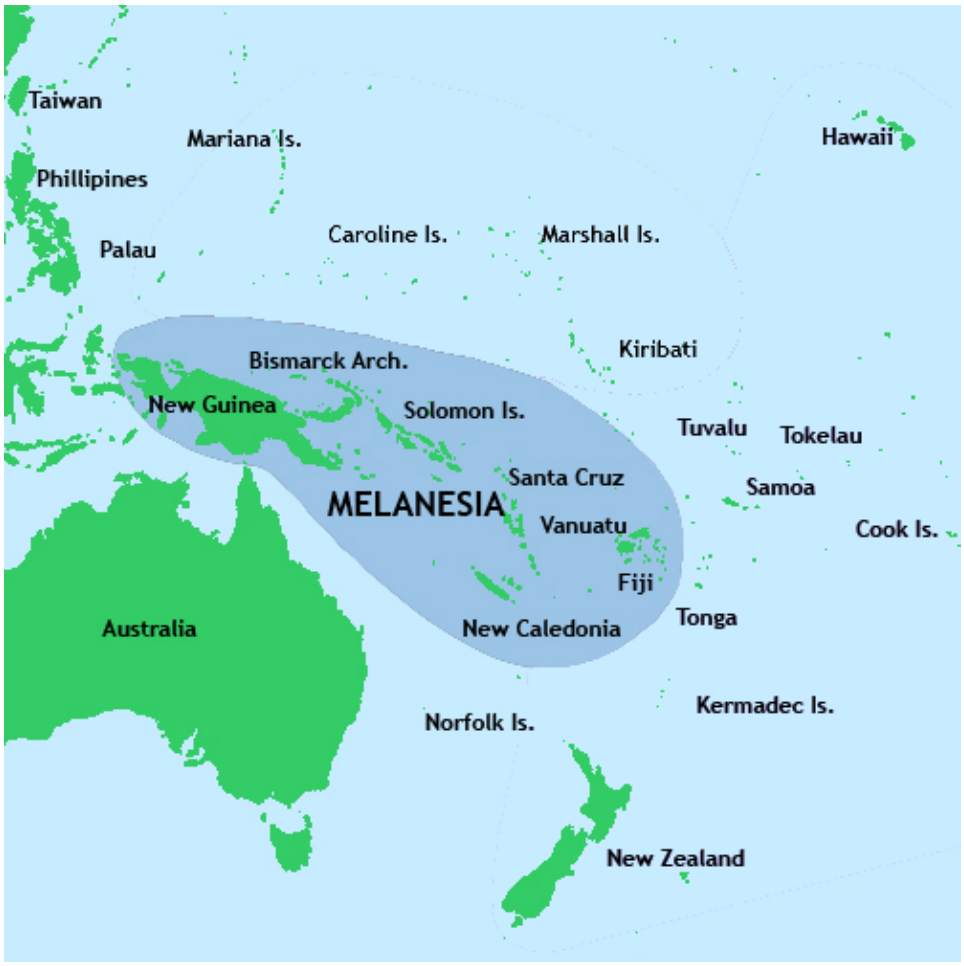

Figure 1: Map of Melanesia adapted from `http://en.wikipedia.org/wiki/Melanesia` accessed 10 July 2011. The countries present in Melanesia are Papua New Guinea, Indonesia, Fiji, France (New Caledonia), Solomon Islands and Vanuatu.

references. This database of references, called LangDoc (Hammarström and Nordhoff 2011), spans the entire world but we restrict it to the Melanesian subset in the present survey.

2 Assessing Status of Description

To assess status of description we first a) collect all relevant bibliographical references, b) annotate them as to (target-)language and type (grammar, wordlist etc), and c) for each language, mark its status of description according to the most extensive or sum description it has.

2.1 Collecting References

Language documentation and description is, and has been, a decentralized activity carried out by missionaries, anthropologists, travellers, naturalists, amateurs, colonial officials, and not least linguists. In order to comprehensively collect all relevant such items, we have, in essence, gone through all handbooks and overviews concerning the Melanesian region[2], in the hope that specialists on families and (sub-)regions have the best knowledge on what descriptive materials actually exist. This is supplemented by a) intensive searching as to (sub-)regions for which there is no recent expert-written handbook/overview paper and b) whole-sale inclusion of relevant existing bibliographical resources such as the WALS, the SIL Bibliography, SIL Papua Guinea Bibliographies, the library catalogue of MPI EVA in Leipzig and so on – see Hammarström and Nordhoff (2011) for a little more detail regarding this procedure and alternatives.

Everything published by a locatable publisher has been included as well as MAs and PhDs since they should, in principle, be findable via the national library or the degree-giving institution. However, field notes, manuscripts, self-published items and items published by a local bible society have not been included since they cannot be located systematically. In our experience, locating manuscripts too often turns out to be a wild goose chase and including them in the current survey would do more harm than good, in particular, give a false picture of the state of (accessible) description. However, we have included a small number of manuscripts and/or fieldnotes where the item in question has been posted on the internet and/or is verified to be located in a

[2]There are over 1 000 handbooks and overview articles relevant for the Melanesian region.

publicly accessible archive (e.g., the KITLV in Leiden), and thus meets the accessibility criterion.

It should be stressed, however, that the amount of original and valuable data sitting in unpublished form is highly significant – to give just a few examples, Capell (1962) cites a large number of missionary manuscripts from the islands east of the Papuan mainland, the archives of the SIL in Jayapura and Ukarumpa (cf. Silzer and Heikkinen-Clouse 1991) hold a huge number of unpublished survey wordlists and/or grammar sketches spanning (in our impression) at least 50% of the languages of Melanesia, and linguists Mark Donohue and William Foley have unpublished field data from Indonesian Papua and the Sepik-Ramu region respectively which is enough for several full grammars and dozens of grammar sketches (p.c. Mark Donohue 2008 and William Foley 2010). If unpublished material is included, the descriptive picture of the languages of Melanesia changes significantly, especially on the breadth side, with far more data on the lesser-known languages (cf. Carrington 1996).

In total, the bibliographical database contains 11 290 references pertaining to Melanesia.

