Saad, George & Arnold, Laura & Peddie, Emma (2025) Late Vernacular Production in Island Southeast Asia and the Pacific. *Glossa Contact* 1(1), 1-55. DOI: 10.82012/glossa.contact.2025.saad.et.al

Late Vernacular Production in Island Southeast Asia and the Pacific

George Saad, University of Amsterdam, the Netherlands and Palacký University Olomouc, Czechia, george.6.saad@gmail.com
Laura Arnold, Australian National University, Australia, laura.arnold@anu.edu.au
Emma Peddie, University of Edinburgh, UK, emmapeddie22@gmail.com

---

In this paper, we define and exemplify Late Vernacular Production (LVP), an under-researched pattern of language acquisition, socialisation, and use. LVP occurs in some minority communities where bilingualism in the local vernacular and the lingua franca is the norm: counter-intuitively, speakers with LVP are active users of only the lingua franca in childhood, going on to become active users of both the vernacular and the lingua franca in late adolescence or early adulthood. After distinguishing LVP from similar patterns of language acquisition and use—simultaneous and sequential bilingual acquisition, heritage language use, and language shift—we exemplify LVP with a case study from Abui (Timor-Alor-Pantar language family), spoken in east Indonesia. We follow this with results from an exhaustive survey of Island Southeast Asia and the Pacific, where we find a total of 32 attestations of LVP. Based on these attestations, we identify the factors that seem necessary for LVP to develop: a particular constellation of attitudes and ideologies which (i) identify the lingua franca as the language of opportunity, (ii) identify the local vernacular as the language of ethnolinguistic expression, and (iii) associate vernacular production with community integration upon reaching late adolescence/early adulthood. Finally, we discuss some of the broader theoretical implications of LVP—specifically, the challenge that LVP poses to current models of bilingual acquisition and use; the potential impact of LVP on contact-induced change; and the significance of LVP for models of endangerment, and by extension also for the design of documentation and revitalisation programmes.

Keywords: Late Vernacular Production, Island Southeast Asia, Pacific, linguistic ecologies, language socialization, language acquisition

1. Introduction

Late Vernacular Production (LVP) is a pattern of language acquisition, socialisation, and use which is attested in minority communities where there is bilingualism in both the local vernacular and the vehicular lingua franca. First identified by Araali & Boone (2011) and Anderbeck (2015), speakers with LVP are active users of only the lingua franca in childhood, thus appearing monolingual; however, speakers begin producing the vernacular at a later lifestage.[1] Based on the studies discussed below, this transition occurs anywhere between 10-25 years of age, particularly around the threshold of late adolescence and early adulthood. As noted by Anderbeck, the transition from production of only the lingua franca to both the lingua franca and the vernacular may be gradual; or it may occur suddenly, correlating with an "initiation into adulthood" (2015: 27).

Here, we present the first overview of LVP, focussing on Island Southeast Asia (ISEA) and the Pacific—both our area of specialism, and the region where LVP is most robustly attested.[2] Until now, no major overview of LVP has been published.[3] Instead, LVP is mostly mentioned in passing in a short paragraph or footnote in a reference grammar, or via personal

1 While Araali & Boone (2011) did not use a specific term for this pattern, Anderbeck (2015: 27) referred to LVP as "adult/delayed vernacular production". Here, we adopt the term *late* to avoid any negative connotations implied by *delayed*; and to include cases where the late production occurs in adolescence, i.e. before adulthood.

2 In this paper, we define ISEA and the Pacific as the area including the Philippine archipelago; the Indonesian archipelago, including Sumatra, Java, Borneo, and Sulawesi, and the smaller islands of the Lesser Sundas and Maluku; Melanesia, including New Guinea, the Solomon archipelago, Vanuatu, New Caledonia, and Fiji; and the islands of Micronesia and Polynesia. The Malay peninsula, sometimes included in ISEA, is for our purposes not included; nor is Australia.

3 Immediately prior to the submission of this paper, David M. Eberhard and Lucija Šimičić informed us that they have a manuscript in preparation looking at attestations of this phenomenon across the world. At the time of submission, we have not seen a draft of their manuscript. We are grateful to them for providing slides from an earlier presentation (Eberhard & Šimičić, 2023); we have incorporated the attestations they mention there, where relevant, into this paper.

communication. While Anderbeck (2015) is the first to flag LVP as a cross-linguistic phenomenon, discussing two case studies in Indonesia, there have been multiple individual attestations of LVP across the globe before and since (see also Eberhard & Šimičić, 2023). Based on an exhaustive survey, in which we consider all available materials written on every language in ISEA and the Pacific from the last 30 years, we identify 32 attestations of LVP in this region. As we argue below, we believe LVP may be even more widespread in the languages of the region (and possibly even the world) than reported—partly because, based on current models of language acquisition and endangerment, late-onset vernacular production is counter-intuitive, and so linguists are not typically primed to notice and document the phenomenon; and partly because we argue that LVP can only be reliably identified with longitudinal observations of language use in a community, which have not been reported for the vast majority of the languages in this area.

Our goals with this paper are four-fold. The primary aim is to more closely define and raise awareness of this little-known phenomenon, which to date has received practically no attention in the theoretical literature. The second is to provide a preliminary hypothesis of the conditions necessary for LVP to develop. Third, we will argue that LVP should be incorporated into models of bilingual acquisition and use, language contact, and endangerment. Finally, we will identify some of the many outstanding questions that LVP raises, which we hope will guide future research programmes to investigate this phenomenon in more detail.

This paper proceeds as follows. In Section 2, we establish LVP as a distinct phenomenon, by distinguishing it from several similar patterns of language acquisition and use. In Section 3, we explore LVP in more detail through a case study of Abui, the language in which LVP has been most comprehensively documented. In Section 4, we provide an overview of LVP in ISEA and the Pacific, summarising all cases that have been attested in this region. Based on the preceding sections, in Section 5, we then discuss the attitudinal conditions that facilitate the

development of LVP in a linguistic ecology; and in Section 6, we outline the theoretical implications of LVP for models of bilingual acquisition and use, language contact, and language endangerment. We conclude in Section 7 with some final remarks, and an outlook for future research programmes.

2. Defining Late Vernacular Production

As mentioned above, speakers with LVP are active users of only the lingua franca in childhood, going on to become active users of both the lingua franca and the vernacular in adolescence or adulthood. However, despite not producing the vernacular until a later lifestage, children with LVP are not monolingual: they acquire a passive command of the vernacular from infancy. Evidence for this comes from two observations. First, as noted by Anderbeck (2015: 27) and recapitulated in some of the sources discussed below, the onset of vernacular production may be very rapid, corresponding with admission into adulthood. This implies that these speakers have a passive command of the vernacular, on which they can draw once the conditions are met for them to become active users. Second, a common theme in the sources is that children in communities where LVP is widespread clearly have a passive command of the vernacular—in that they can, for example, follow orders given or laugh at jokes told in the vernacular.

LVP must be carefully distinguished from several superficially similar patterns of language acquisition and use: simultaneous bilingual acquisition, sequential bilingual acquisition, heritage language use, and language shift.[4] A preliminary point to note is that, while patterns of bilingual acquisition and heritage language use are properties of individual speakers, language shift is a property of the community more widely. While there may be communities in which many or even all speakers have LVP, LVP is also fundamentally a property of the individual.

4 For expository purposes, throughout most of this paper we simplify the picture by considering only bilinguals; although the patterns of acquisition discussed here obviously also apply to speakers of three, four, five or more languages.

For simultaneous and sequential (or successive) bilingual acquisition, we follow the definitions given by Pearson: "…bilinguals can begin both languages at birth simultaneously or learn one first and then after that one is established, learn the next one sequentially (or successively)" (2009: 381). Sequential bilingualism is often further subdivided into early sequential bilingualism, where the bilingual acquires the second language in childhood; and late sequential bilingualism, where the second language is acquired later in life (see e.g. Meisel, 2009; Bi, 2017). To re-cast these distinctions in terms of the linguistic ecologies discussed in this paper, the outcomes in language use of simultaneous bilingualism, early and late sequential bilingualism, and LVP are the same: in all cases, speakers are active users of both the lingua franca and the vernacular later in life. However, they are distinguished by different patterns of language use in childhood. Simultaneous and early sequential bilinguals are active users of both the lingua franca and the vernacular in childhood; and late sequential bilinguals are active users of the lingua franca in childhood, but have no passive command of the vernacular. LVP bilinguals, however, are active users of the lingua franca in childhood, but at the same time have a passive command of the vernacular. In this way, LVP bilinguals are like simultaneous and early sequential bilinguals, in that they acquire both languages at an early lifestage; but they are also similar to late sequential bilinguals, in that they are active users of only a single language in childhood.

Having an active command of the lingua franca and a passive command of a vernacular, of course, is not uncommon. For example, this is frequently seen in heritage speakers. Here we follow the definition given by Polinsky: "…heritage speakers are individuals who were raised in homes where a language other than the dominant community language was spoken and thus possess some degree of bilingualism in the heritage language and the dominant language" (2015: 7–8). There are two main distinctions between LVP bilinguals and heritage speakers. First, LVP bilinguals are raised in a community where the vernacular is still a

dominant language, at least among their parents' generation and older; whereas heritage speakers are raised in a community where the lingua franca is dominant across all generations. Second, heritage speakers may be active users of their parents' language in childhood, but "typically do not reach their parents' or grandparents' level of fluency" (Polinsky, 2015: 7). In adulthood, individuals who identify with the label *heritage speaker* have a wide range of fluencies, from no proficiency through passive competence to an active command. LVP bilinguals, on the other hand, are not active users of the vernacular in childhood, but typically go on to have active proficiency in the vernacular later in life.

An active command of the lingua franca and a passive command of the vernacular is also common in situations of language shift. However, LVP is distinguished from language shift in two ways. First, as mentioned above, LVP is a property of the individual, whereas language shift is a property of the whole community. Second, while younger members of communities undergoing language shift may have passive competence in the vernacular, they do not go on to become active users of the vernacular in later life.

The differences between the patterns of language acquisition and use discussed in this section are summarised in Table 1.

Before moving on, note that we limit our definition of LVP to ecologies involving a vernacular and a lingua franca, as originally intended by Anderbeck (2015). We are aware that there are situations of small-scale multilingualism in which individuals are raised in one vernacular, but become active users of a second later in life. This is the case, for example, with Sui (Tai-Kadai family, southern China, glottocode suii1243),[5] where a child might be brought up speaking his mother's vernacular clanlect but by adulthood has become fully fluent in his father's clanlect (Stanford, 2009). However, instances like these differ in

5 Throughout this paper, when a new language is introduced, its glottocode (and any alternative names) is provided along with language-family and region information. Glottocodes and genealogical affiliations throughout are from Glottolog (Hammarström et al., 2023).

	Property of the	Community dominant in	Childhood		Adulthood	
Late Vernacular Production	Individual	Vernacular	Active: Passive:	Lingua franca Vernacular	Active: Passive:	Lingua franca, vernacular —
Simultaneous/ early sequential acquisition	Individual	Either	Active: Passive:	Lingua franca, vernacular —	Active: Passive:	Lingua franca, vernacular —
Late sequential acquisition	Individual	Either	Active: Passive:	Lingua franca —	Active: Passive:	Lingua franca, vernacular —
Heritage speaker	Individual	Lingua franca	Active: Passive:	Lingua franca (vernacular) (Vernacular)	Active: Passive:	Lingua franca (vernacular) (Vernacular)
Language shift (in progress)	Community	Lingua franca	Active: Passive:	Lingua franca Vernacular	Active: Passive:	Lingua franca Vernacular

Table 1: Features distinguishing late vernacular production from other patterns of language acquisition and use

nature and in their wider implications, so we do not include them as examples of LVP. We address the relationship between LVP and patterns of small-scale multilingualism in more detail in Section 5.1.

3. A case study of Late Vernacular Production: Abui

Abui (abui1241) belongs to the Central Alor branch of the non-Austronesian ('Papuan') Timor-Alor-Pantar family (Schapper & Huber & van Engelenhoven, 2017; Kaiping & Klamer, 2022). It is spoken by around 41,000 people (Charles Grimes, Benediktus Delpada, Daniel Lanma, Waksi Lema, Ance Atalani pers. comm.) on the island of Alor, in East Nusa Tenggara province, east Indonesia.[6] Today, Abui speakers live in western-central Alor, from the northern to the southern coast (shown on the map in Figure 1). Abui is the largest vernacular language spoken on either Alor or the adjacent island Pantar, and there is considerable dialectal variation. There is no consensus on vitality: Ethnologue classifies Abui as stable (Eberhard & Simons & Fennig, 2022), while the Catalogue of Endangered Languages classifies it as threatened (Campbell et al., 2023). LVP has been attested in the Abui community since the

6 In January 2023, at a language documentation and orthography workshop in Kupang, Indonesia, these individuals calculated the newly revised figure of 41,000 Abui speakers, based on administrative and census data. This figure presents a welcome adjustment to the figure of 16,000 cited in most sources (Stokhof, 1984; Kratochvíl, 2007; Klamer, 2017; Saad, 2020; ultimately from Wurm & Hattori, 1981).

early 2000s, brought about by the rise in dominance of the lingua franca, Malay/Indonesian (Austronesian), in particular the regional variety Alor Malay (Austronesian, alor1252).[7] This section discusses some of the socio-historical events that led to this situation, and presents some longitudinal sociolinguistic data documenting language use.

