State-of-the-art 
in the documentation of the Papuan 
languages of Timor, Alor, 
Pantar, and Kisar 
A bibliography

ANTOINETTE SCHAPPER AND JULIETTE HUBER

ABSTRACT
The Timor-Alor-Pantar (TAP) language family has a special place in South-East 
Asian linguistics; its members make up the western extreme of the Papuan 
language sphere. Along with an exhaustive bibliography of works on the 
TAP languages, this paper presents a state-of-the-art review of the ongoing 
documentation of the TAP language family in terms of both linguistic description 
and (pre-)historical reconstruction. The paper concludes with a consideration 
of the prospects for future studies of the TAP languages.

KEYWORDS
Documentation, Papuan languages, Timor, Alor, Pantar, Kisar.

1 INTRODUCTION

At the eastern end of the Minor Sundic Island chain, there is a small Papuan 
language family known as the Timor-Alor-Pantar (TAP) family (Map 1). TAP

1 Schapper’s research was supported by grant 08-EuroBABEL-OP-025 from De 
Nederlandse Organisatie voor Wetenschappelijk Onderzoek. Schapper thanks Emilie T.B. 
Wellfelt who encouraged her to undertake an experiment, and gave much valuable information 
as to culture and history.

ANTOINETTE SCHAPPER took her PhD degree at the Australian National University with 
a grammar of the Papuan language Bunaq spoken in central Timor. She is currently a post- 
doctoral researcher in the Alor-Pantar project at Leiden University, working on several Papuan 
languages of central Alor. Antoinette Schapper may be contacted at: a_schapper@hotmail.com.

JULIETTE HUBER took her PhD degree at Leiden University with a grammar of the Papuan 
language Makalero spoken in East Timor. She is currently a post-doctoral researcher in the 
Language, Cognition and Landscape (LACOLA) project at Lund University, working on the 
Papuan languages of the eastern tip of Timor. Juliette Huber may be contacted at: schueliet@ 
yahoo.com.

© 2012 Faculty of Humanities, University of Indonesia
languages are so-called Papuan “outliers”; far from the New Guinea mainland, they occupy a small corner in a large region which is dominated by languages of the Austronesian family.

The TAP family encompasses around 30 languages. Members of the family predominate on the islands of the Alor archipelago, defined here as consisting of Alor, Pantar and numerous small islands in the Pantar Strait. Other members are found interspersed with Austronesian languages on the islands of Timor and Kisar.

Five TAP languages are spoken in total on Timor and Kisar (Map 2). Bunaq is spoken in Timor by around 80,000 people straddling the border between Indonesian West Timor and independent East Timor. The other three TAP languages on Timor are spoken in a continuous region at the island’s far eastern end. Fataluku has about 30,000 speakers and Makasae about 70,000 speakers. Makalero, with only 6,000 speakers, has only relatively recently been recognized as a language distinct from Makasae. Finally, Oirata, a late off-shoot of a northern Fataluku dialect, is spoken by around 1,000 speakers on the southern side of Kisar.

Roughly, 25 TAP languages (Map 3) are spread across the Alor archipelago. Language communities here are much smaller than on Timor. The largest languages are Abui, spoken on Alor, with around 15,000 speakers and Western Pantar, spoken on Pantar, with around 10,000 speakers. Some of the other languages have up to 6,000 speakers (for example, Kamang), but many have only a few hundred (such as, Hamap, and Reta).

In this paper, we present a state-of-the-art review on the progress of the documentation of the TAP language family. Our treatment is structured as follows. In section 2, we orient the reader by situating the TAP languages within the larger “Papuan” picture; we highlight the special place that the TAP languages have as the western-most Papuan “outliers”. In section 3 and 4, we treat the progress of work on the TAP languages. Section 3 gives an overview of previous and ongoing descriptive materials produced on TAP languages.
Section 4 summarizes the historical linguistic work that has been carried out with respect to TAP languages. In section 5, we consider the prospects for further research into TAP languages, and what particular problems still remain to be explained and elucidated. Finally, we provide an extensive bibliography of materials documenting and treating the TAP languages.²

² Note that the “General references” section of this paper only contains references to general works mentioned in the main text that do not pertain to the Timor-Alor-Pantar languages. These are not mentioned in the Bibliography.
Before moving on to our discussion of TAP languages, we need to clarify what we understand the word “documentation” to mean. This definitional digression is necessitated by the fact that the last fifteen years have seen the creation of a new sub-discipline within linguistics called “language documentation” which is seen as distinct from “language description”. The standard distinction is that of Himmelmann (1998: 166):

The aim of language documentation is to provide a comprehensive record of the linguistic practices characteristic of a given speech community ... This ... differs fundamentally from ... language description [which] aims at the record of a language ... as a system of abstract elements, constructions, and rules.

That is, Himmelmann sees language documentation as having to do with the collection and archiving of primary data including communicative events in the speech community and elicited and experimental data of different kinds. On the basis of such documentary collections, Himmelmann (2012) argues that the generalizations of linguistic descriptions can be made replicable and accountable.

Whilst this distinction is perhaps of some academic, definitional importance, the reality is that, when it comes to endangered, minority languages, such as languages of the TAP family, every scrap of data with or without analysis adds to our knowledge and understanding of that language. As such, in this paper, we use the term “documentation” not in the specific technical sense current in linguistics, but in its most general sense. That is, we use “documentation” as a cover term to mean any materials on a language, be they grammatical descriptions, text collections, literacy manuals or audio recordings.

