Laura Arnold*

Split Inalienable Coding in linguistic Wallacea: typology, origins, spread

Abstract: This paper is a typological survey of inalienable possessive constructions in the linguistic area of Wallacea and its surrounds. In a sample of 189 Austronesian and non-Austronesian languages, 13 have a phenomenon not previously recognised in the theoretical or typological literature: Split Inalienable Coding (SIC), whereby a language has two or more possessive coding strategies that are closely or exclusively associated with expressing inalienable possession. This paper focusses on semantically conditioned splits, where minimally one strategy encodes the possession of body parts, and another the possession of kin terms. Geographically, all of the sampled languages with semantic SIC are located in Wallacea; special attention is therefore given to the development of split inalienables in this region. In most of these languages, SIC has developed very recently. I argue that there have been multiple causes of SIC: Austronesian languages are predisposed to develop SIC, due to the inheritance of a structurally defined class of kin terms that favours the distinction; and contact has also played a role in Northwest New Guinea, with SIC diffusing both across and within genealogical groupings.

Keywords: alienability; attributive possession; inalienable possession; language contact; language typology

1 Introduction

In many languages of the world, there is a formal split in adnominal possessive coding, roughly coinciding with the semantic notion of ‘alienability’: One coding strategy is used to express alienable possession, the other inalienable possession.1

1 Possessive coding may also be conditioned by factors such as phonology, syntax, or information structure. However, these and other factors are not typically associated with alienability (Nichols and Bickel 2013), and are thus not discussed further in this paper. In addition, this study only looks at adnominal possessive strategies, i.e. the NP-internal expression of possession (also called ‘attributive’ possession). While alienability distinctions and Split Inalienable Coding, as defined below, are also

*Corresponding author: Laura Arnold, University of Edinburgh, Edinburgh, UK, E-mail: laura.arnold@ed.ac.uk

Open Access. © 2023 the author(s), published by De Gruyter. This work is licensed under the Creative Commons Attribution 4.0 International License.
This paper investigates coding strategies associated with inalienable possession in the linguistic area of Wallacea, which encompasses much of eastern Indonesia, as well as independent Timor-Leste: the eastern Lesser Sunda islands, Maluku, and parts of western New Guinea (Schapper 2015).

An example of an alienability split is given in (1), with data from Sougb, an East Bird’s Head language of Northwest New Guinea. There are two adnominal possessive coding strategies in Sougb. In both strategies, the possessor NP is ordered before the possessed NP. However, in one, the person and number of the possessor is marked with prefixes on the possessed noun, as in (1a); and in the other, exemplified in (1b), the possessor is marked with prefixes on a possessive marker *an*.\(^2\)

(1) Sougb (East Bird’s Head)

a. dan ind-ums
   1SG 1SG-ear
   ‘my ear’
   (Reesink 2002a: 217)

b. dan ind-an tu
   1SG 1SG-POSS house
   ‘my house’
   (Reesink 2002a: 218)

The choice between the two strategies in Sougb is conditioned by the possessed noun. The prefixing construction is used when the possessed noun is either a kin term or a body part; as will be set out in more detail below, the relationship between the referents of these nouns and their possessors is a classic example of an inalienable relationship, in that the relationship is permanent and not transferable under ordinary circumstances. Applying this terminology to the formal level, the prefixing construction in Sougb can also be referred to as the inalienable construction. The second strategy, with the inflected possessive marker *an*, is used for all other possessive relationships, such as ownership, association, and creation. The relationship between the referents of the possessed and possessor nouns in these constructions is

---

\(^3\)In the examples throughout this paper, the possessor is pronominal. This is for ease of presentation; and because in the sources consulted, pronominal possessors are normally more comprehensively exemplified. While in some languages of the region, possessive marking is dependent on whether the possessor is nominal or pronominal, this is not the case of any of the languages discussed below – the same constructions, including any formal splits, are used regardless. Glossing throughout this paper follows the Leipzig Glossing rules (https://www.eva.mpg.de/lingua/resources/glossing-rules.php).

---
thus alienable, in that it is more temporary and easily severable. This construction can therefore be referred to as the alienable construction.

Linguistic Wallacea is a region of high genealogical diversity, with languages belonging to at least four sub-branches of the Austronesian (An) family, ten unrelated non-Austronesian (NAn) families, and seven NAn isolates. The prevalence of alienability distinctions in the An and NAn languages of Melanesia has long been recognised (e.g. Lévy-Bruhl 1914, cited in Chappell and McGregor 1996: 3). In particular, alienability in and around Wallacea has received a high level of attention when compared with other linguistic features. This is due to the geographic position of Wallacea between an area dominated by An languages to the west, which generally do not have alienability distinctions, and the NAn languages of Melanesia to the east, where alienability distinctions are common. The frequent attestation of alienability distinctions in the An languages from Wallacea eastwards is therefore widely accepted to be due to contact with NAn languages (e.g. Donohue and Schapper 2008; Klamer et al. 2008: 122; Ross 2001: 138). Areal studies by Klamer et al. (2008: 116–122) and Schapper (2015: 108–110) have investigated the geographic and genealogical distributions of alienability distinctions across this region. In addition, comparative work on An languages has shown that the distinction was present in the proto-languages of the South Halmahera-West New Guinea (SHWNG) and Oceanic (Oc) sub-branches (van den Berg 2009; Lichtenberk 1985); whereas in the Central Malayo-Polynesian (CMP) languages, the distinction was innovated multiple times (Donohue and Grimes 2008: 142–143).4

The literature to date has focused on the semantics and formal expression of the distinction between alienable and inalienable possession. However, on closer inspection, some languages of Wallacea have an additional split in possessive coding: two or more strategies that are closely or exclusively associated with the expression of inalienable possession. Henceforth, this phenomenon will be referred to as ‘Split Inalienable Coding’ (SIC). To the best of my knowledge, SIC has not previously been discussed in either the typological or the theoretical literature.

In this paper, I focus on SIC where the split is conditioned by the semantic class of the possessed noun. In semantically conditioned SIC, minimally one coding strategy

---

3 Genealogical classifications throughout this paper are taken from Glottolog (Hammarström et al. 2021). The label ‘non-Austronesian’ (or ‘Papuan’) is defined negatively, in that it refers to all of the genealogically diverse languages spoken in Melanesia which do not belong to the Austronesian family.

4 The status of Central Malayo-Polynesian as a genealogical subgroup of Austronesian is controversial, in that the CMP languages are not considered to be the descendants of a single unified proto-language to the exclusion of other Malayo-Polynesian languages (Donohue and Grimes 2008; Grimes and Edwards, submitted; Ross 1995: 82–84). However, this label will be used here for convenience – it should be read as a descriptive, rather than a classificatory term.
is used when the possessed noun is a body part, and another when the possessed noun is a kin term. As an example of this kind of split, here are some data from Hatam, which is also spoken in Northwest New Guinea. While we saw above that there are two possessive coding strategies in Sough, in Hatam there are three. These are exemplified in (2).

(2) Hatam (Hatam-Mansim)

a. a-de singau
   2SG-POSS knife
   ‘your knife’
   (Reesink 1999: 49)

b. a-ndab
   2SG-hand
   ‘your hand’
   (Reesink 1999: 48)

c. at-nem
   2SG-wife
   ‘your wife’
   (Reesink 1999: 48)

Syntactically, the three coding strategies in Hatam can be divided in two: a construction in which the person and number of the possessor is marked on a prenominal possessive marker de, as in (2a); and one in which the possessor is marked with a prefix on the possessed noun, as in (2b) and (2c). The construction with the marker de is used to code alienable possession, and the prefixing construction is used to code inalienable possession, which in Hatam includes body parts and kin terms. However, there is a further paradigmatic split in the prefixing construction, in that the form of the 2SG prefix differs in (2b) and (2c). This split is systematic and conditioned semantically in Hatam: the a-series is used when the possessed noun is a body part, and the at-series when it is a kin term. This paper surveys similar splits in inalienable coding in the region of Wallacea.

The remainder of this paper is structured as follows. Section 2 provides the terminological and theoretical background to the typological investigation. Following this, Section 3 describes the sampled languages and data set that form the basis of the survey. The results of the survey are presented in Section 4. I show there that 13 genealogically diverse languages in Wallacea are attested with SIC; the genealogical and areal patterning in the formal marking of the distinction is also discussed, as is the antiquity of the distinction in the languages. In Section 5, some possible explanations for the observed geographic, genealogical, and temporal distributions are discussed. There, I argue that A language are predisposed to develop SIC, due to a structurally distinct nominal subclass of kin terms inherited from an earlier stage of
Austronesian; and that contact has also played a role, specifically in the region of Northwest New Guinea. This paper concludes in Section 6, including some suggestions for future lines of research on this little-known phenomenon.

2 Terminology and key concepts

In this section, I provide the theoretical background for the typological survey presented below. This begins in Section 2.1 with an introduction to the terminology used in this paper, in particular that relating to the concept of ‘alienability’. Following this, in Section 2.2, splits in the coding of inalienable possession are discussed in more detail.

2.1 Defining ‘(in)alienability’

This study focusses on formal splits in inalienable possessive coding. In this paper, I use the term ‘split’ in two related ways. The first is in the sense of Corbett (2015), to refer to splits on the level of the lexicon: the “division of a paradigm into parts” (p. 146, fn. 2), such that – in the languages discussed below – the phonological shape of inflectional material differs across different lexemes. The second follows Haspelmath (2017), Ortmann (2018), and Stolz et al. (2008) in extending this idea of competition between two or more formal manifestations of grammatical features to the level of the possessive noun phrase. Both uses of the term are necessary here: in some languages discussed in this paper, the relevant splits are in the lexicon, whereas in others, the splits are on the level of the possessive construction itself.