Contact between Abui and Malay/Indonesian has increased over the last 60 years due to a series of important societal changes. Up until the 1960s, many Abui lived in the inland, mountainous region of Alor. At the time, knowledge of Malay/Indonesian was limited (Du Bois, 1944), although 'egalitarian multilingualism' (see François, 2012) with neighbouring related languages was common. This situation began to change in the 1960s and 70s. Due to increased local government pressure to be situated near schools, churches, markets, and health facilities, many villagers were forced to relocate to the north coast of Alor (Saad & Klamer & Moro, 2019; Saad, 2020). This resettlement brought Abui speakers into contact with speakers of other languages such as Kamang (kama1365), Teiwa (teiw1235), Adang (adan1251), and Kabola (kabo1247) (all of which also belong to the Timor-Alor-Pantar family) as well as Alorese (alor1247) and to a lesser extent Bugis (bugi1245), Binongko (bino1238), and Makasarese (maka1311) (all of which belong to the Austronesian family). Since the early 1900s, Malay/Indonesian had been used as a lingua franca by coastal groups in this region; the Abui incomers entered into this linguistic ecology, and began using Malay/Indonesian for inter-ethnic communication. In

7 We use 'Malay/Indonesian' as a cover term for the different Malay varieties spoken in Indonesia. This includes Standard Indonesian, the national acrolect; and regional colloquial varieties, which themselves vary on lectal clines between mesolect and basilect (see Paauw, 2008; Gil, 2024 for the relationship between different Malay/Indonesian varieties). Despite the varieties often being quite morpho-syntactically distinct, they share much of the lexicon, and are frequently idealised by speakers as a single variety called Bahasa Indonesia. Most of what is spoken on Alor in informal contexts is the regional variety, Alor Malay (Baird & Klamer & Kratochvíl, in prep.) This scenario is somewhat different to other eastern Indonesian regions—such as the Papuan provinces, where speakers sometimes differentiate between the local basilect Papuan Malay and Standard Indonesian (see Morin, 2016).

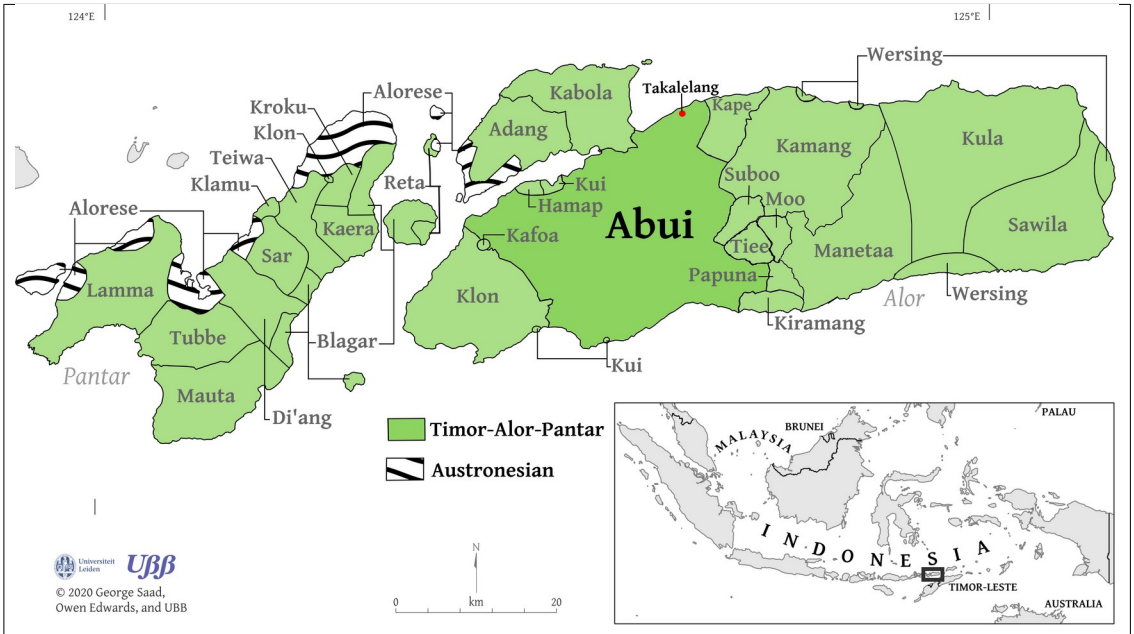

Figure 1: Linguistic map of Alor and Pantar

addition, Malay/Indonesian was the only language used in school settings, as prescribed by national school policies at the time. In the 1970s and 1980s, Abui children with little background in Malay/Indonesian struggled to keep up; and teachers (both non-native and Abui) scolded children for speaking Abui and not being able to converse in the lingua franca. Abui was subsequently banned on school premises, and teachers urged parents through church and community gatherings to raise their children in Indonesian. Parents responded accordingly—although what they used with their children was Alor Malay, rather than the more formal Indonesian. When this generation of children became adults and had children of their own, they typically brought them up using Alor Malay. This trend intensified with time: by the year 2015, most children were raised in Alor Malay (see Saad, 2020, Chapter 2 for further discussion).

LVP in Abui speakers was first observed by Kratochvíl, who stated that in the coastal village of Takalelang "most children are brought up in Malay, though they become more proficient in Abui when they grow up because it is still the main language used between adults" (2007: 4). He also highlighted that this did not apply to inland mountain villages, such as Tifol Afeng, where children were monolingual in Abui until age five or six, just before entering primary school. Follow-up research

conducted in 2015-2016 confirmed that the cohort of young speakers in Takalelang alluded to by Kratochvíl (2007) had by this time become active users of Abui; and they have now fully integrated into Abui-speaking adult life (Saad & Klamer & Moro, 2019; Saad, 2020: 123–125). This research also showed that the patterns of increased Abui proficiency with age noted by Kratochvíl still applied to the upcoming cohorts, 15-20 years later.

To illustrate LVP in Abui, data from a longitudinal sociolinguistic study conducted from 2015-2022 is presented in Figure 2. The study sampled a cohort of 18 speakers, born between 1998 and 2006.[8] The data was collected by a sociolinguistic questionnaire asking speakers to identify which language they typically used with their parents at home during these three intervals. It shows the shift in the language use of these 18 speakers across three lifestages: 1) before primary school (< 7 years), 2) during primary school (7–12 years), and 3) during young adulthood (17–24 years).

As shown in Figure 2, before attending primary school, Abui children predominantly use Alor Malay with their parents (around 90%). Only 10% percent report using either Abui (5%) or both Abui and Alor Malay (5%). However, during this time, children develop passive knowledge of Abui. For example, they overhear adults using the language among themselves; parents typically use Abui to order children to take part in domestic chores (such as serving coffee, cleaning, and fetching firewood), especially in the presence of non-Abui guests; and their grandparents often address them in Abui. Once speakers enter primary school, their use of Abui with their parents increases, with around 45% of speakers claiming they use either Abui (5%) or both languages (40%) with their parents. During young adulthood, this figure rises even further, with 72% percent of speakers claiming they speak either only Abui (22%) or both (50%). In addition, the mean age that these 18 speakers report “feeling comfortable answering back in Abui” is 13.8

8 All speakers who took part in the study filled out a consent form in Indonesian. For speakers under the age of 18, a parent filled out the form.

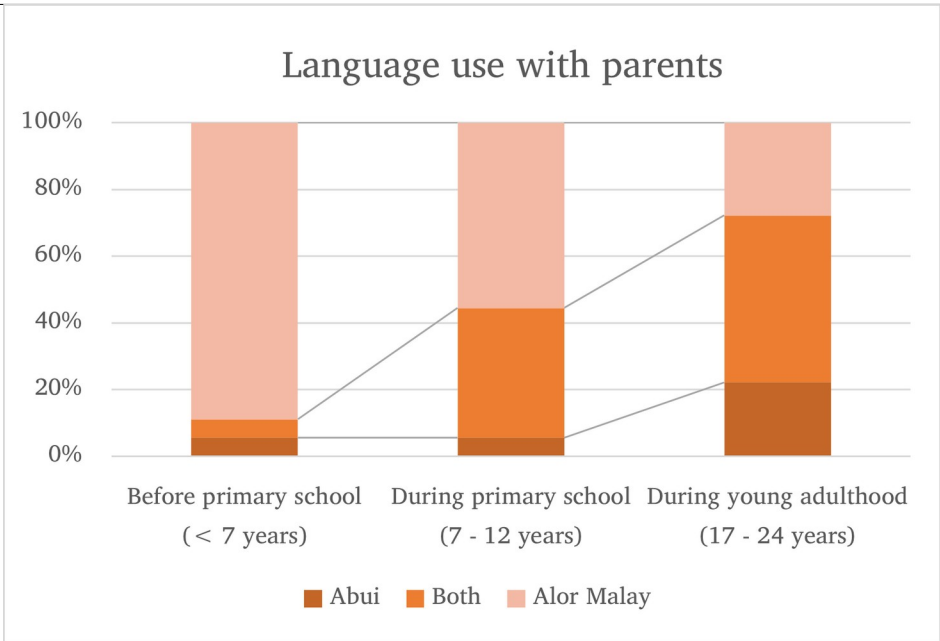

Figure 2: Reported language use with parents by younger Abui speakers (N = 18) across three lifestages

years (SD = 3.98, Range = 6 to 23). Of these, only two speakers gave answers below the age of 11 (viz. 6 years and 7-8 years); interestingly, these two respondents were siblings who were born in Tifol Afeng, the mountain village that Kratochvil (2007) alluded to where children were still raised in Abui. It should be pointed out that while these figures show an increase in Abui language use with parents, such use does not necessarily increase with every interlocutor. For example, siblings typically remain speaking Alor Malay to one another, and speakers and their grandparents typically also remain speaking to one another in Abui. However, children do increase their Abui language usage with their parents, as well as with other adults in the community.

To arrive at these figures, only a subset of a larger sociolinguistic questionnaire was used, namely the questions pertaining to language use with parents across the three lifestages above (the full questionnaire can be found in Saad, 2020: 401). The data presented here is longitudinal: data on the first two intervals, 'before primary school' and 'during primary school', was recorded in 2015-2016 when some speakers were still in primary school – and when speakers were between the ages of 9 – 16 years. Data on the third interval, 'during young adulthood', was collected in 2022 from the same speakers. It must be noted that while

these figures are informative, the data is derived from self-reports, and as such, comes with its own set of limitations. For example, self-reports may suffer from inaccurate self-assessment and problems associated with memory recall, in particular when asking speakers to self-report their behaviour at such early ages. In addition, self-reports are more likely to tell us about what beliefs and norms speakers may have regarding their behaviour, rather than about their actual behaviour itself (Cohen, 1996). However, given the logistic and ethical challenges associated with the collection of more robust types of data, such as (longitudinal) naturalistic observational data (e.g., Beyer, 2010), self-reports presented the most feasible way of obtaining valuable information regarding speakers' language use with their parents across the three lifestages. Besides, sociolinguistic interviews using self-reports have been used as a reliable tool for investigating language use in small-scale multilingual communities (Pakendorf & Dobrushina & Khanina, 2021).

For the purposes of this study, we chose to focus on only one type of interlocutor, namely parents, for two main reasons. Firstly, when conducting interviews in 2015-2016 exploring the root causes of LVP, we found that parents were the main catalysts for initializing the switch from Abui to Alor Malay. Therefore, it made sense to zoom in on language use in parent-child interactions. In addition, due to the design of the questionnaire, we had more data regarding parent-child interactions for the 18 speakers across the three lifestages than we did for other types of interlocutors, such as peers, siblings, and grandparents. This is partly because, for life-stage one (before age 7), the questionnaire only garnered data on language use in parent-child interactions (and not other types of interactions) and also because not all 18 speakers interacted with these interlocutors in a comparable way across the three life-stages; not all speakers had older siblings and grandparents available for all three periods. However, preliminary glances at data regarding "language use with peers" and "language use with siblings" appear to reveal a stable preference for Alor Malay, rather than an increase in Abui with age. In addition, it appears as though

children who had a strong relationship with their grandparents went on to become more proficient in Abui. With large amounts of sociolinguistic data still to be processed, future studies will surely shed light on changing language use patterns with various interlocutors based on a host of variables, such as family cluster, profession, school, and time spent abroad.