2 What are Papuan languages?
When talking about TAP languages, we have repeatedly mentioned that they are “Papuan” languages. Amongst Papuanists, the following represent two typical definitions of a “Papuan” language:

[T]he Papuan languages occupy those areas of New Guinea and adjacent islands not claimed by Austronesian languages (Foley 1986: 1).

[T]he term “Papuan” is a convenient term for the non-Austronesian languages of Papua New Guinea and eastern Indonesia, not all of which are demonstrably related (Tryon 1995: 3).

The languages identified as “Papuan” by such definitions are presented in their regional context in Map 4.

The prominent components of the above definitions are: (i) genealogical, “Papuan” languages are not members of the Austronesian language family, and (ii) geographic, “Papuan” languages occur within the area of the island of New Guinea. The significant emphasis that is, in much of the literature,
placed on the first point, namely, that “Papuan” languages are not part of the Austronesian family, has given rise to “non-Austronesian” as an alternative label to “Papuan”. Yet, the term is misleading; it does not carry with it the geographic restriction to the area of New Guinea which is so crucial to Papuanness. Austronesian languages are in fact in contact with members of multiple other (non-Papuan) language families, including Australian, Austro-Asiatic, Bantu, Tai-Kadai, and Sino-Tibetan.

Map 4. The distribution of Papuan languages.

The downside of the term “Papuan” is that the unified, non-negative label has led to a homogenizing view of Papuan languages, especially amongst non-Papuans. The homogenization has taken place on two levels. Firstly, the genealogical diversity of Papuan languages has been insufficiently recognized. This is the unfortunate side-effect of early “lumping” classifications. The most prominent are Greenberg’s (1971) Indo-Pacific Phylum (encompassing all Papuan languages, the Andamanese languages, and the Tasmanian languages) and the Trans-New Guinea Phylum (consisting of the majority of Papuan languages and running mainly along the highlands of New Guinea). More realistic estimates, such as that of Foley (1986: 3), put Papuan at around 60 families (including isolates), that is around 15 percent of modern language families. A proper recognition of Papuan genealogical diversity, therefore, has significant repercussions for typologists, since the ability to make statistically valid generalizations about the variation in and the frequency of (co-)occurrence of particular structures in the world’s languages hinges to a great extent on our having genealogically balanced samples (Bakker 2010). If Papuan languages do indeed constitute such a large number of families, this
would obviously have a significant statistical impact on typologies (Nichols 1999: 34, 263) and potentially radically alter the view of what is considered “typical” in language.

Secondly, the typology of Papuan languages has been seen as relatively homogeneous (see, for example, Comrie and Cysouw 2006 for a homogenizing view of Papuan languages). The inadequate documentation of the spread of Papuan languages has meant that typologies have been based on relatively few languages. They are concentrated in two regions within the wider Papuan area: (i) languages of the Eastern Highlands and middle Sepik regions of the New Guinea mainland, and (ii) Papuan languages of the Bird’s Head in far west New Guinea. There is still much descriptive work to be done before typologies will be truly able to sample a representative range of Papuan types.

In this paper, we are concerned with a particularly special family of Papuan languages. Languages of the TAP family lie far west of New Guinea, well removed from any other Papuan languages (Map 4). This position at the Papuan perimeter means that TAP languages and their speakers have had a very different history to that of languages on mainland New Guinea. For one, they have closely intermingled with speakers of Austronesian languages for around 2,000 years. It is perhaps unsurprising to learn that TAP languages present a significantly different grammatical profile to that of other Papuan language families (Schapper n.d.). This conclusion has only been able to be reached very recently due to the efforts of several teams of linguists working on TAP languages over the last decade. In the following two sections, we aim to provide the reader with an impression of how rapidly the documentation of TAP languages has progressed in recent years.

3 History of investigation and state of description

The presence of Papuan languages at the eastern end of the Minor Sundic Island chain was first brought to light by De Josselin de Jong’s (1937) description of the Oirata language on Kisar. Earlier reports had taken note of the distinctiveness of the local populations, but focused on racial and cultural characteristics. After De Josselin de Jong’s initial insight, gradually more languages of both Timor as well as Alor-Pantar were identified as Papuan (for instance, Nicolaspeyer (1940) on Abui, and Capell (1943, 1944a, b, c) on Bunaq and Makasae). In the following sections, we provide an overview of the language documentation work that has been completed since these earliest identifications of the Papuan status of the TAP languages.

3.1 Timor languages

Linguistic work on the languages of Timor was kicked off by Capell in WWII refugee camps in Australia, resulting in a number of articles (Capell 1944a, b, c, 1972). It then continued in the period between 1953 and 1975 by António de Almeida, head of the Portuguese Missão Antropológica de Timor, who collected word lists and elicited sentences in most languages of the
Portuguese part of the island. Today these materials are held at the Instituto de Investigação Científica Tropical (IICT) in Lisbon and are partly published in Almeida (1994).

The first more detailed treatment of an individual TAP language, other than Oirata, is found in the pioneering work of Louis Berthe on Bunaq in what is today Indonesian West Timor (Berthe 1959, 1963, 1972). Furthermore, during the period between 1973 and 1999, a series of anthropological papers on a wide variety of aspects of Bunaq culture was published by Claudine Friedberg.

Fataluku was first studied by Alfonso Nacher, a priest at the Missão Salesiana in Fuiloro, who compiled a sizeable dictionary of the language in the period between 1955 and 1968 (published as Nacher 2003, 2004). The first in-depth grammar description of Fataluku followed some years later (Lameiras-Campagnolo and Campagnolo 1979).