Before looking at splits in inalienable coding, we must first define how the terms ‘alienable’ and ‘inalienable’ will be used in this paper. Much has already been written about alienability, and this paper will not break new ground here. As a starting point, I follow e.g. Chappell and McGregor (1996), Heine (1997), Nikolaeva and Spencer (2012), von Prince (2016), and Seiler (2001), by characterising the semantic opposition between inalienable and inalienable possession as follows. Inalienable relationships between the referents of the possessor (PossR) and possessed (PossD) NPs are those in which: (1) the possessive relationship is permanent; (2) the existence of a PossR is entailed by the semantics of the PossD noun; (3) the PossD noun is inherent to the PossR; and (4) the PossR has little or no choice or control in the possessive relationship. Frequently cited examples of inalienable relationships include kinship relations, body part and other part/whole relations, spatial relationships, and inherent attributes. Alienable relationships, on the other hand, are those that are more temporary and severable, over which the PossR has more choice and control – such as relationships of ownership or possession (permanent or temporary), as well as more contextually defined relationships, such as the relationship between a product and its creator.
However, as discussed by Nichols (1988) and Karvovskaya (2018), the terms ‘(in)alienable’ have been used, often implicitly, to refer to a range of different linguistic phenomena, which operate at different levels of the grammar. Building on Karvovskaya (2018: 9–14), the main ways in which these terms have been used in the literature are given in (3).

(3) a. To refer to a split in the formal coding of possession. For example, in a language with two possessive coding strategies conditioned by alienability, one strategy is considered ‘inalienable’, and the other ‘alienable’, depending on the kind of relationship that the respective constructions express.

b. To refer to nominal subclasses, depending on the coding strategy used when a noun is possessed. These subclasses may be semantic, or lexically specified.

c. To refer to nominal subclasses, depending on whether a noun obligatorily occurs in a possessive construction. Again, these subclasses may be semantic, or lexically specified.

d. To refer to the semantic relationship between the referents of the PossR and PossD nouns.

This paper is an investigation of the formal expression of inalienable possession in Wallacea. To facilitate comparison across languages, I thus follow Haspelmath (2017) in using the terms ‘(in)alienable’ to refer to formal properties of possessive constructions, i.e. (3a).

However, as discussed by Dahl and Koptjevskaja-Tamm (2001: 209), rarely is the formal coding entirely predictable on semantic grounds. In languages with an alienability distinction, a subset of nouns can be possessed in the inalienable construction. This subset always includes at least some nouns from either one or both of the semantic classes of body parts and kin terms. However, as will be discussed in more detail below, it may not include all nouns in that class; and sometimes nouns from other semantic classes are also included.

To account for this variation in the semantic contribution to the distinction, the definitions to be used in the rest of this paper are given in (4).

(4) A language with an **alienability distinction** is one that has a formal split in possessive coding, rooted in semantic notions of permanence and inherence of the possessive relationship, lack of control of the PossR over the PossD, and the entailment of a PossR by the semantics of the PossD noun. In a given language, the split may be conditioned by the semantics of or a lexical specification on the PossD noun, or by the semantic relationship between the referents of the PossR and PossD nouns.
a. **Inalienable** possessive coding strategies are those most closely (although not necessarily exclusively) associated with expressing permanent, inherent possessive relationships, such as part/whole relationships, relationships of kinship, and/or spatial relationships.

b. **Alienable** possessive coding strategies are those used elsewhere.

Notably, this approach to ‘(in)alienability’ excludes languages in which there is no formal split in possessive coding – the same strategy is used for all nouns – but in which some nouns are obligatorily coded for possession, in that they are only grammatical if they occur in a possessive construction, i.e. (3c). This definition of alienability means that the analyses in this paper are sometimes at odds with the author of a description. For example, in Moi (West Bird’s Head), Menick (1996: 64) refers to a subclass of nouns including kin terms and body parts as ‘inalienable’; these are distinguished from alienable nouns only by the fact that they are obligatorily possessed. According to the definitions in (4), Moi is not considered to have an alienability distinction, as there is no formal split in possessive coding: the same strategy is used for all nouns, regardless of whether a given noun is obligatorily possessed. As will be described below, split inalienables are only relevant in languages which already have a formal alienability split. For languages such as Moi, where there is no formal split, the question of split inalienables is thus moot.

Some examples will help to further clarify how the terms will be used in this paper. In many languages, the alienability split is binary: using the definitions in (4), one construction can be identified as inalienable, the other as alienable. This is the case, for example, in Wano (Trans-New Guinea > Dani; Burung 2017). In Wano, all possessive constructions are syntactically identical, in that the PossR NP precedes the PossD, and the PossR is marked on the PossD noun with a prefix. However, there are two possessor-marking paradigms: One is used to mark kin terms and body parts, and the other is used elsewhere. This is shown in (5): A 1SG PossR is marked on the PossD kin term *abut* ‘child’ with the prefix *n-*; and on *abui* ‘possum’ with *ne-*.

(5) **Wano** (Trans-New Guinea > Dani)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>n-abut</td>
</tr>
<tr>
<td></td>
<td>1SG-child</td>
</tr>
<tr>
<td></td>
<td>‘my child’</td>
</tr>
<tr>
<td></td>
<td>(Burung 2017: 106)</td>
</tr>
<tr>
<td>b.</td>
<td>ne-abui</td>
</tr>
<tr>
<td></td>
<td>1SG-possum</td>
</tr>
<tr>
<td></td>
<td>‘my possum’</td>
</tr>
<tr>
<td></td>
<td>(Burung 2017: 105)</td>
</tr>
</tbody>
</table>
Of the two paradigms, the first, represented by (5a), is the one used with kin terms and body parts. This construction is therefore closely associated with the expression of permanent and inherent possessive relationships; using the definitions in (4), it is identified as the inalienable paradigm in Wano. The other paradigm, represented by (5b), is used elsewhere. It is thus the alienable paradigm.

The alienability split in Wano is morphological, in that there are two different possessive paradigms. In other languages, the split may be syntactic. For example, in Abun (isolate; Berry and Berry 1999), two different constructions are used to express possession, given in (6). The first is juxtapositional: The PossR precedes the PossD, with no additional marker of possession, as in (6a). This construction is used with body parts, other part/whole relationships, and ‘name’. In the second construction, used elsewhere, the PossR again precedes the PossD; but there is an additional invariant possessive marker bi, as in (6b).

(6) Abun (isolate)
   a. ji syim
      1sg  arm
      ‘my arm’
      (Berry and Berry 1999: 77)
   b. ji bi nggwe
      1sg  poss  garden
      ‘my garden’
      (Berry and Berry 1999: 78)

Although not all semantically inalienable relationships in Abun are expressed with the juxtapositional construction – notably, kin terms are excluded – this construction is more closely associated with the expression of inalienable relationships than the construction with bi. The juxtapositional construction is thus identified as the inalienable construction in Abun. The construction with bi, used elsewhere, is the alienable construction – even though this construction is also used to mark kin terms.

With these preliminaries in place, we now turn to look at languages with multiple inalienable possessive coding strategies.

## 2.2 Splits in inalienable coding

A language is considered to have Split Inalienable Coding if there are three or more distinct possessive coding strategies, at least two of which can be identified as inalienable, following the definitions in (4). In all of the languages surveyed in this paper, the splits are hierarchical: There is a binary syntactic split in possessive
constructions relating to the alienability distinction itself, with a further morpho-
logical or syntactic split in inalienable constructions. As will be described below,
inalienable splits may be semantically conditioned (i.e. predictable from the se-
monic class of the possessed noun), or lexically specified (i.e. semantically unpre-
dictable). For the purposes of this paper, I will focus on semantically conditioned SIC
in the languages of Wallacea. The motivation for this will be explained below.

A preliminary example of SIC was given in (2) above, with data from Hatam. As a
more detailed example, consider the data in (7) from Ambai (AN > SHWNG; Silzer
1983). In Ambai, there is a two-way split in possessive coding: one construction in
which PossR-marking affixes attach to the PossD noun, as in (7a), and another in
which PossR-marking affixes attach to a prenominal possessive marker ne, as in (7b).

(7) Ambai (AN > SHWNG)
   a. awe-ku
      foot-1SG
      'my foot'
      (Silzer 1983: 89)
   b. ne-ku wa
      POSS-1SG canoe
      'my canoe'
      (Silzer 1983: 124)

The affixing construction is used to possess some body parts and kin terms; con-
structions with ne are used elsewhere. While not all body parts and kin terms are
possessed in the affixing construction in Ambai – exceptions include ‘bone’, ‘heart’,
‘older same-sex sibling’, and ‘daughter’ – this construction is more closely associated
with the possession of body parts and kin terms. Thus, by the definitions in (4), the
affixing construction in Ambai is inalienable, and the construction with ne is
alienable.

We find a further split in the Ambai inalienable construction: There are two
possessor-marking paradigms. One paradigm is used when the PossD noun is a body
part, and the other is used when it is a kin term. The two paradigms are shown side-
by-side in Table 1.6

---

5 Section 5.1 presents a cross-linguistic overview of inalienable splits. There, it will be shown that
these splits are not hierarchical in some languages elsewhere in the world. For example, a language
may have three syntactically distinct possessive constructions: one used for the possession of body
parts (inalienable 1), the second for kin terms (inalienable 2), and the third elsewhere (alienable).
6 Many of the languages discussed in this paper distinguish dual number in their possessive para-
digms, and some (including Ambai) also distinguish trial or paucal number. However, paradigmatic
splits in dual and trial/paucal number are rarely attested; and the forms in the surveyed languages
are often transparent, in that they are obvious grammaticalisations involving words for ‘two’ and
As can be seen from Table 1, the difference between the two paradigms is very subtle. For most combinations of person and number, the PossR-marking affixes are identical. However, whereas 3SG PossRs are marked on body parts with -n, they are marked on kin terms with -na. The choice of inalienable paradigm is fully semantically conditioned: Nouns that are licenced to occur in inalienable constructions belong to one of two semantic subclasses, body parts or kin terms. In other words, if one knows that a given kin term or body part can be possessed in the inalienable construction, then the possessor-marking paradigm is entirely predictable from the semantic class of that noun. There is no element of phonological conditioning to this split: For example, *tara* ‘ear’ and *ene* ‘abdomen’ take the body part paradigm, while *tama* ‘father’ and *ina* ‘mother’ take the kin term paradigm.

While the choice of inalienable paradigm in Ambai is completely predictable from the semantics of the PossD noun, in some of the surveyed languages the split has a semantic basis, but there are a handful of lexical exceptions. For example, Sobei (*AN > Oceanic; Sterner and Ross 2011*), like Ambai, distinguishes an alienable construction with a prenominal possessive marker from an affixing inalienable construction; and, like Ambai, has two paradigms used in the inalienable construction. These two paradigms are given in Table 2. Again, the difference is very subtle, being limited to a 3SG PossR.