To conclude, the data in Figure 2 illustrate the increased usage of Abui across the lifespan. As children age, they receive more Abui input, both because they receive access to adult communities of practice, and more generally due to more exposure across a lifespan.[9] Several on-going studies will improve our understanding of how LVP functions in Abui, and the impact of LVP on Abui grammar is currently being investigated by the first author in a panel study. For studies investigating language variation and change across four age-cohorts using an apparent time approach, see Saad & Klamer & Moro (2019), Klamer & Saad (2020), and Saad (2020; 2023). Another on-going pursuit in the Abui context involves studying variation and change in a new cohort of (pre)adolescents. This also allows for a potential exploration of how recent changes in local language policy may affect speakers' Abui language use. These include the recent development of an Abui-based language class at a primary school in Alor via the *muatan lokal (mulok)* "local content" subject, a subject taught through the national curriculum and typically designed to teach students across Indonesia about their local cultures. Furthermore, a number of orthography workshops have been held in the years 2016 and 2024, while the language continues to make strides in the digital world as well (mainly on Facebook). In addition, a number of Abui language festivals and competitions have taken place in Alor in the years 2022-2024, encouraging schoolchildren to write poetry and perform songs and dances in Abui. Future studies can shed light on how these practices may affect LVP. Finally, the first author has worked closely with Abui language activists to integrate the

9 The acquisition and socialisation process of Abui children, adolescents, and young adults is discussed in detail in (Saad, 2020: Chapter 2).

findings of these studies with curriculum development and revitalization efforts in Alor.

4. Late Vernacular Production in Island South-East Asia and Pacific

Cases of LVP similar to that described in Abui are attested elsewhere in ISEA and the Pacific. Peddie (2021), based on a database drawn up by Arnold (forthcoming), provides an overview of these cases, which we build on here. This survey was as exhaustive as possible: we considered all materials discussing language acquisition, socialisation, and use from the last 30 years listed on Glottolog 4.4 (Hammarström et al., 2021) for every language in the ISEA and Pacific region.[10] Information on these topics largely came from sociolinguistic survey reports, as well as grammars and other descriptive materials; in some cases, supplemental information came through communication with fieldworkers and community members. If multiple sources were available, all were consulted. Where none of these sources were available on a given language, information was collected from the 25th edition of Ethnologue (Eberhard & Simons & Fennig, 2022).

Despite our best efforts, the survey is of course still limited in two main ways: (i) by whether materials describing patterns of acquisition, socialisation, and use exist for a given language; and (ii) if so, whether the authors either observed or reported LVP. In most cases, details are scant, and only observations and anecdotal reports suggest there may be LVP. As LVP is so difficult—if not impossible—to conclusively identify without observations of language use across several years, and because such longitudinal studies are so rare, we contend that LVP is likely to be more widespread than presented here.

In total, we identified 32 attestations of LVP in ISEA and the Pacific, nearly all of which are in communities in eastern Indonesia (lingua franca = Malay/Indonesian) and north and east Papua New Guinea

10 Eberhard & Šimičić (2023) discuss one additional source, not listed on Glottolog, which we also incorporate.

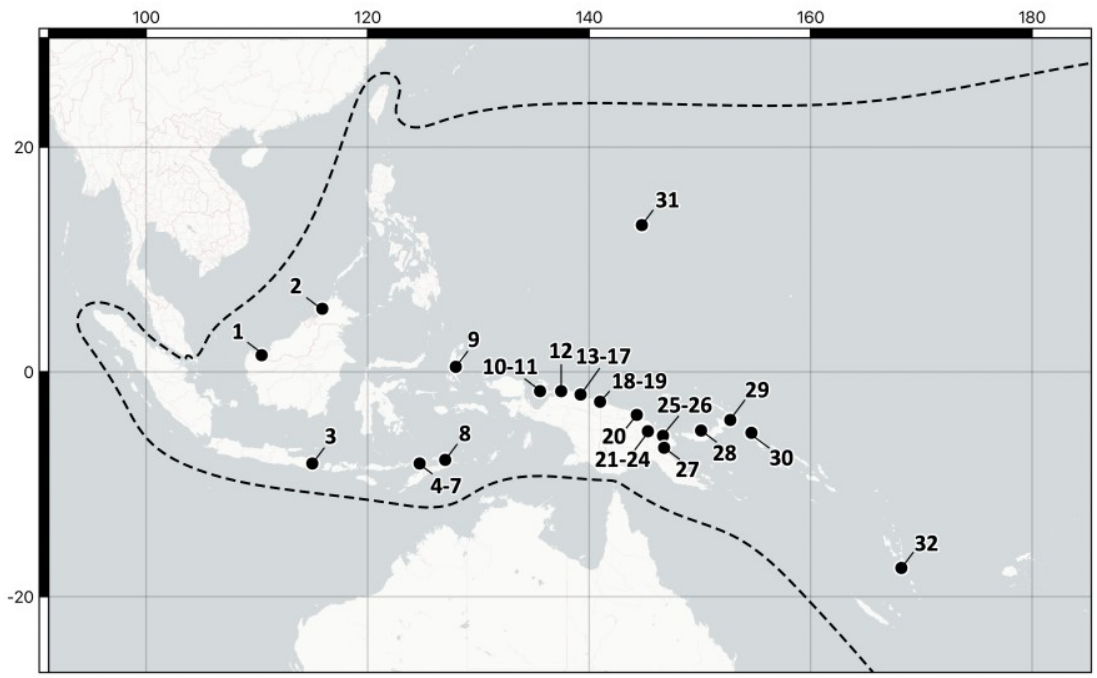

Figure 3: The distribution of LVP in Island Southeast Asia and the Pacific (eastern and southern Pacific not shown). All languages spoken eastwards of the dotted line (as far as 30° North, 109° West, and 48° South) were included in the survey. Numbers mark the following languages: 1. Sebuyau; 2. West Coast Bajau; 3. Balinese; 4-7. Kula, Kafoa, Abui, Kamang; 8. Oirata; 9. Sawai; 10-11. Ansus, Saweru; 12. Warembori; 13-17. Dabe, Vitou, Edwas, Sunum, Kwinsu; 18-19. Sentani, Skou; 20. Tayap; 21-24. Kein, Bau, Uyajitaya, Waube; 25-26. Forak, Sarasira; 27. Labu; 28. Bebeli; 29. Kandas; 30. Petats; 31. Chuukese (community on Guam); 32. Vatlongos (community near Port Vila)

(lingua francas = Tok Pisin and English). The locations of these communities are shown in Figure 3. A full list of all attestations and associated lingua francas, along with the evidence base of each attestation, is given in Appendix 1.

We now take a closer look at each attestation of LVP. For each language, we provide information on the demographics of language use, relevant details about the local linguistic ecology, and the evidence on which the LVP analysis is based.

There are several attestations of LVP in the languages of coastal north New Guinea, around the border of Indonesia and Papua New Guinea, and westwards to Cenderawasih Bay. One well-described case is Skou (Sko family, northern New Guinea, nucl1634). Donohue (2004: 10)

explains that, while school age children are not active users of Skou, they have a passive command due to child-directed vernacular speech from older generations. Soon after leaving school, Donohue reports that teenagers “are suddenly speakers of Skou”. He also mentions in passing similar patterns in three other languages spoken in the wider region: Ansus (Austronesian, South Halmahera-West New Guinea branch, ansu1237), Saweru (Yawa-Saweru, sawe1240), and Warembori (Austronesian, South Halmahera-West New Guinea branch, ware1253). Dwight Hartzler, who has worked with speakers of Sentani (Sentanic, nucl1632) since the 1970s, reports that while some community members never produce the vernacular, the majority of young adults become active users of the vernacular in order to enter Sentani society (pers. comm. cited in Anderbeck, 2015: 27).

Similar longitudinal observations of LVP have been made elsewhere in ISEA and the Pacific. For example, Ron Whisler (pers. comm.) describes the acquisitional trends in Sawai, spoken in north Maluku (Austronesian, South Halmahera-West New Guinea branch, sawa1247): in some villages, children acquire Malay/Indonesian as their L1, and only begin using Sawai regularly upon leaving school. Nazarudin (2021) notes that, until early adolescence (12-13 years), children have only a passive competence in Oirata (Timor-Alor-Pantar, oira1263), going on to become active users in adulthood. In west Borneo, Anonby (2020: 230–232) indicates that many children do not gain fluency in Sebuyau (Austronesian, Malayo-Sumbawan branch, sebu1243) until their late teens. Emelihter Kihleng (pers. comm.) reports that people who grow up in Chuukese (Austronesian, Oceanic branch, chuu1238) communities on Guam and Hawai‘i often do not start producing the language until they are adults. LVP is not restricted to the smaller minority languages of ISEA and the Pacific—it is also reported for urban varieties of Balinese (Austronesian, Bali-Sasak-Sumbawa branch, bali1278) spoken around Denpasar (Suastra, 2022); and inland, speakers only start experimenting with the mid-register of Balinese in their late teens and early 20s, once they transcend a status known as *teruna/teruni* ‘mature yet unmarried

man/woman' and are initiated into higher levels of ritual (Simon Evans, pers. comm.). Finally, the following unsourced Ethnologue entries (Eberhard et al., 2022) indicate that speakers only begin producing the vernacular in adolescence or adulthood, which—if substantiated—would indicate LVP: West Coast Bajau (Austronesian, Basap-Greater Barito branch, west2560), Forak (Trans-New Guinea, Finisterre-Huon branch, fora1245), Bau (also known as Fulumu; Trans-New Guinea, Madang branch, bauu1244), Kein (also known as Bemal; Trans-New Guinea, Madang branch, kein1239), Sarasira (Austronesian, Oceanic branch, sara1323), and Kandas (Austronesian, Oceanic branch, kand1301).

Kulick & Terrill (2019: 17–21) make some particularly remarkable longitudinal observations about language use in the Tayap community (also known as Taiap; isolate, taia1239). Tayap is spoken in a single village in the north of East Sepik province of Papua New Guinea, between the Sepik and Ramu rivers. Language shift is well underway in the community: Tok Pisin is the dominant language of the village and is used by all ages in all domains, with only those aged over 50 using Tayap in day-to-day life. Younger people never spontaneously produce Tayap. However, when pressed to do so, they can produce Tayap narratives, some with "excellent active competence" (*ibid*, p. 20). The Tayap situation differs from all other cases of LVP discussed in this paper, in that there is no increase in the production of the vernacular across the lifespan—while some speakers younger than 50 evidently have an active command of Tayap, they simply do not use it.

In some cases, language attitude surveys indicate that although children are not currently producing the vernacular, community members predict they will become fluent speakers in later life. For example, in at least one village, school age children are not active users of Waube (also known as Kwato; Trans-New Guinea, Madang branch, kwat1244; Lambrecht et al. 2009: 20–21, 31–32); nevertheless, the consensus is that they will become active users of Waube in adulthood (*ibid*, p. 25). Similar attitudes have also been reported by speakers of Kula (Timor-Alor-Pantar, kelo1247; Williams, 2017) and Kamang (Timor-Alor-Pantar,

kama1365; George Saad, fieldnotes); and are mentioned in passing in sociolinguistic surveys of Uyajitaya (alternative name Duduela; Trans-New Guinea, Madang branch, dudu1240; Lambrecht et al., 2008: 32), Bebeli (Austronesian, Oceanic branch, bebe1252; Spencer & Cott & MacKenzie, 2013: 14–16), and Labu (Austronesian, Oceanic branch, labu1248; Landweer, 1998: 66), and an anthropological paper on Petats (Austronesian, Oceanic branch, peta1245; Schneider, 2018: 234). Similarly, Kafoa (Timor-Alor-Pantar, kafo1240) children are active users of only Malay/Indonesian in childhood; but by the age of ten, parents report that children have acquired active fluency of both Kafoa and Malay/Indonesian. In addition, they also acquire fluency of two other local languages spoken in and around Kafoa territory, Klon and Abui (Baird, 2017: 57), suggesting that late production may be expanded to more languages in cases of community-level multilingualism.

In yet other cases, community predictions about the age of fluency in the vernacular indicate possible LVP. This is the case, for example, in one community of Vatlongos speakers (Austronesian, Oceanic branch, sout2859). Vatlongos is traditionally spoken in the southeast of Ambrym Island in Vanuatu. However, following a volcanic eruption in 1951, one village relocated to Mele Maat on the southwest of Efate Island, near the capital Port Vila. While Vatlongos is still acquired from birth on Ambrym, children in Mele Maat are not active users, and are instead dominant in the lingua franca Bislama (Indo-European, Early Melanesian Pidgin branch, bisl1239). However, in a survey by Ridge (2019), respondents in Mele Maat stated that children began producing the language at around 8 or 10; this is supported by observations by Ridge, who notes that “only children and teenagers were comfortable speaking Vatlongos in Mele Maat, though younger children may have already acquired a passive competence in the language” (p. 85).