2.2 Annotation

Bibliographical references are annotated as to identity, i.e., the iso-639-3 code of the language(s) treated, and type of description, i.e., grammar, wordlist etc. As to type, the following hierarchy has been used:

- `grammar`: a extensive description of most elements of the grammar $\sim$ 150 pages and beyond
- `grammar sketch`: a less extensive description of many elements of the grammar 20-150 pages (typically $\sim$ 50 pages)
- `dictionary`: $\sim$ 75 pages and beyond
- `specific feature`: description of some element of grammar (i.e., noun class system, verb morphology etc)
- `phonology`: phonological description with minimal pairs
- `text`: text (collection)
- `wordlist`: $\sim$ a couple of hundred words
- `minimal`: a small number of cited morphemes or remarks on grammar

- `sociolinguistic`: document with detailed sociolinguistic information
- `comparative`: inclusion in a comparative study with or without cited morphemes, e.g., lexicostatistical survey
- `handbook/overview`: document with meta-information about the language (i.e., where spoken, non-intelligibility to other languages etc.)
- `ethnographic`: ethnographic information on the group speaking a language

The hierarchy is an ad-hoc amalgam of existing annotation, automatizability properties and bias towards typologist usage (with grammar at the top at the expense of text and dictionary, and form-function pairs rated higher than sociolinguistic information). It is in many ways imperfect, but it is more informative than nothing. Other existing schemas could not be felicitously adopted, e.g., Moore (2007:33) is similar to the present scheme but credits the existence of various types (scientific articles, dissertations, etc.) rather than their actual content, and AIATSIS (2011:285-297) is also similar to the present scheme but so much more detailed (several hundred categories including vocabulary/animals, vocabulary/body parts, etc.) that it could not be automatized or done by hand within the scope of the present project. Bibliographical references in the present project have been annotated both automatically and by hand. Some examples are shown in Table 2.

Automatic annotation is possible when the title words contain the language name and/or word(s) revealing the type of the document, e.g., "A grammar of Tauya" can be automatically recognized as [tya] and `grammar`. Exactly how this is done and what percentages of correctness are to be expected is described in Hammarström (2008, 2011).

For most references, number of pages is recorded, and is used to rank within categories.

2.3 Status of Description per Language

For each language, the references concerning it are aggregrated and its status of description is straightforwardly assessed as per the annotation hierarchy. In addition, for the purposes of the current presentation, it has been simplified into a more distilled scheme as per Table 3.

There may be missing extant references and manual as well as automatic annotation has gaps and errors. The claim we are able to make is that at least the status of description for every language should be correct. That is, the

Reference	Language	Type	Comment
Lindström, Eva. (2002) Topics in the Grammar of Kuot. Stockholm University doctoral dissertation, 265pp.	Kuot [kto]	`grammar`	although it contains some text and a Swadesh word-list at the end, it counts as `grammar`
Franklin, Karl J. & C. L. Voorhoeve. (1973) Languages near the intersection of the Gulf, Southern Highlands and Western Districts. In Karl J. Franklin (ed.), The linguistic situation in the Gulf District and adjacent areas, Papua New Guinea (Pacific Linguistics: Series C 26), 149-186. Canberra: Research School of Pacific and Asian Studies, Australian National University	Fasu [faa], Foe [foi], Fiwaga [fiw], Kewa [kew]	`overview; comparative; minimal`	There is a discussion of comparative matters and a number of morphemes are given (for each language).
Wirz, Paul. (1924) Anthropologische und ethnologische Ergebnisse der Central Neu-Guinea Expedition 1921-1922. Nova Guinea XVI. 1-148.	Zwart Valley = Dani-Western [dnw]	`ethnographic; grammar sketch`	It contains a grammar sketch in addition to ethnographic data.
Hughes, Jock. (1987) The languages of Kei, Tanimbar and Aru: Lexicostatistic classification. In Soenjono Dardjowidjojo (ed.), Miscellaneous studies of Indonesian and other languages in Indonesia, part 9 (NUSA: Linguistic Studies of Indonesian and Other Languages in Indonesia 27), 71-111. Jakarta: Universitas Katolik Indonesia Atma Jaya.	Mariri [mqi], East Tarangan [tre], Lorang [lrn], Lola [lcd], Koba [kpd], Kompane [kvp], Batuley [bay], Barakai [baj], Karey [kyd]	`overview; comparative`	No actual words or wordlists are included, just results of comparing wordlists.

Table 2: Examples of the annotation scheme used in the present survey.

type	distilled type	numerical value
grammar	grammar	4
grammar sketch	grammar sketch	3
dictionary	phonology/dictionary/specific/text	2
text	phonology/dictionary/specific/text	2
specific feature	phonology/dictionary/specific/text	2
wordlist	wordlist or less	1
minimal	wordlist or less	1
sociolinguistic	wordlist or less	1
comparative	wordlist or less	1
handbook/overview	wordlist or less	1
ethnographic	wordlist or less	1
<type annotation lacking>	wordlist or less	1

Table 3: The full- and distilled description level hierarchy used in the present survey.

outcome has been screened at the language level by an informed human, and inasmuch as errors of omission and annotation remain, they do not alter the (correct) status of description of any language. Thus, for a language which only has a published wordlist to its documentation it may be that there are several wordlists published, but only one of them is accurately reflected in the database (accurately reflecting the others would not change the status of description away from wordlist), and, if a language is given a certain status of description, the claim is that there is, in reality, no other descriptive publication that would give it a higher mark. Of the publications that are the witness to the status of description of a language (the most significant items of description) 95% have been personally inspected by the authors, but, since this was done over a long period of time it is no guarantee of consistency and we are not in a position to assess to quality of a description.

It should be noted again, that the above hierarchy reflects descriptive status and has a bias towards typologist usage. For example, a language that has a grammar, dictionary and text collection will be ranked the same (`grammar`) as a language with only a grammar, even though the former is better overall documented. An index of overall documentation (e.g., with points separately for grammatical-, lexical- and textual documentation) could be computed from the same database. We do not do this for the present survey since we cannot venture the same claim of completeness as with the grammar-oriented scheme above. In other words, the database screening is likely to have missed cases of missing texts and dictionaries for languages which already have a grammar (sketch). The database is released to the public so that others who are more interested in overall documentation can complete the database and compute figures of their own.

The fact that "grammar" is the highest weighted category of description should not be taken to mean that a language with a grammar is completely described – it merely means that it is the highest category of grammatical description that is commonly distinguished by linguists, i.e., there are as yet no descriptions that are called "super-grammars" or the like. However, grammars can be more or less comprehensive and a correlate of this (with validity only on average) may be the number of pages, which is recorded in the present database.

3 Status of Description of Melanesian Languages

Results of the full survey is given in the online appendix, sorted by family, author and language. We review the generalities here.

	Living	Extinct	Total	
grammar	207	0	207	**15.4%**
grammar sketch	245	7	252	**18.7%**
phonology or sim.	107	2	109	**8.1%**
wordlist or less	756	23	779	**57.8%**
			1347	

Table 4: Raw number of languages in Melanesia and their level of description.