Sobei is unlike Ambai, however, in that nouns referring to kin terms straddle the two paradigms in Sobei: *temto* ‘husband’ and *mefne* ‘wife’ are marked with the same paradigm used for body parts, while all other kin terms are marked with the second

<table>
<thead>
<tr>
<th></th>
<th>Paradigm 1</th>
<th>Paradigm 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Body parts</td>
<td>Kin terms</td>
</tr>
<tr>
<td>1SG</td>
<td>-ku</td>
<td>-ku</td>
</tr>
<tr>
<td>2SG</td>
<td>-mu</td>
<td>-mu</td>
</tr>
<tr>
<td>3SG</td>
<td>-n</td>
<td>-na</td>
</tr>
<tr>
<td>1PL.EXCL</td>
<td>ama-...-mi</td>
<td>ama-...-mi</td>
</tr>
<tr>
<td>1PL.INCL</td>
<td>ta-...-mi</td>
<td>ta-...-mi</td>
</tr>
<tr>
<td>2PL</td>
<td>me-...-mi</td>
<td>me-...-mi</td>
</tr>
<tr>
<td>3PL</td>
<td>e-...-mi</td>
<td>e-...-mi</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Paradigm 1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Body parts</td>
</tr>
</tbody>
</table>
| Three (Kamholz 2014: 124, fn. 2). To keep things simple, the data in this paper will be from singular and plural numbers only.
While there is a clear semantic basis for the split in Sobei, the paradigm is lexically specified for the nouns ‘husband’ and ‘wife’.

In yet other languages, there is no obvious semantic basis for the inalienable split. Take, for example, Batuley (AN > CMP; Daigle 2015). In Batuley, as in Ambai and Sobei, there is a two-way alienability split in possessive constructions: some body parts, three kin terms, and at least two emotion and cognition terms are possessed in inalienable constructions where the PossR is marked with suffixes on the PossD noun; otherwise, nouns are possessed in alienable constructions using an inflected prenominal possessive marker. There is then a further split in the inalienable construction, with three different PossR-marking paradigms. These paradigms are given in Table 3.

Table 2: Inalienable paradigms in Sobei (based on Sterner and Ross 2011: 174).

<table>
<thead>
<tr>
<th>Paradigm 1</th>
<th>Paradigm 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Body parts and products; ‘husband’, ‘wife’</td>
<td>Other kin terms</td>
</tr>
<tr>
<td>1SG</td>
<td>-ʔ</td>
</tr>
<tr>
<td>2SG</td>
<td>-m</td>
</tr>
<tr>
<td>3SG</td>
<td>-Ø</td>
</tr>
<tr>
<td>1PL,EXCL</td>
<td>-rim</td>
</tr>
<tr>
<td>1PL,INCL</td>
<td>-rir</td>
</tr>
<tr>
<td>2PL</td>
<td>-rim</td>
</tr>
<tr>
<td>3PL</td>
<td>-ri</td>
</tr>
</tbody>
</table>

Table 3: Inalienable paradigms in Batuley (Daigle 2015: 135).

<table>
<thead>
<tr>
<th>Paradigm 1</th>
<th>Paradigm 2</th>
<th>Paradigm 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Some body parts, ‘younger sibling’</td>
<td>Some body parts, ‘voice’</td>
<td>Some body parts, ‘father’, ‘mother’</td>
</tr>
<tr>
<td>1SG</td>
<td>-ing</td>
<td>-ung</td>
</tr>
<tr>
<td>2SG</td>
<td>-em</td>
<td>-om</td>
</tr>
<tr>
<td>3SG</td>
<td>&lt;e&gt;</td>
<td>&lt;i&gt;</td>
</tr>
<tr>
<td>PL</td>
<td>-in</td>
<td>-un</td>
</tr>
</tbody>
</table>

7 In the rest of this paper, I assume that fully predictable splits, such as the split in Ambai, are historically prior to splits which are largely semantically predictable, but with a few lexical exceptions, such as that in Sobei; and that the less predictable splits arise through the reassignment of a noun in one semantic subfield into the noun class of another (in much the same way as it is thought that nouns tend to drift from inalienable coding to alienable coding; see e.g. van den Berg 2009: 351 on this tendency in SHWNG languages, or Florey 2005 on similar trends in some languages of central Maluku).
However, unlike in Sobei and Ambai, the choice of paradigm is neither wholly nor largely predictable from the semantics of the PossD noun in Batuley: While only body parts and ‘voice’ are possessed with Paradigm 2, both kin terms and body parts are attested with either Paradigm 1 or 3. SIC in Batuley is thus lexically specified, with no semantic basis.

It is semantically motivated SIC of the kind found in Ambai and Sobei with which we are concerned in this paper; lexically specified SIC, like that in Batuley, will not be considered further. This decision was taken for two reasons. First, this mirrors the way in which the alienability distinction itself is conceptualised: While, as mentioned above, languages vary greatly in the factors conditioning alienability distinctions, splits which are not rooted in the semantic distinction of permanent, inherent versus non-permanent, non-inherent possessive relationships are not generally considered to be expressions of alienability. Second, from a diachronic perspective, semantically motivated inalienable splits emerge from a different level of the grammar than lexical splits: While the former carve up the nominal inventory into semantic subclasses, the latter tend to be the result of a diverse array of diachronic phonological, morphological, and/or syntactic processes. By focussing on semantic splits and their development in Wallacea, this paper will thus provide new data on how nominal classification systems emerge and develop. Unless otherwise stipulated, ‘SIC’ should henceforth be understood to refer to semantically motivated SIC.

3 Sources and data set

This study looks at the distribution of SIC in the languages of Wallacea, a linguistic area defined in Schapper (2015) and outlined in Schapper and Gasser (this volume). Wallacea provides the ideal starting point for an initial investigation of SIC. First, as discussed by Himmelmann (2005: 163–165), Holton and Klamer (2018: 600–603), and Klamer et al. (2008: 116–130), there is a high level of formal diversity in possessive constructions, and in particular alienability distinctions, have long interested

8 For example, some An languages in the sample have additional infixing inalienable paradigms (e.g. Ambel, Arnold 2018; Irarutu, van den Berg and Matsumura 2008; Kola, de Winne 2016), which are clearly derived from the fossilisation of earlier compounds. Similarly, the three inalienable paradigms in Batuley given in Table 3 appear to have developed through the reanalysis of root-final vowels as part of the PossR-marking affix. Finally, several An languages belonging to the Timor-Alor-Pantar family have multiple inalienable paradigms; however, additional paradigms in these languages have developed from a formerly free possessive marker, which has become more tightly bound with the PossD noun. I am grateful to Antoinette Schapper for pointing out the historical processes involved in the developments of the Batuley and the Timor-Alor-Pantar paradigms.
scholars in the area, including recent investigations of the role of inheritance and contact in the development of alienability in the region by van den Berg (2009), Donohue and Grimes (2008: 142–143), Donohue and Schapper (2008), Kamholz (2014: 128–135), and Klamer et al. (2008: 116–122). Possessive constructions are often discussed in detail in grammars and other works on the languages of Wallacea, thus providing strong source material on which to base this investigation. Finally, as will be shown below, SIC has an interesting areal and genealogical distribution in the area. The historical implications of this distribution will be returned to in Section 5.

This study is based on a sample of 189 languages for which descriptive materials on possessive coding were available, supplemented by personal communication with fieldworkers for those languages where information was either unavailable or unclear. As many languages of the area are under- or undescribed, the sample was necessarily one of convenience. Geographically, the sample is centred on Wallacea; in order to capture areal tendencies, languages spoken to the west of Wallacea, as far as Sulawesi and Sumbawa, and to the east, as far as the border between Indonesian and Papua New Guinea, are also included.9

Genealogically, 113 of languages in the sample are Austronesian, representing nine sub-branches. The remaining 76 languages are non-Austronesian, comprising 18 families (including six branches of Trans-New Guinea) and seven isolates. A breakdown of the language families represented in the survey is given in Table 4.10

For each language, a data entry was created.11 This entry includes the name of the language, the ISO 639-3 code, and information on genealogical affiliation, the region in which the language is spoken, and the source(s) consulted. The language was then coded for the presence or absence of a formal alienability distinction, as defined in (4). A schematic representation of the possessive construction(s) was included; and, if the language is attested with an alienability distinction, notes on the factors conditioning this distinction were made.

For those languages with an alienability distinction, the number of attested coding strategies that could be considered ‘inalienable’, again following the definitions in (4), was recorded. If there was more than one inalienable coding strategy – i.e., if the language was attested with Split Inalienable Coding – then notes were made

---

9 A map showing the distribution of the sampled languages, along with a full list of the sampled languages, ISO 639-3 codes, genealogical affiliation, region spoken, and sources consulted, can be found at doi.org/10.5281/zenodo.5503780.

10 As can be seen from Table 4, the lack of descriptive materials for many languages of the region means that the sample is not at all balanced: For example, while the survey includes 19 of the two dozen or so Timor-Alor-Pantar languages, only 16 languages belonging to branches of the large and very diverse Trans-New Guinea family are included. Some families of the area are completely unrepresented, for example Pauwasi.

11 The full data set is available at doi.org/10.5281/zenodo.5503780.
on the formal differences between the inalienable strategies; the nouns or semantic classes of noun associated with each of the inalienable strategies; and, from this, it was determined whether the split had a semantic basis. Only those languages in which SIC is semantically motivated were analysed further. Finally, there was additional space to make freeform notes regarding the coding of possession in each language.

As many of the published sources are somewhat minimal in their description, it should be stressed that a language was not coded for the presence versus absence of an alienability distinction, nor the presence versus absence of SIC; rather, the approach taken was one of attestation versus non-attestation. As noted for Ambai and Sobei in Section 2.2, the split is very subtle in some of the surveyed languages,

Table 4: A genealogical breakdown of the sampled languages (TNG = Trans-New Guinea).