Lee & Wambaliau (2004) and Lee & Sawi (2005) report similar trends in several Tor-Orya and Austronesian (Oceanic branch) languages spoken in north-west New Guinea, where Malay/Indonesian is the lingua franca. Again, community predictions about the age at which children

Village	Language	Glottocode	Language family	Age children fully understand vernacular	Age children fluently produce the vernacular	Source
Keder Lama	Keijar (aka Keder)	kede1239	Tor-Orya	10+	10+	Lee & Wambaliau (2004)
Dabe	Dabe	dabe1239	Tor-Orya	10	20	Lee & Wambaliau (2004)
Takar	Vitou (aka Takar)	vito1235	Tor-Orya	5-10	10-20	Lee & Wambaliau (2004)
Beneraf	Edwas (aka Bonerif)	bone1255	Tor-Orya	5	15	Lee & Wambaliau (2004)
Betaf	Tena (aka Betaf)	beta1253	Tor-Orya	5	10	Lee & Wambaliau (2004)
Ansudu	Kwinsu	kwin1242	Tor-Orya	10	17-20	Lee & Wambaliau (2004)
Yamna Pulau	Sunum (aka Yamna)	yamn1237	Austronesian > Oceanic	17	30	Lee & Sawi (2005)
Yamna Darat	Sunum (aka Yamna)	yamn1237	Austronesian > Oceanic	10	20	Lee & Sawi (2005)
Podena Darat	Fedan (aka Podena)	pode1237	Austronesian > Oceanic	18	20	Lee & Sawi (2005)
Anus Darat	Korur (aka Anus)	anus1237	Austronesian > Oceanic	5-6	7-8	Lee & Sawi (2005)

Table 2: Community-reported age of passive and active use of several Tor-Orya and Austronesian (Oceanic) communities (summarised from Lee & Wambaliau 2004 and Lee & Sawi 2005).

understand and produce the vernacular indicate a possible delay between passive and active use. As neither of these surveys are publicly available, the findings are summarised in Table 2.[11]

Table 2 shows a large lag between the community-reported age at which younger generations understand and produce the vernacular in the following villages: Dabe (where Dabe is spoken), Takar (Vitou), Beneraf (Edwas), Ansudu (Kwinsu), Yamna Pulau (Sunum), and Yamna Darat (Sunum). This suggests there may be LVP. However, these results are only preliminary. Notably, community members were asked the age at which children speak the vernacular "fluently"; but local attitudes towards what counts as a "fluent" speaker are often quite different from how fluency is measured in the acquisition literature. For example, from our own fieldwork in the region, our experience is that when a language

11 Many thanks to Randy Lebold at SIL for permission to reproduce these findings here.

is undergoing contact-induced change, older members of the community may not consider children acquiring the modified variety to be “fluent”. Further investigation is required to determine whether these communities have LVP.

While these community predictions around lifestage and language use may well be borne out, this is not always the case. Here, the case of Ulwa (also known as Yaul; Keram, yaul1241), spoken in the Sepik-Ramu basin of Papua New Guinea, serves a cautionary tale. Barlow (2018: 23–26) reports that some adults are optimistic about the language use of younger generations, “assuming that they will naturally become speakers of Ulwa once they become older” (p. 26). However, these positive sentiments are at odds with the present-day demographics of language use: today there are barely any fluent speakers of Ulwa under 40, and only some under 20 with even a passive comprehension. Perhaps historically, Tok Pisin-dominant children began producing Ulwa as they entered adulthood (thus accounting for community expectations of LVP). However, the generations born after 1980 evidently did not become active users of Ulwa. As the youngest generation of active users of Ulwa has grown older, the community now appears to be undergoing language shift.

From this overview, we see that LVP occurs in multiple communities across ISEA and the Pacific. In all attestations in our database, the lingua franca is a recent addition to the local linguistic ecology, and is an official language at the national level: Malaysian in Malaysia, Malay/Indonesian in Indonesia, Tok Pisin and English in Papua New Guinea, Bislama in Vanuatu, and English on Guam and Hawai’i. This patterning suggests that LVP in this region is likely a relatively recent phenomenon, which has been developing over the last few decades as colonial and postcolonial lingua francas spread and take hold. There is one attestation of an LVP-like phenomenon in a more traditional society of inland New Guinea: Janet Bateman describes how young speakers of Iau (Lakes Plain, iauu1242) in at least one village grow up speaking both Iau and the local trade language Elopi, before switching to Iau as their

primary language on reaching marriageable age (around 14-15 years; pers. comm. cited in Donohue, 2004: 10). While we do not consider this to be an example of LVP, as Iau speakers actively use the vernacular from birth, we note it here as a possible parallel situation of language use in a more traditional, small-scale ecology.

In the following section, we identify some common factors of the linguistic ecologies in which LVP is attested, to begin to determine what conditions are necessary for this acquisitional pattern to occur.

5. Factors conditioning Late Vernacular Production

Based on the attestations discussed above, two major factors are necessary for LVP to become widespread in a community. In this section, we discuss these in turn: positive attitudes towards the lingua franca, primarily as the language of education and social and economic opportunity (Section 5.1); coupled with a strong in-group community identity, positive attitudes towards the vernacular, and certain local linguistic ideologies (Section 5.2). Taken together, these two factors explain why the lingua franca is transmitted to children first; and why speakers begin producing the vernacular later in life. In Section 5.3, we discuss some other supplementary ideologies that are common in communities with LVP, and that may serve to reinforce its development.

5.1. Attitudes towards the lingua franca as the language of opportunity

As mentioned above, in all attestations of LVP in ISEA and the Pacific the lingua franca is an official national language. As such, these languages are widely used in education systems, and are seen as advantageous for gaining access to work opportunities.

Formal education in ISEA and the Pacific has had a serious impact on local communities in a number of different ways (Foley, 2005; Landweer, 2012; Anderbeck, 2015), and plays a significant role in the development of LVP. Several sources in our database discuss language use in educational settings (e.g. Abui, Bebeli, Labu, Waube); for these

communities, the lingua franca is the only language used as the medium of instruction. Coupled with this, the lingua franca is often enforced and the vernacular stigmatised in schools. For example, as mentioned in Section 3, speaking Abui was frequently punished in the 1970s-90s. Abui children experienced language trauma through the harsh policies of their teachers, including scoldings and beatings; interviews with schoolteachers who carried out these policies revealed that national school grades in Alor had been among the lowest in Indonesia (Saad, 2020: 82). In Papua New Guinea, Lambrecht et al. (2009: 9) report that Waube children from grade three upwards are punished for using anything other than English on school grounds. These factors mean that there is both overt and covert pressure on parents to use the lingua franca in child-directed speech, to ensure their children are fluent in the lingua franca before entering school—so that they neither fall behind in their studies, nor receive punishment for speaking the vernacular.

In addition to the dominance of the lingua franca in the education system, local job opportunities are often scarce, particularly those for which knowledge of the vernacular is an advantage. This means that if community members wish to increase their economic standing in the world, they typically must seek work outside of the local area, necessitating knowledge of the lingua franca. For example, in their description of the local economy of the Tor-Orya languages in north New Guinea, Lee & Wambaliau (2004: 7) explain that the majority of the population sustain themselves through hunting, sago harvesting, small-scale agriculture, and fishing; and that if someone wishes to earn more money, they have to leave the village in search of employment. Similarly, Saad reports that: “Because… economic opportunities are sparse in Alor, it is common for Abui people to go abroad [to other parts of Indonesia] in search of work” (2020: 31).

These pressures mean that, in all cases where information is available, the lingua franca is considered the language of education and economic and social opportunity, and there is thus a perceived benefit for parents to raise their children monolingually in the lingua franca. These

attitudes are summarised by Spencer & Cott & MacKenzie (2013: 22), who report that Bebeli parents "want their children to learn English and Tok Pisin because those languages are good for education"; Lambrecht et al. (2009: 25), who explain that in the Waube community "Tok Pisin is valued as a language of wider communication, a necessary skill for education and economic advancement"; and Lambrecht et al. (2008: 31–32), who state that: "The fact that [Uyajitaya] adults speak more Tok Pisin than vernacular to their children seems to indicate that they see significant advantage to their children learning to speak Tok Pisin. This could be due to educational opportunities, as well as economic and social benefits related to being able to interact with outsiders".

Kulick & Dobrin (forthcoming) explore why such attitudes are so widespread in Melanesia, especially Papua New Guinea. Rather than simply attributing it, as others have, to linguistic neo-colonialism (e.g. Dixon, 1991; Mühlhäusler, 2006), they propose that the same forces that underlie the vast linguistic diversity of Papua New Guinea have led to the prioritization of the lingua franca. Specifically, they suggest that the concept of exchange—in this case exchange of language as an emblematic practice—was historically key in maintaining language diversity and multilingualism; but nowadays, this very same concept is causing the demise of local languages. Just as in the past, "a command of other people's language was a possession many Papua New Guineans […] actively sought to acquire" (p. 7), today, they are "continuing their long-practiced habit of treating their languages as exchange objects: they are allowing their traditional languages to languish as they depreciate in exchange value, while acquiring one they regard as a more valuable commodity" (p. 9).

Together, these attitudes explain why the lingua franca is overwhelmingly used in child-directed speech in communities with LVP. However, the trends reported in this section are not sufficient for LVP to occur—they are found in countless communities around the world undergoing language shift. In the following section, we discuss the

attitudinal and ideological factors that motivate the transition to active production of the vernacular later in life.

5.2. Attitudes towards the vernacular as the language of in-group identity

A crucial factor for LVP to develop appears to be a particularly strong sense of local ethnic and linguistic pride, interwoven with specific local ideologies around lifestage, community integration, and language use. Together, these attitudes and ideologies seem to be necessary conditions for speakers to produce the vernacular at a later lifestage.

A robust, positive community identity, constructed and strengthened through use of the vernacular, is reported for several languages with LVP. For example, as described in Section 4, younger generations of Vatlongos speakers raised near Port Vila, far from the traditional territory on Ambrym Island, develop an active command of their parents' vernacular in adolescence. Ridge describes their "need to have a language connected to their island identity" as a primary motivation for this change in language use (2019: 85). We additionally know of two very clear cases of a strong community identity in languages with LVP from outside the ISEA/Pacific region: speakers of Bale-Dha (also known as Lendu; Central Sudanic, lend1245) in the northeast of the Democratic Republic of Congo have a strong in-group/out-group boundary, which is based on whether someone is a proficient user of the language (Araali & Boone, 2011); and the increase in Gurindji-derived elements across the lifespan in Gurindji Kriol (guri1249), a mixed language spoken in northern Australia, has been linked to the strong in-group identity of speakers, resulting in part from the key role played by Gurindji speakers in the Aboriginal rights movement (Sloan & Meakins & Algy, 2022: 4–5). Finally, this strong sense of identity and local pride is frequently seen in the preservation of traditional practices: for example, the maintenance of traditional marriage practices of the Abui (Saad, 2020: 33–35); or the observation by Donohue that "life in Skou [society] has not drastically changed compared to the way it proceeded, say, fifty or eighty years ago" (2004: 14).

As a result of this strong sense of ethnolinguistic pride, positive attitudes towards the vernacular as the language of the 'in-group' are frequently reported. For example, while there was a prevailing attitude in the 1970s and 1980s that Abui was not an appropriate language to use with children, the prestige of Abui for older generations never waned (Saad, 2020: 99). Similarly, respondents in all nine of the Uyajitaya villages surveyed by Lambrecht et al. (2008: 31) reported that they would like the vernacular to be used to write books and tell stories; in contrast, respondents in only four villages reported they wanted to hear stories in Tok Pisin, and none said that they wanted to hear stories in English. In their survey of Waube, Lambrecht et al. (2009: 26) report that community members "hold a positive attitude toward both their language and toward Tok Pisin, the language of wider communication." Importantly, to the best of our knowledge, there is only one example of a community with LVP where there are negative or neutral attitudes towards their vernacular: Labu, which is described as having "negative status both inside and outside the community" (Landweer, 1998: 66).[12]

However, positive attitudes towards the vernacular are not sufficient for LVP to develop—again, there are many communities around the world where attitudes to the vernacular are positive but shift is underway. The key factor appears to be the manifestation of these positive attitudes through local ideologies linking language use, lifestage, and status in the community. Typically, children's speech is not policed in communities with LVP; for example, Spencer & Cott & MacKenzie report that in the Bebeli community, "children's speech is not considered important" (2013: 15). Similarly, Donohue suggests that younger members of the

12 However, it should be noted that LVP in the Labu community appears less stable than the other examples discussed in this paper: most speakers are (or were, when the observations were made in 1988) simply shifting to Tok Pisin. The situation is described thus: "The… generation of children [in 1988] were learning Tok Pisin first and acquiring the vernacular (*if at all*) in later childhood or teenage years" (Landweer, 1998: 66; our emphasis).

Skou community use Malay/Indonesian in order to construct their identity in opposition to older generations of Skou speakers (2004: 10).

However, on reaching maturity, there is a strong expectation that younger generations will integrate more fully into community life, including in their language use. Again in the Skou community, the sudden switch from Indonesian to Skou by school leavers:

> "…reflects their status now not as wards of the state educational system, immune from prosecution for any violations of village conduct because of their requirement to fulfil governmental requirements, but as members of the village community. As such… [they] now adopt a more traditional lifestyle, including gardening, hunting, fishing, and speaking the language of their ancestors." (Donohue 2004: 10)

As discussed in Saad (2020: 90–96), a strong command of Abui is necessary to enter the adult sphere, which includes marriage and accumulation of wealth, and is linked to a marked increase in social status. A similar attitude is found amongst the Bebeli, where "functioning as an adult necessitates the local vernacular" (Spencer et al., 2013: 15).