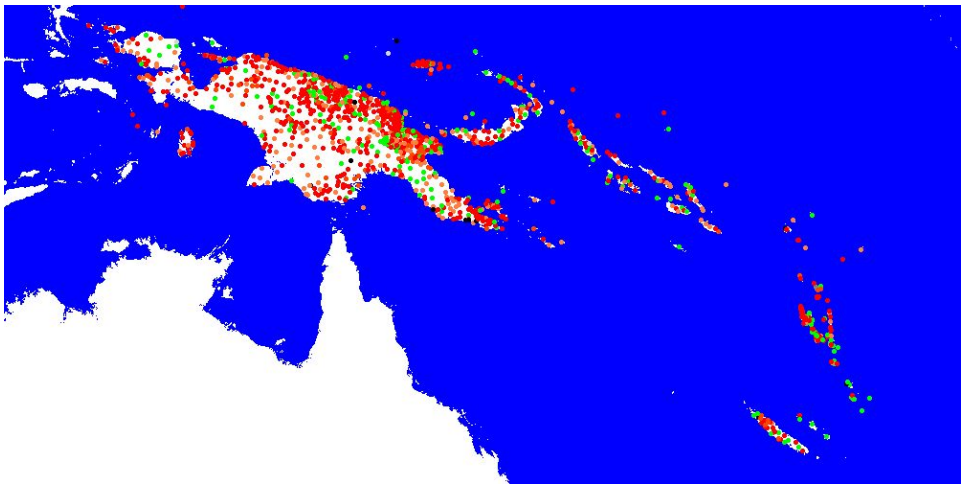

Figure 2: The location and description level of Melanesian languages. The colour coding is grammar = `green`, grammar sketch = `orange` or `light gray` (if extinct), phonology of sim. = `orange red` or `slate gray` (if extinct), wordlist or less = `red` or `black` (if extinct).

Raw numbers of languages described to various degrees are shown in Table 4 and a map is shown in Figure 2. The numbers speak for themselves, yet the most conspicuous fact is that more than half of the languages of Melanesia have only a wordlist or less of published descriptive material. Any non-trivial generalizing statement concerning the grammar of languages of Melanesia can only be at most half-fully grounded empirically. For example, (Wurm 1954), drawing on data and experience from Capell, was acquainted with all Melanesian languages described at the time, and lists some 20 tone languages – survey's of tone on New Guinea half a century later (Cahill 2011, Donohue 1997) turn up far more and far different tonal languages in Melanesia.

Historically speaking, early wordlists were catalogued superbly[3] by Ray (1893, 1912, 1914, 1919, 1920, 1923, 1926, 1929, 1938b,a) for the entire Melanesian area, and the history of research has been adequately surveyed qualitatively by area experts (Beaumont 1976, Chowning 1976, Dutton 1976, Grace 1976, Haudricourt 1971, Healey 1976, Hooley 1976, Laycock 1975, 1976, Laycock and Voorhoeve 1971, Lincoln 1976, Lithgow 1976, Lynch and Crowley 2001, Schütz 1972, Taylor 1976, Tryon and Hackman 1983, Voorhoeve 1975b, Z'graggen 1976). We supplement these with some quantitative results in Figure 3. As can be seen, language description in Melanesia takes off in the second half of the 19th century with travellers, colonial officers, and missionaries producing wordlists From there description increases at a steady pace, due mostly to missionaries and German scientists. A sharp rise in the number of items produced every year, and a corresponding (but less sharp) increase in the overall descriptive status, happens after 1950, presumably due to the establishment of the SIL in Papua New Guina (Hooley 1968, Foley 1986:13). The pace has since been kept up mainly by SIL missionaries and academic linguists in Australia and other western countries. Very little has so far been produced by Melanesians themselves; notable exceptions include (Flassy 2002, Nekitel 1985, Sumbuk 1999). There are more than a dozen languages whose correspondig ethnic groups have a monograph-length ethnographic description, yet the languages are not described beyond a wordlist, e.g., Gnau [gnu] (Lewis 1975) or Banaro [byz] (Juillerat 1993).

In the early times, languages near the coast were much better known than inland languages. At the present time, this correlation is much diluted.

[3]We have never really found an example of a language missed by Ray, except possibly the 9 word vocabulary of Berau given to Marsden (1834) by non-native speakers (the second earliest vocabulary of Papuan language from the New Guinea mainland). However, this vocabulary could not be confirmed as a Papuan language in Ray's times since the first independent Berau data appeared only in Galis (1955).