<table>
<thead>
<tr>
<th>Austronesian</th>
<th>No. of languages</th>
<th>Non-Austronesian</th>
<th>No. of languages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sangiric</td>
<td>3</td>
<td>Timor-Alor-Pantar</td>
<td>19</td>
</tr>
<tr>
<td>Minahasan</td>
<td>1</td>
<td>North Halmahera</td>
<td>6</td>
</tr>
<tr>
<td>Gorontalo-Mongondow</td>
<td>3</td>
<td>West Bird’s Head</td>
<td>3</td>
</tr>
<tr>
<td>Celebic</td>
<td>19</td>
<td>Konda-Yahadian</td>
<td>1</td>
</tr>
<tr>
<td>South Sulawesi</td>
<td>7</td>
<td>South Bird’s Head</td>
<td>3</td>
</tr>
<tr>
<td>Central Malayo-Polynesian</td>
<td>43</td>
<td>East Bird’s Head</td>
<td>3</td>
</tr>
<tr>
<td>Malayic</td>
<td>4</td>
<td>Hatam-Mansim</td>
<td>2</td>
</tr>
<tr>
<td>South Halmahera-West New Guinea</td>
<td>30</td>
<td>West Bomberai</td>
<td>2</td>
</tr>
<tr>
<td>Oceanic</td>
<td>3</td>
<td>Mairasic</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Yawa-Saweru</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Geelvink Bay</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Greater Kwerba</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tor-Orya</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lakes Plain</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nimboranic</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sentanic</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TNG &gt; Paniai Lakes</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TNG &gt; Asmat-Kamoro</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TNG &gt; Dani</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TNG &gt; Mek</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TNG &gt; Ok-Oksapmin</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TNG &gt; Awyu</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Anim</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bulaka River</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Kolopom</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mombum-Koneraw</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Language isolates</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>113</td>
<td></td>
<td>76</td>
</tr>
</tbody>
</table>
with differences of a single phonological segment limited to a single cell of a possessive paradigm – something that may easily be missed if a researcher only had limited time with a speech community. Only those languages for which SIC is explicitly attested were therefore coded as such. In addition, in some cases the semantic contribution to the split was unclear. For example, Matbat (An → SHWNG; Remijsen 2010) is described with six inalienable paradigms. However, only one noun is attested in each paradigm. This means it is not possible to tell the extent to which the choice of paradigm is semantically motivated, and what other factors – for example, phonological conditioning, and/or lexical specification – may be involved. While further investigation may reveal that Matbat does in fact have semantically motivated SIC, for the purposes of this paper this language was coded as ‘SIC not attested’.

4 Split Inalienable Coding in Wallacea

Of the 189 languages in the sample, 105 are attested with a formal alienability distinction (58 An, 47 NAn). Of these 105 languages, 13 are attested with semantically motivated SIC (10 An, 3 NAn). The results from the survey are given in Map 1.

As can be seen from this map, alienability distinctions are largely absent in the languages of Sulawesi, as well as in the languages of the Lesser Sundas west of Flores.

(as also discussed by Himmelmann 2005: 163–165 and Klamer et al. 2008: 116–122). The distribution of the alienability distinction is also patchy on Timor, Halmahera, the Bomberai peninsula, and in central New Guinea.

There are three points to note on alienability distinctions in the surveyed languages. First, the definitions given in (4) gloss over a great deal of morphosyntactic variation in the coding of alienability across the region; for more on the formal coding of possession in the surveyed languages, see Himmelmann (2005: 163–165), Holton and Klamer (2018: 600–603), and Klamer et al. (2008: 116–130). Second, the definitions in (4) also gloss over the variation in the lexical and semantic contribution to the alienability distinction. For example, in some of the surveyed languages, only body parts are possessed in the inalienable construction, whereas kin terms are possessed in alienable constructions (e.g. Burmeso, isolate; Donohue 2001); similarly, in other languages, only kin terms are possessed in the inalienable construction, with body parts possessed in alienable constructions (e.g. Moni, TNG > Paniai Lakes; Larson and Larson 1958). Furthermore, some languages additionally possess some non-permanently or non-inherently possessed nouns in the inalienable construction (e.g. Warembori, AN > SHWNG; Donohue 1999b). Finally, whereas in some languages possessive coding operates in a fixed system in which every noun in the lexicon is specified for possession in one and only one possessive construction (e.g. Tehit, West Bird’s Head; Flassy 1991), in others, the system is flexible, in that one and the same noun may be possessed either in an inalienable or an alienable construction, depending on whether the intended reading of the relationship between the possessor and possessed noun is inalienable (e.g. Ambel, AN > SHWNG; Arnold 2018).

As can be seen from Map 1, most of the 13 languages with SIC are spoken in the north and northwest of New Guinea: the Raja Ampat archipelago, the Bird’s Head peninsula, the coasts and islands of Cenderawasih Bay, and on the north coast of central New Guinea. There are additionally two AN outliers in the Lesser Sundas. The following sections examine SIC in Wallacea in more detail. As the formal manifestation of the distinction patterns genealogically, so too does this discussion. I thus begin with a description of SIC in the AN languages, before turning to the NAn languages.

### 4.1 Split Inalienable Coding in Austronesian languages

The 10 sampled Austronesian languages with SIC are given in Table 5. These languages belong to the South Halmahera-West New Guinea and Oceanic sub-branches; three Central Malayo-Polynesian languages are also attested with SIC.12

---

12 The genealogical affiliation of Irarutu has long been unclear (see Jackson 2014: 187–244 for discussion). However, the most recent research suggests it is an East Bomberai language (Grimes and Edwards, submitted), and can thus be grouped with the CMP languages.
The alienable constructions in these An languages are somewhat variable in their structure. However, the inalienable constructions are structurally similar, in that all of the languages use PossR-marking morphology on the PossD noun. With the exception of Salawati, which will be discussed below, the inalienable splits in these languages are morphological, in that different paradigms are used, depending on the semantic class of the PossD noun.

In five of these languages – Central Lembata, Ambel, Wamesa, Ambai, and Sobei – the split is limited to a single cell in the paradigm, when the PossR is 3SG. This subtle split was shown for Ambai and Sobei in Section 2.2. As another example, the paradigm for Central Lembata is given in Table 6. A 3SG PossR is marked on kin terms with the suffix -nu, whereas if the PossD noun is a body part, the marker is phonologically conditioned: Consonant-final nouns are unmarked, and there is vowel lengthening on vowel-final nouns.

In the other An languages with SIC, the split is found in more than one cell in the paradigm. For example, in the Kotos variety of Amarasi, the two inalienable paradigms are completely distinct, with no overlapping phonological form (Edwards 2020). These two paradigms are given in Table 7.14

Table 5: Sampled Austronesian languages with semantically motivated Split Inalienable Coding (WNG = West New Guinea).

<table>
<thead>
<tr>
<th>Language</th>
<th>Subbranch</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central Lembata</td>
<td>[lvu]</td>
<td>CMP (Flores-Lembata)</td>
</tr>
<tr>
<td>Amarasi</td>
<td>[aa]</td>
<td>CMP (Rote-Meto)</td>
</tr>
<tr>
<td>Irarutu</td>
<td>[ir]</td>
<td>CMP (East Bomberai)</td>
</tr>
<tr>
<td>Salawati</td>
<td>[xm]</td>
<td>SHWNG &gt; RASH</td>
</tr>
<tr>
<td>Ambel</td>
<td>[wgo]</td>
<td>SHWNG &gt; RASH</td>
</tr>
<tr>
<td>Wamesa</td>
<td>[wad]</td>
<td>SHWNG &gt; Western Yapen</td>
</tr>
<tr>
<td>Roon</td>
<td>[rnn]</td>
<td>SHWNG &gt; Biakic</td>
</tr>
<tr>
<td>Biak</td>
<td>[bhw]</td>
<td>SHWNG &gt; Biakic</td>
</tr>
<tr>
<td>Ambai</td>
<td>[amk]</td>
<td>SHWNG &gt; Western Yapen</td>
</tr>
<tr>
<td>Sobei</td>
<td>[sob]</td>
<td>Oc &gt; WOc &gt; North New Guinea</td>
</tr>
</tbody>
</table>

The alienable constructions in these An languages are somewhat variable in their structure. However, the inalienable constructions are structurally similar, in that all of the languages use PossR-marking morphology on the PossD noun. With the exception of Salawati, which will be discussed below, the inalienable splits in these languages are morphological, in that different paradigms are used, depending on the semantic class of the PossD noun.

In five of these languages – Central Lembata, Ambel, Wamesa, Ambai, and Sobei – the split is limited to a single cell in the paradigm, when the PossR is 3SG. This subtle split was shown for Ambai and Sobei in Section 2.2. As another example, the paradigm for Central Lembata is given in Table 6. A 3SG PossR is marked on kin terms with the suffix -nu, whereas if the PossD noun is a body part, the marker is phonologically conditioned: Consonant-final nouns are unmarked, and there is vowel lengthening on vowel-final nouns.

In the other An languages with SIC, the split is found in more than one cell in the paradigm. For example, in the Kotos variety of Amarasi, the two inalienable paradigms are completely distinct, with no overlapping phonological form (Edwards 2020). These two paradigms are given in Table 7.14

13 Whereas Central Lembata, Wamesa, Ambai, and Sobei all have two paradigms, Ambel has three: one used with body parts, a second with kin terms, and a third with nouns which have a close association with their PossR, for example ‘shadow’ and ‘name’.