These ideologies appear to play a pivotal role in creating the conditions for LVP to occur: the drive for community integration signalled by language use and the concomitant rise in community status provides a compelling motivation for a late adolescent individual to produce the vernacular. This may explain the situation in the Tayap community, where the younger generation has an active command of the language but never use it: as "the overwhelming majority of the social, cultural and ritual practices that link [the Tayap] to their history have vanished" (Kulick & Terrill, 2019: 3), there is no incentive for younger generations to speak Tayap. All of this points to an attitude towards the vernacular as being a skill that is developed over time and well into adulthood—much like hunting, farming, and acquiring wealth.[13]

13 We thank Don Kulick for raising this point.

5.3. Other language ideologies

In addition to these positive attitudes towards the lingua franca and the vernacular, the association of vernacular use with particular life stages, and a liberal attitude to children's language use, there are some broader language ideologies shared by many communities with LVP.

One is a family language policy that favours sequential bilingualism over simultaneous bilingualism. Among the Abui, very few parents intend, or even believe it is possible, to raise their children as simultaneous bilinguals (George Saad, fieldnotes). They almost always assert that children should learn Malay/Indonesian first, and then Abui once they reach late adolescence. When asked to elaborate, many parents do not feel that splitting their child-directed speech between Abui and Indonesian was a viable option (apart from uttering a few requests in Abui, in the presence of visitors). It is unclear whether parents actually feel that bilingual child-directed speech would hamper their linguistic development; however, what was clear was that parents felt that they had to prioritise Indonesian to prepare children for school.[14]

In addition, perceptions on the difficulty of a given language may contribute to LVP. For example, adult Gurindji/Gurindji Kriol speakers in Kalgaringi and Daguragu, Northern Territory, Australia explain that children learn Gurindji Kriol first because it is easier, and only gradually learn Gurindji—considered more difficult—as they grow older (McConvell, 1991; Sloan & Meakins & Algy, 2022).[15]

Finally, another force which might prevent children from using the vernacular is a sense of shame that comes from speaking the vernacular imperfectly: while, as mentioned in the preceding section, language choice is not policed, children’s proficiency in the vernacular is. This is

14 There does not seem to be any evidence suggesting that this is part of an ideology directed against code-mixing, as found in some other small-scale multilingual communities (cf. Pakendorf & Dobrushina & Khanina, 2021).

15 Similar patterns have been observed in numerous bilingual communities, even those that do not necessarily experience LVP. For example, a speaker of Rutul (Nakh-Daghestan, rutu1240) in Dagestan (Russia) reports that her child speaks Russian instead of Rutul because Russian is easier to learn (Nina Dobrushina, pers. comm.).

frequently aggravated by the behaviour of adults, who may mock their children's mistakes or blame their children for not speaking the vernacular. Consequently, in some cases children blame their parents for not raising them in the vernacular (Kulick & Terrill, 2019; Saad, 2020).

While the focus of this paper is LVP in ISEA and the Pacific, we expect factors similar to those discussed in these sections to play a role in conditioning LVP elsewhere in the world. Going forward, we encourage those from or working with speech communities with LVP to pay careful attention to how local ideologies—particularly towards the lingua franca, the vernacular, and the relationship of language use to lifestage—interact with patterns of language acquisition and use, to develop a more comprehensive picture of LVP ecologies across the globe.

6. Theoretical implications

As a newly-identified pattern of acquisition, socialisation, and use, LVP is theoretically significant for several sub-fields of linguistics. In this section, we discuss the implications for models of bilingual acquisition and use (Section 6.1), language contact (Section 6.2), and language endangerment (Section 6.3).

6.1. Models of bilingual acquisition and use

LVP has several important implications for theories of bilingual acquisition and use. In this section, we discuss how LVP bilingualism can be integrated into typologies of both bilingual speakers, and patterns of multilingualism in small-scale communities. In addition, as we show throughout this section, further study of this little-known phenomenon will respond to calls in the literature to diversify the languages and acquisition scenarios studied in bilingualism research, to better understand the role that social environment plays.

First, we address how individuals with LVP fit into typologies of acquisition. As discussed in Section 2, a distinction is often made between simultaneous and sequential bilingualism. Simultaneous bilinguals are those who acquire more than one language from birth,

and are considered to have two L1s—the term birth often implies that two languages were learnt before age 3 (McLaughlin, 1978; De Houwer, 1995; Paradis, 2007), which is typically assumed to be the age at which basic language development is established (Lust, 2006; Kim & Park & Lust, 2016). Sequential bilinguals, on the other hand, acquire their L1 from birth, while the L2 is learned sometime after age 3 (Paradis, 2007). This includes "early bilinguals (individuals who are exposed to two languages from a very early age) and late bilinguals (people who acquire a second language later in life)" (Bi, 2017: iii). While there is some debate as to the age cut-off point between early and late sequential bilinguals, Gathercole & Moawad (2010), for example, treat those who begin learning an L2 at 6 as early sequential bilinguals, and those learning an L2 at age 12 as late sequential bilinguals.

Furthermore, it is well established that children and adults have different acquisition strategies, and that different domains of language are affected by age of acquisition. In terms of phonological acquisition, for example, early sequential bilingual adults (Chinese L1 and Spanish L1 learners of English L2) still had a perceivable foreign accent in English, having learned English in an English-speaking environment from ages 5-8 (Flege & Fletcher, 1992). In contrast, the acquisition of vocabulary has been shown to increase with proficiency and so is less affected by age than other aspects of language (Green, 2003). Features that are difficult for adults to acquire include: 'i) intonation, ii) phonological distinctions among lexical items that are not made in their dialect, iii) new grammatical patterns'. However, adults may acquire: 'i) new vocabulary, ii) new pronunciations of known words, iii) for some speakers, new phonetic realisations of a given phoneme' (Ross 2013:20; Kerswill 1996).

In all these models, there is a general assumption that the onset of exposure corresponds to an onset of production. However, for LVP bilinguals, there is a prolonged gap between onset of exposure and consistent production; as discussed in Section 2, this means LVP bilinguals pattern with both simultaneous bilinguals and late sequential

bilinguals. There are no detailed studies investigating acquisition where there is a gap of some 10-20 years between onset of exposure and production. The consequences of this lag between exposure and production for learner proficiency in their vernacular are currently unknown (see Section 6.2).

We propose two stages of LVP in the individual. For LVP bilinguals who do not yet produce the vernacular, we use the term ‘active-passive bilingual’, coined by Kulick & Terrill (2019: 21) to describe the proficiency of those who have an active competence in a language, but do not spontaneously produce it. The second stage comes once an individual begins to orient themselves to community participation, where local ideologies place high value on active use of the vernacular; at this stage, speakers begin to produce the vernacular more regularly.

It is no secret that the field of language acquisition has historically focussed on major world languages (Adamou, 2021): for example, Anand et al. (2011) show that 85% of over 4,000 psycholinguistic conference and journal abstracts focussed on just 10 widely-used national languages such as English, German, and Japanese; and Kidd & Garcia (2022) have recently shown that a similar percentage of language acquisition studies are of Indo-European languages, with a heavy skew towards monolingual acquisition. With this foundation in place, we are now in the fortunate position to explore whether theories developed through the study of language of the Global North apply universally, in speech communities with vastly different learning environments. As pointed out by several authors, it is crucial to collect acquisition data from as broad a typological range of languages and learning environments as possible, in order to enhance our understanding of how humans learn language; however, small-scale, indigenous languages continue to be underrepresented in the psycholinguistic literature (Bowerman, 2011; Kelly et al., 2015; Meakins & Green & Turpin, 2018: 216–219; Adamou, 2021; Kidd & Garcia, 2022).

Research has shown that the social environments involved in which language is acquired in small-scale societies differ significantly from

those in Western, Educated, Industrialised, Rich, and Democratic (WEIRD; Henrich & Heine & Norenzayan, 2010) communities. One such example is in child-directed speech. Numerous studies on acquisition in WEIRD communities demonstrate the value of child-directed speech as linguistic input to enhance lexical and syntactic development (Goodman & Bates, 1997; Shneidman & Goldin-Meadow, 2012; Casillas & Brown & Levinson, 2021; *inter alia*). However, recent investigations have shown that child-directed speech is largely absent in at least some small-scale communities: for example, among the Tseltal (Mayan, Tzeltalan group, tzel1254) in Mexico (Casillas & Brown & Levinson, 2020) and the Yélî Dnye (isolate, yele1255) in Papua New Guinea (Casillas & Brown & Levinson, 2021). This trend was also anecdotally observed for the Abui both in the early 1930s (Du Bois, 1944: 47) and the 2010s (Saad, 2020: 92–93). In small-scale communities, where caregivers are preoccupied with social and domestic commitments which involve frequent conversing with other adults, much of what an infant is exposed to is overheard speech, rather than child-directed speech (Brown & Gaskins, 2014; Casillas et al., 2021). Despite this differing input, the linguistic development of Tseltal and Yélî Dnye children was shown to be on par with children in WEIRD societies (for example, in the achievement of milestones such as single- and multi-word utterances). This suggests that children in small-scale communities derive word meanings from hearing others use words, as opposed to being directly engaged in joint attention (Brown & Gaskins, 2014). The claim here is that infants in such communities may develop more advanced strategies to decipher overheard speech when compared with infants in societies that are engaged in child-centred face-to-face interaction with their caregivers.

While the Tseltal and Yélî Dnye studies only considered L1 acquisition, the finding that overheard speech plays a crucial role in language acquisition in small-scale communities has implications for LVP. Recall from Section 5.1 that child-directed speech in communities with LVP is generally carried out in the lingua franca, but that the vernacular is the dominant language of adult-to-adult speech. As discussed in Section 4,

LVP bilinguals typically become fluent users of the vernacular in later life; this suggests that this overheard vernacular speech provides sufficient input for this level of proficiency to develop later in life.

Finally, the identification of LVP has implications for our understanding of patterns of multilingual language use more broadly. Thus far, we have been describing LVP in terms of a binary tension between the lingua franca and a single vernacular. However, the traditional linguistic ecologies of ISEA and the Pacific are often highly multilingual: individuals are proficient in multiple local and indigenous repertoires, and there is no great imbalance of power between the different varieties (e.g. Sankoff, 1980; Thurston, 1992; Clifton, 1994; François, 2012; Kashima, 2020: 54–61). Similar contexts of small-scale multilingualism are attested across the world, and the theoretical study of these linguistic ecologies is a rapidly emerging field (for useful overviews, see Lüpke, 2016; 2021; Vaughan & Singer, 2018; Di Carlo & Good & Ojong Diba, 2019). In a recent typology of small-scale multilingualism, Pakendorf & Dobrushina & Khanina (2021) note that individuals in these ecologies are often also competent in a colonial lingua franca. Closer study is, however, required to understand how LVP is reshaping the relationship between local vernaculars and the more widespread colonial and postcolonial languages, and the extent to which these processes are similar across different communities. More broadly, a dynamic model of multilingual language use in these and similar contexts is required to capture the changing identity-formation processes across the lifespan of individuals with LVP: from a more global, outwardly oriented identity in childhood and early adolescence, to a more community-based, locally oriented identity as one enters adulthood.

6.2. Models of language contact

A language with widespread LVP is necessarily in intense contact with the lingua franca, and thus a candidate for contact-induced change. It is well known that the outcomes of contact depend in part on the age at which bilinguals acquire their languages. For example, there is wide

agreement that simultaneous bilinguals acquire and enhance complexity, while adult learners generally decrease complexity; and that, while adults easily acquire new vocabulary, they struggle to acquire new phonological oppositions and grammatical patterns such as inflectional paradigms (see e.g. Kerswill, 1996; Trudgill, 2009; Thomason, 2010; Ross, 2013; O'Shannessy & Davidson, 2020). However, as discussed in the preceding section, acquisition in LVP bilinguals is distinct from that of both simultaneous and sequential bilinguals. The effects of LVP on language change are therefore currently unknown. A major question for future research is the extent to which the vernacular produced by LVP bilinguals varies from that of their parents, i.e. to what extent speakers produce a 'native' (L1) variety once their vernacular language use stabilises in adulthood; or whether the late production introduces L2-like changes. As a preliminary hypothesis, if L2-like change is introduced, we do not expect all linguistic variables to be affected in the same way: based on the literature on bilingual acquisition and sociolinguistic variation in bilingual settings, we would expect to primarily see increased morphosyntactic and phonological simplification with each successive generation of LVP bilinguals, owing to the difficulty for adults to acquire complexity in these domains.

In addition, there may be considerable variation among speakers with LVP, calling for a nuanced approach examining the sociolinguistic variables. This can include within-cohort variation, conditioned by various sociolinguistic variables such as gender, clan, and social networks. It can also include inter-generational differences among LVP bilinguals, whereby cohorts can pattern quite differently from one another. This is based on data from Abui, where it was found that the first generation of LVP bilinguals (born in the 1980s) went on to produce a variety that was much closer to their parents' generation than that of the two younger generations of LVP bilinguals (born in the 1990s and 2000s; Saad, 2020). In the absence of extensive longitudinal data, it is impossible to explain exactly how this might have come about. However, differences in both quantity and quality of Abui input both during

childhood and during adulthood are likely to have played a role (Saad & Klamer & Moro 2019: 21–24). Specifically, the first generation of LVP bilinguals is likely to have received a significant amount of input from their parents' generation of L1 Abui speakers, while still exhibiting some effects of incomplete acquisition. In contrast, the subsequent age-groups are likely to have received more variable input from adjacent age-groups, who themselves would have exhibited effects of incomplete acquisition.