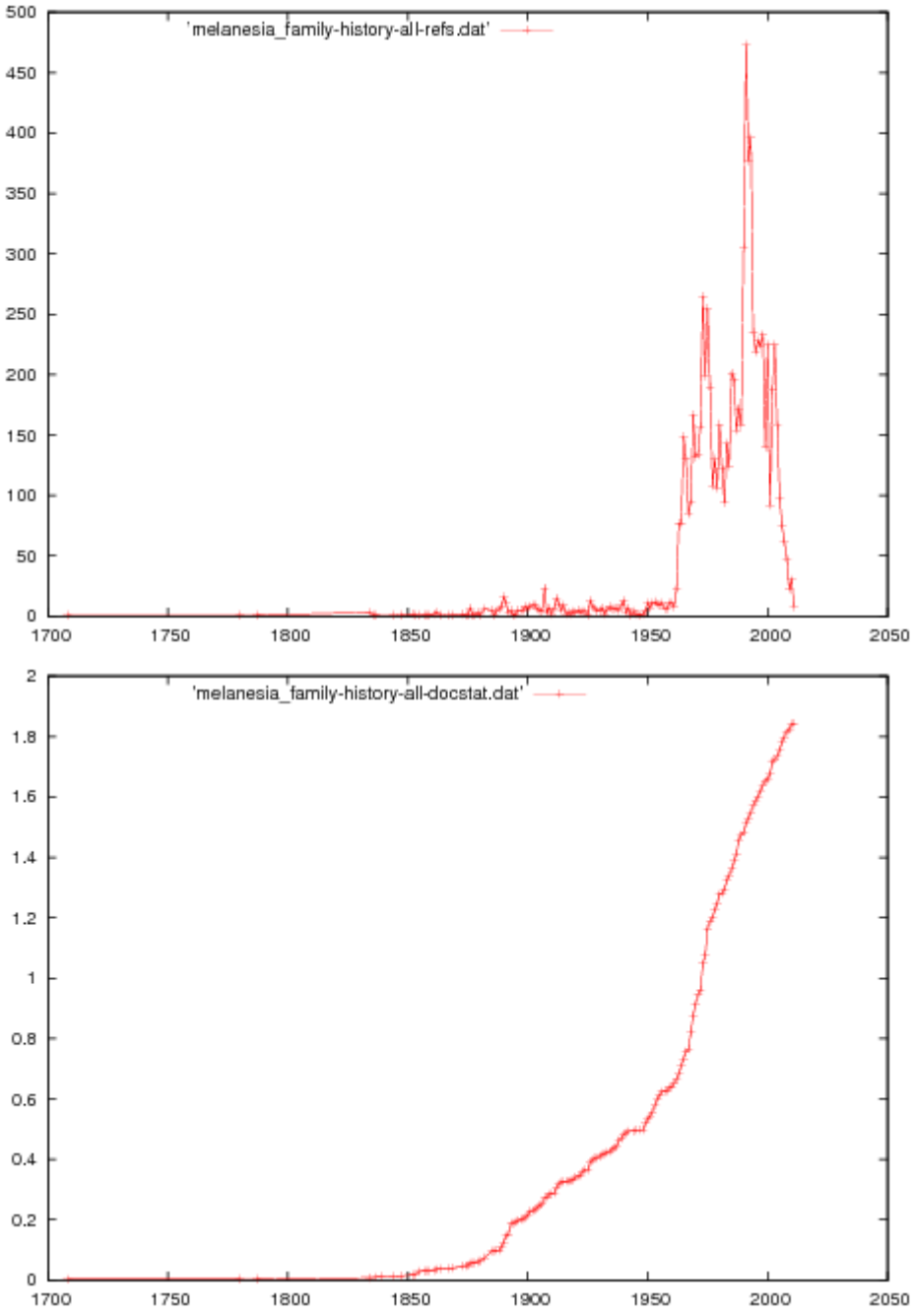

Figure 3: The upper diagram shows the raw number of publications per year concerning languages of Melanesia. The lower diagram shows the average description level as it increases through time.

	Avg. dist. to coast	$p \approx$	Median dist. to coast	$p \approx$
grammar	44.91 kms	0.340	14.97 kms	**0.026**
grammar sketch	46.84 kms	0.463	17.90 kms	0.462
phonology or sim.	46.71 kms	0.466	16.75 kms	0.346
wordlist or less	46.85 kms	0.373	20.09 kms	0.133
overall	46.51 kms		17.95 kms	

Table 5: Average and median distance (as the crow flies) for languages of various levels of description. Significance testing is by selecting 1000 random subsets of the corresponding size from the total pool of 1347 languages and checking how many of those have an average/median distance lower viz. higher than the distance to be tested.

Table 5 shows the median and average distances (as the crow flies) to the coast for the various levels of description, which shows little difference[4]. The slight tendency for grammars to be written of languages nearer to the coast is not statistically significant for average distances, but it is so for median distances. This means that half of languages with grammars are within 14.97 kms to the coast whereas half of the languages of other categories are 10-15% further away, and that languages with grammars that are not near the coast (the exceptions) are so far away that they blur the tendency on average. This overall lack of a stronger trend must be taken to mean that flight and river access inland, balances the amount of neglected languages on the coast and immediate coastal hinterlands.

As is well-known, the language of Melanesia divide into two classes, the Austronesian languages (522 languages) and the non-Austronesian languages (825 languages). The Austronesian languages are more coastal (average 12.79 kms and median 9.92 kms from the coast) than the Papuan ones (average 67.92 kms and median 44.66 kms), but since there is only a weak or no trend that favours the description of coastal languages, we can check fairly easily if there is a bias towards the description of Austronesian or non-Austronesian languages. Figure 4 shows that, historically, there was a long time during which AN languages were better described on average (presumably due to being coastal) and in recent times the slightly higher level has been regained. The current average level of description for AN languages in Melanesia is 2.04 against 1.84 for non-AN languages. The difference is slight but highly

[4]For all computations we have used a database of language centrepoints emanating from Lewis (2009) and the coastlines implied by the SRTM topography database (downloadable from `http://www-radar.jpl.nasa.gov/srtm/` accessed 4 May 2011).

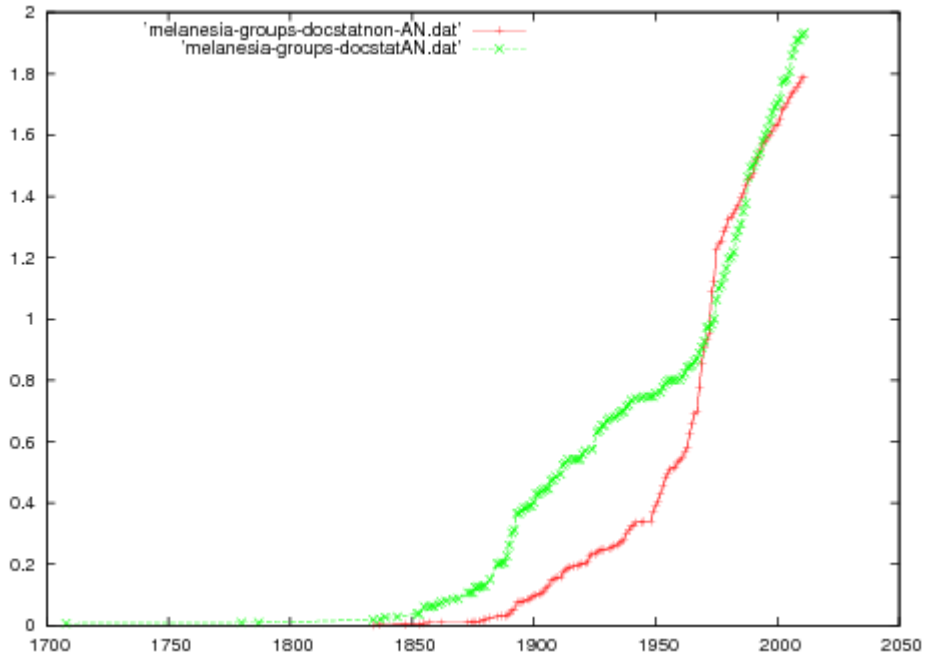

Figure 4: The average description level for Austronesian (AN, green) and non-Austronesian (non-AN, red) languages through time.

significant $p \approx 0.002$. The difference is hardly due to the tendency for full grammars to be coastal, as the AN languages have higher representation at all levels (beyond wordlist) as per Table 6. We do not know what the reason for this bias is.