14 In the body part paradigm, Kotos Amarasi additionally has what Edwards (2020: 449) refers to as the “0 person” suffix -f, used to mark the noun when the PossR is irrelevant or unknown. The variety of Kotos Amarasi spoken in Koro’oto hamlet differs, in that -f is used to mark all persons and numbers in the kin paradigm. The Ro’is variety of Amarasi does not have SIC: The same inalienable suffixing paradigm is used to mark both body parts and kin terms. Kin terms may also optionally occur...
The two inalienable paradigms in Irarutu (van den Berg and Matsumura 2008) are given in Table 8: In one (Paradigm 1), the affixes are mainly circumfixing, whereas the other (Paradigm 2) is comprised of double-stacked prefixes. The choice of inalienable paradigm in Irarutu has a semantic basis, but there are two lexical exceptions: Paradigm 1 is used with body parts plus the kin terms *wa* ‘spouse’ and *rü* ‘cross-sibling’, whereas Paradigm 2 is used to mark all other kin terms.\(^\text{15}\)

The two inalienable paradigms in Irarutu (van den Berg and Matsumura 2008) are given in Table 8: In one (Paradigm 1), the affixes are mainly circumfixing, whereas the other (Paradigm 2) is comprised of double-stacked prefixes. The choice of inalienable paradigm in Irarutu has a semantic basis, but there are two lexical exceptions: Paradigm 1 is used with body parts plus the kin terms *wa* ‘spouse’ and *rü* ‘cross-sibling’, whereas Paradigm 2 is used to mark all other kin terms.\(^\text{15}\)

---

\(^{15}\)Only forms for singular PossRs are given in van den Berg and Matsumura (2008). Several additional comments can be made about the data in Table 8. First, the *û* in Paradigm 1 is only present if the root is consonant-final. Second, van den Berg and Matsumura (2008) describe possession for the East Arguni dialect of Irarutu. The semantic basis of the split in the Fruata dialect, described in Jackson (2014: 119–123), differs slightly: Paradigm 1 is additionally used with ‘widow’, ‘sibling’, and ‘plate’. This presumably reflects the drift from a semantic subclass into the noun class of another, described in footnote 7. Finally, possession in Irarutu is actually more complex than presented here, with many irregularities.
In Biak (van den Heuvel 2006), there are three inalienable paradigms: one for kin terms; one for body parts normally occurring in pairs, such as legs and eyes (and including na ‘tooth’ and snoni ‘nose’); and one for other body parts.\(^{16}\) Formal differences are found across all markers of a singular PossR for all three paradigms, as shown in Table 9.\(^{17}\)

The Butlih variety of Salawati (own fieldnotes 2019) differs from all of the other AN languages with SIC: The distinction is not paradigmatic, but in the optionality of an inflected possessive marker ni.\(^{18}\) This inflected possessive marker is used in both alienable and inalienable constructions for most PossRs in Butlih Salawati; whereas it is the only marker of possession in alienable constructions, it co-occurs with PossR-marking affixes in inalienable constructions. The inalienable constructions are split, in that while this particle is optional for body parts, it is obligatory for kin terms for all persons and numbers except 2SG. This is shown in Table 10.\(^{19}\)

---

\(^{16}\) Roon, which also has SIC, is closely related to Biak: Both are primary branches of the Biakic sub-branch of SHWNG (Kamholz 2014: 138). The Roon data are still preliminary (see Gil, this volume), so they are not presented here. However, from what is available, some forms marking a 2SG PossR are phonologically similar to the Biak forms marking 2SG PossRs of kin terms and non-paired body parts (David Gil pers. comm.) – they may therefore be cognate.

\(^{17}\) The paradigm for non-singular PossRs of kin terms is not attested in van den Heuvel (2006). Note also the =i enclitic in Paradigm 1 marks a singular PossD noun; dual PossD nouns are marked instead with =su, and trial with =sko.

\(^{18}\) Detailed data on the possessive constructions in other varieties of Salawati are not available. It is also unknown whether SIC in Butlih Salawati is fully semantically predictable, or whether there are exceptions.

\(^{19}\) 1SG possessive constructions are a juxtaposition of the possessed noun and the 1SG pronoun. Diacritics mark tone; 3SG and 3PL PossRs are additionally marked suprasegmentally, in that the PossD noun is toneless. Note that the obligatoriness of the possessive particle for kin terms in Butlih Salawati is different from the obligatory coding for possession described in Section 2.1: The question here is not whether kin terms and body parts are obligatorily coded for possession, but whether or not one of the markers of possession is obligatory.
SIC does not appear to have a great antiquity in any of the AN languages: There is no conclusive evidence to reconstruct SIC higher than the attested languages. This can be most clearly seen in the SHWNG languages. For example, both Ambel and Butlih Salawati belong to the Raja Ampat-South Halmahera (RASH) sub-branch of SHWNG. The forms marking SIC are not cognate in these two languages: In Ambel, the difference is paradigmatic (3SG PossRs are marked on kin terms with -i, and are unmarked on body parts), whereas in Butlih Salawati, as just described, the distinction is marked by the optionality of the prenominal particle. Furthermore, SIC is not attested in any of the other Raja Ampat languages (own fieldnotes, 2019–2020), nor in any languages belonging to the South Halmahera sub-branch of RASH (Rødvand this volume). SIC thus cannot be reconstructed to proto-RASH. Similarly, both Wamesa and Ambai belong to the Western Yapen sub-branch of SHWNG. Here, too, the forms are not cognate: In Ambai, 3SG kin terms are marked with -na and body parts with -n, and in Wamesa 3SG kin terms are marked with -ni, and body parts are marked with a determiner enclitic (typically of the form =pa). As the only Western Yapen languages attested with SIC, there is thus no evidence to reconstruct to proto-

<table>
<thead>
<tr>
<th>Paradigm 1</th>
<th>Paradigm 2</th>
<th>Paradigm 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kin terms</td>
<td>Body parts (paired, ‘tooth’, ‘nose’)</td>
<td>Body parts (else)</td>
</tr>
<tr>
<td>1SG</td>
<td>=i</td>
<td>-si</td>
</tr>
<tr>
<td>2SG</td>
<td>-m=i</td>
<td>-m-si</td>
</tr>
<tr>
<td>3SG</td>
<td>-r=i</td>
<td>-si</td>
</tr>
<tr>
<td>1PL.EXCL</td>
<td>n/a</td>
<td>nko-…-s-na</td>
</tr>
<tr>
<td>1PL.INCL</td>
<td>n/a</td>
<td>ko-…-s-na</td>
</tr>
<tr>
<td>2PL</td>
<td>n/a</td>
<td>mko-…-m-s-na</td>
</tr>
<tr>
<td>3PL</td>
<td>n/a</td>
<td>si-…-s-na</td>
</tr>
</tbody>
</table>

Table 9: Biak inalienable paradigms (based on van den Heuvel 2006: 238–243; n/a = ‘not attested’).

<table>
<thead>
<tr>
<th>Construction 1</th>
<th>Construction 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kin terms</td>
<td>Body parts</td>
</tr>
<tr>
<td>tabū ‘grandchild’</td>
<td>tabu ‘eye’</td>
</tr>
<tr>
<td>1SG</td>
<td>tabū</td>
</tr>
<tr>
<td>2SG</td>
<td>(ni-m) tabū-m</td>
</tr>
<tr>
<td>3SG</td>
<td>ni tabū</td>
</tr>
<tr>
<td>1PL.EXCL</td>
<td>am-ni-m tabū-m</td>
</tr>
<tr>
<td>1PL.INCL</td>
<td>at-ni-m tabū-m</td>
</tr>
<tr>
<td>2PL</td>
<td>me-ni-m tabū-m</td>
</tr>
<tr>
<td>3PL</td>
<td>ha-ni tabu</td>
</tr>
<tr>
<td>1PL.EXCL</td>
<td>(am-ni-m) tá-m</td>
</tr>
<tr>
<td>1PL.INCL</td>
<td>(at-ni-m) tá-m</td>
</tr>
<tr>
<td>2PL</td>
<td>(me-ni-m) tá-m</td>
</tr>
<tr>
<td>3PL</td>
<td>(ha-ni) ta</td>
</tr>
</tbody>
</table>

Table 10: Butlih Salawati inalienable constructions (own field notes, 2019).
Western Yapen. One possible exception to this trend is in the Biakic sub-branch of SHWNG: In footnote 16, it was noted that the forms marking SIC may be cognate in Biak and Roon. If correct, this would imply SIC was present in proto-Biakic, the most recent common ancestor of these two languages (Kamholz 2014: 138). However, owing to a lack of data, this inheritance hypothesis is far from conclusive. Finally, as there are many branches of SHWNG in which SIC is unattested, including all of the branch-level isolates, there is no evidence that SIC is a retention from proto-SHWNG.

Outside of SHWNG, the CMP languages are the only ones attested in their respective branches with SIC, preventing reconstruction to an ancestor language. With regards to Sobei, there is at least one other Oceanic language with SIC outside Wallacea: Nadrogā, a variety of Western Fijian (see Section 5.1). However, once again, the forms are not cognate, and the distribution far too patchy, to reconstruct SIC any higher in Oceanic than the attested languages – and it is not mentioned, for example, in the discussion and reconstruction of proto-Oceanic possession in Lynch et al. (2011: 75–80).

The distribution of SIC in the An languages thus appears to be due to multiple independent and recent innovations. I will return to this point in Section 5, in the discussion of the origins and spread of SIC in Wallacea.

4.2 Split Inalienable Coding in non-Austronesian languages

Three non-Austronesian languages of Wallacea are attested with semantically conditioned SIC, given in Table 11. These three languages are spoken in close proximity to one another, in the Arfak mountains in the east of the Bird’s Head peninsula in Northwest New Guinea.

The three NaN languages with SIC belong to two small unrelated families: Hatam is one of two languages comprising the Hatam-Mansim family; and Meyah and Moskona together comprise the Meax sub-branch of the three-language East Bird’s Head (EBH) family. In all three languages in Table 11, inalienable constructions have PossR-PossD order, and the PossR is marked on the PossD noun with prefixes. This is in contrast to alienable constructions, which also have PossR-PossD order, but in which the PossR is marked by an inflected particle which occurs before the PossD noun.

All three languages have two prefixing inalienable paradigms: One is used with kin terms, the other with body parts. An example of SIC in Hatam was given in (2) above; the full inalienable paradigms are given in Table 12.20

20 The archiphonemic element /T/ of the kin paradigm is /t/ before /n/, /p/ before /m/, and is deleted preceding non-nasals.
SIC is more complex in Meyah and Moskona. In both languages, the body part paradigms are straightforward: They are identical with the verbal subject-marking paradigms. However, there is no unified kin term paradigm in either of the two languages, owing to morphological borrowing, suppletion, and phonologically conditioned developments in Meyah; and, in Moskona, a radical erosion of inalienable morphology for some PossRs of kin terms, and the extension of the markers from some cells to others. Arnold (Forthcoming), based on the data and discussion in Gravelle (2004) and Gravelle (2010), reconstructs the pre-Meyah and pre-Moskona inalienable paradigms. These are given in Table 13.

Table 11: Sampled non-Austronesian languages with semantically motivated Split Inalienable Coding (WNG = West New Guinea).