6.3. Models of language endangerment

First noted by Araali & Boone (2011) and Anderbeck (2015: 27–28), LVP additionally has implications for models of language endangerment, as it could muddy the waters when quantifying the vitality status of a language. In a nutshell, an apparent time snapshot of language use in a community with LVP looks identical to a community in the initial stages of language shift: while the older generations are active users of the vernacular, the younger generations are not producing the language. In these circumstances, it would be easy to misunderstand the generational asymmetry in language use as a break in intergenerational transmission —when in fact this asymmetry is due to widespread LVP.

Quantifications of vitality derived from frameworks such as the Expanded Graded Intergenerational Disruption Scale (EGIDS; Lewis & Simons, 2010), the UNESCO framework (UNESCO Ad Hoc Expert Group on Endangered Languages, 2003), and the Language Endangerment Index (LEI; Lee & Van Way, 2016; 2018) have a wide range of applications within and beyond our discipline. These three models all privilege, either implicitly or explicitly, intergenerational transmission as the primary measure of vitality. This is most clearly seen in the LEI framework. In this model, a language is assigned a score for each of several factors, including intergenerational transmission; overall vitality is then quantified using a formula that gives the intergenerational transmission score twice as much weight than the others (Lee & Van Way, 2018). As stated explicitly by Lee & Van Way, the Catalogue of

Endangered Languages (for which the LEI formula was developed) "identifies intergenerational transmission as the most critical factor in assessing level of endangerment" (2016: 280). In the UNESCO framework, nine factors are used to assess vitality—again including intergenerational transmission—and the authors stress that all nine should be considered together to give a holistic picture of vitality (2003: 17). However, the framework does not give an overall score to compare vitality between languages; in practice, this is typically achieved by comparing only the transmission component (e.g., Lewis & Simons, 2010: 5; Moseley, 2010: 11–12). Finally, the EGIDS framework (as described by Lewis & Simons, 2010) is a decision tree which takes into account various factors in the assessment of vitality. Two critical questions when quantifying vitality using this model are: (1) whether all parents are transmitting the language to their children; and (2) if not, what the youngest generation is with at least some proficient speakers. If the answer to the first question is negative, the language is considered threatened to some degree; the level of threat is then measured by the age of the youngest proficient speakers.

Furthermore, all three frameworks promote an apparent time approach in the quantification of vitality, in that the descriptions of each level of endangerment refer to the age of the youngest fluent speakers. For example, for a language to be given a score of 2 'Threatened' in the intergenerational transmission scale of LEI, the following observation must hold: "Most adults in the community are speakers, but children generally are not" (Lee & Van Way, 2016: 281). In the UNESCO framework, a similar rating on the intergenerational transmission scale is described as: "The language is used by some children in all domains" (2003: 8). Finally, in the EGIDS framework, a level of 6b 'Threatened' is given if the answer to the question "What is the youngest generation that has some proficient speakers?" is "[some] children" (Lewis & Simons, 2010: 114).

These descriptors assume that individual language use stabilises in adolescence, and that if an individual is not producing a language by

this point, they will not acquire an active proficiency—in other words, intergenerational asymmetries in language use are used as a proxy for intergenerational transmission. However, as we have shown above, this is not necessarily the case: LVP bilinguals are passive users of the vernacular in childhood, only becoming active users in adolescence or adulthood. Where a linguist only has limited time (weeks, days, even hours) in a community to assess vitality, and children have only a passive competence of the vernacular, this means that there is often not enough evidence to determine whether the community is undergoing language shift or whether there is widespread LVP.[16] This is because in such a scenario, the linguist must rely on limited observations, and thus the apparent time construct, to infer the extent of intergenerational transmission. The linguist may additionally seek information on language use from community members: however, if the community reports that children typically become active users later in life, this may accurately reflect LVP (as in the case of Abui), or this may be wishful thinking (as reported for Ulwa). Similarly, there may be communities where LVP occurs but community members are not aware of it, and so do not comment on it—at present, we have no information on such situations. In a nutshell, only longitudinal observations can conclusively distinguish between language shift and widespread LVP in these contexts.

Before concluding, there are a few more points on the interaction between LVP and language endangerment that we would like to flag. First, based on the issues discussed in this section, we contend that at least some languages currently considered endangered may, on closer longitudinal inspection, turn out to have LVP. Having said this, our second point is that by defining and discussing LVP, we do not wish to detract from or undermine the serious situation of language endangerment that the world is facing today (Hale et al., 1992;

16 As mentioned in Section 2, LVP is the property of an individual, and shift a property of the community. The two are therefore not mutually exclusive: the community as a whole may be shifting to the lingua franca, but some individuals may be LVP bilinguals.

Bromham et al., 2021). We are acutely aware of the state of the field: the first and second authors both have first-hand experience carrying out documentation, engagement, training, and revitalisation activities with marginalised communities in east Indonesia. At present, we have no firm idea of the extent of LVP (either in ISEA and the Pacific, or worldwide). However, the fact that the phenomenon is only now being identified suggests that it is comparatively rare—at least when compared with the large number of languages which are obviously undergoing language shift. Rather, our goal is to encourage a more nuanced picture of endangerment that takes LVP bilingualism into account, so that linguists, policy makers, and stakeholders can more accurately evaluate the vitality of a language, decide how best to distribute finite resources, and develop more nuanced strategies in documentation and revitalisation projects, by integrating contextual information from the local linguistic ecology.

7. Conclusion and outlook

In this paper, we have shown that Late Vernacular Production (LVP), an underdocumented pattern of language acquisition and use, is attested in multiple communities throughout ISEA and the Pacific. As just discussed, we suspect that LVP may be more widespread than reported here—as it is so difficult to distinguish from language shift, and as the pattern is counter-intuitive, it is quite likely that it has gone unnoticed in other communities around the world. We imagine that at least some readers will be familiar with the patterns of LVP discussed above from their own communities, or the communities with which they work. If this is the case, we encourage documentation and description of the acquisition patterns, attitudes, and ideologies in these small-scale multilingual ecologies, to provide the data necessary to deepen our understanding of this little-known phenomenon. With a more comprehensive understanding of LVP, we hope that insights can be integrated into future language maintenance and revitalisation programmes.

We have raised far more questions than we have answered here. LVP bilingualism is fertile ground for investigation, which we hope future research programmes will pick up on and explore further. With regards to the implications of LVP for theories of bilingual acquisition and use, key questions include: (i) Given that LVP bilinguals pattern with both simultaneous and sequential bilinguals, is their eventual vernacular production more L1-like (as with simultaneous bilinguals), more L2-like (as with late sequential bilinguals), or neither? (ii) Which linguistic features can be acquired through an early period of passive learning followed by a later period of active production? (iii) Related to this, which linguistic features can only be acquired through active production in early childhood? (iv) How can insights from LVP be integrated into bilingual curricula, for example by shedding light on which areas of grammar might need extra attention at a given stage (e.g. Meakins & Green & Turpin, 2018: 219)? (v) How might changes in language policy with an increased amount of vernacular instruction impact LVP?

In the field of contact, outstanding questions include: (i) Does LVP itself introduce innovations into a speech community—primarily into the vernacular, but also the lingua franca? (ii) If so, how are these propagated and spread? (iii) Are any changes introduced similar to those expected from childhood bilinguals, adult learners, or quantitatively and/or qualitatively different from either? (iv) If changes are introduced through LVP, are these maintained through the lifespan, or are they overturned through increased exposure to the vernacular in adulthood?

Finally, another current unknown is whether a linguistic ecology involving LVP can be stable—by which we mean, can persist with little fluctuation in the demographics of language use for several generations or more—or whether a language with widespread LVP is necessarily on the path to language shift. In principle, we are inclined to believe that such situations are potentially stable. This is suggested by the attestation of patterns similar to LVP in more traditional contexts, where the language in which the younger generations are dominant is a local lingua franca (for example, the case of Iau described at the end of

Section 4, where youngsters prefer the local trade language Elopi until reaching adolescence). In practice, however, it seems likely that any community with LVP in today's rapidly globalising world will, without intervention, become endangered within the next few generations. We therefore encourage further study of both these underdocumented linguistic ecologies, and the languages themselves, while they are still vital.

Author contributions

George Saad was responsible for data collection, analysis, and drafting Section 3. Laura Arnold and Emma Peddie were responsible for data collection, analysis, and drafting Section 4. All authors share responsibility for paper concept, drafting sections 1, 2, 5, 6, and 7, and paper revisions.

Acknowledgements

We would like to acknowledge our collaboration with the Abui community, who have been working with the first author in trying to understand the root causes of LVP as well as work towards implementing solutions to mitigate some of its effects. In addition, we would like to extend our sincere thanks to Harald Hammarström and Randy Lebold for generously supplying many unpublished sources that were used in the survey; Simon Evans, Emelihter Kihleng, and Ron Whisler, for providing additional information on local language use; participants at the Risk and Renewal in the Pacific seminar series at the University of Cambridge, the 51st Poznań Linguistic Meeting, the 2022 OCSEAN International Conference, the 14th International Symposium on Bilingualism, the Sixth Workshop on the Languages of Papua, the 15th Austronesian and Papuan Languages and Linguistics conference, the 56th Annual Meeting of the Societas Linguistica Europaea, and the Bilingualism and Developmental Linguistics Research Group at the University of Edinburgh, where aspects of this research were presented;

František Kratochvíl and Don Kulick for kindly offering such helpful comments on earlier drafts of this paper; and, at *Glossa Contact*, three anonymous reviewers and editors Nina Dobrushina and Igor Yanovich for their constructive criticism. All errors and misinterpretations are our own.

Abbreviations

EGIDS: Expanded Graded Intergenerational Disruption Scale

ISEA: Island South East Asia and the Pacific

L1: First language

L2: Second language

LEI: Language Endangerment Index

LVP: Late Vernacular Production

WEIRD: Western, Educated, Industrialised, Rich, and Democratic

Data accessibility statement

The data presented in Section 3 comes from the first author's field research. It consists of recorded sociolinguistic interviews as well as a spreadsheet file that charts the responses to the questions. Access to the data can be obtained by contacting the first author. Most of the data provided in Section 4 comes from published or unpublished manuscripts as cited.

Ethics and consent

The first batch of data from 2015-2016 was collected in compliance with the ethics provisions of Leiden University, while the second batch of data collected in 2022 was collected in compliance with the ethics provisions of Palacký University in Olomouc. Both universities have a comprehensive Ethics Codex and a rigorous internal approval process for research involving human subjects. At both universities, data managers and data stewards were consulted regarding the design of the study, the nature of the data, and the contents of the consent forms. The consent

forms were approved before the execution of the study. For the participants that took part in the study in section 3, consent was collected, both orally and in the form of consent forms. Guardians were consulted for participants under the age of 18.

Funding information

This work was supported by the OP JAC Project “MSCA Fellowships at Palacký University I." CZ.02.01.01/00/22_010/0002593, hosted at Palacký University Olomouc, the Czech Republic, “Verb class analysis accelerator for low-resource languages – RoboCorp”, 20-18407S, also hosted at Palacký University Olomouc, the VICI research project “Reconstructing the past through languages of the present: The Lesser Sunda Islands” awarded to Marian Klamer by the Netherlands Organization for Scientific Research (NWO), project number 277-70-012, hosted at Leiden University, and a British Academy Postdoctoral Fellowship, PF19\100004, hosted at the University of Edinburgh.

Competing interests

The authors have no competing interests.

References

Adamou, Evangelia (2021). *The adaptive bilingual mind: Insights from endangered languages*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/9781108884266

Anand, Pranav & Chung, Sandra & Wagers, Matthew (2011). *Widening the net: Challenges for gathering linguistic data in the digital age*. Response to NSF SBE 2020: Future Research in the Social, Behavioral & Economic Sciences. Retrieved from: https://people.ucsc.edu/%7Eschung/anandchungwagers.pdf, accessed 2023-10-25.

Anderbeck, Karl, (2015). Portraits of language vitality in the languages of Indonesia. In Arka, I Wayan & Seri Malini, L. N. & Puspani, I.A.M (eds.), *Language documentation and cultural practices in the Austronesian world: Papers from ICAL 12 (A-PL 019)*, 19–47. Canberra: Pacific Linguistics.

Anonby, Stanley (2020). *Prolonged multilingualism among the Sebuyau: An ethnography of communication*. PhD dissertation, Simon Fraser University.

Araali, Bagamba & Boone, Douglas (2011). *Challenges to applying the EGIDS in northeastern Democratic Republic of Congo*. Presented at the colloquium *Ecological perspectives on language endangerment: Applying the sustainable use model for language development*; American Association of Applied Linguistics, Chicago, IL.

Arnold, Laura (forthcoming). Island Southeast Asia and the Pacific. In Mosley, Christopher (ed.), *Encyclopedia of the world's endangered languages, second edition*. Routledge.