Very little work on the languages of Timor appeared during the era of the Indonesian occupation of East Timor (1975-1999): The Indonesian Pusat Pembinaan dan Pengembangan Bahasa (recently called Badan Bahasa, Language Centre) produced a general survey on the languages of Timor (Sudiartha et al. 1994), a sketch grammar of Bunaq (Sawardo et al. 1996) and one on Makasae (Sudiartha et al. 1998). Apart from these error-ridden materials, the only work to appear during this period was by Timorese expatriates, namely a Portuguese-language learner grammar of Makasae (Marques 1990) and a study of the Makasae affinal alliance system submitted as an MA thesis at the University of Melbourne (Guterres 1997).³

It was not until the beginning of the 21st century that new materials on the Papuan languages of Timor began to emerge. Initially, basic data on a number of the languages appeared in Hajek and Tilman (2001) and Hull (2004b), but more comprehensive documentation followed soon after for all languages.

A reference grammar of the Lamaknen dialect of Bunaq (Schapper 2010) was submitted as a doctoral dissertation at the Australian National University. Since then, Schapper has written and co-authored a number of academic papers on that Bunaq dialect, which is spoken in Indonesian West Timor, both descriptive (2011a; Schapper and San Roque 2011) and historical (Schapper 2011d, e; Klamer and Schapper Forthcoming) in nature. Furthermore, a range of school readers is currently being produced in East Timor to support Bunaq literacy (for instance, Fundasaun Alola 2011). This, along with the older work listed above, probably makes Bunaq the best documented of the Papuan languages of Timor. Nevertheless, much descriptive work remains to be done on its many, often highly divergent, dialects spoken in East Timor. Schapper (Forthcoming c) represents a first step towards this goal.

Since the East Timorese independence, Makasae has garnered attention from a range of corners. Following a field methods course on the Ossu dialect, two descriptive theses on specific language domains came out of the

³ In 2007, the author of this work was appointed Plenipotentiary Ambassador Extraordinaire to the Holy See for the Democratic Republic of Timor-Leste. At present he appears to be serving a ten-year prison sentence for embezzling state money.
Australian National University; the first of these (Brotherson 2003) treats Makasae deixis and space, while the other (Carr 2004) discusses a variety of traditional Makasae genres. Around 2005, two dialects of Makasae were studied at the Language Documentation Training Centre at the University of Hawaii, resulting in mini-sketches, word lists and some text material appearing online. The language has also been the subject of a detailed phonological study (Fogaça 2011), a short dictionary (Hull and Correira 2006), and two grammar sketches, one on the Baucau dialect (Hull 2005a) and another on the Ossu dialect (Huber 2005, later published as Huber 2008a). In 2002, the Timor Loro Sa’e-Nippon Culture Center brought out a short Makasae-language text on local traditions (Ximenes and Menezes 2002). However, to the present day, there is no comprehensive reference grammar of Makasae, even though, in terms of speakers, it is one of the main languages of the country. Also, an investigation into Makasae dialects remains to be done.

Previously considered one of several dialects of Makasae, Makalero has been given the status of a separate language by the East Timorese Instituto Nacional de Linguística (see, for instance, Hull 2004a). Despite its brief history as a language of its own, Makalero has been the subject of a reference grammar, which was submitted as a PhD dissertation at Leiden University (Huber 2011). The same author has also published a short paper on agreement morphology (Huber 2008b) and given a number of talks on various grammatical properties of Makalero (see, for instance, Huber 2010). More descriptive materials (Huber Forthcoming b, c), as well as a comparative grammar sketch of Makalero and Makasae (Huber Forthcoming a) are underway. Furthermore, the Timor Loro Sa’e Nippon Culture Center has overseen the production of a number of materials by native speakers, among them a monolingual Makalero dictionary (Pinto 2004), a Makalero language course with Tetum and English translations (Pinto 2007) and a text on the local traditions (Seixas 2004) in Makalero, with translations into Tetum and Portuguese.

Fataluku is the subject of extensive and ongoing development and documentation driven by an engaged speaker community (see Langford 2011). Local materials produced with the support of the ubiquitous Timor Loro Sa’e-Nippon Culture Center include a monolingual dictionary (Valentim 2002), lists of words and phrases (Valentim 2001b, 2004a) and short texts in and on Fataluku (Valentim 2001a, 2004b). Furthermore, several school readers are being produced with the help of involved native speakers (for instance, Langford 2012). Between 2005 and 2008, an academic documentation project on Fataluku funded by the Endangered Languages Programme of De Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) was run.

4 The materials on Makasae Fatumaka are online at: http://www.ling.hawaii.edu/ldtc/languages/makasae_fatum/. Those on Makasae Osoroa can be retrieved at: http://www.ling.hawaii.edu/ldtc/languages/makasae_osor/.
5 This text is available online at: http://www.geocities.jp/hkbts/hakerek/Hakerek06/ginigni.htm.
out of the University of Leiden. This project resulted in several descriptive papers on the eastern dialect (for instance, Stoel 2007a, b; Van Engelenhoven 2009, 2010a, b), and a dictionary online at the project’s webpage. Also, one of the documentation projects conducted at the University of Hawaii’s Language Documentation Training Centre was on Fataluku, and a sketch grammar is being prepared (Van Engelenhoven Forthcoming). Despite the relative wealth of materials, much descriptive work remains to be done, particularly with respect to the different Fataluku dialects.

In summary, the last decade has seen dramatic advances in our knowledge and understanding of the character and composition of the Papuan languages on Timor, even though much still remains to be done in terms of documentation of all speech domains and dialects.