<table>
<thead>
<tr>
<th>Language</th>
<th>Classification</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hatam</td>
<td>[had]</td>
<td>Hatam-Mansim</td>
</tr>
<tr>
<td>Moskona</td>
<td>[mtj]</td>
<td>East Bird’s Head</td>
</tr>
<tr>
<td>Meyah</td>
<td>[mej]</td>
<td>East Bird’s Head</td>
</tr>
</tbody>
</table>

Table 12: Hatam inalienable paradigms (based on Reesink 1999: 48).

<table>
<thead>
<tr>
<th>Paradigm 1</th>
<th>Paradigm 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Body parts</td>
<td>Kin terms</td>
</tr>
<tr>
<td>1SG</td>
<td>di-</td>
</tr>
<tr>
<td>2SG</td>
<td>a-</td>
</tr>
<tr>
<td>3SG</td>
<td>ni-</td>
</tr>
<tr>
<td>1PL.EXCL</td>
<td>ni-</td>
</tr>
<tr>
<td>1PL.INCL</td>
<td>i-</td>
</tr>
<tr>
<td>2PL</td>
<td>ji-</td>
</tr>
<tr>
<td>3PL</td>
<td>i-</td>
</tr>
</tbody>
</table>

Table 13: Pre-Meyah and pre-Moskona paradigms (Arnold Forthcoming, based on Gravelle 2004, 2010).

<table>
<thead>
<tr>
<th>Pre-Meyah</th>
<th>Pre-Moskona</th>
</tr>
</thead>
<tbody>
<tr>
<td>Body parts</td>
<td>Kin terms</td>
</tr>
<tr>
<td>1SG</td>
<td>*di-</td>
</tr>
<tr>
<td>2SG</td>
<td>*bi-</td>
</tr>
<tr>
<td>3SG</td>
<td>*Ø-</td>
</tr>
<tr>
<td>1PL.EXCL</td>
<td>*me-</td>
</tr>
<tr>
<td>1PL.INCL</td>
<td>*mi-</td>
</tr>
<tr>
<td>2PL</td>
<td>*i-</td>
</tr>
<tr>
<td>3PL</td>
<td>*ri-</td>
</tr>
</tbody>
</table>

352 Arnold
From Table 13, we can see that the body part paradigms are cognate in pre-Meyah and pre-Moskona, as are the kin term paradigms; see Arnold (Forthcoming) for a reconstruction of these two paradigms in the ancestral language, proto-Meax. However, the third member of EBH, Sougb, has a single inalienable paradigm for both kin terms and body parts. The lack of SIC in Sougb means that the distinction cannot be reconstructed to proto-EBH. Also worth noting here is that the kin term paradigms of Meyah and Moskona are more archaic than the body part paradigms, in that they are cognate with the single possessive-marking paradigm in Sougb. In addition, the fact that the body part paradigms are identical with the verbal subject-marking paradigms in both languages (and are also cognate with the subject-marking paradigm of Sougb) suggests that the same source material was used to derive both paradigms in the Meax languages.

5 Discussion: the origins and spread of SIC in Wallacea

In this section, I consider how the distribution of languages with SIC in Wallacea may have emerged. There are four hypotheses that could account for the observed distribution, given in (8).

(8) The presence of SIC in at least some of the languages of Wallacea can be explained by:
   a. inheritance of the distinction from a common ancestor;
   b. purely coincidental, independent innovations;
   c. parallel innovations in structurally similar languages; or
   d. contact-induced innovation.

Hypothesis (8a) was pre-empted above, in Section 4. There it was shown that, with the exception of the Meax sub-branch of East Bird’s Head, and possibly the Biakic sub-branch of SHWNG, there is no evidence that SIC has been inherited in the languages of Wallacea. This means that for most languages, we are dealing with recent innovations. In this section, I address the remaining three hypotheses. This begins in Section 5.1 with hypothesis (8b): Here, I show that, since SIC is very rare cross-linguistically, it is unlikely to have been innovated completely independently many times in Wallacea. In Section 5.2, I address hypothesis (8c) as the explanation for SIC in multiple geographically disparate AN languages. Finally, in Section 5.3, I consider hypothesis (8d), specifically focussing on the role of contact in the development of SIC in Northwest New Guinea. The conclusion of the latter two sections is that there were multiple causations of SIC: Both inherited structural similarities and contact have conspired in Wallacea to give the distribution of SIC observed today.
5.1 Split Inalienable Coding: a cross-linguistic overview

All else being equal, the easier a feature is to innovate, the more cross-linguistically common that feature should be. At first blush, it would be reasonable to assume that SIC is a logical feature to innovate. For example, Dahl and Koptjevskaja-Tamm (2001: 209) discuss how the classes of body parts and kin terms are semantically and syntactically distinct: While both classes are similar in that they entail a PossR, body parts are less animate than kin terms and thus have lower discourse status and are more likely to occur as objects or adverbials. Furthermore, it is well-known that body parts and kin terms are often treated differently when it comes to the alienability distinction itself – with, for example, only one class or the other patterning as inalienable (see e.g. Nichols 1988), or with the two classes following different diachronic pathways (Florey 2005; Schuster 2019). As the demarcation of body part and kin term classes in the lexicon is well-motivated, one may expect SIC to be easy to innovate and thus fairly common in the languages of the world.

While a systematic cross-linguistic survey has not yet been carried out, a first glance through the literature suggests that SIC is in fact a rare feature worldwide. In what follows, I consider only semantically motivated SIC, for the reasons given at the end of Section 2.2. I also take the broad formal definition of SIC introduced there: If the strategies used to code the possession of body parts and kin terms are distinct both from each other, and from the coding of other possessive relationships, the language was considered to have SIC. With this definition, around a dozen other languages worldwide have so far been attested with SIC.

As mentioned in Section 4.1, an Oceanic language spoken outside of Wallacea is attested with SIC: the Nadrogā variety of Western Fijian (wyy; Geraghty 2011). In the Nadrogā inalienable construction, the person and number of the PossR is marked with affixes on the PossD noun; however, while the PossR-marking morphology is suffixing for kin terms, it is prefixing for body parts.

SIC is also attested in Australia: the non-Pama-Nyungan language Nunggubuyu (nuy; Heath 1984), the Worrorran language Ungarinjin (ung; Rumsey 1978), and the languages of the closely related Mantharta and Kanyara branches of Pama-Nyungan (Austin 2015; Peter Austin pers. comm.). In Nunggubuyu, possession is generally expressed with a relative case suffix. However, many parts of wholes (especially parts of plants and other inanimate wholes) are possessed in constructions involving noun-class harmony, in which the PossD noun takes a prefix to derive a noun as the

---

21 This preliminary survey was carried out by following up on sources cited in the key typological works on alienability (e.g. Chappell and McGregor 1996; Nichols 1988; Nichols and Bickel 2013), and through the many helpful responses by colleagues to a call for information on the LingTyp mailing list.
same class as the PossR; and the PossR of kinship terms is marked on the PossD noun, with a complex variety of affixes or stem alternations. Ungarinjin has three possessive constructions: an alienable construction, in which possession is marked with possessive pronouns and/or a genitive postposition on the PossR NP; a construction used with some body parts, in which the PossD noun is inflected with PossR-marking prefixes; and a construction used with kin terms, in which PossR-marking suffixes attach to the PossD noun. In the Mantharta and Kanyara languages, most possessed nouns occur in a construction in which PossR is marked with the dative suffix. However, kin terms are possessed with PossR-marking suffixes on the PossD noun; and body parts are possessed in appositive constructions, in which they are inflected for the same case as the PossR.

Finally, there are two attestations of SIC in the Americas: Trumai (tpy; Guirardello 1999), an isolate of central Brazil; and Northern Pomo (pej; O’Connor 1987), a Pomoan language of northern California. In Trumai, adnominal possessive constructions have the order PossR-PossD. While in alienable constructions, a genitive marker attaches to the PossR, inalienable constructions are generally juxtapositional, with no special possessive morphology. However, there is a split in inalienable constructions when the PossR is third person: Kin terms may optionally be marked with an anaphoric prefix, and body parts with case-marking enclitics. In Northern Pomo, kin terms occur in constructions with PossR-marking prefixes; body parts in constructions in which the PossR is raised to become a dependent of the predicate; and for all other types of possession, the PossR is marked with oblique case.

This preliminary survey suggests that semantically motivated SIC is a cross-linguistically rare feature: Presuming that the presence in both the Mantharta and Kanyara branches of Pama-Nyungan is either due to inheritance, or to contact between these geographically proximate languages, we are looking at no more than a half dozen independent innovations outside of Wallacea. This sporadic attestation worldwide is reflected in the lack of literature on the theory or typology of SIC noted in Section 1. While a more comprehensive and systematic survey is required to confirm the cross-linguistic rarity of this phenomenon, it does seem unlikely from this overview that all 13 languages with SIC in Wallacea innovated the distinction completely independently from each other.

5.2 Internally motivated change: inherited structural pressures in Austronesian

Given how rare the feature is worldwide, the presence of SIC in several geographically disparate AN languages – from the Lesser Sundas as far east as Fiji – would seem to be more than just a coincidence. While these languages have not inherited SIC
itself, there is evidence that they have inherited a pressure that favoured the
development of SIC: namely, a structurally distinct lexical class of kin terms.

Cross-linguistically, kin terms often have special grammatical properties, in
many cases patterning with proper names (Dahl and Koptjevskaja-Tamm 2001).
However, there is evidence from across Austronesian that kin terms in this family
behave as a class separate from all others in the grammar. For example, Blust (1979)
discusses various morphological processes which distinguish referential and voca-
tive uses of kin terms; these processes, some of which he reconstructs to proto-
Austronesian, are specific to this lexical class. In several languages of the northern
Philippines, kin terms are obligatorily possessed (Schapper and McConvell, forth-
coming); as are some kin terms in many languages of central and western Borneo
(Blust 2013: 67). In Tukang Besi, some kin terms are distinguished from other nouns in
that they are optionally possessed with a possessive marker *mai*, which functions to
“emphasise the inalienability of the object” (Donohue 1999a: 346–347). And in Ambel,
most kin terms occur in alienable possessive constructions, with an inflected pos-
sessive particle; however, these kin terms are distinguished from all other alienable
nouns by differences both in the form of the particle, and in the inflectional paradigm
(Arnold 2018: 292–300). Diachronically, kin terms and body parts have also been
shown to follow different developmental pathways in Austronesian: In her study of
the drift from inflectional inalienable to analytic alienable constructions in central
Maluku, Florey (2005) finds that kin terms are shifting from inalienable to alienable
constructions in Soahuku; and that in Alune, this same shift is happening for body
parts. Similarly, Wilson (1982: 35–40) reconstructs a lexical class in proto-Polynesian
containing six kin terms, which retained the suffixing possessive morphology
inherited from Proto-Oceanic when a system of ‘dominant’ versus ‘subordinate’
possession was innovated elsewhere in the lexicon.