It is difficult to say which is the best described language of Melanesia as would require a quality judgment that we are not in a position to make. However, the description with largest number of pages is Lichtenberk (2008)'s 1409-page grammar of To'aba'ita (an Oceanic Austronesian language). In fact, it is also the longest grammar of any lesser-known language in the world, in terms of number of pages devoted to grammatical description. The second longest grammar of a language of Melanesia is Aikhenvald (2008)'s 727-page grammar of Manambu (a Ndu language). As far as can be told with documents accessible to us, the least described languages whose existence seems certain enough, are Kehu [khh] and Kembra [xkw], two seemingly isolated languages in Indonesian Papua. Kehu is known from from two unpublished minuscule wordlists (Moxness 1998, Whitehouse no date) at least one of

	Austronesian		non-Austronesian		Total	
grammar	93	**17.82%**	114	**13.82%**	207	**15.37%**
grammar sketch	104	**19.92%**	148	**17.94%**	252	**18.71%**
phonology or sim.	55	**10.54%**	54	**6.55%**	109	**8.09%**
wordlist or less	270	**51.72%**	509	**61.70%**	779	**57.83%**

Table 6: Numbers and proportions of Austronesian (AN) and non-Austronesian (nAN) languages at different levels of description.

which is from a non-native speaker, and Kembra is known from a minuscule wordlist taken up from a transient speaker by Doriot (1991) attributed to a village named Kembra near the confluence of the Sobger and Nawa (Kiambra appears at the right place on a colonial map, Hoogland 1940).

Arguably the most prolific author of descriptive work on Melanesian languages is the Dutch Catholic priest Petrus Drabbe (Voorhoeve 2000) who can count to his name no less than 4 languages with grammars, another 19 with grammar sketches and wordlists for 6 more[5] spanning a range of different families. Linguist Terry Crowley wrote 6 grammars and 9 grammar sketches of Austronesian languages before his premature death in 2005. Linguists such as Arthur Capell, Stephen Wurm, Sidney Ray, Malcolm Ross, J. C. Anceaux, J. A. Z'Graggen, Darrell Tryon and C. L. Voorhoeve have published wordlists (or similar bits of information) of several hundred languages, either collected themselves or by others.

A current discussion among linguists as to priorities for documentation – the context being that time is running out – is whether to describe an undescribed isolated language or whether to describe an undescribed language from a family with other described languages. At present, we count 45 language isolates for the Melanesian region (see Hammarström 2010a,b:appendix for a justification of this figure). The 45 isolates have an average description level of 2.20 and the 1 298 non-isolates have 1.91. The difference, however, is not statistically significant at conventional levels of significance ($p \approx 0.070$). That is, there is no overall principle at work that favours the description of isolates rather than non-isolates. Nevertheless, there is a conspicuously large *absolute* number of underdescribed isolates and small families in the Melanesian region, especially lowland New Guinea – see Hammarström (2010b) for details.

[5] In addition, a grammar of Fordata (Drabbe 1926) which falls just outside the Melanesian region, a manuscript sketch grammar of Tamagário cited by Voorhoeve (1975a) as well as a number of dictionaries and scriptural text material.

	Africa	Australia	Eurasia	N America	Papua+AN	S America
grammar	780 [20]	94 [28]	537 [40]	264 [25]	415 [1]	260 [25]
grammar sketch	483 [35]	30 [22]	135 [18]	67 [32]	428 [10]	82 [28]
phonology or sim.	120 [5]	15 [2]	109 [2]	44 [9]	157 [2]	31 [17]
wordlist or less	603 [77]	45 [53]	684 [112]	105 [39]	978 [40]	30 [125]
Total	1986 [137]	184 [105]	1465 [172]	480 [105]	1978 [53]	403 [195]
Average desc.	2.68	2.69	2.31	2.91	2.12	2.88
% grammar	37.68%	42.21%	35.25%	49.40%	20.48%	47.66%
% living undoc.	28.40%	15.57%	41.78%	17.95%	48.15%	5.02%

Table 7: The number of languages at various levels of description broken up by macro-areas. The numbers outside brackets refer to strictly living languages and those within brackets refer to extinct. The last row gives the proportion of living languages with only a wordlist of less.

4 Melanesian Languages in Relation to the Rest of the World

The bibliographical database LangDoc spans the entire world in a fairly uniform way, allowing us to compare Melanesia to other conventional macro-areas of the world. The total database contains over 160 000 references collected and annotated in much the same way as the Melanesian subpart (Hammarström and Nordhoff 2011). Although the Eurasian, Australian and Meso-American sections have not been screened as thoroughly as the other areas yet, the general trends of the comparisons with should still be trustworthy. For this section, we will consider all Papuan-Austronesian languages together, not just the Melanesian ones, in order to appropriately cover all of the world's languages. This entails that the Eurasia figures do not include the Austronesian languages of South East Asia, the Philippines and Indonesia. Figures are shown in Table 7.

In absolute terms, Papua+Austronesian has the largest number of languages with only a wordlist to their documentation. In relative terms, Papua+Austronesian has the lowest proportion of grammars, the highest proportion of languages with only a wordlist or less, and the lowest average level of documentation. The Melanesia subpart scores slightly lower on all relative accounts. Therefore, Papua+Austronesian, and the languages of Melanesia in particular, can rightly be called the linguistically least known area of the world.

5 21st Century Challenges for Documentation

As is clear from the figures above, a formidable challenge for linguistic science is to provide descriptions of the vast number of un(der)described languages in the Melanesian region before it is too late.

On the optimistic side, a) the trend from the past century predicts a continued large production of grammatical descriptions and, b) it seems, impressionistically, that people from a wider array of countries of the world are taking interest in the Melanesian languages, and c) infrastructure in Melanesia is making it easier to reach and live in otherwise remote areas.

On the pessimistic side, a) at the same pace as infrastructure is developing the languages become endangered, b) violence, tropical diseases, visa/permit-matters and lack of funding continue to deter Westerners from in situ fieldwork, c) development of local talent and interest has been almost entirely lacking, and d) large amounts of descriptive work never reaches the scientific community, as if such materials had no scientific merit.

A few comments are in order.

The failure of local interest develop into active descriptive work is not endemic to Melanesia per se, but is widespread in all of the language-rich countries of the world. However, exceptions such as Brazil and Ethiopia show that it is possible for local universities and communities to take a productive interest in local languages.

In addition to unpublished materials alluded to above, many valuable descriptive works are difficult to access, in particular, a large number of unpublished PhD and MA-theses. PhD and MA theses are in many instances the most extensive description there is of a language. Many universities (for instance, the Australian National University) that regularly keep MA-theses do not allow interlibrary loans of them precisely because theirs is the only copy. Other universities, including the convenors of the 3L Language Documentation school, i.e., Leiden University, Université Lumière Lyon II and SOAS, do not regularly keep awarded MA theses at all. Perhaps the most blatant example of a university in antipathy of its scientific production actually being used is Université Libre de Bruxelles, as the first author experienced personally after making the trip to Bruxelles to read the presumably only library copy of Levy (2002)'s PhD grammar of Nubia-Awar – by far the most extensive description of that language. According to regulations, nobody – be it registered library card holders or visitors – is allowed to *read* this thesis (let alone borrow or photocopy from!) without the written consent of the author. Similarly, finished documents and reports from SIL Papua New Guinea and SIL Indonesia cannot be systematically accessed, although

many items have been made accessible in publication series and other outlets. Dissemination is a scientific principle, and scholarly institutions – be they missionary organizations or universities – that actively or passively restrict access to, or effectively let scientifically valuable documents be thrown away, do not fully merit the label 'scientific institution'. If descriptive work is continues to be disvalued in the above exemplified ways, there is less incentive for more descriptive work to be produced.