3.2 Kisar language

A single Papuan language is spoken on the island of Kisar. Oirata, as the language and village where it is spoken are known, is the result of a late migration from Timor. According to Dutch historical records, the Oirata arrived in Kisar in 1721, having fled from Loikera in the Fataluku-speaking area (Riedel 1886: 403). Still today the northern Fataluku dialect is claimed to be, at least partly, mutually intelligible with Oirata (Katrina Langford personal communication), though other Fataluku dialects appear to be too divergent (Aone van Engelenhoven personal communication).

Oirata was the subject of a monumental monograph by De Josselin de Jong (1937); it was in this work that the presence of Papuan languages at the eastern end of the Minor Sundic Island chain was first brought to light. Most subsequent treatments and discussions of Oirata have all been based on this pioneering work. For instance, Cowan (1965) argued for the inclusion of Oirata, along with Bunaq and Makasae, into the West Papuan phylum; and Donohue and Brown (1999) used Oirata data to demonstrate that it is possible to find ergative syntax in a language with wholly accusative morphology, in contradiction to earlier claims about a correlation between syntactic ergativity and ergative morphology. Faust (2005) provided interlinearization of De Josselin de Jong’s (1937) text, and on this basis, was able to describe several grammatical patterns not been mentioned in the original treatment by De Josselin de Jong. Most recently, Mandala (2010, expanding on initial work presented in Mandala 2003), undertook to examine the relationship between Oirata and the Eastern Timor languages (Fataluku and Makasae, to the exclusion of Makalero) based on lexical data gathered by himself in Kisar and Timor. Both works are in Indonesian. Mandala argues for a Fataluku-Oirata subgroup and a Makasae subgroup in Eastern Timor and reconstructs 172 proto-forms (in which, however, are included numerous loan words). His findings are summarized in English in Mandala et al. (2011).

---

7 See www.fataluku.com.
8 The resulting mini-sketches, word list and some text material are online at: http://www.ling.hawaii.edu/ldtc/languages/fataluku/.
3.3 Alor Languages

A long silence followed the initial anthropological and linguistic work on Abui. Not until the 1970s did our knowledge of Alor languages progress again with the work of Wim Stokhof. Stokhof (1975) represents the first range of lexical materials on Alor languages; it contains seventeen word lists taken from a range of locations across the island, as well as some rudimentary presentation of their morphology. Several further publications by Stokhof expanded our knowledge of specific Alor languages. Collectively, Stokhof (1976, 1977, 1978, 1979, 1982, 1983) represent the first significant body of materials on any Alor language, in this case, Kamang, albeit under the name Woisika. These works encompass ethnographic description, preliminary phonological analysis, collections of texts and riddles and a school reader. Stokhof also published short individual texts in Abui (Stokhof 1984) and Kabola (Stokhof 1987).

The next linguists to venture to Alor did not arrive until the mid-1990s. Mark Donohue (personal communication) conducted several weeks fieldwork on Alor; he collected information on Kui in west Alor, and on Kula and Wersing in east Alor. Results from this work have been circulated in one published article on Kula (under the name Tanglapui; Donohue 1996), a talk handout on the different alignment systems he encountered in the three Alor languages named above (Donohue 1997, some of which is reproduced in Donohue 2008: 66-68), and a manuscript on Wersing phonology (under the name of Kolana; Donohue n.d.). Kula is also the subject of Haan and Johnston (n.d.) and Johnston (n.d.). In 1999, Doug Marmion (personal communication) conducted a scouting trip to Alor in preparation for a PhD project. He collected word lists and basic elicitation sentences for Klon and Kui. Unfortunately, Marmion did not pursue doctoral research on Alor due to the volatile situation created by sectarian conflict across much of the eastern Indonesia region.

In the early 2000s, there was a surge of Indonesian work on Alor languages: The Pusat Pembinaan dan Pengembangan Bahasa produced survey word lists of the languages of Alor (Martis et al. 2000). This was followed up by the first full-length grammar of an Alor language: In 2001 John Haan, a native speaker of Adang, completed his PhD grammar of his language at the University of Sydney (Haan 2001).

In early 2003, Louise Baird and František Kratochvíl arrived in Alor as part a NWO project hosted at Leiden University. Over the next three years, Baird conducted extensive surveying in west Alor, focussing in particular on Klon (Baird 2005, 2008, 2010) and Kafoa (Baird n.d.). Kratochvíl’s fieldwork was part of a PhD project centring on the Abui language spoken in north-central Alor. His work has thus far resulted in a grammar (Kratochvíl 2007), a dictionary (Kratochvíl and Delpada 2008a), a book of stories (Kratochvíl

---

9 This name is not used here as it is recognized neither by the Kamang speakers themselves nor by other Alorese ethnolinguistic groups as a name for a language or people. ‘Woisika’ means ‘Stone Stack’ (< *woi* ‘stone’, *sika* ‘stack’) in Kamang and refers to a village within the Kamang area. Today, Woisika is a desa officially known as ‘Waisika’ due to an early spelling mistake on the part of Indonesian administrators from Java.
and Delpada 2008b) as well as several articles (Klamer and Kratochvíl 2006, 2010; Kratochvíl 2011a, b). Most recently, Benny Delpada, an Abui speaker trained in linguistics by Kratochvíl, completed an impressive survey of some 400 expressions for emotions and cognitive processes involving body part nouns (Delpada 2012).