Baird, Louise (2017). Kafoa. In Schapper, Antoinette (ed.), *The Papuan languages of Timor, Alor, and Pantar: Volume 2*, 55–108. Boston/Berlin: Mouton de Gruyter. DOI: https://doi.org/10.1515/9781614519027-002

Baird, Louise & Klamer, Marian & Kratochvíl, František (in prep.) *Alor Malay: A grammar sketch of an emerging Eastern variety*. Manuscript, Leiden University.

Barlow, Russell (2018). *A grammar of Ulwa*. PhD dissertation, University of Hawai'i at Mānoa.

Beyer, Klaus (2010). Language contact and change: A look at social factors in an African rural environment. *Journal of Language Contact* 3(1), 131–152.
DOI: https://doi.org/10.1163/19552629-90000008

Bi, Jin (2017). *Bilingualism and cognitive control: A comparison of sequential and simultaneous bilinguals*. PhD dissertation, University of Utah.

Bowerman, Melissa (2011). Linguistic typology and first language acquisition. In Song, J. J. (ed.), *The Oxford handbook of linguistic typology*, 591–617. Oxford: Oxford University Press.
DOI: https://doi.org/10.1093/oxfordhb/9780199281251.013.0028

Bromham, Lindell & Dinnage, Russell & Skirgård, Hedvig & Ritchie, Andrew & Cardillo, Marcel & Meakins, Felicity & … Hua, Xia (2021). Global predictors of language endangerment and the future of linguistic diversity. *Nature Ecology and Evolution* 6, 163–173.
DOI: https://doi.org/10.1038/s41559-021-01604-y

Brown, Penelope & Gaskins, Suzanne (2014). Language acquisition and language socialization. In Enfield, N. J. & Kockelman, Paul & Sidnell, Jack (eds.), *The Cambridge handbook of linguistic anthropology*, 187–226. Cambridge: Cambridge University Press.
DOI: https://doi.org/10.1017/CBO9781139342872.010

Campbell, Lyle & Lee, Nala Huiying & Okura, Eve & Simpson, Sean & Ueki, Kaori (2023). *The catalogue of endangered languages (ElCat)* [Data set]. http://endangeredlanguages.com/userquery/download/, accessed 2023-10-19.

Casillas, Marisa & Brown, Penelope & Levinson, Stephen C. (2020). Early language experience in a Tseltal Mayan village. *Child Development 91(5)*, 1819–1835. DOI: https://doi.org/10.1111/cdev.13349

Casillas, Marisa & Brown, Penelope & Levinson, Stephen C. (2021). Early language experience in a Papuan community. *Journal of Child Language 48(4)*, 792–814. DOI: https://doi.org/10.1017/S0305000920000549

Clifton, John M. (1994). Stable multilingualism in a small language group: The case of Kaki Ae. *Language and Linguistics in Melanesia* 25, 107–124.

Cohen, Andrew D. (1996). Verbal reports as a source of insights into second language learner strategies. *Applied Language Learning 7(1)*, 11–27.

De Houwer, Annick (1995). Bilingual language acquisition. In Fletcher, Paul & MacWhinney, Brian (eds.), *The handbook of child language*, 219–250. Oxford: Blackwell.
DOI: https://doi.org/10.1111/b.9780631203124.1996.00009.x

Di Carlo, Pierpaolo & Good, Jeff & Ojong Diba, Rachel (2019). Mulitlingualism in rural Africa. In *Oxford research encyclopedia of linguistics.* Oxford: Oxford University Press. Retrieved from https://doi.org/10.1093/acrefore/9780199384655.013.227

Dixon, R.M.W. (1991). The endangered languages of Australia, Indonesia and Oceania. In Robins, Robert H. & Uhlenbock, Eugenius M. (eds.), *Endangered languages*, 229–256. Oxford/New York: Berg.

Donohue, Mark (2004). *A grammar of the Skou language of New Guinea*. Retrieved from:
https://pure.mpg.de/pubman/faces/ViewItemOverviewPage.jsp?itemId=item_402710, accessed 2023-11-09.

Du Bois, Cora (1944). *The people of Alor: A social-psychological study of an East Indian island*. Minneapolis: The University of Minnesota Press.

Eberhard, David M. & Šimičić, Lucija (2023). *Delayed Vernacular Acquisition among minoritized language speakers: Preliminary findings from Sãotomense, Angolar, and Molise Croatian communities*. Presented at the Presented at the 5th Variation in Language Acquisition Conference, Brussels, Belgium, June 1-2.

Eberhard, David M. & Simons, Gary F. & Fennig, Charles D. (eds.) (2022). *Ethnologue: Languages of the world, twenty-fifth edition*. Dallas, Texas: SIL International. Retrieved from http://www.ethnologue.com

Flege, James Emil & Fletcher, Kathryn L. (1992). Talker and listener effects on degree of perceived foreign accent. *The Journal of the Acoustical Society of America* 91(1), 370–389. DOI: https://doi.org/10.1121/1.402780

Foley, William A. (2005). Linguistic prehistory in the Sepik-Ramu basin. In Pawley, Andrew & Attenborough, Robert & Golson, Jack & Robin Hide (eds.), *Papuan pasts: cultural, linguistic and biological histories of Papuan-speaking peoples*, 109–144. Canberra: Pacific Linguistics.

François, Alexandre (2012). The dynamics of linguistic diversity: Egalitarian multilingualism and power imbalance among northern Vanuatu languages. *International Journal of the Sociology of Language* 214, 85–110. DOI: https://doi.org/10.1515/ijsl-2012-0022

Gathercole, Virginia C. Mueller & Moawad, Ruba Abdelmatloub (2010). Semantic interaction in early and late bilinguals: All words are not created equally. Bilingualism: Language and Cognition 13(4). 385–408. DOI: https://doi.org/10.1017/S1366728909990460

Gil, David (2024). Multilingualism. In Adelaar, Alexander K. & Schapper, Antoinette (eds.), *The Oxford guide to the Malayo-Polynesian languages of Southeast Asia*, 205-223. Oxford: Oxford University Press. DOI: 10.1093/oso/9780198807353.003.0016

Goodman, Elizabeth & Bates, Judith C. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia and real-time processing. *Language and Cognitive Processes* 12(5-6), 507–584. DOI: https://doi.org/10.1080/016909697386628

Green, David W. (2003). The neural basis of the lexicon and the grammar in L2 acquisition. In van Hout, Roeland & Hulk, Aafke & Kuiken, Folkert & Towel, Richard (eds.), *The interface between syntax and*

the lexicon in second language acquisition, 197–208. Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/lald.30.10gre

Hale, Ken & Krauss, Michael & Watahomigie, Lucille J. & Yamamoto, Akira Y. & Craig, Colette & Jeanne, LaVerne Masayesva & & England, Nora C. (1992). Endangered Languages. *Language* 68(1), 1–42. DOI: https://doi.org/10.2307/416368

Hammarström, Harald & Forkel, Robert & Haspelmath, Martin & Bank, Sebastian (2021). *Glottolog 4.4*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Hammarström, Harald & Forkel, Robert & Haspelmath, Martin & Bank, Sebastian. 2023. *Glottolog 4.8*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Henrich, Joseph & Heine, Steven J. & Norenzayan, Ara (2010). *Most people are not WEIRD*. *Nature* 466, 29.
DOI: https://doi.org/10.1038/466029a

Kaiping, Gereon A. & Klamer, Marian (2022). The dialect chain of the Timor-Alor-Pantar language family: A new analysis using systematic Bayesian phylogenetics. *Language Dynamics and Change* 12(2), 274–326. DOI: https://doi.org/10.1163/22105832-bja10019

Kashima, Eri (2020). Language In my mouth: Linguistic variation in the Nmbo speech community of southern New Guinea. PhD dissertation, Australian National University.

Kelly, Barbara F. & Forshaw, William & Nordlinger, Rachel & Wigglesworth, Gillian (2015). Linguistic diversity in first language acquisition research: Moving beyond the challenges. *First Language* 35(4–5), 286-304. DOI: https://doi.org/10.1177/0142723715602350

Kerswill, Paul (1996). Children, adolescents, and language change. *Language Variation and Change* 8(2), 177–202.
DOI: https://doi.org/10.1017/S0954394500001137

Kidd, Evan & Garcia, Rowena (2022). How diverse is child language acquisition research? *First Language* 42(6), 703–735. DOI: https://doi.org/10.1177/01427237211066405

Kim, Ah-Young (Alicia) & Park, Anne & Lust, Barbara (2016). Simultaneous vs. successive bilingualism among preschool-aged children: A study of four-year-old Korean–English bilinguals in the USA. *International Journal of Bilingual Education and Bilingualism* 21(2), 164–178. DOI: https://doi.org/10.1080/13670050.2016.1145186

Klamer, Marian (2017). The Alor-Pantar languages: Linguistic context, history and typology. In Klamer, Marian (ed.), *The Alor-Pantar languages: History and typology, 2nd edition*, 1–48. Berlin: Language Science Press.

Klamer, Marian & Saad, George (2020). Reduplication in Abui: A case of pattern extension. *Morphology* 30(4), 311–346. DOI: https://doi.org/10.1007/s11525-020-09369-z

Kratochvíl, František (2007). *A grammar of Abui: A Papuan language of Alor*. PhD dissertation, Leiden University.

Kulick, Don & Dobrin, Lisa M. (forthcoming). Exchanging language: Explaining language shift in Papua New Guinea. In Evans, Nicholas & Fedden, Sebastian (eds.), *The Oxford Guide to the Papuan Languages*. Oxford: Oxford University Press.

Kulick, Don & Terrill, Angela (2019). A grammar and dictionary of Tayap: The life and death of a Papuan language. Boston/Berlin: Mouton de Gruyter. DOI: https://doi.org/10.1515/9781501512209

Lambrecht, Philip & Kassell, Alison & Potter, Margaret & Tucker, Sarah (2008). The sociolinguistic situation of the Uyajitaya [duk] language. Retrieved from https://www.sil.org/resources/archives/9137, accessed 2023-10-25.

Lambrecht, Philip & Kassell, Alison & Potter, Margaret & Tucker, Sarah (2009). A sociolinguistic profile of the Waube [kop] language group. Retrieved from

https://www.sil.org/resources/publications/entry/9140, accessed 2023-10-25.

Landweer, M. Lynn (1998). Indicators of ethnolinguistic vitality: Case study of two languages: Labu and Vanimo. In Ostler, Nicholas (ed.), *Proceedings of the second FEL conference: Endangered languages – what role for the specialist? (Edinburgh, Scotland, September 25-27, 1998)*, 64–72. Bath: Foundation for Endangered Languages.

Landweer, M. Lynn (2012). Methods of language endangerment research: A perspective from Melanesia. *International Journal of the Sociology of Language* 214, 153–178. DOI: https://doi.org/10.1515/ijsl-2012-0024

Lee, Hope & Sawi, Agustina (2005). Draft survey report on the Sobei related languages in Sarmi county, northeastern Papua, Indonesia. Unpublished manuscript, SIL Electronic Survey Reports.

Lee, Nala H. & Van Way, John R. (2016). Assessing levels of endangerment in the Catalogue of Endangered Languages (ElCat) using the Language Endangerment Index (LEI). *Language in Society* 45(2), 271–292. DOI: https://doi.org/10.1017/S0047404515000962

Lee, Nala H. & Van Way, John R. (2018). Assessing degrees of language endangerment. In *The Oxford handbook of endangered languages*, 48–65. Oxford: Oxford University Press.
DOI: https://doi.org/10.1093/oxfordhb/9780190610029.013.5

Lee, Sang-kem & Wambaliau, Theresia (2004). Draft survey report on the Kwanem languages in northeastern Papua, Indonesia. Unpublished manuscript, SIL Electronic Survey Reports.

Lewis, M. Paul & Simons, Gary F. (2010). Assessing endangerment: Expanding Fishman's GIDS. *Revue Roumaine de Linguistique* 55(2), 103–120.

Lüpke, Friederike (2016). Uncovering small-scale multilingualism. *Critical Multilingualism Studies* 4(2), 35–74.

Lüpke, Friederike (2021). Patterns and perspectives shape perceptions: Epistemological and methodological reflections on the study of small-scale multilingualism. *International Journal of Bilingualism* 25(4), 878–900. DOI: https://doi.org/10.1177/13670069211023145

Lust, Barbara C. (2006). Child language: Acquisition and growth. Cambridge: Cambridge University Press.
DOI: https://doi.org/10.1017/CBO9780511803413

McConvell, Patrick (1991). Understanding language shift: A step towards language maintenance. In Romaine, Suzanne (ed.), *Language in Australia*, 143–155. Cambridge: Cambridge University Press.
DOI: https://doi.org/10.1017/CBO9780511620881.010

McLaughlin, Barry (1978). *Second language acquisition in childhood.* Hillsdale, NJ: Lawrence Erlbaum.

Meakins, Felicity & Green, Jennifer & Turpin, Myfany (2018). *Understanding linguistic fieldwork*. London/New York: Routledge.
DOI: https://doi.org/10.4324/9780203701294

Meisel, Jürgen M. (2009). Second language acquisition in early childhood. *Zeitschrift für Sprachwissenschaft* 28, 5–34.
DOI: https://doi.org/10.1515/ZFSW.2009.002

Morin, Izak (2016). Marginalizing and revaluing Papuan Malay: The impact of politics, policy and technology in Indonesia. In *Margins, Hubs, and Peripheries in a Decentralizing Indonesia*, 101–111.