The argument here is thus that kin terms form a distinct, robustly defined lexical
subclass in Austronesian. This may be marked in various areas of the grammar – for
example, with dedicated morphology, or in differential treatment by the alienability
distinction. In those languages with an alienability distinction, the inalienable
morphology is additionally available to demarcate this lexical subclass. In this way,
this inherited structural pressure may have facilitated parallel developments of SIC
in the A\(n\) languages.

Similar parallel developments in other areas of the grammar have been reported
elsewhere in Austronesian. For example, Crowley (1991) looks at a widespread
pattern in central Vanuatu, in which the initial consonants of verbal roots mutate in
certain grammatical contexts. He shows that these mutations have not been
inherited; rather, they are parallel innovations in response to a phonological
asymmetry introduced in a common ancestor. Similarly, Blust (2017) suggests that
apparently independent developments of unusually conditioned sound changes in
disparate Austronesian subgroups is due to inherited phonological pressures; this is taken up by Arnold (2020b), who hypothesises that independent tonal developments conditioned by vowel height in four SHWNG and Oceanic languages may be explained by inherited phonetic disparities in the realisation of vowels. Outside of Austronesian, analogous developments can be observed in parallel grammaticalisation pathways of several features in different subgroups of Tibeto-Burman (LaPolla 1994); and in the multiple innovations of cross-linguistically rare wh-relatives in Indo-European languages (Gisbourne and Truswell 2018).

5.3 Externally motivated change: contact in Northwest New Guinea

The internally motivated development of SIC through inherited structural pressures, however, is not the whole story. First, it cannot account for the presence of SIC in genealogically diverse languages in Wallacea. Second, note that the geographic distribution of AN languages with SIC is not equal: Of the 11 AN languages with SIC (including Nadroga), eight are spoken in Northwest New Guinea (NWNG), in the region of the Bird’s Head peninsula. In this section, I will argue that, while AN languages are predisposed to develop SIC, in NWNG this development was facilitated by contact. In the scenario presented below, contact has been both inter-familial, i.e. between AN and NAN languages, and the two NAN families; and intra-familial, with most AN languages having developed the distinction through contact with the AN language Biak. The scenario presented here thus accounts both for the clustering of languages with SIC in NWNG, as well as the genealogical diversity of these languages.

First, some theoretical preliminaries. As we have seen throughout this paper, the forms used to express SIC are distinct in the various languages; where contact has been involved, we are therefore dealing with an instance of pattern replication (i.e., structural change), rather than replication of matter (i.e., borrowing of form; see Matras and Sakel 2007). Following standard terminology in the field (e.g. Weinreich 1963), I will therefore use the terms ‘replica’ and ‘model’ languages: the former referring to the language which undergoes contact-induced structural change, the latter to the language on which the change is based. Considering the three types of structural change discussed in Ross (2013) – lexical calquing, grammatical calquing, and metatypy – the development of SIC can be argued to be an example of grammatical calquing, in that it split a grammatical category in the replica language (‘inalienable nouns’) into two or more separate categories (‘inalienable: body parts’ and ‘inalienable: kin terms’). The complexification of the grammatical systems of the replica languages thus suggests a prior period of intense contact between the model
and the replica communities, involving early-childhood bilingualism within the replica community.

As discussed in Section 4.2, SIC in the East Bird’s Head family is reasonably old, as the distinction can be reconstructed to proto-Meax. Let us therefore first address the diffusion of SIC across the genealogical boundary between the two NAN groups, East Bird’s Head and Hatam-Mansim. Hatam and the three EBH languages have been in close contact for at least the past few hundred years. According to oral histories, these languages were once spoken near the headwaters of the Meyof, Rawara, and Wasian rivers in the southeastern interior of the Bird’s Head (Pouwer 1999: 476). At some point, speakers of Meyah, Sough, and Hatam migrated north and east to their present-day locations; these migrations were probably caused by increasing slave raids along the rivers of the south coast of the Bird’s Head, beginning in the 15th and 16th centuries (Reesink 2002b: 23). Speakers of Moskona remained close to the proto-Meax homeland; the Meyah migrations therefore likely caused the split of proto-Meax. This timescale for the break-up of proto-Meax is supported by the shallow relationship between Meyah and Moskona: With roughly 80% cognate vocabulary (Gravelle 2010: 7), the languages probably split a few hundred years ago.

As well as being geographically proximate, both historically and in the present day, speakers of Hatam and the EBH languages have had particularly close and non-antagonistic relationships over at least the last few centuries: Reesink (1999, 2002b) reports extensive intermarriage between these groups, and up to 30–40% of marriages were intertribal in the middle of the twentieth century (Pouwer 1958, cited in Reesink 2002b: 24–25). It seems clear that these interactions provided the ideal conditions for the kind of contact-induced structural change described here. Indeed, Reesink (1998) notes another linguistic correlate of this contact between Hatam and the EBH languages: the use of a verbal prefix to mark instruments on transitive verbs. He also discusses other features shared by the NAN languages of the Bird’s Head region more widely, such as spatial nouns.

What then is the relationship between SIC in these NAN languages and the AN languages of Northwest New Guinea? As mentioned above, SIC in most of the AN languages is remarkably recent; and these AN languages are spoken over a wide area, from Ambel and Salawati in the Raja Ampat archipelago, to Ambai on Yapen, to Sobei on the north coast of New Guinea. There is no reason to believe that Meyah, Moskona, or Hatam have been in contact with all of these AN languages over the last few hundred years – at least, not to the extent required for a feature like SIC to diffuse. Instead, I suggest SIC diffused between these NAN languages and the AN language Biak (or an ancestor of Biak), with which there is evidence for contact. Unlike the NAN...
languages, Biak has had a recent and very strong influence over the coasts and islands of NWNG; once it developed SIC, there was thus the opportunity for it to act as a vector of spread to the other local AN languages.

It is unknown whether SIC originated in the NAn languages, which then served as a model for Biak; or in an ancestor of Biak, which then served as a model for the NAn languages. The direction of diffusion depends on the antiquity of SIC in the Biakic languages. As mentioned in Section 4.1, SIC may be reconstructable to proto-Biakic, the most recent common ancestor of Biak and Roon. The time depth of the divergence of proto-Biakic is unknown. However, as will be returned to below, Biak has been the language of regional prestige in NWNG since at least the fifteenth century; proto-Biakic almost certainly diverged before that. If SIC is reconstructable to proto-Biakic, and if the rough dates given above for the split of proto-Meax a few hundred years ago are accurate, this would mean that SIC in proto-Biakic was historically prior to proto-Meax. This suggests that the spread was from proto-Biakic into proto-Meax (possibly via Hatam as an intermediary).

However, this scenario is unsatisfactory from both a geographic and a socio-linguistic point of view. At that point, proto-Meax was still spoken in the interior of the Bird’s Head, and proto-Biakic in Wandamen Bay some 150 km to the southeast. While the An languages have been influential on the NAn languages spoken on the coasts and islands of NWNG, this influence wanes towards the interior – for example, unlike those spoken in coastal regions, the NAn languages of the Bird’s Head interior (including Moskona) did not develop a clusivity distinction through contact with An languages (Klamer et al. 2008: 114–115). The likelihood that speakers of proto-Biakic and proto-Meax were in contact – either direct or indirect – to the extent required for the diffusion of SIC is thus low.

Alternatively, if SIC cannot be reconstructed in proto-Biakic – i.e. if SIC developed separately in Biak and Roon – then SIC is historically prior in the NAn languages. These languages could then have acted as the model for Biak. A disadvantage of this hypothesis is that this would necessitate a coincidental independent innovation of SIC in the NAn languages of the Bird’s Head; as seen in Section 5.1, the cross-linguistic data suggests this is unlikely. However, this hypothesis does account for the presence of SIC in proto-Meax. In this scenario, it is unclear which of the three NAn languages first innovated SIC,23 nor is it clear which served as the model for Biak. However, there is evidence for a particularly strong link between speakers of Biak and speakers of the Hatam-Mansim languages, suggesting the model may have been

\[23\] Indeed, SIC need not have originated in any of the extant languages: The change may have been innovated through contact with another NAn language, which is either extinct, or the descendants of which have subsequently lost the distinction.
Hatam. For example, the Hatam myth of the ancestral figure Digomang ends with the marriage of one of the Hatam clans with Biak speakers (Reesink 1999: 5); oral history from Raja Ampat records speakers of Biak taking wives from Hatam communities (own fieldnotes 2019); and intermarriage and friendly relations between the Biak and the Mansim in the region of present-day Manokwari have been documented since the end of the eighteenth century (Reesink 2002c). These kinds of interactions between speakers of Biak and Hatam-Mansim would have created the conditions necessary for the diffusion of SIC. 

I remain agnostic as to which of these two scenarios is more accurate, pending comparative work on Biakic, and further data from other languages in NWNG. For now, suffice it to say that SIC developed in a contact zone involving the Meax languages, Hatam, and Biak(ic) before it developed in any of the other languages of the region. Reesink (2002b: 21–22) describes another linguistic correlate specific to these languages: productive calques of adverb-like elements. He describes this as a “mini-areal feature”, which originated in the NAn languages and spread into Biak. Similarly, Reesink (2002c: 300–304) outlines lexical and morphological similarities between Biak and the two Hatam-Mansim languages, some of which may be due to contact.