Kratochvíl’s work documenting Abui continues today (personal communication) and has since 2008 been supplemented by a research project on Sawila in far eastern Alor. Whilst no materials on Sawila are yet published, a number of talks given by Kratochvíl (for instance, Kratochvíl et al. 2011) contain information on Sawila. In addition, draft manuscripts of a Sawila grammar sketch (Kratochvíl Forthcoming) and a short dictionary (Kratochvíl n.d.) are in circulation.10

In 2009, Asako Shiohara followed up Marmion’s and Baird’s work with a survey of coastal Kui (some results reported on in Shiohara 2010), but also did not continue to more in-depth study. However, as we write this, a student from the University of Sydney is preparing to undertake a first fieldtrip to Kui with a view to a larger project.

In 2010, three more linguists began work on Alor in connection with the research project “Alor-Pantar Languages; Origins and Theoretical Impacts”.11 Nicholas Williams began a PhD project looking at the encoding space and place in the Kula language of east Alor (previously treated under the name Tanglapui in Donohue 1997). Laura Robinson began work with John Haan and other speakers of Adang to extend our knowledge of that language (results in Robinson and Haan Forthcoming). Antoinette Schapper began building on Stokhof’s early work on Kamang with a dictionary (Schapper and Manimau 2011) and a short grammar (Schapper Forthcoming b). She also collected new data on Abui and has produced several detailed treatments of topics in Kamang and Abui grammar among other languages (for example, Fedden et al. Forthcoming; Schapper 2011a, Forthcoming a, c; Schapper and San Roque 2011; Schapper and Klamer 2011). Furthermore, she has recently taken up work on Wersing (Schapper and Hendery Forthcoming). Previous materials for this language include a reader produced by a native speaker to support the local content curriculum (Langko n.d.), as well as a translation of the gospel of Mark (Malaikosa n.d.).

3.4 Pantar and Pantar Strait languages
The early work of Wim Stokhof in Alor was paralleled by Hein Steinhauer on Pantar and in the Pantar Straits (as reported on in Steinhauer and Stokhof 1976). Throughout the 1970s, Steinhauer conducted extensive fieldwork on the Blagar

10 Kratochvíl had been collaborating significantly with Anderias Malaikosa, an Alor native fluent in Sawila, Wersing, and Kula. The sudden death of Anderias in Kalabahi in late 2011 was a great loss. He was a talented linguist and a committed champion of Alor languages and cultures.

11 The project is part of the programme “Better Analyses Based on Endangered Languages“ (EuroBABEL) run by the European Science Foundation. The project website is: www.alor-pantar.org.
language; most of his time was spent working on the Dolabang dialect of Pura Island, but Steinhauer also collected materials on the Bukalabang dialect of Blagar on Pantar (described in Steinhauer 1995). It has taken several decades for details of Steinhauer’s work on Blagar to emerge. In addition to a few published academic papers on Blagar (Steinhauer 1977, 1991, 1993, 1999, 2010, 2012), we now have available: a sketch grammar (Steinhauer Forthcoming), a sizeable dictionary (Steinhauer n.d.), and a school reader (Gomang et al. 1976). The next work to appear on a language of Pantar was a rather unsatisfactory sketch of Lamma produced by Pusat Pembinaan dan Pengembangan Bahasa (Nitbani et al. 2001). This was, luckily, quickly followed up by more careful work on the part of Gary Holton from the University of Alaska using a pan-dialectal name for the language, Western Pantar. Holton’s documentation of Western Pantar is ongoing (Holton 2004), but has already resulted in a few descriptive papers (Holton 2010, 2011), an introductory dictionary (Holton and Lamma Koly 2008) and a brief sketch (Holton Forthcoming).

At around the same time as Holton, Marian Klamer of Leiden University began work on two languages of Pantar, Teiwa and Kaera. Teiwa was already known by name from a list in Stokhof (1975), but Kaera had not yet been discovered and still today has no ISO 639-1 code. Klamer’s productivity has been prodigious. On Teiwa, a grammar (Klamer 2010a) and extended word list (Klamer 2011) have been published, along with a suite of papers (Klamer 2010b, c, 2012a; Klamer and Kratochvíl 2006). On Kaera, descriptive work is still in progress, but a sketch grammar is already imminent (Klamer Forthcoming a).

4 HISTORICAL AND PREHISTORICAL STUDIES

Apart from the descriptive materials already discussed, TAP languages have featured extensively in (pre)historical studies of the insular eastern Indonesian region and Timor-Leste. The high degree of (pre)historical interest in them is largely a function of their special place as the western-most Papuan outlier.

Among the questions that have been debated are the following: How and when did the TAP languages get to be in their current outlier position? Are they the remnant of the earliest human migrations into the region or a later back migration from New Guinea? Are all TAP languages related? If so, what is the internal structure of the TAP family?

In the following sections, we review the discussion of such questions in the literature dealing with TAP and their role in the (pre)history of their region.

4.1 HIGH-LEVEL AFFILIATIONS

Despite the – until recent – lack of in-depth materials for TAP languages, speculation about their high-level affiliations was rampant almost from the time of their first mention in print. The question remains unresolved and is still a matter of debate today.

Capell (1944c) was the first to note similarities between the non-Austronesian languages of Timor and those of the only other Papuan outlier in the eastern Indonesian archipelago, spoken in North Halmahera. He did
not, however, assert any genetic relationship. It was twenty years later that
this was attempted by Cowan (1963), who argued that Bunaq and the other
Timor-Kisar languages were related to the languages of the Bird’s Head of
New Guinea. Anceaux (1973) suggested the inclusion of several Alor languages
in Cowan’s (1957) West Papuan group. Stokhof (1975) suggested yet another
relationship upon comparing some Alor-Pantar languages to languages of the
South Bird’s Head of New Guinea.