Moseley, Christopher (ed.) (2010). *Atlas of the world's languages in danger, 3rd ed*. Paris: UNESCO Publishing.

Mühlhäusler, Peter (2006). Naming languages, drawing language boundaries and maintaining languages with special reference to the linguistic situation in Papua New Guinea. In Cunningham, Denis & Ingram, David E. & Sumbuk, Kenneth (eds.), *Language diversity in the Pacific: Endangerment and survival*, 24–39. Clevedon: Multilingual Matters Ltd. DOI: https://doi.org/10.21832/9781853598685-006

Nazarudin (2021). Review of “Austronesian undressed.” Presented at the book discussion “Austronesian undressed: How and why languages become isolating.” Retrieved from https://www.youtube.com/watch?v=2J2qCc7SFKY, accessed 2023-11-09.

O’Shannessy, Carmel & Davidson, Lucinda (2020). Language contact and change through child first language acquisition. In Hickey, Raymond (ed.), *The handbook of language contact, second edition*, 67–91. Hoboken, NJ: Wiley-Blackwell.
DOI: https://doi.org/10.1002/9781119485094.ch3

Paauw, Scott H. (2008). *The Malay contact varieties of eastern Indonesia: A typological comparison*. PhD dissertation, State University of New York, Buffalo.

Pakendorf, Brigitte & Dobrushina, Nina & Khanina, Olesya (2021). A typology of small-scale multilingualism. *International Journal of Bilingualism* 25(4), 835–859.
DOI: https://doi.org/10.1177/13670069211023137

Paradis, Johanne (2007). Early bilingual and multilingual acquisition. In Auer, Peter & Wei, Li (eds.), *Handbook of multilingualism and multilingual communication*, 15–44. Berlin: Mouton de Gruyter.
DOI: https://doi.org/10.1515/9783110198553.1.15

Pearson, Barbara Zurer (2009). Children with two languages. In Bavin, Edith L. (ed.), *The Cambridge handbook of child language*, 379–397. Cambridge: Cambridge University Press.
DOI: https://doi.org/10.1017/CBO9780511576164.021

Peddie, Emma (2021). Delayed vernacular production in Melanesia: A case study-based analysis. MA dissertation, University of Edinburgh.

Polinsky, Maria (2015). Heritage languages and their speakers: State of the field, challenges, perspectives for future work, and methodologies. *Zeitschrift für Fremdsprachwissenschaft*. Retrieved from https://dash.harvard.edu/handle/1/37108756

Ridge, Eleanor (2019). Variation in Vatlongos verbal morphosyntax: speaker communities in Southeast Ambrym and Mele Maat. PhD dissertation, University of London (SOAS).

Ross, Malcolm (2013). Diagnosing contact processes from their outcomes: The importance of life stages. *Journal of Language Contact* 6(1), 5–47. DOI: https://doi.org/10.1163/19552629-006001002

Saad, George (2020). Variation and change in Abui: The impact of Alor Malay on an indigenous language of Indonesia. PhD dissertation, Leiden University.

Saad, George (2023). Semantic variation in a bilingual speech community. In Klamer, Marian & Moro, Francesca (eds.), *Traces of contact in the lexicon: Austronesian and Papuan studies*, 392–420. Leiden: Brill. DOI: https://doi.org/10.1163/9789004529458_012

Saad, George & Klamer, Marian & Moro, Francesca (2019). Identifying agents of change: Simplification of possessive marking in Abui-Malay bilinguals. *Glossa: A Journal of General Linguistics* 4(1), 1–29. DOI: https://doi.org/10.5334/gjgl.846

Sankoff, Gillian (1980). Multilingualism in Papua New Guinea. In Sankoff, Gillian (ed.), *The social life of language*, 95–132. Pennsylvania: University of Pennsylvania Press.
DOI: https://doi.org/10.9783/9781512809589-008

Schapper, Antoinette & Huber, Juliette & van Engelenhoven, Aone (2017). The relatedness of Timor-Kisar and Alor-Pantar languages: A preliminary demonstration. In Klamer, Marian (ed.), *The Alor-Pantar languages: History and typology, 2nd edition*, 99–154. Berlin: Language Science Press.

Schneider, Katharina (2018). Matrilineal kinship at sea in Bougainville, PNG. *Humaniora* 30(3), 223–236.
DOI: https://doi.org/10.22146/jh.39083

Shneidman, Laura A. & Goldin-Meadow, Susan (2012). Language input and acquisition in a Mayan village: How important is directed speech?

Developmental Science 15(5), 659–673.
DOI: https://doi.org/10.1111/j.1467-7687.2012.01168.x

Sloan, Bodean & Meakins, Felicity & Algy, Cassandra (2022). Intergenerational changes in Gurindji Kriol: Comparing apparent-time and real-time data. *Asia-Pacific Language Variation* 8(1), 1–31. DOI: https://doi.org/10.1075/aplv.21001.slo

Spencer, Juliann & Cott, Sara Van & MacKenzie, Bonnie (2013). A sociolinguistic survey of the Amio-Gelimi language area. Retrieved from https://www.sil.org/resources/archives/54624,
accessed 2023-11-09.

Stanford, James N. (2009). Clan as a sociolinguistic variable: Three approaches to Sui clans. In Stanford, James N. & Preston, Dennis R. (eds.), *Variation in indigenous minority languages*, 463–484. Amsterdam/Philadelphia: John Benjamins.
DOI: https://doi.org/10.1075/impact.25.23sta

Stokhof, Willem Arnoldus Laurens (1984). Annotations to a text in the Abui language (Alor). *Bijdragen Tot de Taal-, Land- En Volkenkunde* 140, 106–162. DOI: https://doi.org/10.1163/22134379-90003430

Suastra, Made (2022). *Roadmap penelitian bahasa ibu*. Denpasar: Suasta Nulus.

Thomason, Sarah (2010). Contact explanations in linguistics. In Hickey, Raymond (ed.), *The handbook of language contact*, 31–47. Oxford: Blackwell. DOI: https://doi.org/10.1002/9781444318159.ch1

Thurston, William R. (1992). Sociolinguistic typology and other factors effecting change in northwestern New Britain, Papua New Gunea. In *Culture change, language change: Case studies from Melanesia*, 123–139. Canberra: Pacific Linguistics.

Trudgill, Peter (2009). Sociolinguistic typology and complexification. In Sampson, Geoffrey & Gil, David & Trudgill, Peter (eds.), *Language complexity as an evolving variable*, 98–109. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/oso/9780199545216.003.0007

UNESCO Ad Hoc Expert Group on Endangered Languages (2003). *Language vitality and endangerment*. Retrieved from https://ich.unesco.org/doc/src/00120-EN.pdf, accessed 2025-04-09.

Vaughan, Jill & Singer, Ruth (2018). Indigenous multilingualisms past and present. *Language & Communication* 62(B), 83–90. DOI: https://doi.org/10.1016/j.langcom.2018.06.003

Williams, Nicholas J. (2017). Kula. In Schapper, Antoinette (ed.), *The Papuan languages of Timor, Alor, and Pantar: Volume 2*, 185–266. Boston/Berlin: Mouton de Gruyter.
DOI: doi.org/10.1515/9781614519027-004

Wurm, Stephen Adolphe & Hattori, Shiro (1981). Language atlas of the Pacific area: New Guinea area, Oceania, Australia; 2, Japan area, Taiwan-Formosa, Philippines, Mainland and insular South-East Asia. Canberra: Pacific Linguistics.

Appendix 1

Abbreviations: PNG = Papua New Guinea, TNG = Trans-New Guinea, AN = Austronesian, SHWNG = South Halmahera-West New Guinea

	Vernacular	Alt. name	Glottocode	Family	Country	Region	Lingua franca	Source	Type of evidence
1.	Sebuyau	--	sebu1243	AN > Malayo-Sumbawan	Malaysia	Sarawak	Malaysian	Anonby (2020)	Longitudinal observation
2.	West Coast Bajau	--	west2560	AN > Basap-Greater Barito	Malaysia	Sabah	Malaysian	Eberhard et al. (2022)	Unknown
3.	Balinese	--	bali1278	AN > Bali-Sasak-Sumbawa	Indonesia	Bali	Malay/ Indonesian	Suastra (2022), Simon Evans (pers. comm.)	Longitudinal observation (anecdotal)
4.	Kula	--	kelo1247	Timor-Alor-Pantar	Indonesia	Nusa Tenggara Timur	Malay/ Indonesian	Williams (2017)	Community report
5.	Kafoa	--	kafo1240	Timor-Alor-Pantar	Indonesia	Nusa Tenggara Timur	Malay/ Indonesian	George Saad (fieldnotes), Baird (2017)	Community report
6.	Abui	--	abui1241	Timor-Alor-Pantar	Indonesia	Nusa Tenggara Timur	Malay/ Indonesian	Kratochvíl (2007), Saad et al. (2019), Saad (2020)	Longitudinal observation (systematic)
7.	Kamang	--	kama1365	Timor-Alor-Pantar	Indonesia	Nusa Tenggara Timur	Malay/ Indonesian	George Saad (fieldnotes)	Commmunity report
8.	Oirata	--	oira1263	Timor-Alor-Pantar	Indonesia	Maluku	Malay/ Indonesian	Nazarudin (2021)	Community report
9.	Sawai	--	sawa1247	AN > SHWNG	Indonesia	North Maluku	Malay/ Indonesian	Ron Whisler (pers. comm.)	Longitudinal observation (anecdotal)
10.	Ansus	--	ansu1237	AN > SHWNG	Indonesia	Papua	Malay/ Indonesian	Donohue (2004)	Unknown
11.	Saweru	--	sawe1240	Yawa-Saweru	Indonesia	Papua	Malay/ Indonesian	Donohue (2004)	Longitudinal observation (anecdotal)
12.	Warembori	--	ware1253	AN > SHWNG	Indonesia	Papua	Malay/ Indonesian	Donohue (2004)	Unknown
13.	Dabe	--	dabe1239	Tor-Orya	Indonesia	Papua	Malay/ Indonesian	Lee & Wambaliau (2004)	Community report
14.	Vitou	Takar	vito1235	Tor-Orya	Indonesia	Papua	Malay/ Indonesian	Lee & Wambaliau (2004)	Community report
15.	Edwas	Bonerif	bone1255	Tor-Orya	Indonesia	Papua	Malay/ Indonesian	Lee & Wambaliau (2004)	Community report
16.	Sunum	Yamna	yamn1237	AN > Oceanic	Indonesia	Papua	Malay/ Indonesian	Lee & Sawi (2005)	Community report
17.	Kwinsu	--	kwin1242	Tor-Orya	Indonesia	Papua	Malay/ Indonesian	Lee & Wambaliau (2004)	Community report
18.	Sentani	--	nucl1632	Sentanic	Indonesia	Papua	Malay/ Indonesian	Dwight Hartzler, cited in Anderbeck (2015)	Longitudinal observation (anecdotal)
19.	Skou	--	nucl1634	Sko	Indonesia	Papua	Malay/ Indonesian	Donohue (2004)	Longitudinal observation

									(anecdotal)
20.	Tayap	Taiap	taia1239	Isolate	PNG	East Sepik	Tok Pisin	Kulick & Terrill (2019)	Longitudinal observation (systematic)
21.	Kein	Bemal	kein1239	TNG > Madang	PNG	Madang	Tok Pisin	Eberhard et al. (2022)	Unknown
22.	Bau	Fulumu	bauu1244	TNG > Madang	PNG	Madang	Tok Pisin	Eberhard et al. (2022)	Unknown
23.	Uyajitaya	Duduel a	dudu1240	TNG > Madang	PNG	Madang	Tok Pisin	Lambrecht et al. (2008)	Community report
24.	Waube	Kwato	kwat1244	TNG > Madang	PNG	Madang	Tok Pisin	Lambrecht et al. (2009)	Community report
25.	Forak	--	fora1245	TNG > Finisterre-Huon	PNG	Madang	Tok Pisin	Eberhard et al. (2022)	Unknown
26.	Sarasira	--	sara1323	An > Oceanic	PNG	Morobe	*(Not mentioned)*	Eberhard et al. (2022)	Unknown
27.	Labu	--	labu1248	An > Oceanic	PNG	Morobe	Tok Pisin	Landweer (1998)	Community report
28.	Bebeli	--	bebe1252	An > Oceanic	PNG	West New Britain	Tok Pisin	Spencer & Cott & MacKenzie (2013)	Community report
29.	Kandas	--	kand1301	An > Oceanic	PNG	New Ireland	*(Not mentioned)*	Eberhard et al. (2022)	Unknown
30.	Petats	--	peta1245	An > Oceanic	PNG	Bougainvill e	Tok Pisin	Schneider (2018)	Unknown
31.	Chuukese	--	chuu1238	An > Oceanic	Communities on Guam and Hawaii		English	Emelihter Kihleng (pers. comm.)	Longitudinal observation (anecdotal)
32.	Vatlongos	--	sout2859	An > Oceanic	Vanuatu	Community near Port Vila	Bislama	Ridge (2019)	Community report