Once Biak had developed SIC, there was then ample opportunity for the language to serve as the model for other AAnn languages in NWNG. For at least the past few centuries, Biak speakers have been very influential throughout the area: for example, they had strong socio-political ties with the powerful Tidore sultanate in Halmahera between the 15th and late 19th centuries (Andaya 1993: 104; Huizinga 1998); and since the early 20th century have been prominent in the church of the now-predominantly Christian Papua (van den Heuvel 2006: 2–3). As a result of this, Biak has historically been the main lingua franca of the region, particularly around Cenderawasih Bay and Raja Ampat (Arwam and Sawaki 2020).

Coupled with this top–down political and socio-cultural influence, the Biak are well-known throughout NWNG as explorers, wanderers, and settlers (Kamma 1972). They have a strong tradition of intermarriage with other local coastal groups, from Halmahera and Raja Ampat in the west, as far as the north coast of central New Guinea in the east (Arnold 2020a; Lee and Sawi 2005; own fieldnotes 2019). In particular, this intermarriage would have created the conditions of childhood

---

24 As SIC is only attested in Hatam, it cannot be reconstructed to proto-Hatam-Mansim. However, this may be due to the lack of data from Mansim: Reesink (2002c: 284) describes the data on body part possession in Mansim as “confusing”, and no mention is made of kin terms.

25 Contacts of a perhaps less intense nature between Hatam, East Bird’s Head, and other AAnn languages are also documented: For example, the Hatam hero Anggos was said to have visited the Wamesa area (Reesink 1999: 167); and Sougb genealogies indicate an ancestral link with the Irarutu (Ger Reesink, pers. comm.).
bilingualism necessary for the other AN languages to replicate the SIC distinction. Additional linguistic, archaeological, and cultural correlates of recent contact between Biak and other AN groups of NWNG abound: For example, shared ritual artefacts, mythological elements, and messianic movements (Corbey 2017; van der Leeden 1989; Mawene 2004); musicological similarities, such as the tifa drum-driven wor dance-songs (Kunst 1967: 124–125); and copious shallow lexical and grammatical loans from Biak into other local AN languages (e.g. Arnold 2018; Gasser 2019).

The contact-induced development of SIC discussed in this section is not at odds with the hypothesis proposed in the previous section, in which inherited pressures contributed to the development in the AN languages. Quite the opposite: There may be multiple causations of language change; and, as features are more easily transferred between typologically similar systems, inherited similarities indeed favour contact-induced developments. Similar examples from elsewhere in the world include convergences in the pronominal, verbal agreement, and TAM systems in Babylonian and Aramaic, facilitated by structural traits inherited from an earlier stage of Semitic (Beaulieu 2013); the loss of infinitives in Romanian and Greek due to an inherited competition between infinitives and finite structures, reinforced by contact in the Balkan area (Joseph 1983); and the mutual development of periphrastic perfect constructions in Latin and Greek based on pre-existing structural similarities (Bruno 2012). Enfield (2002) discusses a related scenario, whereby unrelated languages come to share structural pressures through contact, rather than inheritance; these contact-induced similarities then drive parallel innovations of a feature. He illustrates this with the independent but non-coincidental developments of structurally similar resultative constructions in genealogically diverse languages in mainland Southeast Asia. In fact, the combination of an inherited class of kin terms and the contact scenario presented above neatly accounts for the relative density of the AN languages with SIC in NWNG: while the AN languages were predisposed to develop SIC, Biak provided a conspicuous model for the change, thus galvanising the innovations in these languages.

6 Conclusions

In this paper, we have discussed Split Inalienable Coding, a feature that has not previously been discussed in the typological and theoretical literature. Based on a sample of 189 genealogically diverse languages, we have investigated the geographic distribution of semantically motivated SIC in the linguistic area of Wallacea and its surrounds. Of these sampled languages, 105 have a formal alienability distinction; of these, 13 are attested with SIC. All 13 of these languages are spoken within Wallacea. SIC cross-cuts genealogical boundaries in the area, occurring in several branches of
Austronesian and two unrelated non-Austronesian families; in most cases, SIC cannot be reconstructed any higher than the language in which it is attested, meaning we are looking at multiple recent innovations.

In Section 5, some explanations for the observed geographic and temporal distributions were explored. There it was shown that, since semantically motivated SIC is a rare phenomenon worldwide, the clustering of languages with SIC within Wallacea is unlikely to be a coincidence. It was argued that An languages are predisposed to develop SIC, due to the inheritance of a structurally defined lexical subclass of kin terms; if a language has a formal alienability distinction, the inalienable morphology is thus available to demarcate this subclass. This internal motivation explains the large geographic range of An languages with the distinction. It was argued that contact has also played a role in Northwest New Guinea. Based on the available evidence, it seems that SIC originated in a contact zone involving the Nan languages Meyah, Moskona, Hatam, and the An Biak; and that SIC subsequently diffused from Biak into the other An languages – a change that was facilitated by the pre-existing structural conditions. The external motivation for SIC thus accounts for both the genealogical diversity of languages with SIC in NWNG, as well as the relative frequency of An languages with the distinction in this region. This paper has therefore contributed to theoretical discussions on parallel developments in related languages; the relationship between internally and externally motivated language change; and the development of nominal classification systems. It has also added to what is known about the history of contact and population movements in NWNG.

However, several outstanding questions remain. With regards to the diffusion of SIC across Northwest New Guinea, the ultimate origins are unknown: What mechanisms were involved in the first innovation in the region? Second, something commented on at several points in this paper is that in several An languages with SIC, the split is limited to a single cell of the paradigm, when the PossR is 3sg. However, in all of the Nan languages with the distinction, as well as Biak, the split is found at least for all singular PossRs. If Biak did serve as the model for the other An languages of NWNG, why was the split limited to the 3sg cell in the replica languages? With regards to the typology of SIC, a more systematic survey is recommended, both in the languages spoken further to the east of the sampled area, to confirm that SIC is a feature specific to Wallacea, and not one of Linguistic Melanesia more widely; and worldwide, to verify that semantically motivated SIC is indeed a cross-linguistically rare phenomenon. Finally, it was commented on above that body parts and kin terms are semantically and syntactically distinct. These distinctions manifest in the possessive constructions of languages around the world, both synchronically and diachronically: For example, in those languages where body parts and kin terms are treated differentially by the alienability distinction itself, with only one class or the other
patterning as inalienable. The final outstanding question is thus: If body parts and kin terms are indeed so distinct, why is SIC not more frequently attested in the languages of the world?

**Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1st person</td>
</tr>
<tr>
<td>2</td>
<td>2nd person</td>
</tr>
<tr>
<td>3</td>
<td>3rd person</td>
</tr>
<tr>
<td>AN</td>
<td>Austronesian</td>
</tr>
<tr>
<td>CMP</td>
<td>Central Malayo-Polynesian</td>
</tr>
<tr>
<td>EBH</td>
<td>East Bird's Head</td>
</tr>
<tr>
<td>EXCL</td>
<td>exclusive</td>
</tr>
<tr>
<td>INCL</td>
<td>inclusive</td>
</tr>
<tr>
<td>NAn</td>
<td>non-Austronesian</td>
</tr>
<tr>
<td>NP</td>
<td>noun phrase</td>
</tr>
<tr>
<td>NWNG</td>
<td>Northwest New Guinea</td>
</tr>
<tr>
<td>Oc</td>
<td>Oceanic</td>
</tr>
<tr>
<td>PL</td>
<td>plural</td>
</tr>
<tr>
<td>POSS</td>
<td>possessive</td>
</tr>
<tr>
<td>POSSD</td>
<td>possessed</td>
</tr>
<tr>
<td>POSSR</td>
<td>possessor</td>
</tr>
<tr>
<td>RASH</td>
<td>Raja Ampat-South Halmahera</td>
</tr>
<tr>
<td>SG</td>
<td>singular</td>
</tr>
<tr>
<td>SHWNG</td>
<td>South Halmahera-West New Guinea</td>
</tr>
<tr>
<td>SIC</td>
<td>Split Inalienable Coding</td>
</tr>
<tr>
<td>TAM</td>
<td>tense/aspect/mood</td>
</tr>
<tr>
<td>TNG</td>
<td>Trans-New Guinea</td>
</tr>
<tr>
<td>WNG</td>
<td>West New Guinea</td>
</tr>
<tr>
<td>WOC</td>
<td>Western Oceanic</td>
</tr>
</tbody>
</table>

**Acknowledgements:** I am grateful to the editors, Antoinette Schapper and Emily Gasser, as well as Owen Edwards, Michael Franjieh, Harald Hammarström, and an anonymous reviewer, all of whom gave valuable feedback on earlier drafts of this paper. My thanks also to the following, for providing additional data, sources, and/or discussion: Peter Austin, Jürgen Bohnemeyer, Don Daniels, Marcel Erdal, Alex François, David Gil, Harmut Haberland, Martin Haspelmath, Jeffrey Heath, David Kamholz, František Kratochvíl, Randy LaPolla, Neil Myer, Åshild Næss, Sebastian Nordhoff, Maia Ponsonnet, Bert Remijsen, Linn Iren Sjānes Rødand, Jean Rohleder, Poppy Siahaan, Eline Visser, Ron Whisler, Jeroen Willemsen, and Philip Yampolsky. Terminological discussions with Greville Corbett, Matthew Dryer, and Hedvig Skirgård were particularly helpful. Thanks also to participants at the Second Workshop of the Languages of Wallacea, and the 10th, 12th, and 13th Austronesian and Papuan Languages and Linguistics conferences, where aspects of this work were presented. All errors and misunderstandings are my own.
References


Austin, Peter. 2015. A reference grammar of the Mantharta Languages, Western Australia. Available at: https://www.academia.edu/15938656/A%20Grammar%20of%20the%20Mantharta%20Languages%20of%20Western%20Australia (accessed 17 November 2021).


Jackson, Jason. 2014. A grammar of Irarutu, a language of West Papua, Indonesia, with historical analysis. Honolulu: University of Hawai‘i at Mānoa (PhD dissertation).


Karvovskaya, Lena. 2018. The typology and formal semantics of adnominal possession. Utrecht: LOT.


Reesink, Ger P. 2002b. The eastern Bird’s Head languages compared. In Ger P. Reesink (ed.), Languages of the eastern Bird’s Head (Pacific Linguistics 524), 1–44. Canberra: Research School of Pacific and Asian Studies, Australian National University.