Apart from first-hand descriptive fieldwork, there are less obvious ways in which one can contribute to the description of Melanesian languages. A non-trivial number of languages of Melanesia have scripture translations, i.e., bodies of text with translation, but no published grammatical descriptions. The languages for which scripture translations are said to exist are given in Lewis (2009)[6]. Partial but substantial analyses of grammar can be done on the basis of text data from scripture translations, without fieldwork in situ. Comparative and typological work on languages of Melanesia can help generate interest in producing more detailed descriptions of Melanesian these languages. The digital era allows for tools on management, annotation and interoperability of language resources which can free up time for strictly human-needed analysis for language description. And, if nothing else, publishing or making available legacy resources is a valuable contribution, notably Smits and Voorhoeve (1992a,b, 1994, 1998)'s publication of Anceaux's gigantic wordlist collection from Indonesian Papua and PARADISEC digitization of Arthur Capell and Donald Laycock's fieldnotes from Papua New Guinea are prime examples.

Finally, we may point out a few somewhat subjective items of interest concerning language description in the Melanesian area. From a typological point of view, the languages with alliterative noun class concord of the Upper Sepik (Healey 1964, Laycock and Z'Graggen 1975) or the very isolating languages of Kolopom island (Drabbe 1949) or the Yuat river (Foley 2005); from a socio-historical point of view, the large number of small families and isolates between the Sepik and the Mamberamo which may constitute the world's linguistically most diverse region; from a contact-linguistics perspective, the almost undocumented pidgins Arira trade pidgin (Etherington 2002:3) and Duvle-Wano pidgin (Hammarström and Kamholz 2010) between the Lakes Plain and the Irian Highlands or yet more Austronesian-Papuan contact situations along the north coast of Indonesian Papua; from a

[6]Since it is non-trivial to find out the exact bibliographical reference to such scripture translations, they are not systematically included in the present reference database, unless, of course, the exact reference is cited in handbooks or the like.

historical-comparative perspective the Austronesian subgrouping and settlement of the Bird's Neck or the Kol language of New Britain and its possible relation to the Baining languages.

6 Conclusion

150 years of language description in Melanesia has produced at least some grammatical information for almost half of the languages of Melanesia, almost evenly spread among coastal/non-coastal, Austronesian/non-Austronesian and isolates/large families. Nevertheless, only 15.4% of these languages have a grammar and another 18.7% have a grammar sketch. Compared to Eurasia, Africa and the Americas, the Papua-Austronesian region is the region with the largest number of poorly documented languages and the largest proportion of poorly documented languages.

Acknowledgements

The authors wish to thank two anonymous reviewers for helpful comments.

References

AIATSIS, . 2011. AIATSIS Subject Thesaurus. Document posted at `http://www1.aiatsis.gov.au/thesaurus/data/SubjectThesaurus.pdf` accessed 10 June 2011.

Aikhenvald, Alexandra Y. 2008. *The Manambu language of East Sepik, Papua New Guinea*. Oxford University Press.

Bakker, Peter & Mikael Parkvall. 2010. Catalogue of Pidgin languages. Paper presented at the second APiCS conference 11-14 Nov, 2010.

Beaumont, Clive H. 1976. History of Research in Austronesian Languages: New Ireland. In Stephen A. Wurm (ed.), *New Guinea Area Languages and Language Study Vol 2: Austronesian Languages* (Pacific Linguistics: Series C 39), 171-177. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Blust, Robert. 1996. The linguistic position of the Western Islands, Papua New Guinea. In John Lynch & Pat Fa'afo (eds.), *Oceanic Studies: proceedings of the First International Conference on Oceanic Linguistics* (Pacific Linguistics: Series C 133), 1-46. Canberra: Australian National University.

Cahill, Michael. 2011. Tonal Diversity in Languages of Papua New Guinea. *SIL Electronic Working Papers* 2011-008. 23.

Capell, Arthur. 1962. *Linguistic Survey of the South-Western Pacific (New and revised edition)* (South Pacific Commission Technical Paper 136). Noumea: South Pacific Commission.

Carrington, Lois. 1996. *A linguistic bibliography of the New Guinea area* (Pacific Linguistics: Series D 90). Canberra: Research School of Pacific and Asian Studies, Australian National University.

Chowning, Ann. 1976. History of Research in Austronesian Languages: New Britain. In Stephen A. Wurm (ed.), *New Guinea Area Languages and Language Study Vol 2: Austronesian Languages* (Pacific Linguistics: Series C 39), 179-195. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Corris, Miriam. 2005. A grammar of Barupu, a language of Papua New Guinea. University of Sydney doctoral dissertation.

Cowan, H. K. J. 1953. *Voorlopige Resultaten van een Ambtelijk Taalonderzoek in Nieuw-Guinea.* 'S-Gravenhage: Martinus Nijhoff.

Crowley, Terry. 2006a. *Nese: A diminishing speech variety of Northwest Malakula (Vanuatu)* (Pacific Linguistics 577). Canberra: Research School of Pacific and Asian Studies, Australian National University.

Crowley, Terry. 2006b. *Tape: A declining language of Malakula (Vanuatu)* (Pacific Linguistics 575). Canberra: Research School of Pacific and Asian Studies, Australian National University.

Donohue, Mark. 1997. Tone systems in New Guinea. *Linguistic Typology* 1. 347–386.

Donohue, Mark & Melissa Crowther. 2005. Meeting in the middle: interaction in North-Central New Guinea. In Andrew Pawley, Robert Attenborough, Jack Golson & Robin Hide (eds.), *Papuan Pasts: Studies in the Cultural, Linguistic and Biological History of the Papuan-speaking Peoples* (Pacific Linguistics 572), 167-184. Canberra: Research School of Pacific and Asian Studies, Australian National University.

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

Drabbe, Peter. 1926. *Spraakkunst der Fordaatsche Taal* (Verhandelingen van het Koninklijk Bataviaasch Genootschap van Kunsten en Wetenschappen LXVII:Eerste Stuk). Batavia.