On the basis of typological similarities, Capell (1975) proposed that the TAP
languages should be grouped with the West Papuan phylum languages of the
Bird’s Head and Bomberai peninsular. However, the absence of identifiable
lexical correspondences forced Capell to argue for a major split between the
“Alor-Timor languages” and the rest of the West Papuan Phylum.

In the same volume as Capell (1975), Voorhoeve (1975) made an argument
for the inclusion of the languages of the southern Bird’s Head and Bomberai
peninsular belonged to the Trans-New-Guinea (TNG) phylum. The corollary
of this was that all Papuan languages in the Timor area must also be TNG.
This claim was made explicit by Wurm, Voorhoeve, and McElhanon (1975),
albeit with the caveat that “whichever way they are classified, they contain
strong substratum elements of the other … phyla involved” (1975: 318). This
cautions note is repeated by Pawley (2005: 73).

In recent years, several authors have supported a link with Bomberai
languages. Hull (2004b) suggested a genetic relationship between the TAP
languages and the languages of the Bomberai peninsula on the basis of
some rather tenuous lexical similarities. Ross (2005) used a comparison of
independent pronoun forms as evidence to propose a classification of the
Timor-Kisar languages with Bomberai languages, making them part of a large
‘western linkage’ within TNG.

The classification of the TAP languages as TNG has since been taken as a
given in much of the literature, even though none of the proposals in question
is backed by the adequate methodological rigour. Only Pawley (2001, 2011,
2012) provides 98 Proto-Trans-New-Guinea (PTNG) reconstructions; of these,
nine are given with reflexes in TAP languages. The connection of TAP to TNG
forms seems to be based on impressionistic similarities, rather than any sound
correspondences. Indeed, in his to-date most elaborated reconstruction of
PTNG, Pawley (2012) makes no attempt to identify or posit any regular sound
changes that might have given rise to these TAP forms. Holton et al. (2012),
based on a cursory comparison between their reconstruction of Proto-Alor-
Pantar (PAP) and Pawley’s PTNG forms conclude that, even though evidence
from pronouns and typology may suggest a connection between Alor-Pantar
and TNG, the lexical data do not.

Most recently, Donohue (2008) has revived the possibility of a connection
between West Papuan and TAP. He suggests a history of contact between a
West Papuan source language and a pre-TAP language spoken in the Bomberai
peninsula or in the South Bird’s Head area. His proposal raises the possibility
that the linguistic prehistory of the Alor-Pantar languages may evade a simple
cladistic classification.
While these proposals assume that the modern TAP languages are the result of a back migration from New Guinea, some researchers have advocated a different view. For instance, Dunn and Reesink (2007: 1) discuss a scenario according to which today’s speakers of TAP languages could be stay-behind descendants of a “migration of speakers of a TNG precursor”. This would perhaps account for the parallels that have been observed between TAP and TNG languages, whilst at the same time explaining the radically different lexicons. However, it is ultimately little but conjecture; moreover, it is conjecture for which there is never likely to be proof available one way or another.

In short, claims relating the TAP languages to other Papuan language families on the New Guinea mainland remain highly speculative. At this stage, a cursory inspection of the limited data claimed to support various wider relationships is sufficient to raise doubts in even the most well-disposed reader.

4.2 Low-level affiliations

The literature dealing with the high-level classification of the TAP languages has typically assumed that the TAP languages were all related to one another. This assumption was in some ways natural in that the languages all occupied a relatively restricted geographic area and had obvious parallels in their clause structure that clearly set them apart from the surrounding Austronesian languages. Yet, their exact relationships have remained to be demonstrated through the rigorous application of the comparative method.

A range of interrelationships has been proposed. Stokhof (1975: 24) argued for a closer affinity between Makasae and the Alor-Pantar group. In the 2000s, the emergence of materials on Makasae (see Section 3.1), however, showed that Makasae could be clearly grouped within the other languages of Eastern Timor, Makalero, Fataluku, and Oirata. Ross (2005), based on his large-scale comparison of free pronoun forms in TNG languages, groups Bunaq with Alor-Pantar, as opposed to an East Timor group. This is contradicted by Schapper (2010), who argues that morphological evidence from the Timor languages suggests that the Papuan languages of Timor and Kisar, including Bunaq, may be more closely related to one another than to the Alor-Pantar languages (Schapper 2010: 21, 346).

Only now are we seeing the emergence of results of reconstructive work on the TAP languages. The groundwork was recently laid in a preliminary paper of Holton et al. (2012). Using the comparative method, they show that the Alor-Pantar (AP) languages are all related to one another. They reconstruct a set of 97 Proto-AP lexemes and identify a few sound changes which point to some rudimentary subgrouping possibilities.

This work on Alor-Pantar has been paralleled by Schapper, Huber, and Van Engelenhoven’s (2012) reconstruction of the Timor-Kisar languages. In this, the authors identify cognate vocabulary with regular sound correspondences across the five Papuan languages of Timor and Kisar, concluding that they are all also related. 93 proto-lexemes could be reconstructed and several
low-level subgroups are firmly identified on the basis of shared innovations. Most significantly, an Eastern Timor subgroup made up by the four eastern languages, Makasae, Makalero, Fataluku, and Oirata, is recognized, with Bunaq being on its own in a distinct subgroup.

While both of these works clearly demonstrate the relatedness of the Alor-Pantar languages and the Timor-Kisar languages, respectively, the relation of these two groups to one another still awaits final confirmation. A set of cognates and preliminary consonant correspondences are identified in Schapper, Van Engelenhoven, and Huber (2012). A preliminary formulation of the relatedness of the TAP languages using the comparative method is underway and proceeding apace (Schapper, Huber, and Van Engelenhoven Forthcoming).

