

# The dispersal of Austronesian languages in Island South East Asia: Current findings and debates

Marian Klamer 

Centre for Linguistics, Faculty of Humanities, Leiden University, Leiden, The Netherlands

## Correspondence

Marian Klamer, Leiden University Centre for Linguistics (LUCL), Faculty of Humanities, Leiden University, Van Wijkplaats 4, Leiden 2311 BX, The Netherlands.

Email: m.a.f.klamer@hum.leidenuniv.nl

## Abstract

This paper reviews the “standard” view of the Austronesian language family tree in connection with the archeological “farming/language dispersal” hypothesis of Neolithic populations moving into Island South East Asia (ISEA) and beyond. It focuses on what is currently known about the dispersal history of the ~650 languages spoken in ISEA (Malaysia, the Philippines, Indonesia, and Timor-Leste) that belong to the Malayo-Polynesian branch of Austronesian and points out where the topology of the MP branch is agreed upon and where it is contested. The conclusion is that historical linguistics is currently not in the position to provide information about higher order temporal and spatial relations between speaker groups within ISEA, unlike that which the language/farming dispersal hypothesis suggests. It also reviews some claims that can be heard in support of this hypothesis and concludes (i) that the expansion of MP languages into ISEA was less monolithic than often suggested, but rather that their lexical and structural diversity suggests multiple migrations of different groups, in many different directions, at different points in time; (ii) that the history of MP languages very likely involved long-term, intense contact in multilingual communities where

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2019 The Authors Language and Linguistics Compass Published by John Wiley & Sons Ltd

newcomers and autochthonous people lived together for centuries if not millennia; (iii) that the original populations of Island SE Asia were not (only) hunter-gatherers but had sea-faring groups and agriculturalists among them; and (iv) that the histories reflected in languages, archeological findings and human genetics do not always converge. Simple macro-level models like the standard Austronesian tree and the farming/language dispersal hypotheses are unable to catch the linguistic history of ISEA, with its complex geography, human networks, and migrations. To deepen our understanding of this history, the focus is currently shifting from macro-level hypotheses to more detailed bottom-up investigations of regional MP language groups and their speakers.

## 1 | INTRODUCTION: THE HOMELAND AND DISPERSAL OF AUSTRONESIAN SPEAKERS

The Austronesian family comprises some 1,200 languages which cover half of the globe (Figure 1). Similarities in their vocabularies suggest that they descended from one common ancestor language proto-Austronesian (proto-AN).

The spread of a single language family across such a large area of islands raises intriguing questions about the nature of the dispersal of the speakers of these languages. Island South East Asia (ISEA) as a region covers over 4,000 kilometers west–east, 3,000 km north–south, comprises some 15,000 islands of all kinds and sizes, wet and fertile or dry and barren. It has a history of volcanic eruptions, earthquakes, tsunamis, floods and draughts obliterating populations, setting them on the move, and attracting new waves of settlers.

Speakers of languages prior to Austronesian had arrived in the region from SE Asia about 50,000 years ago (Bellwood, 2017, 86–87). How did the Austronesian languages become so



**FIGURE 1** Spread of Austronesian languages in the world

dominant across Island South East Asia (ISEA) and Melanesia, while in ISEA most of the languages of the previous inhabitants disappeared?

The classic model of the expansion of Austronesian language speakers, referred to as the “Out of Taiwan” model or the “Early farming dispersal hypothesis” (Bellwood, 1997, 2005, 2011), which will be referred to here as the “language/farming dispersal” hypothesis, has addressed questions like these. This model was originally proposed to account for the pioneering settlement of the southwest Pacific. This area started to be colonized by people around 3,000 years ago, and the archeological sites from this period contain a material cultural “package” referred to as Lapita. The package contains plain or red-slipped pottery in a range of simple forms, which sometimes bear decorations, and also polished stone adzes, shell artifacts, tattooing chisels, fishhooks, bark cloth beaters, and stone net sinkers (Bellwood, 1997, 219–30, 2002, 26). As all languages of remote Oceania belong to a single branch of the Austronesian family, it was hypothesized that the Lapita culture represented an Austronesian intrusion into Oceania.

Originally, a hypothesis about the origin of people and the approx. 500 Austronesian languages in Oceania, the language/farming dispersal hypothesis, was subsequently connected to the linguistic observations that Austronesian languages developed in Taiwan (Blust, 1999). The linguistic evidence for this Austronesian homeland is extensive, and most Austronesianists today agree that proto-Austronesian was indeed spoken in Taiwan some 5,500 years ago (Bellwood, 1985; Blust, 1995) where it diverged over time into a number of major subgroups. Most of these subgroups remained in Taiwan, while one of them, proto-Malayo-Polynesian (MP), moved out of the homeland (Blust, 1977, 1978).

By correlating the reconstructed vocabulary of proto-MP (e.g., terms for pottery and canoes) with material evidence in the archeological record (Bellwood, 1985), the expansion of speakers of proto-MP from Taiwan has been dated (cf., Bellwood, 1997, Figure 1) to have started some 4,500 years ago, in the Neolithic, the period from the innovation of agriculture to the introduction of metal tools. In Bellwood's “early farming” model, a change in economy in Taiwan took place around 4,500–4,000 years ago and triggered a demographic increase which led to a demic migration of MP speakers across Island SE Asia and Oceania. According to the model, the migration was enabled by the agricultural technologies of these people, who were growing primarily rice and millet, and brought along domesticated animals (pigs, dogs) as well as the Lapita package of material cultural traits mentioned above. With their new food-producing technologies, the MP migrators were able to colonize and replace the preexisting hunter–gatherer (forager) populations of ISEA, bringing along their cultural assemblages. From Taiwan, Neolithic migrants moved across the Batanes Strait into Luzon, northern Philippines between 4,500–3,500 BP. From there, one dispersal went into Borneo, Sulawesi, and towards Java, Sumatra, and Peninsular Malaysia between 3,500–2,500 years ago. Another migration of MP-speaking groups took place from the northern Philippines southeast into the Moluccas and Lesser Sundas, where they arrived in the Timor area around 3,500 years ago (Pawley, 2005). Additional movements of people speaking MP languages between 4,000–3,500 years ago, went via Halmahera in the north Moluccas to the east, passing the north coasts of New Guinea, to the Bismarck Archipelago, where at ~3,300 BP the Lapita Cultural Complex appeared (Bellwood, 1997, 2005; Bellwood, 2017; Diamond, 2001; Posth et al., 2018). In this paper, I focus on the dispersal of the ~650 Austronesian languages of ISEA, which are spoken in Malaysia, the Philippines, Indonesia, and Timor-Leste. By way of summary, Table 1 lists the approximate dates for the expansion of Austronesian settlement that are relevant for the discussion of language dispersal in ISEA. The movements into Oceania started from the Bismarck Archipelago and are beyond the scope of this paper.

**TABLE 1** Approximate dates for the expansion of Austronesian settlement relevant for ISEA

5,500 BP	Proto-Austronesian spoken in Taiwan
4,500–4,000 BP	Change in economy in Taiwan triggers demographic increase and migration of proto-MP speakers into ISEA
4,500–3,500 BP	Proto-MP speakers move across the Batanes Strait into Luzon, northern Philippines
3,500–2,500 BP	(Descendants of) proto-MP speakers move into Borneo, Sulawesi, Java, Sumatra, Peninsular Malaysia
~3,500 BP	Arrival of (descendants of) proto-MP migrations in the Timor area
4,500–3,500 BP	(Descendants of) proto-MP speakers migrate via north Moluccas and north coast of Papua New Guinea to Bismarck Archipelago
~3,300 BP	Lapita culture complex appears in Bismarck Archipelago

This paper reviews what is currently known about the patterns of language dispersal in the region of ISEA and that which is under debate. In Section 1, I review the literature on MP subgrouping and evaluate the use of the language tree as a model for the dispersal of MP languages in ISEA, pointing out which parts are (relatively) uncontested as representations of linguistic divergences and which parts are less certain or unclear. For the ~650 MP languages spoken in ISEA in particular, there is much uncertainty about how their micro-groupings connect to each other into the macro subgroupings of the Austronesian family tree. I argue that this uncertainty is partly due to the limited and narrow data sets used in earlier classifications and partly to the fact that in Austronesian comparative studies; the focus has long been on studying how the languages can be united rather than how they vary.

In Section 2, I review linguistic work which has shifted the attention from lexical conservatism to studying lexical residues, borrowings, and grammatical diversities. Rather than unity, the latter types of work suggest much diversity, both in the historical trajectories of individual languages and in the mutual relationships between language groups. The diversity is so complex that it is unlikely to be captured in a single model. I also revisit some popular ideas about the Austronesian languages which can be heard in connection with the farming/language dispersal but lack general support in the field.

The first idea is the claim that MP languages are relatively monolithic and homogeneous, especially when compared to the “substantial heterogeneity” of Austronesian languages within Taiwan (Bellwood, 2017, 188), so that they have “a common core” (Pereltsvaig, 2012, 146) where they share much vocabulary and many grammatical features.

The second idea is that migrations of MP speakers involved settlers (or traders or warriors, e.g., Comrie, Matthews, & Polinsky, 2003: 93) who overtook and/or dominated the original populations of hunter–gatherer or foraging population in ISEA by their technological superiority. This was achieved by their technological superiority by way of explaining why the expansion of the MP languages was so monolithic and hardly left any traces of the languages of previous populations.

Finally, in Section 3, I present a brief look beyond linguistics, referring to work in the archeology and molecular anthropology of ISEA. Studies in archeology call for caution in connecting the material evidence in ISEA directly to the dispersal of people and languages; and from studies in human genetics, we learn about the complexities of the evolutionary processes in ISEA and the current inconclusiveness of the data with regard to directions of migrations.

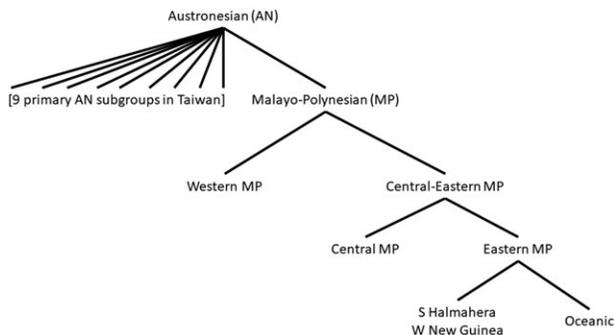
In fact, it would be odd to expect that a simple macro-level model like the language/farming dispersal hypothesis can account for the spread of languages, people, and cultures in the vast, ecologically unstable insular region of ISEA. In this region, knowledge about past language dispersals starts with bottom-up micro-level investigations of regional languages and language groups.