Drabbe, Peter. 1949. Bijzonderheden uit de Talen van Frederik-Hendrik-Eiland: Kimaghama, Ndom en Riantana. *Bijdragen tot Taal-, Land- en Volkenkunde* 105. 1–24.

Dunn, Michael & Malcolm Ross. 2007. Are Kazukuru languages really Austronesian?. *Oceanic Linguistics* 46(1). 210–231.

Dutton, Tom. 1999. From Pots to People: Fine-tuning the prehistory of Mailu Island and Neighbouring Coast, South-East Papua New Guinea. In Roger M. Blench & Matthew Spriggs (eds.), *Archaeology and Language, III* (One World Archaeology 34), 90-108. London & New York: Routledge.

Dutton, Tom E. 1976. History of Research in Austronesian Languages: Eastern Part of South-Eastern Mainland Papua. In Stephen A. Wurm (ed.), *New Guinea Area Languages and Language Study Vol 2: Austronesian Languages* (Pacific Linguistics: Series C 39), 129-140. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Etherington, Paul Anthony. 2002. Nggem Morphology and Syntax. Darwin: Northern Territory University MA thesis.

Fabritius, G. J. 1855. Anteekeningen omtrent Nieuw-Guinea. *Tijdschrift voor Indische Taal-, Land- en Volkenkunde* IV. 209–215.

Flassy, Don Augusthinus Lamaech. 2002. *Toror: A Name Beyond Language and Culture Fusion.* Jakarta: Balai Pustaka.

Foley, William A. 1986. *The Papuan languages of New Guinea* (Cambridge language surveys). Cambridge University Press.

Foley, William A. 2005. Linguistic prehistory in the Sepik-Ramu Basin. In Andrew Pawley, Robert Attenborough, Jack Golson & Robin Hide (eds.), *Papuan Pasts: Studies in the Cultural, Linguistic and Biological History of the Papuan-speaking Peoples* (Pacific Linguistics 572), 109-144. Canberra: Research School of Pacific and Asian Studies, Australian National University.

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

Grace, George W. 1976. History of Research in Austronesian Languages of the New Guinea Area: General. In Stephen A. Wurm (ed.), *New Guinea Area Languages and Language Study Vol 2: Austronesian Languages* (Pacific Linguistics: Series C 39), 55-72. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Hammarström, Harald. 2008. Automatic annotation of bibliographical references with target language. In *Proceedings of MMIES-2: Workshop on Multi-source, Multilingual Information Extraction and Summarization*, 57-64. ACL.

Hammarström, Harald. 2010a. A full-scale test of the language farming dispersal hypothesis. *Diachronica* XXVII(2). 197–213. Plus appendix 351pp.

Hammarström, Harald. 2010b. The Status of the Least Documented Language Families in the World. *Language Documentation & Conservation* 4. 177–212.

Hammarström, Harald. 2011. Automatic Annotation of Bibliographical References for Descriptive Language Materials. In Pamela Forner, Julio Gonzalo, Jaana Kekäläinen, Mounia Lalmas & Maarten de Rijke (eds.), *Proceedings of the CLEF 2011 Conference on Multilingual and Multimodal Information Access Evaluation* (LNCS 6941), 62-73. Berlin: Springer.

Hammarström, Harald & David Kamholz. 2010. A note on Duvle-Wano pidgin. Paper presented at the second APiCS conference 11-14 Nov, 2010.

Hammarström, Harald & Sebastian Nordhoff. 2011. LangDoc: Bibliographic Infrastructure for Linguistic Typology. *Oslo Studies in Language* 3(2). 31–43.

Haudricourt, André G. 1971. New Caledonia and the Loyalty Islands. In Thomas A. Sebeok (ed.), *Linguistics in Oceanea* (Current Trends in Linguistics 8), 359-396. Berlin: Mouton de Gruyter.

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

Healey, Alan. 1976. History of Research in Austronesian Languages: Admiralty Islands Area. In Stephen A. Wurm (ed.), *New Guinea Area Languages and Language Study Vol 2: Austronesian Languages* (Pacific Linguistics:

Series C 39), 223-231. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Hoogland, J. 1940. Memorie van Overgave van de Onderafdeling Hollandia. Nationaal Archief, Den Haag, Ministerie van Koloniën: Kantoor Bevolkingszaken Nieuw-Guinea te Hollandia: Rapportenarchief, 1950-1962, nummer toegang 2.10.25, inventarisnummer 24.

Hooley, Bruce A. 1968. SIL Research in New Guinea. *Kivung* 1(2). 63–70.

Hooley, Bruce A. 1976. History of research in Austronesian Languages: Morobe Province. In Stephen A. Wurm (ed.), *New Guinea Area Languages and Language Study Vol 2: Austronesian Languages* (Pacific Linguistics: Series C 39), 115-128. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Juillerat, Bernard. 1993. *La révocation des Tambaran: les Banaro et Richard Thurnwald revisités* (CNRS Ethnologie). Paris: CNRS.

Laufer, Carl. 1959. P. Futschers Aufzeichnungen über die Butam-Sprache (Neubritannien). *Anthropos* 54. 183–212.

Laycock, Donald C. 1975. A Hundred Years of Papuan Linguistic Research: Eastern New Guinea Area. In Stephen A. Wurm (ed.), *New Guinea Area Languages and Language Study Vol 1: Papuan Languages and the New Guinea linguistic scene* (Pacific Linguistics: Series C 38), 43-116. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Laycock, Donald C. 1976. History of Research in Austronesian Languages: Sepik Provinces. In Stephen A. Wurm (ed.), *New Guinea Area Languages and Language Study Vol 2: Austronesian Languages* (Pacific Linguistics: Series C 39), 73-93. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Laycock, Donald C. & C. L. Voorhoeve. 1971. History of Research in Papuan languages. In Thomas A. Sebeok (ed.), *Linguistics in Oceanea* (Current Trends in Linguistics 8), 509-540. Berlin: Mouton de Gruyter.

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

Levy, Catherine. 2002. A tentative description of Awar phonology and morphology: lower Ramu family, Papua-New Guinea. Université Libre de Bruxelles doctoral dissertation.

Lewis, Gilbert. 1975. *Knowledge of illness in a sepik society: a study of the Gnau, New Guinea.* London: Athlone Press.

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

Lichtenberk, Frantisek. 2008. *A grammar of Toqabaqita* (Mouton Grammar Library 42). Berlin: Mouton de Gruyter. 2 vols.