5 Prospects for future research

We now turn to a brief consideration of future directions for research into TAP languages.

First and foremost for the documentary linguist is the preservative imperative. That is, there is an urgent need to document the most fragile of the TAP languages before their loss. Language endangerment exists most acutely in Alor and Pantar. Whilst some languages like Abui and Western Pantar remain vital, many others, such as Kamang and Teiwa, languish; their extinction is probable in the next two generations. For others, such as Nedebang in northern Pantar, the war has already been lost; only a few semi-competent, elderly speakers remain. As already discussed above, grammatical descriptions (albeit in some cases just brief sketches) are forthcoming for the following languages: Blagar, Fataluku, Kaera, Kamang, Wersing, and Western Pantar (Schapper Forthcoming d).

Whilst the appearance of these sketches will add substantially to our knowledge of the TAP languages, there are other, as yet undocumented, languages which appear to be of considerable interest. For instance, preliminary data from Kui indicate that it is rather divergent from other languages of the family (see, for instance, its subgrouping: Holton et al. 2012, and numeral formation: Klamer and Schapper 2012). From an academic linguist’s perspective, it would be useful to have information on languages like Kui, so we can develop a fuller picture of the whole expanse of grammatical variation in the TAP languages. Similarly, we know very little of the full extent of dialect variation in the languages of Timor. For instance, there appear to be several intermediate language varieties between Makasae and Makalero. Yet we do not know whether they are different enough from both to be considered separate languages or not. Comparable questions exist for Bunaq and Fataluku varieties.

The unusually heterogeneous morphological profiles of the TAP languages (as described in Schapper Forthcoming d) also warrant investigation. TAP languages range from extremely isolating in word structure (for instance, Makasae) to morphologically complex (for instance, Kamang). The
reconstructive evidence points to their common ancestor, proto-TAP, as having been close to isolating. The question is then: How did the morphological differences between these languages develop? And why have some languages reduced proto-TAP morphology, whilst others show extensive innovative morphologicalization? This raises challenging questions about time-depth and rates of historical change in relation to lexical complexity and the factors that drive morphological change.

There are also significant questions of wider position and history of TAP languages in the region. In particular, the relationship between TAP languages and Papuan languages of West Bomberai deserves serious attention following more complete reconstructions of the two proto-languages. A comparison of the reconstructions will test the TNG hypothesis at its lowest level and bring some serious historical perspective –most importantly the application of the comparative method– to the validity of TNG at its claimed fringes.

The dynamics of cultural and linguistic interactions between speakers of TAP and Austronesian languages that have resulted in the current distribution of TAP languages have only recently begun to be explored. For example, it is still uncertain whether TAP groups pre- or post-dated initial Austronesian settlement in the Timor region (see, for instance, O’Connor 2007 and McWilliam 2007a). Careful comparative analysis of linguistic forms and structures, place names, material culture, and anthropological enquiry would go a long way to working out the relative timing of the TAP arrival.

Should the settlement of the TAP languages in the Timor area have taken place before the Austronesian arrival (as argued by Schapper 2011e), we are left with a puzzle: how did the speakers of TAP languages resist the incoming tide of Austronesians? The conventional understanding of Austronesian history involves the expansion of the agriculturalist Austronesians across the eastern Indonesian archipelago, overwhelming pre-existing sparse populations of Papuan hunter-gatherers and transforming them into speakers of Austronesian languages (Bellwood 1998). Yet the TAP languages survive today, and we may speculate that their speech communities possessed some characteristics that allowed them to maintain themselves in the face of the Austronesian arrival. Until recently, we had no clue as to what this advantage might have been. However, new excavations in Timor provide unambiguous evidence of fishhook manufacture and pelagic fishing at 42,000 BP (O’Connor, Ono, and Clarkson 2011). These finds show that the early modern human inhabitants of Timor carried out deep-sea maritime expeditions requiring a high level of organisation, planning and skill. A community with such technology and skill may well have been able to stand firm against the Austronesian wave. However, in order to make this a convincing argument for the TAP resistance it will be necessary to show that there was technological continuity between the earliest humans capturing pelagic fish on Timor and the TAP speech communities presumed to have been present in Timor at the Austronesian

12 Reconstructive work on West Bomberai languages is currently being carried out by Mark Donohue at The Australian National University.
arrival some 39,000 years later. We await more archaeological research to resolve this question.

Conclusions
In this paper, we have outlined the existing and ongoing documentation of Papuan languages belonging to the TAP family. We have highlighted the great advances that have been made in our knowledge of these languages in the past ten years. At the same time, we draw attention to the many gaps that still remain and impede our full appraisal of the TAP languages; indeed, the recent discovery of Kaera (Klamer Forthcoming a) shows that gaps may still exist at the most basic level. We conclude by identifying areas where further research is particularly needed. On the one hand, there is the need for basic description and documentation: the acute endangerment of several of the small languages makes such efforts imperative. On the other hand, this enigmatic group of languages presents a variety of diachronic and historical conundrums. These range from the development of wildly different morphological profiles to their origin, high-level affiliations to other Papuan language groups and the interaction and nexus with neighbouring Austronesian languages in the area. We stress the importance of findings from other academic fields such as archaeology and anthropology in enriching our understanding of the character and history of TAP languages and their speakers.

General references
Comrie, Bernard and Michael Cysouw. 2006. “New Guinea through the eyes of WALS”. [Unpublished manuscript; Online: cysouwNEWGUINEAtext.pdf.]

**Bibliography of the Timor-Alor-Pantar languages**

The bibliography is divided into five sections: (1) early works, (2) language documentation and description: monographs, (3) language documentation and description articles, papers, and non-academic material, (4) historical and areal linguistics and prehistory, and (5) anthropology, ethnography, and history.