## 2 | THE UNCERTAIN TOPOLOGY OF THE MALAYO-POLYNESIAN TREE

All of the Austronesian languages spoken in ISEA belong to the Malayo-Polynesian (MP) branch. If the dispersal of the MP languages is assumed to represent the dispersal of speakers with their material culture, as in the language/farming dispersal hypothesis, then the peopling of ISEA is expected to be reflected in the structure of the MP subgroup. The “standard theory” of MP branching is presented in Figure 2:

The “standard” tree features in overviews of language families such as Ruhlen (1987), *Ethnologue* (Simons & Fennig, 2018), and *Glottolog* (Hammarström, Forkel, & Haspelmath, n.d.).<sup>1</sup> Sources like these are consulted by scholars who are not specialists in Austronesian historical linguistics; for instance, by archeologists who use the spread of the ISEA Neolithic as a proxy for Austronesian expansion (cf., Spriggs, 2011, 511), or by linguistic typologists who draw from existing linguistic classifications to create genetically balanced language samples or by descriptive linguists who describe a language within the context of the subgroup it belongs to. The standard tree has also entered the common reservoir of academic knowledge through textbooks such as Bellwood and Hiscock (2005; 281), Pereltsvaig (2012; 148), and Campbell (2013; 179).

Since the 19th century, historical linguists have been aware of the limitations of representing a language family as a tree with successive bifurcations. It suggests a history where different languages arise through in-group differentiation: There is a single ancestral group, which separates in space and/or in time and then develops different dialects which (given enough distance in space and time) become different languages. The lines branching from nodes schematically represent those divergences that led to discrete and irreversible splits. But it has been often highlighted, also in Austronesian linguistics, that this is an idealization of how languages diverge, as noted by Andrew Pawley:



**FIGURE 2** The “standard theory” of Malayo-Polynesian branching (Blust, 2009, 37; 2014, Table 7)

*“The family tree model of linguistic relationships has proved far too valuable to discard, but the seductive simplicity of its images and terminology at times makes us glib and lazy in our analysis of linguistic relationships. We should be alive to its limitations and should seek to refine and supplement it.” (Pawley, 1999, 129)*

Among linguists, there is little doubt that the proto-Austronesian homeland is in Taiwan (Blust, 1999). Taiwan is favored as the Austronesian homeland as the proto-AN reconstructed flora and fauna vocabulary locates the homeland to the west of the Wallace line, and Taiwan is the region with most primary subgroups of Austronesian. The dispersal of the Malayo-Polynesian languages started with proto-MP (Blust, 1977; Dahl, 1973). Much has been reconstructed at the level of proto-MP: Apart from a segment inventory and a vocabulary of several thousands of words (Blust & Trussell, n.d.), we also know about the pronominal paradigms (Reid, 2009), verbal paradigms, clausal syntax (Ross, 2002), and derivational morphology (Blust, 2013).

It is also an established fact that languages descending from proto-MP are spoken throughout ISEA, and that proto-MP was ancestral to proto-Oceanic from which all the Oceanic languages in Melanesia, Micronesia, and Polynesia sprang. Given the geographical locations of Taiwan, ISEA, and Polynesia, there is thus no doubt that the dispersal of MP languages took place in the direction north to south and west to east from Taiwan, across ISEA to Polynesia—and not the other way round.

In an ideal world, the branches in the MP family tree should represent “interstage” languages that are reconstructible as single proto-languages, with recursive branching matching the progress of MP languages through ISEA. For the MP tree in Figure 2, those branches applying to languages in ISEA are not supported in this way. This does not mean that we do not know anything about the affiliations of languages in this area. It has been perfectly possible to establish lower level subgroups, where the languages covering part of an island, or some adjacent islands, form a clearly motivated subgroup that derives from MP. For example, the MP subgroups discussed in Adelaar (2005) include six subgroups located on Borneo, three on Sumatra, seven in Sulawesi, and one spreading from Sumatra, via Java, Madura, Bali, and Lombok to Sumbawa (Adelaar 2005, 15–16). That is, there is good evidence that in the north-western part of Indonesia, MP has more than a dozen first-order branches. However, to date, there is no evidence of defining innovations that would connect these lower level subgroups *to each other* and allow the reconstruction of a higher level interstage language such as “Western MP.” It is very clear from the literature over the last decades (Ross, 1995; Tryon, 1995, 25) that Western MP is not considered a reconstructed language and the languages comprised by it<sup>2</sup> do not form a single subgroup but many distinct ones.

Furthermore, the suggestion of the tree in Figure 2 that Central-Eastern MP (CEMP) is an innovation-defined subgroup has not generally been accepted among linguists. Blust (1982, 1983–1984, 1993) has argued for its unity, but Adelaar (2005) notes that one morphosyntactic piece of evidence for CEMP, the innovation of using proclitic subject markers on the verb, is not reflected in cognate forms so could also be the result of a convergent development. Moreover, as Blust admits, subject proclitics are also found in Sulawesi languages, as well as in Barrier Island languages, so that the phenomenon of proclitic subject marking is not unique to CEMP. The second innovation defining CEMP, the morphological distinction between alienable and inalienable possession, is considered as stronger evidence by Blust. However, this distinction is not unique to CEMP either, and there are significant differences in possessive forms and constructions (Adelaar 2005, 25–26). Donohue and Grimes (2008) present evidence indicating that proto-CEMP did not exist as an interstage language because some of the phonological innovations on which it is based

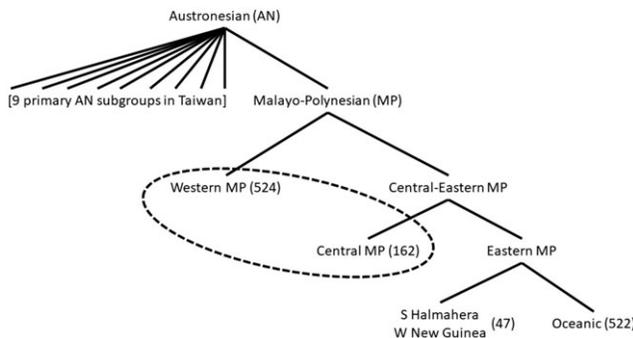
also occur in some western MP languages. Blust (2009) responded that unique innovations remain but also stresses that CEMP “poses some of the most complex and challenging subgrouping issues that are found in [Austronesian]” (Blust, 2009, 75).

Third, in Figure 2, the CEMP node splits into two groups: Central Malayo-Polynesian (CMP) languages and Eastern Malayo-Polynesian (EMP). For over 30 years, it has been known that for neither of these groups can a single ancestor language be established (Ross, 1995; Adelaar 2005; Ross, 2008). The partly overlapping distribution of various innovations weakens the argument for a CMP subgroup (Ross, 1995, 82). Instead, Blust (1993) suggests that CMP languages descended from a dialect chain or network that would necessarily have been hundreds of miles long and evolved when MP languages spread through eastern Indonesia very rapidly. However, the evaluation of this hypothesis requires more bottom-up reconstruction work on a larger amount of geographically balanced lexical materials than has hitherto been used (on the scarcity of data, see also below).

The two remaining subgroups are South Halmahera-West New Guinea (SHWNG) and Oceanic. Both of these subgroups are defined by a clear set of sound changes (for a recent proposal, see Kamholz, 2014). Oceanic is the most clearly defined of all Austronesian subgroups, having sound changes accompanied by other kinds of innovations (Ross, 1995). However, within the Oceanic subgroup, the structure of the tree is also very rake-like, with nine first-order branches (Lynch, Ross, & Crowley, 2002; Ross, 2017), but as Oceania is beyond the region of ISEA it will not be further discussed here.

Instead of the standard tree in Figure 2, the tree that is more commonly accepted in the field is the one in Figure 3, where a distinction is made between actually reconstructable proto-languages and names for groups of languages that do not derive from a single proto-language; the latter are shown in italics in Figure 3. For example, Pawley (Pawley, 2007, 21) refers to the “Western Malayo-Polynesian language groups” and the “Central Malayo-Polynesian linkage.” “Linkage” is used to refer to a grouping for which no proto-language can be reconstructed. In Figure 3, the branches within the dotted circle are located in ISEA; the South Halmahera West New Guinea languages are partly located on the Papuan mainland.

In other words, the historical reconstruction data available at present do not allow us to say that proto-MP branched out in a few daughter languages (such as “Western MP,” “Central Eastern MP” or “Central MP”) from which the lower subgroupings of languages in ISEA



**FIGURE 3** More commonly accepted Malayo-Polynesian branching where branches in italics represent a linkage, not a proto-language (Pawley, 2007, 21). Numbers in brackets are the number of languages that go into the grouping according to *Glottolog* (Hammarström et al., n.d.).

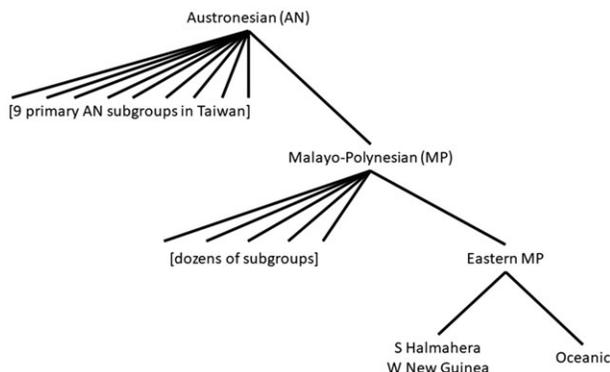
Note that *Glottolog* does not recognize Western MP, instead it has 25 subgroups of MP. It does recognize CEMP, CMP, EMP, SHWNG, and Oceanic

derived. What we know at the moment is that under the proto-MP node there exist dozens of hierarchically unordered clades whose history cannot be modeled with this tree. This may be represented in a rake-like family tree as in Figure 4 (see Ross, 1995, 2005; Adelaar, 2005; Donohue & Grimes, 2008).

This observation is relevant for the language/farming dispersal hypothesis: as long as the lower-level language groupings of the Philippines, Malaysia, Borneo, Sulawesi, western and eastern Indonesia cannot be linked to each other at a higher level; historical linguistics is not in a position to provide information about the temporal and spatial relations between these groups. For example, we cannot say that the northern MP languages as spoken in the Philippines derive from “higher” or “older” branches of the family tree than the eastern MP groups spoken in the Lesser Sunda islands, unlike what is suggested by the trees in figures such as Figures 1 and 2. This is important to note, because it means that for more than half of the ~1200 Austronesian languages (524 languages in the western grouping and 162 languages in the central grouping (see Figure 3), historical linguists are uncertain about their higher order affiliations and, therefore, their temporal and spatial relations.