Lincoln, Peter C. 1976. History of Research in Austronesian Languages: Bougainville Province. In Stephen A. Wurm (ed.), *New Guinea Area Languages and Language Study Vol 2: Austronesian Languages* (Pacific Linguistics: Series C 39), 197-222. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Lithgow, David. 1976. History of Research in Austronesian Languages: Milne Bay Province. In Stephen A. Wurm (ed.), *New Guinea Area Languages and Language Study Vol 2: Austronesian Languages* (Pacific Linguistics: Series C 39), 157-170. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Lynch, John & Terry Crowley. 2001. *Languages of Vanuatu: A New Survey and Bibliography* (Pacific Linguistics 517). Canberra: Research School of Pacific and Asian Studies, Australian National University.

Marsden, William. 1834. On the Polynesian, or East-Insular languages. In *Miscellaneous Works*, 65-65. London: Parbury, Allen and Co.

Moore, Denny. 2007. Endangered languages of lowland tropical South America. In Matthias Brenzinger (ed.), *Language Diversity Endangered* (Trends in Linguistics: Studies and Monographs), 29-58. Berlin: Mouton de Gruyter.

Anonymous (Max Moszkowski). 1913. Wörterverzeichnisse von Papua-Sprachen aus holländisch-Neuguinea. *Anthropos* VIII. 254–259.

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

Nekitel, Otto. 1985. Sociolinguistic Aspects of Abu', a Papuan Language of the Sepik Area, Papua New Guinea. Canberra: Australian National University doctoral dissertation.

Ray, Sidney H. 1893. The Languages of the New Hebrides. *Journal and Proceedings of the Royal Society of New South Wales* XXVII. 101–167, 469–470.

Ray, Sidney H. 1912. Notes on Languages in the East of Netherlands New Guinea. In A. F. R. Wollaston (ed.), *Pygmies and Papuans: the Stone Age to-day in Dutch New Guinea*, 322-345. London: Smith, Elder & Co.

Ray, Sidney H. 1913-1914. The Languages of the Papuan Gulf District, Papua. *Zeitschrift für Kolonialsprachen* IV. 20–67.

Ray, Sidney H. 1919. The Languages of Northern Papua. *Journal of the Royal Anthropological Institute of Great Britain and Ireland* 49. 317–341.

Ray, Sidney H. 1919/1920. The Polynesian Languages in Melanesia. *Anthropos* 14/15. 46–96.

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

Ray, Sidney H. 1926. *A Comparative Study of the Melanesian Island Languages.* Cambridge University Press.

Ray, Sidney H. 1929. The Languages of the Central Division of Papua. *Journal of the Royal Anthropological Institute of Great Britain and Ireland* 59. 65–96.

Ray, Sidney H. 1937/1938b. The Languages of the Eastern Louisiade Archipelago. *Bulletin of the School of Oriental and African Studies* IX(2). 363–384.

Ray, Sidney H. 1938a. The Languages of the Eastern and South-Eastern Division of Papua. *Journal of the Royal Anthropological Institute of Great Britain and Ireland* 68. 153–208.

Reesink, Ger P. 2002. Mansim, a lost Language of the Bird's Head. In Ger P. Reesink (ed.), *Languages of the Eastern Bird's Head* (Pacific Linguistics 524), 277-340. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Remijsen, A. C. L. 2002. Word-Prosodic Systems of Raja Ampat Languages. Rijksuniversiteit te Leiden doctoral dissertation.

Robidé van der Aa, Pieter Jan Baptist Carel. 1885. Reizen van D. F. van Braam Morris naar de noordkust van Nederlandsch Nieuw-Guinea eerste vaart op de Amberno- of Rochussen-Rivier. *Bijdragen tot de Taal-, Land- en Volkenkunde van Nederlandsch-Indië, 4e volg., Deel X* 34. 73–114.

Schütz, Albert J. 1972. *The languages of Fiji.* Oxford: Clarendon Press.

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

Smits, Leo & C. L. Voorhoeve. 1992a. *The J. C. Anceaux collection of wordlists of Irian Jaya languages A: Austronesian languages (Part I)* (Irian Jaya Source Material No. 4 Series B 1). Leiden-Jakarta: DSALCUL/IRIS.

Smits, Leo & C. L. Voorhoeve. 1992b. *The J. C. Anceaux collection of wordlists of Irian Jaya languages A: Austronesian languages (Part II)* (Irian Jaya Source Material No. 5 Series B 2). Leiden-Jakarta: DSALCUL/IRIS.

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

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

Stebbins, Tonya N. 2010. The Papuan languages of the Eastern Bismarcks: migration, origins and connections. In Bethwyn Evans (ed.), *Discovering history through language: Papers in honour of Malcolm Ross* (Pacific Linguistics 605), 223-243. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Sumbuk, Kenneth Memson. 1999. Morphosyntax of Sare. University of Waikato doctoral dissertation.

Taylor, A. J. 1976. History of Research in Austronesian Languages: Western Part of South-Eastern Mainland Papua. In Stephen A. Wurm (ed.), *New Guinea Area Languages and Language Study Vol 2: Austronesian Languages* (Pacific Linguistics: Series C 39), 141-155. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Tryon, Darrell T. & B. D. Hackman. 1983. *Solomon Islands Languages: An Internal Classification* (Pacific Linguistics: Series C 72). Canberra: Research School of Pacific and Asian Studies, Australian National University.

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

Voorhoeve, Bert. 2000. Een moeilijk begin: Zending, missie en taalwetenschap in Nieuw-Guinea. In Willem van der Molen & Bernard Arps (eds.), *Woord en schrift in de Oost: de betekenis van zending en missie voor de studie van taal en literatuur in Zuidoost-Azië* (Semaian 19), 184-199. Leiden: Universiteit Leiden.

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

Voorhoeve, C. L. 1975b. A Hundred Years of Papuan Linguistic Research: Western New Guinea Area. In Stephen A. Wurm (ed.), *New Guinea Area Languages and Language Study Vol 1: Papuan Languages and the New Guinea linguistic scene* (Pacific Linguistics: Series C 38), 117-142. Canberra: Research School of Pacific and Asian Studies, Australian National University.

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

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

Wilkes, J. R. Adams. 1926. Appendix B: Vocabulary of native languages. *Australian Report on the Administration of New Guinea* 1924-1925. 75–78.

Wurm, Stefan. 1954. Tonal Languages in New Guinea and the Adjacent Islands. *Anthropos* 49(3/4). 697–702.

Z'graggen, John A. 1975. *The Languages of the Madang Districty, Papua New Guinea* (Pacific Linguistics: Series B 41). Canberra: Research School of Pacific and Asian Studies, Australian National University.

Z'graggen, John A. 1976. History of Research in Austronesian Languages: Madang Province. In Stephen A. Wurm (ed.), *New Guinea Area Languages and Language Study Vol 2: Austronesian Languages* (Pacific Linguistics: Series C 39), 95-114. Canberra: Research School of Pacific and Asian Studies, Australian National University.