Early works consists of works dating from the colonial period up to its end (in Portuguese Timor) in 1975. These come from a variety of academic fields, including linguistics, ethnography and anthropology. All following sections contain post-colonial materials only.

The second section, linguistic documentation- monographs, includes descriptive academic works on individual languages. All materials in this section are of a length of at least 100 pages and represent substantial contributions to TAP linguistics.

The third section, language documentation and description- articles, papers and non-academic material, contains descriptive and analytic material on individual TAP languages or sub-sets of TAP languages, as well as general survey work. The term “documentation” is understood in a broad sense, and a variety of materials of a non-academic nature are also listed here, including school readers, dictionaries and word lists, course books, stories and texts in TAP languages. Many of these materials are produced by or with the help of native speakers and include a wealth of information relevant to the descriptive linguist.

The fourth section, historical and areal linguistics and prehistory, contains a variety of papers that are concerned with diachronic aspects of individual languages and language groups and comparative work aiming at the establishment of the low-level affiliations of the TAP languages. Also, it lists academic work that places TAP in a wider perspective, both in terms of phylogenetics and areal linguistics.

The fifth section, archaeology, anthropology, ethnography and history, includes academic works of a primarily non-linguistic purpose. Several of them include a wealth of language material and are thus of interest to the study of the TAP languages.

**Early works (until 1975)**


Nicolspeyer, Martha Margaretha. 1940. *De sociale structuur van een Aloereesche bevolkingsgroep*. Rijswijk: Kramers.


**Language documentation and description-monographs (grammars and dictionaries)**


**Language documentation and description–articles, papers, and non-academic material**


Donohue, Mark. n.d. “Kolana; Syllables, (small) words and CV”. [Manuscript, Australian National University, Canberra.]


Fedden, Sebastian, Dunstan Brown, Greville Corbett, Gary Holton, Marian Klamer, Laura C. Robinson, and Antoinette Schapper. Forthcoming. “Conditions on pronominal marking in the Alor-Pantar languages”.


Huber, Juliette. Forthcoming b. “Landscape in East Timor Papuan”.

Huber, Juliette. Forthcoming c. “Places and non-places; Evidence from Makalero.”


Malikosa, Anderias. n.d. ”Yesus Sakku Geleworo Kana”. [Unpublished manuscript, UBB-GMIT Kupang, Indonesia.]


Pinto, Tomas Mariano. 2007. Ma’u Fi Makalero Lolo. Dili: Timor Loro Sa’e-Nippon Culture Center.


Schapper, Antoinette (ed.). Forthcoming d. Papuan languages of Timor-Alor-Pantar; Sketch grammars Volume I and II. Asia-Pacific Linguistics, Mouton de Gruyter.


Shiohara, Asako. 2010. “Verb types and semantic roles of person markers of Kui, Alor”. Talk handout, Workshop on the Languages of Papua 2, Melanesian Languages on the Edge of Asia; Past Present and Future, Universitas Negeri Papua, 8-12 February.


Steinhauer, Hein. 1995. “Two varieties of the Blagar language (Alor, Indonesia)”, in: Connie Baak, Mary Bakker, and Dick van der Meij (eds), Tales from a concave world; Liber amicorum Bert Voorhoeve, pp. 269-296. Leiden: Projects Division, Department of Languages and Cultures of South-East Asia and Oceania.


Steinhauer, Hein. 2010. “Pura when we were younger than today”, in: Artem Fedorchuk and Svetlana Chlenova (eds), Studia Anthropologica; A festschrift in honour of Michael Chlenov/Сборник статей в честь М.Ф.Членова, pp. 261-283. Москва/Иерусалим: Мосты Культуры/Gesharim.


*(Pre-)Historical and areal linguistics*


Klamer, Marian and Antoinette Schapper. Forthcoming. “The history of ‘give’ constructions in the Papuan languages of Timor-Alor-Pantar”.


Pawley, Andrew. 2005. “The chequered career of the Trans New Guinea hypothesis; Recent research and its implications”, in: Andrew Pawley,


Schapper, Antoinette, Juliette Huber, and Aone van Engelenhoven. Forthcoming. “The relatedness of Timor-Kisar and Alor-Pantar languages; A preliminary demonstration”, in: Marian Klamer (ed.), *Alor-Pantar
languages; History, lexicon and typology. Leiden: Brill.


Archaeology, Anthropology, Ethnography, and History


(4): 548-562. [Special Issue; C. Barraud and J.D.M. Platenkamp (eds), Rituals and socio-cosmic order in eastern Indonesian societies, Part 1: Nusa Tenggara Timur.]


O’Connor, Sue, Rintaro Ono and Chris Clarkson. 2011. “Pelagic fishing at 42,000 years before the present and the maritime skills of modern humans”, *Science* 334: 1117. [Http://www.sciencemag.org/content/334/6059/1117.full.pdf.]


Sousa, Lúcio. 2010a. “This is the beginning of the relationship’; Material


Wellfelt, Emilie. Forthcoming a. “The anthropologist as heroine – contemporary interpretations of memory and heritage in an Indonesian valley”, in: Alistair Thomson et al. (eds), Historical fragments in Southeast Asia; At the interfaces of oral history, memory and heritage.

Wellfelt, Emilie. Forthcoming b. “Heritage in Alor; Sustaining local identity in a globalized world”, in: Susan Legêne and Henk Schulte-Nordholt (eds), Sites, bodies and stories; Indonesia, India, and (post)colonial heritage formation.