One reason for these uncertainties may be that migration hypotheses so far have been established on family wide (or “macro”) comparisons rather than micro-comparisons of languages that are geographically close (Kikusawa, 2015, 665). Many parts of ISEA are linguistically still very much underexplored. The data sets on which earlier classifications of languages in this region have been based are very limited and narrow. For example, of the 208 languages that *Glottolog* (Hammarström, Forkel, & Haspelmath, n.d.) lists as spoken in eastern Indonesia,<sup>3</sup> less than 50% (102) are represented in the major reference work on Austronesian reconstructions, the *Austronesian Comparative Dictionary* (*ACD*; Blust & Trussell, n.d.). Of those, about 50% (52) features in the database with less than 10 words, and 25 of them with only one (!) word. For the languages of western Indonesia, Sulawesi, Borneo, the Philippines, and Malaysia, the figures are only slightly better. *Glottolog* lists 524 languages spoken there, of which 412 are in the *ACD*. However, the lexical data available per language are equally limited: about 50% (221) of the north-western languages feature in the *ACD* with less than 10 words and some 22% (90) are present with just one word. Clearly, affiliations based on so little data collected from such a vast insular region comprising thousands of islands leave a lot of room for uncertainties.

More recently, (parts of) basic vocabulary word lists of many (but by no means all) MP languages have been made publically available through online lexical databases such as the *Austronesian Basic Vocabulary Database* (*ABVD*; Greenhill, Blust, & Gray, 2008) and



**FIGURE 4** Rake-like tree of Austronesian languages (Adelaar, 2005; Donohue & Grimes, 2008; Ross, 1995, 2005)

*LexiRumah* (Kaiping & Klamer, 2018).<sup>4</sup> However, traditional comparative work on such data sets remains limited, mostly because the available (funding of) manpower for this labor intensive work is scarce. Using the computer to assist in this comparison seems promising. Gray, Drummond, and Greenhill (2009) compared the basic vocabulary from 400 Austronesian languages in the *ABVD* to investigate the topology of the Austronesian tree with phylogenetic methods used in biology. Their methodology confirmed many individual lower level subgroups that had been traditionally identified and also confirmed the lack of support for the higher order groupings of CMP and WMP. In addition, the inferred sequence of the “pulses” between the start point in Taiwan and the end point in Oceania is so compressed that many groupings are presumably based on a weak signal that could also be compatible with a dispersal across Island SEA that was more “haphazard and network-like” (Heggarty 2015, 613).

Indeed, for the micro-groups that have been studied reasonably well (cf., Adelaar 2005), the absence of evidence for intermediate ancestral nodes connecting them suggests a complex dispersal scenario. There may have been just a few MP groups migrating into ISEA, but then each of these groups did not spend a sufficiently long period of time together as unified speech communities after their split from MP. Because if they had, they would have innovated words in their lexicon or developed different pronunciations which would have been inherited in their daughter languages and could now be used to define the subgroup as such by its unique lexical and phonetic innovations. There may also have been not a few but *many* different MP communities migrating at different times and in different directions throughout history. Overall, the lack of shared inherited features suggests that most of the MP languages that dispersed into ISEA shared a relatively short local micro-history. In contrast, the shared inherited features of, for example, proto-Oceanic, indicate that this group did remain a unified speech community for a period that was sufficiently long to create such defining innovations and pass them on to their daughter languages. In sum, there is no evidence that the dispersal of MP speakers in ISEA followed simple trajectories; instead, it is more likely that multidirectional processes of contact and language admixture have been at work for millennia.

### 3 | DIVERSITY IN LEXICON AND GRAMMAR

The previous section evaluated the evidence from historical linguistics regarding the dispersal of the languages of ISEA. Here, I discuss evidence from linguistic typology and contact linguistics showing that the multifaceted dispersal of the languages of ISEA is reflected in the diversity of lexical and grammatical forms and structures they exhibit.

In Austronesian comparative studies, the focus has long been on what unites the languages: Finding shared sound changes in cognate sets (reflexes of proto-forms), so that affiliations between languages could be established. This has led to the reconstruction of a proto-MP vocabulary of thousands of words (Blust & Trussell, n.d.). Some have suggested that MP languages show “remarkable conservatism,” as most of them are argued to have a retention rate of the proto-MP basic vocabulary of 30% or more (Donohue & Denham, 2010; 229). This idea is, however, not based on empirical evidence. Since the rejection of classic glottochronology (Bergsland & Vogt, 1962; Lees, 1953), there is no known constant rate of change for basic vocabulary in terms of elapsed time across different language families. Moreover, supposedly “constant” retention rates ignore the often observed fact that different cognate sets have

different rates of change. Certain (basic) vocabulary words are replaced more quickly than others; for example, numerals are typically quite stable, while words expressing certain particular activities (e.g., “to squeeze”) show more variation over time (Dyen, James, & Cole, 1967). Also, there is variation in retention of basic vocabulary not only between languages but also between families (Blust, 2000).

And, finally, any calculations of retention rates within families by definition overestimate retention, as the family relations themselves are already based on retained basic vocabulary (i.e., the words used to reconstruct proto-forms). That is, a modern language that has (almost) no lexical similarity with any other MP language would not be considered as potentially affiliated to the MP group in the first place. This would not only exclude all non-MP languages from the calculations but also those languages that were originally MP but went through a stage where their MP basic vocabulary was (largely) replaced. In other words, the potentially major witnesses to a high rate of replacement in MP languages would not be considered in the calculation of the overall MP retention rate.<sup>5</sup> In sum, as long as objective measurements of lexical conservatism across language families are lacking, we do not know whether MP languages are lexically conservative or not, and neither do we know whether there is anything remarkable about the outcome, given the time depth of the MP group.

An approach that may be more fruitful in tracing the history of languages in ISEA would be to shift the focus from studying sound changes in reflexes of the reconstructed ancestor language (i.e., in cognate sets) to (also) systematically investigating lexical innovations, residues, (taboo) replacements, and borrowings. The question to address would be what the patterns attested in these non-cognate lexical inventories suggest about the history and mutual relationships between languages and language groups. An interesting example of this type of approach is Edwards (2016), who describes the MP language Uab Meto<sup>6</sup> on Timor as having two parallel lexicons, each with their own set of *regular* sound correspondences: one containing reflexes of proto-MP lexemes, the other containing lexemes for which no MP origin has been found. The sheer size of the non-MP vocabulary (including basic vocabulary) of pre-Uab Meto, and the fact that it has restructured the phonological system of the language, points to a prolonged period of intense and intimate language contact between one or more incoming MP language(s) and one or more non-MP languages that were spoken in the region before their arrival. It is quite likely that future research will find similar witnesses of histories of bilingual contact between MP and non-MP languages in ISEA, but we first have to start looking for them.<sup>7</sup>

The complex dispersal history of the languages of ISEA is also reflected in the diversity of grammatical structures and forms they have. While the Oceanic languages share many typological characteristics (for a detailed overview, see Lynch et al., 2002; chapter 3), the group of MP languages of ISEA is so large and so variegated that few if any features characterize it as a whole.<sup>8</sup> Negative structural characteristics include the absence of tonal contrast, the scarcity of plural affixes on nouns,<sup>9</sup> and the absence of tense marking on verbs. Positive features present in most of the Austronesian family and shared among the MP languages of ISEA include reduplication, the distinction between inclusive and exclusive pronouns, and the usual presence of morphological causatives (Himmelmann, 2005; 110). At the same time, the languages (even very closely related ones) differ along many dimensions, such as:

- word order (verb-initial vs. verb-second or verb-final: “eat he yam” vs. “he eat yam” or “he yam eat”); possessors following or preceding the possessed noun (“house my” vs. “my house”); negators preceding or following the predicate (“not go” vs. “go not”);

- the expression of alienable vs. inalienable possession (so that my house may mark its possessor different than “my foot” in some languages but not others);
- the morphological expression of voice alternations (the famous many different “passive” voices in the languages of the Philippines are either simplified in various ways or completely absent in languages in Indonesia);
- the use of numeral classifiers (which are not used in the Philippines but frequently used everywhere else);
- the use of plural words (which are used in the Philippines and eastern Indonesia but absent in Malaysia and western Indonesia (Wu (2016)); and so on.

Austronesian languages show an enormous grammatical diversity, and it is a misconception that they have “many common grammatical features” (Pereltsvaig, 2012; 149–55). In the next two subsections, I discuss these differences from two perspectives: focusing on linguistic differences that grew out of adding certain features (3.1) and diversity created by the loss or simplification of features (3.2).

### 3.1 | Diversity through adding grammatical features

It has been observed for a very long time, by many different scholars (e.g., Grimes, 1991; Himmelmann, 2005; Klamer & Ewing, 2010; Klamer, Reesink, & van Staden, 2008; Reesink, 2002; Reesink & Dunn, 2017; Schapper, 2015) that some of the salient features of MP languages spoken in the eastern part of Indonesia and in the vicinity of New Guinea must be due to contact with non-MP (or Papuan)<sup>10</sup> languages. For example, the order “possessor precedes noun” is almost universal in Papuan languages and a major pattern in many MP languages of the Lesser Sundas, Central and South Moluccas, Halmahera and the Cenderawasih Bay, but not in languages of western ISEA.<sup>11</sup> So this suggests that it is not an inherited MP structure. It could have been the result of spontaneous independent developments in the languages of these different regions, but that would not explain why these developments frequently occurred in eastern and not in western languages. Since they are confined to the region where we know that Papuan languages are, or have been spoken it seems plausible that the structure was borrowed into MP languages from Papuan languages spoken in their vicinities. Other features that appear to have leaked from Papuan languages into MP languages include the use of a post-predicate negator instead of, or in addition to, a pre-predicate one (Reesink, 2002; Klamer et al., 2008, 130–34; Florey, 2010; Fricke, 2017); and making a formal distinction between nouns that are alienably or inalienably possessed (Ross, 2001, 138; Klamer et al., 2008, 116–122).<sup>12</sup>

When languages have atypical and additive grammatical features such as these and there are other indications that speakers may have been in contact with speakers of languages that possess these features (e.g., the presence of lexical borrowings), then the atypical grammatical features may be hypothesized to be remnants of contact with those other language(s). In order for contact between speakers of language A and B to lead to the addition of features in language B, it must be long-term, intense, and multipurpose. That is, language B should not just be used in circumscribed contexts such as trading, ritual events, or songs, but in a wide array of social domains (Trudgill, 2010, 304, 315; Ross, 2013). Social situations that lead to such additive changes in language B occur when it is spoken in bi/multilingual communities by adults as well as pre-adolescent children (cf., Ross, 2013).

### 3.2 | Diversity through loss or simplification of grammatical features

Regarding the simplification of the proto-MP verb morphology mentioned above, since Ross (2002) reconstructed proto-MP with a suit of tense, aspect, and mood morphemes as summarized in Table 2 below, it is possible to see to what extent this original rich system has been maintained in the MP languages in ISEA.

The verbal structure of proto-MP has been largely retained in the languages of the Philippines, Sabah, North Sulawesi, and Madagascar. However, in many Austronesian languages spoken in Malaysia and Indonesia, the proto-MP voice system is reduced to a simple opposition between actor and undergoer voice (Adelaar 2005, 6–8), or has been lost completely. A quick glance at the verbal morphologies of individual languages spoken in ISEA, for example, Javanese (Ogloblin, 2005, 600), Sasak (Wouk, 2002, 299), Kambera (Klamer, 1996, 1998), Rongga (Arka, 2016), and Kéo (Baird, 2002), already indicates that their verbal morphology is significantly simpler than that of proto-MP in Table 2.

Such morphological simplification may be due to independent, language-internal evolutionary processes. For example, one can imagine a scenario where a reflex of the proto-MP patient suffix *\*-ən* develops an allomorph *-n* for vowel-final roots, which subsequently becomes reanalyzed as a segment that is part of the root, analogous to other root-final consonants. However, morphological simplification can also be caused by language contact. Inflectional morphology in particular is known to be one of the most vulnerable areas of linguistic knowledge in contact situations, because it straddles the interface between syntax, semantics, and pragmatics (see, e.g., Montrul, 2004; 126). In known cases where language contact has led to loss of morphological complexity, it typically involves adults as second language learners who simplify non-native morphological structures, as for instance in Afrikaans (den Besten, 1989) and in adult second language Dutch (Blom, Polis̃enská, & Weerman, 2006). Morphological simplification through adult second language learners also happens in small-scale preindustrial societies in ISEA, though it has not yet been researched much; an example is Alorese (Moro, In press; Klamer, 2012), discussed below.

In order for the simplified patterns to stabilize, the contact must involve a community of bilinguals with a large number of second language speakers, and the contact must be long-term, intense, and multi-purpose (Moro, In press; Kusters, 2003; Trudgill, 2011). It may be that the simplifying second language was (originally) used as a trade language or lingua franca, but

**TABLE 2** Proto-MP voice, mood and aspect morphemes based on Ross (2002, 49) where the root *káRaw* “scratch” is used

Voice/Mood/Aspect	Argument role			
	Actor	Patient	Location	Circumstantial
Indicative				
Neutral	<um > R	R-ən	R-an	i-R
Perfective	<umin > R	<in > R	<in > R-an	i- <in > R
Imperfective	<um > RDP-R	RDP-R-ən	RDP-R-an	i-RDP-R
Non-Indicative				
Atemporal	R	R-a	R-i	R-an
Projective	R-a	(R-aw)	R-ay	-

Abbreviations: <...>, infix; R, verbal root; RDP, reduplicated initial syllable of root.

for any changes to become entrenched in it, it must have been used as a second language (L2) in wider communicative contexts. This second language may be the language of a technologically, politically, or culturally dominant group that the speakers of other languages wish to communicate or associate with.<sup>13</sup> However, it may also be the language of a community that is incorporating many foreign adults (such as spouses or slaves) with different linguistic backgrounds. A language that is spoken as an L2 can become a shifted language when the L2 speakers are a minority in the community and die out, while their offspring grows up speaking the community language as L1. However, if the number of L2 speakers in a community is sufficiently large, e.g., constituting half or more of the population, as in the case of Alorese (Moro, In press), and if there is a constant influx of new L2 speakers during many generations, then stable bilingual communities can exist for centuries without shifting to either of the languages. In other words, a community where a simplifying second language is spoken does not automatically shift as a whole to that language, losing the other (first) language(s). And even when a bilingual community does (eventually) shift completely to language B, that still does not imply that language A necessarily becomes extinct. Not all speakers of language A are necessarily part of the bilingual community shifting to language B; language A may still have its own monolingual community elsewhere and/or language A may be spoken in other bilingual communities together with languages C or D. In short, if language shift occurs, it is commonly preceded by a period of bilingualism which can continue for a very long time given the right social circumstances; and language shifts do not necessarily imply the death of the first language(s).

Most small-scale language groups in ISEA today are bi/multilingual with neighboring groups, e.g., because of marriage exchanges or cultural connections. Small-scale migrations of bands of people moving to different locations on the same or a neighboring island where other languages were already spoken have been taking place in historical times and are still taking place today. If situations that can still be observed today prevailed from the early Neolithic onwards, then there must have been thousands or even millions of different sociocultural and historical micro-contexts of contact between speaker groups in ISEA, all with their own time paths and effects. In fact, even a single language may have had different types and stages of contact, in different locations, as indicated by traces of different strata of contacts.

Klamer (2012) is a case study where the language ancestral to modern Lamoholot and Alorese (two closely related MP languages spoken in Lesser Sunda) have acquired a suite of typological features that are seen as typical of the Papuan languages of the region—including post-predicate negation, the marking of possessors, noun-locational order in locative constructions, the presence of a focus particle, and the absence of a passive verb form. This “Papuanisation” of proto-Lamoholot took place in the Flores-Lembata region, under conditions of long-term stable contact involving preadolescents acquiring the complexities of both Papuan and MP languages and melding them into a new system.

In a second phase, following the migration of pre-Alorese speakers to the island of Pantar and the separation this entailed from their Lamoholot cousins, a series of further changes occurred. Alorese contrasts drastically with Lamoholot in terms of morphological complexity. Where Lamoholot has two sets of subject affixes to the verb, Alorese relies on free pronouns with all but a few frequent verbs which retain fossilized agent prefixes. And where Lamoholot has a number of derivational affixes (some productive, some lexicalised), Alorese has no derivational morphology at all—reduplication is its only productive word formation process. These differences suggest a radical process of morphological simplification in the passage from Lamoholot to Alorese. Historical records indicate that initially there were only a few Alorese-speaking communities on Pantar, and that they were demographically small and geographically

scattered. In order to survive, they would have needed marriage as well as trade connections with the speakers of Papuan languages of the inland. The almost complete loss of morphology in Alorese is likely due to contacts with Papuan language speakers who used Alorese as a second language. This hypothesis is confirmed in an experimental study by (Moro, In press) investigating the use of subject agreement prefixes in six Alorese first language speakers and 12 Alorese second language speakers. The study shows that the second language speakers make significantly more errors than the first language speakers, and that they have only a single default subject agreement marker.

### 3.3 | Linguistic diversity and the dispersal scenario

Case studies like the ones discussed above illustrate a type of multiphase contact scenario likely to have been played out between groups of MP and non-MP language speakers in many parts of ISEA at different stages over the last millennia. Contact may lead to the addition of features or the loss of them. Such opposite outcomes of contact reflected in the linguistic and typological diversity in the region are a reminder of the social and linguistic complexity that must have existed between groups who would have been demographically equally stable and interdependent in many ways. The diverse features of modern MP languages suggest that numerous changes, mixes, and shifts have occurred throughout the history of ISEA, involving all imaginable kinds of contact situations and migrations, occurring over hundreds of generations. What they definitely do *not* suggest is rapid and total replacement of languages by a limited number of ancestor groups.

Yet the fact remains that the larger part of ISEA today is inhabited by speakers of MP languages, while there are far fewer traces of the preexisting non-MP languages. This has been explained by a scenario where the MP speakers dominated the earlier non-MP speakers technologically, culturally, economically, or otherwise so that the latter gave up their languages and shifted to using MP languages. Dominance of MP speakers is likely to have played a role in the disappearance of any preexisting non-MP languages. However, linguistic dominance does not necessarily involve conquest or replacement. An incoming group can come to be dominant over preexisting local languages without having significantly higher speaker numbers. If the incoming language enables these speakers to extend their network beyond their traditional territorial range and if these speakers want to exploit this possibility, then the incomer's language will gain currency and prestige and become a lingua franca. If subsequently it gets used in a wider range of contexts, it can become the first language of the offspring of the initial second language speakers.

For example, in the past, non-MP speaking groups of the Negritos in the Philippines have shifted to MP languages, which is clear from the non-MP lexicon that has been retained in these languages (Reid, 1994). However, language shift is not always simply from language A to language B and may involve multiple layers, some of which have left few traces. For example, prehistorical language shift has occurred in Borneo, where the MP Land Dayak languages of Borneo have striking lexical and phonological similarities in common with non-MP Aslian languages (Adelaar, 1995). This could be interpreted as Land Dayak originating as the result of a language shift from Aslian. However, Adelaar argues that it is more likely that both Land Dayak and Aslian have in common a substrate from an unknown *third* (non-MP) language. So here, we have a shift from languages A to B plus a shift from languages A to C, where B survives and

shares features with C, so that it looks as if C speakers shifted to B, while in fact the shared features between B and C originally come from A.

Also, the region of ISEA is immense and islands must have been populated quite sparsely when the first migrations of MP speakers occurred after 4,500 BP, as the global population is estimated to have been around 5 million around that time (even today, with a global population of more than 7 billion, many of the islands in ISEA are scarcely populated). In many cases, migrating MP speakers may have settled on previously uninhabited (parts of) islands. Much of the spread of MP languages need not have involved any contact with non-MP speakers, let alone language dominance or shift.

In addition, throughout the past millennia, frequent volcanic eruptions, earth quakes, and tsunamis must have wiped out numerous groups of speakers, both MP and non-MP, leaving no trace of their languages.<sup>14</sup> Areas that were depopulated as a result of such natural disasters could have been repopulated by incoming MP speakers from neighboring islands, giving rise to the current situation where the languages in some regions, especially in western ISEA, have very few traces of non-MP languages. In other words, the fact that today most of the languages in ISEA are of the MP family does not imply that MP speakers always displaced or conquered preexisting non-MP languages.

There are many possible scenarios to explain the current distribution of MP languages in much of the region. They include the economic or technical (e.g., sea faring) superiority, or cultural dominance, of MP speakers, leading to the development mentioned above, where a regional lingua franca over time became used in a wider array of contexts and got nativized. However, they may also have involved marriage practices where unions between couples of different linguistic backgrounds lead to the generational transfer of only one of the parent's languages. None of these scenarios necessarily involves major population movements.

At the same time, it should not be forgotten that there are dozens of MP and non-MP languages that have coexisted in relatively small spaces for many hundreds, perhaps thousands of years. There are numerous cases in eastern Indonesia where MP and non-MP speakers came to share certain islands until today, and have lived in peaceful, long-term coexistence, with both language types surviving. This is not only the case in New Guinea, but also, for example, on Timor where the south-west is MP and the north-east is non-MP; Alor and Pantar, where the MP language Alorese is spoken in coastal pockets, surrounded by dozens of non-MP languages; on Halmahera where the south is MP and the north is non-MP; and on Makian where on the east coast the MP language Taba is spoken and on the west coast non-MP Moi (Holton & Klamer, 2017). A simple scenario where MP languages have generally dominated and/or obliterated the earlier non-MP languages in ISEA does not account for these non-MP enclaves. It is important to mention this, as the literature on the Austronesian dispersal tends to take a macro-view on ISEA where the complex linguistic situation of eastern ISEA is glossed over or minimized,<sup>15</sup> thus missing crucial clues in the reconstruction of the linguistic past of the region.

The hypothesis that MP speakers demographically dominated non-MP groups also does not account for the fact that there are many non-MP languages in eastern ISEA which picked up certain hallmark MP features, such as a pronominal distinction between inclusive and exclusive plural or a head-initial word order (Klamer et al., 2008). This suggests situations of linguistic dominance and admixture where the dominant language was non-MP (McWilliam, 2007). In conclusion, the fact that today most of the languages in ISEA are MP is likely due to many different reasons, shift from non-MP languages being just one of them. If we focus on investigating the differences between MP languages rather than on what they have in common, much

variation in both their vocabulary and grammar will be found. It is this variation that is worth further investigation.

For the farming/language dispersal scenario, a variationist view implies that (1) it is highly unlikely that the ancestors of the low-level MP language groups in ISEA came with an identifiable number of migrations that can be historically and geographically defined; (2) it is more likely that the low-level MP language groups have grown from multiple different migrations, in various directions, of probably small bands of people in various time spans; and (3) there is no reason to assume that the processes by which immigrant MP languages became established in ISEA involved swift “replacement” of original inhabitants and their non-MP languages. Rather, it is more likely to have involved millennia-long periods of sequences of disasters and migrations, and often intense and stable multi or bilingual contact.

## 4 | INTERDISCIPLINARY EVIDENCE

In this section, I briefly review work in the archeology (4.1) and the molecular diversity (4.2) of ISEA, addressing the question to what extent archeological and genetic data can be linked to the spread of MP languages according to the language/farming dispersal hypothesis.

### 4.1 | Archeological evidence for networks of culture and trade

Much archeological research in ISEA argues that the spread of MP languages cannot be linked to the spread of a single type of culture-material, economic, or otherwise or the use of a single maritime trade network. For this, I have made use of references presented in Donohue and Denham (2010), Hung, Nguyen, Bellwood, and Carson (2013), Specht, Denham, Goff, and Terrell (2014), O'Connor (2017), and the overview in Galipaud (In press).

The Lapita “culture” belongs to a prehistoric Pacific Ocean people, and the emergence of Lapita culture on the Mussau islands in the Bismarck Islands is now dated at 3,470–3,250 BP (Denham, Bronk Ramsey, & Specht, 2012). Archeologists believe that the Lapita is the ancestor of prehistoric cultures in Polynesia, Micronesia, and the areas of Melanesia where Oceanic (Austronesian) languages are spoken. To link the Oceanic cultures to their Austronesian ancestors, archeology has looked for ancestors of Lapita people mainly in the region linking Taiwan (as the Austronesian homeland) with the Pacific.

However, recent pottery finds from the northern Luzon in the Philippines, dated at around 3,500 BP, suggest that the pre or initial Lapita culture was not in Taiwan but in the northern Philippines (Hung et al., 2011). There is also increasing evidence of a direct link between these northern Luzon sites and the first settlements by 3,500 BP of the Marshall islands more than 2,000 km to the east. Bellwood (2011, 2017) now agrees that the stamped pottery which will later become Lapita has an ultimate origin in SE Asia rather than Taiwan. This is because the pottery found in Taiwan is of a particular red-slip type that is not very widespread in ISEA, is not well-dated, and only becomes abundant in the region at a later stage, some 2,500 BP (Paz, 2006).

Another objection that has been raised against the earlier hypothesis is that the Lapita assemblages of “similar” items of material culture, which are claimed to derive from expanding MP populations, are in fact not similar at all when examined more closely. For instance, the shell artifacts attested across ISEA are diverse in form, and different ways of producing them have been used contemporaneously in different local areas (Szabó & O'Connor, 2004, 624; Szabó,

2010); all of them differing from techniques and types of the Lapita shell bead-working practices.

But even if there *is* similarity in material culture that does not necessarily mean, it has the same ancestry, being transmitted vertically, through the generations, thus suggesting a familial relationship between the practitioners of that culture, let alone a linguistic one. Superficial typological similarity of material cultures can—just like structural typological similarity between languages—be due to various developments: (i) inheritance; (ii) similar evolutionary pressures leading to similar typological features independently (“homoplasy” without “homology”); or (iii) similarity that is the result of horizontal transmission through trade or exchange, technological transfer, diffusion of an idea, or indeed transmission from different already settled groups (Szabó & O'Connor, 2004).

The idea that the original populations of ISEA were (mostly) hunter-gatherers or foragers that were overwhelmed by newcomers with a farming technology has also been debated. Pelagic fish hook finds in eastern Timor dated to  $9741 \pm 60$  BP (O'Connor & Veth, 2005), that is, at least 5,000 years before the MP expansion, suggest that in pre-MP times, there were seagoing societies in the region. In addition, original populations could also have been agriculturalists, or populations mixing vegeculture and arboriculture, of which it is hard to find archeological traces (Latinis, 2000; Oliveira, 2008, 248). If there is little archeological data in the area to clearly support the hypothesis of a tuber economy prior to the cereal one that might be due to the difficulty in obtaining such evidence from the archeological context, it does not prove that such an economy did not exist.

Finally, there is good evidence that maritime trade and diffusion of goods already took place in ISEA from the end of the Pleistocene (10,000 BP) onwards, between mainland Asia (Vietnam) and ISEA (Bulbeck, 2008). Many archeological findings attest to a complex regional network of closely related traditions which cannot be directly related to the Pacific colonization process (Specht et al., 2014). During the Metal Age (~3,000–1,500 BP), traces of trade networks become more visible. For example, Galipaud (In press) mentions three types of links: one linking south Vietnam and Thailand and the Indonesian islands, the other linking north-west Vietnam, Laos and Yunnan to Sumatra, eastern Indonesia and New Guinea, and the third linking south Vietnam with the Phillipines and Serawak.

South Vietnam and Thailand are linked with the Indonesian islands way into the east, to the Lesser Sunda islands Flores, Lembata, Pantar, and Alor by the existence of Dong-Song drums and the existence of cemeteries with burials where bodies were put in certain types of jars. The transport of some Dong-Song drums into ISEA is one of the clearest traces of a regional connection between the SE Asia mainland and the islands and is dated after the dispersal of MP languages. The cemeteries with urn containers are always coastal and therefore suggestive of mobile communities connected by a vast network of social relations extending possibly as far as the Southeast Asian mainland. In a recent excavation of the cemetery of Pain Haka in eastern Flores (Galipaud et al., 2016), the skeletons were dated and revealed that urn burials were already in use by 2,800 BP in Pain Haka, and that the cemetery remained in use for more than 500 years, without any evidence of imported metal. Pain Haka excavations provide clear evidence of burial practices similar to those documented in other parts of SE Asia. Chief among these is the use of pottery jars alongside other forms of container for the interment of the dead. The dating of the site combined with the fact that this burial practice is present over such a wide geographic area suggests a widespread belief system during the Neolithic period across much of SE Asia. In other words, burial traditions linking ISEA with the Southeast Asian mainland started well before the Metal Age, i.e., before the dispersal of MP languages.

The link between north-west Vietnam, Laos, and Yunnan to Sumatra, eastern Indonesia, and New Guinea is the dispersal of a specific domesticated pig which is found across this entire region (cf., Cucchi, Fujita, & Dobney, 2009). Finally, there is also a link between the Philippines, Serawak, and South Vietnam through the circulation of nephrites ornaments (Galipaud, In press). To sum up, complex maritime networks existed long *before* as well as *after* the dispersal of MP languages. Directions of these networks were diverse. And importantly, archeological evidence does not confirm the existence of a single identifiable network that encompasses the entire region where MP languages originated and dispersed.

The archeological evidence for the view of the introduction of a foreign cultural package from Taiwan and ISEA which was subsequently adopted in its entirety throughout the Pacific is thus very weak. The spread of MP languages cannot be linked to the spread of a single type of culture, material, economic or otherwise, or the use of a single maritime trade network.

## 4.2 | Genetic evidence points to complex admixture evolutionary processes

The spread of languages in ISEA has also been connected with the spread of human genes (DNA molecules), where the history of the region's peoples is reconstructed on the basis of inherited genetic characters. Here, I summarize some of the findings; a more detailed overview of molecular anthropological work on the population dispersal in ISEA is given by Murray Cox in Bellwood (2017; 107–16). However, it is important to keep in mind that a crucial difference between genes and languages is that genes are transmitted only vertically, while languages are transmitted both horizontally and vertically. This implies that it cannot be assumed that DNA molecules and words reflect the same population histories.

Genetic research confirms that during the Holocene period in ISEA, groups with Asian ancestry spread through ISEA. Although the data on ISEA are still very sparse, there is presently statistical support that Taiwan played an instrumental role in Holocene dispersals across ISEA, but the data do not clearly support a *majority* genetic input from there. Perhaps a fifth of the maternally inherited mtDNA lineages reflect movements from Taiwan (Brandão, Eng, et al., 2016), while other lineages reflect movements within and between island groups in Indonesia and the Philippines (Tumonggor et al., 2013). Hudjashov et al. (2017) show that during the period in question (<4,000 years) there is very limited evidence for autosomal gene transfer to Indonesian populations, and that which is tends to be from the Philippines. Moreover, there is no evidence of gene transfer to western Indonesia and Taiwan does not seem to have a direct role. Overall, the speakers of Austronesian languages today are very diverse in biological terms, and the genetic evidence indicates that the evolutionary processes in ISEA have been very complex and have involved a great deal of admixture. Finally, the genetic data are often inconclusive about the directions of these dispersals (Lansing et al., 2011; 265).

## 5 | CONCLUSIONS

The main conclusions of this review are as follows. There is universal consensus that the homeland of the Austronesian languages is Taiwan, and that the family spread from Taiwan into the Pacific. For the ~600 languages spoken in ISEA, many of the lower level, micro-affiliations have been worked out, but there is also much that is still unsure, both in the

lower level relations between languages as well as in how they relate within higher orders. Thus, apart from the fact that speakers of MP languages must have moved through ISEA on their way from Taiwan to the Pacific, linguistics has not many details to provide about their dispersal history within ISEA.

The MP languages of ISEA are both lexically and grammatically very much diverse. This variation can at least partly be explained with a dispersal scenario that involved multiple migrations in many different directions and at different points in time. If the processes and contact situations that are observed in small-scale societies in ISEA today also prevailed in historical and prehistorical times, then the history of many MP languages more likely consisted of long-term, intense contact in multilingual communities where newcomers and autochthonous people lived together for many centuries, rather than being determined by external traders or warriors (although such contacts may also have existed). The original populations of ISEA were not (only) hunter-gatherers but had among them sea-faring groups and agriculturalists, or populations mixing vegeculture and arboriculture.

Before any more work can be done on the higher level linguistic classifications of the ISEA languages, it is mandatory to work out the low-level reconstructions of the many understudied MP language groups: "Simply finding two languages at a 'safe' distance (that is, far enough away from each other so that contact and borrowing are less likely) and proposing a high level reconstruction on the presence of an apparently cognate lexeme in the two, is not a safe policy. Without the low-level reconstructions and subgroupings, we cannot hope to fast-track our way to high-level conclusions..." (Donohue & Grimes, 2008; 153).

In addition, we need detailed studies of local languages, each within their own historical context: investigations of the linguistic behavior of bi or multilingual speakers, and studies of substrate effects and contact-induced language change or shift in both MP and non-MP languages. Such studies may shed light on the sociocultural circumstances that shaped the changes which must have occurred in the past. In archeology, similar calls to focus on bottom-up research are being voiced: "The spatial and temporal complexities manifest in the archaeological record call for caution in the assignment of cultural expressions to monolithic categories." "The patterning of archaeological cultures needs to be built up from rigorous comparative analysis." (Szabó & O'Connor, 2004; 626).

The enormous diversity of languages and cultures in ISEA is a direct outcome of the region's complex geography, ecological disasters, history of migrations and settlement. Many of the histories reflected in the languages, archeological findings and DNA molecules do not converge. To what extent can we expect a simple dispersal model like the farming/language hypothesis to account for this complexity? In order to move forward, the focus should perhaps shift from developing macro-perspectives to doing the necessary ground work: collecting more comparative data at the level of regional language groups and societies with their own dispersal and contact histories. The mosaics that will result when we combine these micro-histories may then provide us with a glimpse of the past of larger regions within ISEA.

## ACKNOWLEDGEMENT

I would like to thank Owen Edwards, Phillip Endicott, Hanna Fricke, Russell Gray, Harald Hammarström, Gary Holton, Tom Hoogervorst, Gereon Kaiping, Francesca Moro, George Saad, and two anonymous reviewers for their comments on earlier versions of this article. While they may not always agree with the opinions expressed here, I am grateful for their suggestions and critical remarks; all errors are my own. This research was supported by the VICI research

project “Reconstructing the past through languages of the present: the Lesser Sunda Islands” awarded to Marian Klamer and funded by the Netherlands Organisation for Scientific Research (NWO), project number 277-70-012.

## ENDNOTES

- <sup>1</sup>*Glottolog* and *Ethnologue* deviate from this tree by not recognizing a Western MP subgroup (*Glottolog* has instead 25 subgroups of MP), more discussion of this issue is presented in Section 1. Both *Glottolog* and *Ethnologue* do recognize the subgroups CEMP, CMP, EMP, SHWNG, and Oceanic.
- <sup>2</sup>These are the languages of the Philippines and West Indonesia, including Bali, Lombok, West Sumbawa, Sulawesi, Banggai islands, Tukang Besi and Muna-Buton islands, as well as Charnorro, Palauan, Malagasy and Chamic (Adelaar 2005, 9).
- <sup>3</sup>That is, those that are classified as belonging to CMP and Greater SHWNG in the database.
- <sup>4</sup>Unlike the *ABVD*, which only has basic (Swadesh) vocabulary lists, the languages in *LexiRumah* contain lexemes for up to 615 concepts per language, including nonbasic lexicon, terms for flora and fauna, cultural artifacts, and so on.
- <sup>5</sup>Unless the family relation would be established on something else in addition, e.g., morphology. Thanks to Harald Hammarström for sharing these thoughts on lexical retention rates.
- <sup>6</sup>Also known as Dawan, Timorese, or Atoni.
- <sup>7</sup>Note also that, while lexicon can be transferred in bilingual situations, there can also be cultural constraints prohibiting or avoiding it (Aikhenvald, 2003, 2007; Rice, 2004), so that absence of transferred lexicon does not necessarily imply lack of contact.
- <sup>8</sup>Unlike, e.g., the typological features characterizing some of the languages in Europe (Haspelmath, 1998).
- <sup>9</sup>In a proportionally balanced sample of 128 Austronesian languages, only 11.7% has a plural suffix, prefix, or infix. More frequent strategies of coding plurality are reduplication and using a plural word (Wu, 2016).
- <sup>10</sup>The term “Papuan” refers to those languages spoken in New Guinea or its vicinity that are non-MP. Papuan is not a genetic unit: Papuan languages group into numerous different language families.
- <sup>11</sup>A reviewer asked for information whether the syntactic category of the possessor plays a role in the order of possessor and possessed, e.g., when the possessor only precedes the possessed when it is a pronoun, as in Polynesian languages. MP languages with possessor-possessed order show variable patterns in this respect. A free pronoun may or may not combine with a noun that has a possessor suffix, and the possessor pronoun may either precede or follow the noun, while ordering restrictions on possessor nouns vs pronouns may apply (for example, the possessive constructions in three closely related varieties Lamaholot-Lamhalera, Lamaholot-Lewoingu and Alorese show variable patterns, Klamer, 2012, 80–82).
- <sup>12</sup>The formal distinction between nouns that are alienably or inalienably possessed had been observed earlier but was then analyzed as an innovation in the CEMP group of Austronesian languages (Blust, 1978, 1993, 258; Lichtenberk, 1985). Ross (2001; 138) hypothesized that the distinction may have entered proto-Oceanic or an immediate precursor through Papuan contact. In line with this hypothesis, Klamer et al. (2008) have argued that a subset of the CEMP languages in eastern Indonesia adopted the distinction as a result of contact with Papuan languages.
- <sup>13</sup>This was probably the case for proto-Oceanic, which is reconstructed as morphologically transparent, possibly reflecting its use as a lingua franca being spoken over a large area (Pawley, 2008).
- <sup>14</sup>For instance, Bellwood notes that archeological sites become less clear towards western Indonesia (Java and Sumatra) because “sites along the former northern coastlines are now likely to be buried under many meters of alluvium and beneath the water table” (Bellwood, 1997, 231).
- <sup>15</sup>For instance, Bellwood (2017, 183) writes “Almost all the indigenous peoples of Island SE Asia today thus speak languages within the Austronesian family, except in pockets of eastern Indonesia close to New Guinea where a few Papuan languages are spoken.” The question is: What is ‘a few’? In eastern Indonesia, at least 35 non-MP

languages are spoken; more languages than in the whole of Europe. And what is 'close' to New Guinea? Timor-Alor-Pantar are approx. 1,000 kilometers away from New Guinea.

## ORCID

Marian Klamer  <https://orcid.org/0000-0002-2905-7161>

## REFERENCES

- Adelaar, A. (1995). Borneo as the cross-roads for comparative Austronesian linguistics. In P. Bellwood, J. Fox, & D. T. Tryon (Eds.), *The Austronesians: Historical and comparative perspectives* (pp. 75–95). Canberra: Pacific Linguistics.
- Adelaar, A. (2005). The Austronesian languages of Asia and Madagascar: A historical perspective. In A. Adelaar, & N. Himmelmann (Eds.), *The Austronesian languages of Asia and Madagascar* (pp. 1–42). London New York: Routledge.
- Aikhenvald, A. K. (2003). *Language contact in Amazonia*. Oxford: Oxford University Press.
- Aikhenvald, A. K. (2007). Semantics and pragmatics of grammatical relations in the Vaupés linguistic area. In *Grammars in contact: A crosslinguistic typology* (pp. 237–266).
- Arka, I. W. (2016). *Bahasa Rongga: Deskripsi, Teori Dan Tipologi*. Jakarta: Penerbit Universitas Katolik Indonesia Atma Jaya.
- Baird, Louise. 2002. "A grammar of Kéo: An Austronesian language of East Nusantara." PhD Thesis, Australian National University.
- Bellwood, P. (1985). *Prehistory of the Indo-Malaysian archipelago*. Sydney: Academic Press.
- Bellwood, P. (1997). *The prehistory of the Indo-Pacific archipelago* (2nd ed.). Honolulu: University of Hawai'i Press.
- Bellwood, P. (2005). *First farmers: The origins of agricultural societies*. Oxford: Blackwell.
- Bellwood, P. (2011). Holocene population history in the Pacific region as a model for worldwide food producer dispersals. *Current Anthropology*, 52, S363–S378. <https://doi.org/10.1086/658181>
- Bellwood, P. (2017). *First islanders*. Wiley-Blackwell: Prehistory and Human Migration in Island Southeast Asia. <https://doi.org/10.1002/9781119251583>
- Bellwood, P., & Hiscock, P. (2005). Australia and the Austronesians. In C. Scarre (Ed.), *The human past* (pp. 265–305). London: Thames & Hudson Ltd.
- Bergsland, K., & Vogt, H. (1962). On the Validity of Glottochronology. *Current Anthropology* 3:115–153. *Current Anthropology*, 3, 115–153. <https://doi.org/10.1086/200264>
- Blom, E., Polisčenská, D., & Weerman, F. (2006). Effects of Age on the Acquisition of Agreement Inflection. *Morphology*, 16, 313–336.
- Blust, R. A. (1977). The Proto-Austronesian pronouns and Austronesian subgrouping. *Working Papers in Linguistics, University of Hawaii*, 9(2), 1–15.
- Blust, Robert A. 1978. "Eastern Malayo-Polynesian: A subgrouping argument." In Second International Conference on Austronesian Linguistics Proceedings., edited by Stephen Wurm and Lois Carrington, Fascicle 1:181–234. Pacific Linguistics Series 61. Canberra: Dept. of Linguistics, Research School of Pacific Studies, Australian National University.
- Blust, R. A. (1982). The linguistic value of the Wallace line. *Bijdragen Tot Taal-, Land- En Volkenkunde*, 138, 231–250.
- Blust, R. A. (1983). More on the position of the languages of eastern Indonesia. *Oceanic Linguistics*, 22–23(1984), 1–28.
- Blust, R. A. (1993). Central and Central-Eastern Malayo-Polynesian. *Oceanic Linguistics*, 32, 241–293. <https://doi.org/10.2307/3623195>

- Blust, R. A. (1995). The prehistory of the Austronesian-speaking peoples: A view from language. *Journal of World Prehistory*, 9(4), 453–510. <https://doi.org/10.1007/BF02221119>
- Blust, R. A. (1999). Subgrouping, circularity and extinction: Some issues in Austronesian comparative linguistics. In E. Zeitoun, & P. J.-k. Li (Eds.), *Selected papers from the eighth international conference on Austronesian linguistics* (Vol. 1) (pp. 31–94). Taipei, Taiwan: Institute of Linguistics, Academia Sinica.
- Blust, R. A. (2000). Why lexicostatistics doesn't work: The 'universal' constant hypothesis and the Austronesian languages. In C. Renfrew, A. McMahon, & L. Trask (Eds.), *Time Depth in Historical Linguistics* (pp. 311–331). Cambridge: McDonald Institute for Archaeological Research..
- Blust, R. A. (2009). The position of the languages of Eastern Indonesia: A reply to Donohue and Grimes. *Oceanic Linguistics*, 48(1), 36–77. <https://doi.org/10.1353/ol.0.0034>
- Blust, R. A. (2013). *The Austronesian languages*. Canberra: Pacific Linguistics. <https://digitalcollections.anu.edu.au/handle/1885/10191>
- Blust, Robert A., and Trussell Stephen. n.d. "The austronesian comparative dictionary." <http://www.trussell2.com/acd/>.
- Brandão, A., Eng, K., et al. (2016). Quantifying the legacy of the Chinese neolithic on the maternal genetic heritage of Taiwan and Island Southeast Asia. *Human Genetics*, 135, 363–376. <https://doi.org/10.1007/s00439-016-1640-3>
- Bulbeck, D. (2008). An integrated perspective on the Austronesian diaspora: Colonisation of Island SE Asia. *Australian Archaeology*, 67, 31–51. <https://doi.org/10.1080/03122417.2008.11681877>
- Campbell, L. (2013). *Historical linguistics* (Third ed.). Edinburgh: Edinburgh University Press.
- Comrie, B., Matthews, S., & Polinsky, M. (2003). *The atlas of languages* (Revised ed.). London: Quarto Publishing.
- Cucchi, T., Fujita, M., & Dobney, K. (2009). New insights into pig taxonomy, domestication, and human dispersal in Island South East Asia: Molar shape analysis of sus remains from Niah Caves, Sarawak. *International Journal of Osteoarchaeology*, 19, 508–530. <https://doi.org/10.1002/oa.974>
- Dahl, O. C. (1973). *Proto-Austronesian. Vol. 15. Scandinavian institute of Asian studies monograph series*. Lund, Sweden: Studentlitteratur.
- den Besten, H. (1989). From Khoekhoe foreigner-talk via Hottentot Dutch to Afrikaans: The creation of a novel grammar. In *Wheels within Wheels; Papers of the Duisburg Symposium on Pidgin and Creole Languages* (pp. 207–249). Frankfurt am: Main.
- Denham, T., Bronk Ramsey, C., & Specht, J. (2012). Dating the appearance of Lapita Pottery in the Bismarck archipelago and its dispersal to remote Oceania. *Archaeology in Oceania*, 47, 39–46. <https://doi.org/10.1002/j.1834-4453.2012.tb00113.x>
- Diamond, J. (2001). Polynesian origins: Slow boat to melanesia? *Nature*, 410(6825), 167. <https://doi.org/10.1038/35065523>
- Donohue, M., & Denham, T. (2010). Farming and language in Island Southeast Asia: Reframing Austronesian history. *Current Anthropology*, 51(2), 223–256. <https://doi.org/10.1086/650991>
- Donohue, M., & Grimes, C. E. (2008). Yet more on the position of the languages of Eastern Indonesia and East Timor. *Oceanic Linguistics*, 47(1), 114–158. <https://doi.org/10.1353/ol.0.0008>
- Dyen, I., James, A. T., & Cole, J. W. L. (1967). Language divergence and estimated word retention rate. *Language*, 43, 150–171. <https://doi.org/10.2307/411390>
- Edwards, O. (2016). Parallel sound correspondences in Uab Meto. *Oceanic Linguistics*, 55(1), 52–86. <https://doi.org/10.1353/ol.2016.0008>
- Florey, M. (2010). Negation in Moluccan languages. In *East Nusantara: Typological and areal analyses* (pp. 227–250). Canberra: Pacific Linguistics.
- Fricke, H. (2017). The rise of clause-final negation in Flores-Lembata, Eastern Indonesia. *Linguistics in the Netherlands*, 34, 47–62. <https://doi.org/DOI>. <https://doi.org/10.1075/avt.34.04fri>
- Galipaud, J.-C. (In press). The dynamics of human expansion and cultural diversification in Southeast Asia and Oceania during the neolithic: An archaeological perspective. In M. Crevels, & P. Muysken

- (Eds.), *Language Dispersal, Diversification, and Contact: A Global Perspective*. Oxford: Oxford University Press.
- Galipaud, J.-C., Kinaston, R., Halcrow, S., Foster, A., Harris, N., Simanjuk, T., ... Buckley, H. (2016). The Pain Haka burial ground on Flores: Indonesian evidence for a shared Neolithic Belief System in Southeast Asia. *Antiquity*, 90(354), 1505–1521. <https://doi.org/10.15184/ay.2016.185>
- Gray, R. D., Drummond, A. J., & Greenhill, S. J. (2009). Language phylogenies reveal expansion pulses and pauses in pacific settlement. *Science*, 323(5913), 479–483. <https://doi.org/10.1126/science.1166858>
- Greenhill, S. J., Blust, R., & Gray, R. D. (2008). The Austronesian basic vocabulary database: From Bioinformatics to Lexomics. *Evolutionary Bioinformatics*, 4, 271–283.
- Grimes, Charles E. 1991. The Buru language of Eastern Indonesia. Canberra: PhD thesis Australian National University.
- Hammarström, Harald, Robert Forkel, and Martin Haspelmath, eds. n.d. “Glottolog 3.0.” [glottolog.org](http://glottolog.org).
- Haspelmath, M. (1998). How young is standard average European? *Language Sciences*, 20(3), 271–287. [https://doi.org/10.1016/S0388-0001\(98\)00004-7](https://doi.org/10.1016/S0388-0001(98)00004-7)
- Heggarty, P. (2015). Prehistory through language and archaeology. In C. Bower, & B. Evans (Eds.), *The Routledge Handbook of Historical Linguistics*. London New York: Routledge.
- Himmelman, N. P. (2005). The Austronesian languages of Asia and Madagascar: Typological characteristics. In *The Austronesian languages of Asia and Madagascar* (pp. 110–181). London/New York: Routledge: Routledge Language Family Series.
- Holton, G., & Klammer, M. (2017). The Papuan languages of East Nusantara and the bird’s head. In B. Palmer (Ed.), *The languages and linguistics of the New Guinea area* (pp. 569–640). Berlin/New York: Mouton de Gruyter. <https://doi.org/10.1515/9783110295252-005>
- Hudjashov, G., Karafet, T. M., Lawson, D. J., Downey, S., Savina, O., Herawati, S., ... Cox, M. P. (2017). Complex patterns of admixture across the Indonesian Archipelago. *Molecular Biology and Evolution*, 34(10), 2439–2452. <https://doi.org/10.1093/molbev/msx196>
- Hung, H.-c., Carson, M. T., Peter, B., Campos, F. Z., Piper, P. J., Dizon, E., & Bolunia, M. J. L. A. (2011). The first settlement of Remote Oceania: The Phillipines to the Marianas. *Antiquity*, 85, 909–926. <https://doi.org/10.1017/S0003598X00068393>
- Hung, H.-c., Nguyen, K. D., Bellwood, P., & Carson, M. T. (2013). Coastal connectivity: Long-term trading networks across the South China Sea. *The Journal of Island and Coastal Archaeology*, 8(3), 384–404. <https://doi.org/DOI>. <https://doi.org/10.1080/15564894.2013.781085>
- Kaiping, G., & Klammer, M. (2018). LexiRumah: An online lexical database of the lesser Sunda islands. *PLoS ONE*, 13(10), e0205250. <https://doi.org/10.1371/journal.pone.0205250>
- Kamholz, David. 2014. Austronesians in Papua: Diversification and change in South Halmahera–West New Guinea. PhD Dissertation. Berkeley: University of California.
- Kikusawa, R. (2015). The Austronesian language family. In C. Bower, & B. Evans (Eds.), *Routledge handbook of historical linguistics* (pp. 657–674). Oxford: Routledge.
- Klammer, M. (1996). Kampera Has No Passive. In M. Klammer (Ed.), *Voice in Austronesian. NUSA-linguistic studies of Indonesian and Other Languages in Indonesia* 39. (pp. 12–30). Jakarta: Universitas Atma Jaya.
- Klammer, M. (1998). *A grammar of Kampera*. Berlin/New York: Mouton de Gruyter. <https://doi.org/10.1515/9783110805536>
- Klammer, M. (2012). Papuan-Austronesian language contact: Alorese from an Areal Perspective. In N. Evans, & M. Klammer (Eds.), *Melanesian languages on the edge of Asia: Challenges for the 21st century. Language Documentation & Conservation Special Publication* 5. (pp. 72–108). Honolulu: University of Hawai’i Press. <http://scholarspace.manoa.hawaii.edu/handle/10125/4561>
- Klammer, M., & Ewing, M. C. (2010). The Languages of East Nusantara: An introduction. In *East Nusantara: Typological and areal analyses* (Vol. 618) (pp. 1–24). Canberra: Pacific Linguistics.

- Klamer, M., Reesink, G. P., & van Staden, M. (2008). Eastern Indonesia as a linguistic area. In P. Muysken (Ed.), *From linguistic areas to areal linguistics* (pp. 95–149). Amsterdam: Benjamins. <https://doi.org/10.1075/slcs.90.03kla>
- Kusters, Christiaan Wouter. 2003. “Linguistic complexity: The influence of social change on verbal inflection.” PhD Thesis, Utrecht: LOT Publications: Leiden University.
- Lansing, J. S., Cox Murray, P., De Vet, T. A., Downey, S. S., Hallmark, B., & Sudoyo, H. (2011). An ongoing austronesian expansion in Island Southeast Asia. *Journal of Anthropological Archaeology*, 30(3), 262–272. <https://doi.org/10.1016/j.jaa.2011.06.004>
- Latinis, D. K. (2000). The development of subsistence system models for Island Southeast Asia and Near Oceania: The nature and role of arboriculture and arboreal-based economies. *World Archaeology*, 32(1), 41–67. <https://doi.org/10.1080/004382400409880>
- Lees, R. B. (1953). The basis of glottochronology. *Language*, 29, 113–127. <https://doi.org/10.2307/410164>
- Lichtenberk, F. (1985). Possessive constructions in Oceanic languages and in Proto-Oceanic. In A. Pawley & L. Carrington (Eds.), 93–140. Canberra: Pacific Linguistics. In A. Pawley, & L. Carrington (Eds.), *Austronesian linguistics at the 15th Pacific science congress* (pp. 93–140). Canberra: Pacific Linguistics.
- Lynch, J., Ross, M. D., & Crowley, T. (Eds.) (2002). *The oceanic languages* (Vol. 1). Curzon Language Family Series. Curzon: Richmond.
- McWilliam, A. (2007). Austronesians in linguistic disguise: Fataluku cultural fusion in East Timor. *Journal of Southeast Asian Studies*, 38(2), 355–375. <https://doi.org/10.1017/S0022463407000082>
- Montrul, S. (2004). Subject and object expression in Spanish heritage speakers: A case of morphosyntactic convergence. *Bilingualism: Language and Cognition*, 7(2), 125–142. <https://doi.org/10.1017/S1366728904001464>
- Moro, F. (In press). Loss of morphology in Alorese (Austronesian): Simplification in adult language contact. *Journal of Language Contact*, 12
- O'Connor, S. (2017). Rethinking the neolithic in Island Southeast Asia, with particular reference to the archaeology of Timor-Leste and Sulawesi. *Archipel [Online]*, 90, 15–47.
- O'Connor, S., & Veth, P. (2005). Early holocene shell fish hooks from Lene Hara Cave, East Timor establish complex fishing technology was in use in Island Southeast Asia five thousand years before Austronesian settlement. *Antiquity*, 79, 249–256. <https://doi.org/10.1017/S0003598X0011405X>
- Ogoblin, Alexander K. 2005. “Javanese.” In *The Austronesian languages of Asia and Madagascar*, 590–624. Oxon: Routledge.
- Oliveira, Nuno Vasco. 2008. Subsistence archaeobotany: Food production and the agricultural transition in east Timor. Canberra: PhD thesis Australian National University.
- Pawley, A. (1999). Chasing rainbows: Implications of the rapid dispersal of Austronesian languages for subgrouping and reconstruction. In E. Zeitoun, & P. J.-k. Li (Eds.), *Selected papers from the eight international conference on Austronesian linguistics* (pp. 95–138). Taipei: Institute of Linguistics, Academia Sinica.
- Pawley, A. (2005). The chequered career of the trans New Guinea hypothesis: Recent research and its implications. In A. Pawley, R. Attenborough, J. Golson, & R. Hide (Eds.), *Papuan pasts: Cultural, linguistic and biological histories of Papuan-speaking peoples* (Vol. 572) (pp. 67–108). Canberra: Pacific Linguistics.
- Pawley, A. (2007). The origins of early Lapita culture: The testimony of historical linguistics. In S. Bedford, C. Sand, & S. Connaughton (Eds.), *Oceanic explorations: Lapita and western Pacific settlement* (Vol. 26) (pp. 17–49). Terra Australis. Canberra: ANU ePress.
- Pawley, A. (2008). Where and when was Proto Oceanic spoken? Linguistic and archaeological evidence. In Y. A. Lander, & A. K. Ogoblin (Eds.), *Language and text in the Austronesian world: Studies in honour of Ülo Sirk*. Munich: Lincom Europa.
- Paz, V. (2006). The Philippine Islands and the discourse on the Austronesian dispersal. In T. Simanjuntak, I. H. E. Pojoh, & M. Hisyam (Eds.), *Austronesian diaspora and the ethnogenesis of people in the Indonesian archipelago* (pp. 279–298). Jakarta: Indonesian Institute of Sciences; International Center for Prehistoric and Austronesian Studies.

- Pereltsvaig, A. (2012). *Languages of the world: An introduction*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139026178>
- Posth, C., Nägele, K., Collieran, H., Valentin, F., Bedford, S., Kami, K. W., ... Kinaston, R. etc.(2018). Language continuity despite population replacement in remote Oceania. *Nature Ecology & Evolution*, 2(4), 731–740. <https://doi.org/10.1038/s41559-018-0498-2>
- Reesink, G. P. (2002). Clause-final negation: Structure and interpretation. *Functions of Language*, 9, 239–268.
- Reesink, G. P., & Dunn, M. (2017). Contact phenomena n Austronesian and Papuan languages. In B. Palmer (Ed.), *The languages and linguistics of the New Guinea Area*. Berlin/New York: Mouton de Gruyter.
- Reid, L. A. (1994). Possible Non-Austronesian Lexical Elements in Philippine Negrito Languages. *Oceanic Linguistics*, 33(1), 37–72. <https://doi.org/10.2307/3623000>
- Reid, L. A. (2009). The reconstruction of a dual pronoun to Proto Malayo-Polynesian. In B. Evans (Ed.), *Discovering history through language* (pp. 461–477). Canberra: Pacific Linguistics.
- Rice, K. (2004). Language contact, phonemic inventories, and the Athapaskan language family. *Linguistic Typology*, 8, 321–383.
- Ross, M. (1995). Some current issues in Austronesian languages. In D. T. Tryon (Ed.), Pt. 1 fasc *Comparative Austronesian dictionary: An introduction to Austronesian studies* (Vol. 1) (pp. 45–120). Berlin: Mouton de Gruyter.
- Ross, M. (2001). Contact-induced change in Oceanic languages in North-West Melanesia. In A. Y. Aikhenvald, & R. M. W. Dixon (Eds.), *Areal diffusion and genetic inheritance* (pp. 134–166). Oxford: Oxford University Press.
- Ross, M. (2002). The history and transitivity of western Austronesian voice and voice-marking. In F. Wouk, & M. Ross (Eds.), *The history and typology of western Austronesian voice systems* (pp. 17–62). Canberra: Pacific Linguistics.
- Ross, M. (2005). Pronouns as preliminary diagnostic for grouping Papuan languages. In A. Pawley, R. Attenborough, J. Golson, & R. Hide (Eds.), *Papuan pasts: Cultural, linguistic and biological histories of Papuan-speaking peoples*. *Pacific Linguistics* 572. (pp. 15–65). Canberra: Research School of Pacific & Asian Studies, Australian National University.
- Ross, M. (2008). The integrity of the Austronesian family: From Taiwan to Oceania. In *Past human migrations in East Asia* (pp. 161–181). London/New York: Routledge.
- Ross, M. (2013). Diagnosing contact processes from their outcomes: The importance of life stages. *Journal of Language Contact*, 6(1), 5–47. <https://doi.org/10.1163/19552629-006001002>
- Ross, Malcolm. 2017. “Linguistic evidence for prehistory: Oceanic examples.” In *Issues in Austronesian Historical Linguistics*, edited by Hsiu-chuan Liao. Journal of the Southeast Asian Linguistics Society Special Publication No. 1. University of Hawaii Press.
- Ruhlen, M. (1987). *A guide to the world's languages*. Stanford: Stanford University Press.
- Schapper, A. (2015). Wallacea, a Linguistic Area. *Archipel*, 90, 99–151.
- Simons, G. F., & Fennig, C. D. (Eds.) (2018). *Ethnologue: Languages of the world, twentieth edition*. Dallas, Texas: SIL International. <http://www.ethnologue.com>
- Specht, J., Denham, T., Goff, J., & Terrell, J. E. (2014). Deconstructing the Lapita Cultural Complex in the Bismarck archipelago. *J Archaeol Research*, 22(2), 89–14. <https://doi.org/10.1007/s10814-013-9070-4>
- Spriggs, M. (2011). Archaeology and the Austronesian expansion: Where are we now? *Antiquity*, 85, 510–528. <https://doi.org/10.1017/S0003598X00067910>
- Szabó, K. (2010). Shell artefacts and shell-working within the Lapita cultural complex. *Journal of Pacific Archaeology*, 1(2), 115–127.
- Szabó, K., & O'Connor, S. (2004). Migration and complexity in Holocene Island Southeast Asia. *World Archaeology*, 36, 621–628. <https://doi.org/10.1080/0043824042000303809>
- Trudgill, P. (2010). Contact and sociolinguistic typology. In R. Hickey (Ed.), *The handbook of language contact* (pp. 299–319). Chichester, UK; Malden, MA: Wiley-Blackwell.
- Trudgill, P. (2011). *Sociolinguistic typology: Social determinants of linguistic complexity*. Oxford: Oxford University Press.

- Tryon, D. T. (1995). Proto-Austronesian and the major Austronesian subgroups. In P. Bellwood, J. J. Fox, & D. T. Tryon (Eds.), *The Austronesians: Historical and comparative perspectives* (pp. 17–38). Canberra: Dept. of Anthropology as part of the Comparative Austronesian Project, Research School of Pacific and Asian Studies, Australian National University.
- Tumonggor, M. K., Karafet, T. M., Brian, H., Stephen Lansing, J., Sudoyo, H., Hammer, M. F., & Cox, M. P. (2013). The Indonesian archipelago: An ancient genetic highway linking Asia and the Pacific. *Journal of Human Genetics*, 58, 165–173. <https://doi.org/10.1038/jhg.2012.154>
- Wouk, F. (2002). Voice in the languages of Nusa Tenggara Barat. In *The history and typology of western Austronesian voice systems* (pp. 285–309). Canberra: Pacific Linguistics.
- Wu, J. (2016). *Plural words in Austronesian languages: Typology and history* Leiden University Research MA Thesis. <https://openaccess.leidenuniv.nl/handle/1887/45798>

## AUTHOR BIOGRAPHY

**Marian Klamer** (1965) is Professor of Austronesian and Papuan Linguistics at Leiden University. After her PhD in 1994 at the VU University in Amsterdam, she held various research fellowships, including one with the Royal Netherlands Academy of Arts and Sciences (KNAW). Since 2003, she has been teaching at Leiden University, where she was also the PI of a VIDI research project funded by the Netherlands Organisation for Scientific Research (NWO), a EuroBABEL project funded by the European Science Foundation, and an archiving project funded by CLARIN-NL. Klamer has worked on Austronesian and Papuan languages in Indonesia for over 20 years, publishing grammars of two Austronesian languages (Kambera, 1998; Alorese, 2011) and two Papuan languages (Teiwa, 2010; Kaera, 2014). She has published extensively on a wide range of topics including morphology, typology, grammaticalization, language contact, and historical reconstruction of languages in Indonesia. Currently, she is leading the NWO-VICI project “Reconstructing the past through languages of the present: The Lesser Sunda Islands” (2014–2019).

**How to cite this article:** Klamer M. The dispersal of Austronesian languages in Island South East Asia: Current findings and debates. *Lang Linguist Compass*. 2019;13:e12325. <https://doi.org/10.1111/lnc3.12325>