Alessandra Barotto and Andrea Sansò*

When possessor and argument indexes coincide. A cross-linguistic survey

https://doi.org/10.1515/flin-2024-0013
Received August 29, 2024; accepted January 16, 2025; published online February 24, 2025

Abstract: Cross-linguistically, it is not uncommon to find the same person indexes in noun and verb phrases, referring, respectively, to the possessor of the noun and to various (combinations of) arguments (A, S and P). This phenomenon has never been the object of a dedicated typological survey. This article seeks to address this gap by investigating cases in which possessor and argument indexes coincide in a variety sample of 225 languages. The parameters of analysis include the (combinations of) arguments indexed, the type of coincidence (total, covering all persons in a given index set, or partial, limited to some persons), the segmental similarity between coinciding argument/possessor indexes and independent pronouns, the restriction of coinciding indexes to specific TAM and/or voice constructions, and the type of possessor indexes involved (alienable vs. inalienable). The typological survey is further complemented by a discussion of diachronic scenarios that may explain these patterns of coincidence, including (but not limited to) parallel developments from independent pronouns and the reinterpretation of nominalized forms of the verb (where possessor indexes correspond to an argument of the full verb) as predicative forms.

Keywords: argument indexes; possessor indexes; nominalization; source-oriented typology

1 Introduction[1]

In various languages of the world, the same person indexes are used on both noun and verb phrases. On noun phrases, they refer to the possessor of the noun, whereas on verb phrases, they can refer to different arguments: (i) the only argument of a

1 This work is the result of the collaboration between the two authors. Specifically, Alessandra Barotto wrote Sections 2, 3.1, 3.3, 3.4, 3.6, 3.7, 3.8, and 4.1, while Andrea Sansò wrote Sections 1, 3.2, 3.5, 4.2, and 5. We would like to thank two anonymous reviewers for their comments on the first version of the paper.

***Corresponding author: Andrea Sansò**, Dipartimento di Studi Umanistici, Università di Salerno, Via Giovanni Paolo II, 132 – 84084 Fisciano, Salerno, Italy, E-mail: asanso@unisa.it
Alessandra Barotto, Dipartimento di Scienze Teoriche e Applicate, Università dell'Insubria, Via J.H Dunant 3, I-22100, Varese, Italy, E-mail: alessandra.barotto@uninsubria.it

monovalent predicate (henceforth S), (ii) the most agentive argument of a bivalent predicate (henceforth A), (iii) the least agentive argument of a bivalent predicate (henceforth P), or (iv) various combinations of (i)-(iii). An example is found in Mono: in the construction in (1a), the S is indexed on a postverbal imperfective particle following the reduplicated uninflected verb. The bound person form *-na* is the same that is used in (1b) to index the possessor, attached to a possessor-indexing host (*sa-*):

(1) Mono [mte] (Oceanic; Palmer 2011: 687)[2]

a. *Soipa* *ma-mate* ***e-na***
S. RED-die IPFV-3SG.SBJ
'Soipa was dying.'

b. *kai-gu* *sa-**na*** *auau*
brother-1SG.POSSR GENP-3SG.POSSR dog
'my brother's dog'

In some languages, coincidence between possessor/argument indexes is restricted to specific voice or tense-aspect-modality (TAM) constructions, while in others, it extends throughout the entire verbal paradigm. In Mono, for instance, the coincidence illustrated in (1) is only found in realis past forms of the verb. In contrast, in Ineseño the possessor-indexing prefixes on nouns exemplified in (2a) also index the A and S arguments across all tenses, aspects and moods of the verbal complex (cf. (2b-c)):

(2) Ineseño [inz] (Central Chumashan; Applegate 1972: 167, 308)

a. ***k**-ičtiṅ* ***p**-pu*
1-child 2-arm
'my child' 'your arm'

b. ***k**-axšiš*
1-call
'I call him.'

c. *ʔinina-**p**-kuy*
PST.INT-2-take
'You were going to take it.'

These correspondences in the person indexes occurring on nouns and verbs have been explained in terms of semantic or functional affinities between specific arguments and the possessor. Seiler (1983: 22), for example, argues that the coincidence

This work was supported by the University of Insubria through a Junior Post-Doc Grant (year 2021; Principal Investigator: Andrea Sansò). All remaining errors and shortcomings are our own.

2 On their first occurrence, the languages cited in this article are followed by their ISO 639-3 code. Names and genealogical classifications of the languages are based on Glottolog (Hammarström et al. 2023, last accessed December 2023).

between possessor indexes and P indexes depends on the conceptualization of the possessor as an inactive patient (especially in the case of inalienable possession). Other scholars (e.g. Hofling 1990) suggest that affinities between the possessor and the A are frequent, especially in languages with ergative alignment, because the possessor in prototypical possessive relationships (e.g. kinship, body parts) "is prototypically animate, given information, characteristics it shares with transitive subjects" (Hofling 1990: 558; see the discussion in Siewierska 1998: 1–4). Source-oriented explanations of these correspondences have also been proposed, either based on the historical development of individual languages (Palmer 2011; Starosta et al. 2009 [1982], among others) or on universal diachronic tendencies (Givón 1976, among others). When possessor and argument indexes coincide, both may derive from independent pronouns through parallel (and independent) processes of grammaticalization. Since Givón's (1976) seminal paper, it has been widely recognized that bound person indexes typically derive from the grammaticalization of anaphoric independent pronouns (see Siewierska 2004: 262; see also Seržant 2021 for a critical perspective). This possibility is illustrated by Moskona, where argument and possessor indexes display segmental similarities with independent pronouns, suggesting a diachronic pathway from independent pronouns to bound person indexes:

(3) Moskona [mtj] (East Bird's Head; Gravelle 2010: 106, 186)

	A/S and inalienable possessor prefixes	Independent pronouns
1SG	*di-*	*dif*
2SG	*bi-*	*bua*
1PL	*mi-*	*mif*
2PL	*yi-*	*yua*

In other cases, there is little or no segmental similarity between argument/possessor indexes and independent pronouns, making the scenario of a parallel grammaticalization of the latter into the former less plausible or implausible. Lopnor Uyghur is a case in point:

(4) Lopnor Uyghur [uig] (Turkestan Turkic; Abdurehim 2014: 154)

	A/S prefixes (past/habitual) and inalienable possessor prefixes	Independent pronouns
1SG	*-m*	*män*
2SG	*-ŋ*	*sän*
1PL	*-q*	*biz*
2PL	*-ŋla(r)*	*silä(r)*

The aim of this article is to provide a cross-linguistic survey of the patterns of coincidence between argument and possessor indexes. The survey is based on a variety sample of 225 languages, in which 105 cases of coincidence have been identified. Through both

quantitative and qualitative analysis, the survey examines the prevalence and distribution of this phenomenon across languages, addressing the following questions:

(a) which arguments (A, S, P) and combinations of arguments are more likely to be indexed by possessor indexes?
(b) Are there any correlations between (combinations of) arguments indexed by possessor indexes and other typological features of the language (e.g. alignment type)?
(c) In which verb constructions (in terms of voice or TAM) is the coincidence between argument and possessor indexes more likely to occur?
(d) In those languages where there is a grammatical distinction between alienable and inalienable possession, which types of possessor indexes (alienable/inalienable) are more likely to coincide with argument indexes? Are there any correlations between (in)alienability and the argument(s) indexed?

The answers provided by this article will at times be inconclusive and, in certain cases, even negative: some synchronic correlations discussed in the literature are simply not supported when a sufficiently large language sample is considered. Nonetheless, addressing the questions posed above will also lead us to explore the explanatory potential of diachronic scenarios proposed to account for the coincidence between argument and possessor indexes other than the one mentioned above, i.e. the parallel development from independent pronouns. As will be discussed below (Section 4), some of these alternative scenarios point to a diachronic functional shift from nominal to verbal marking, in which a nominalized form of the verb, with possessor indexes corresponding to an argument or a combination of arguments of the full verb, has been reinterpreted, taking on a predicative function. In other diachronic scenarios, nominalization is not involved, and the origin of possessor-like indexation of arguments draws on different mechanisms altogether.

The article is organized as follows. In Section 1.1, we will provide some definitions and delimitations of the study. Section 1.2 is devoted to a discussion of a few previous studies dealing with this phenomenon. The sample is described in Section 2, while in Section 3 each parameter of classification of the patterns of coincidence found in the sample will be discussed in detail. In Section 4, the diachronic scenarios explaining the coincidence patterns will be introduced, and finally in Section 5 some concluding remarks will be presented.

1.1 Terminological and methodological issues

We adopt the notion of *index* from Haspelmath (2013). *Index* is a cover term for clitic and affixal person forms that allows us to avoid using terms such as pronominal or

agreement markers, whose delimitation may be problematic in cross-linguistic comparison. Indexes are typically organized into sets: an *index set* is a set of indexes used with the same syntactic function. For instance, a set of clitic/bound forms exclusively referring to subjects of transitive verbs (A) can be labeled *ergative index set*. Similarly, a set of forms that indexes objects of transitive verbs (P) may be referred to as an *accusative index set*.

This article will survey all the index sets in each language of the sample that are used at the same time on verbs to refer to A, S, P, or any combination of them, and on nouns to refer to the possessor. All these sets can be considered as cases of *coincidence* between argument and possessor indexes (see below for the definition of coincidence).

A few methodological choices should be clarified before discussing the survey results. First, we have included in our sample a few index sets that consist, either wholly or partially, of forms defined as free in the grammatical description of the language. This inclusion expands Haspelmath's (2013) definition of indexes, which only covers affixal and clitic forms, drawing instead on Creissels' (2005) more fine-grained typology of pronominal reference. This typology is based on African languages, but its criteria have broader applicability: according to Creissels (2005: 50), there are grounds to consider the morphemes termed subject pronouns and object pronouns in African languages as bound morphemes, regardless of how they are categorized in the grammatical descriptions of these languages. Creissels proposes three specific criteria for identifying indexes ("pronominal markers" in his terminology): (a) they are obligatory; (b) their position is distinct from that of the corresponding noun phrase; (c) they show a relatively high degree of morphophonological interaction with TAM and polarity markers, unlike nouns or their equivalents, which do not generally display similar interactions with these markers (Creissels 2005: 50). In Creissels' classification, cases in which only one or two of the criteria (a)–(c) are met qualify as Stage I indexes. Stage I indexes are typically in a complementary distribution with the corresponding noun phrases, but can also cooccur with the noun phrase, especially if the latter is topicalized in a dislocated construction, similarly to Romance clitic pronouns (Creissels 2005: 44). Moreover, they represent the first stage in an ideal evolutionary path "whose starting point is the cliticization of free pronouns" (Creissels 2005: 44).

Based on Creissels' criteria, we have considered as Stage I indexes, and thus included in our survey, cases like the following:

(a) The index set composed of forms that are defined as free is *positionally less free than another set of independent pronouns* in the same language.

Consider the following data from Tigon Mbembe:

(5) Tigon Mbembe [nza] (Benue-Congo; Richter 2014: 186-189)

	Independent		Subject pronouns		Object pronouns	
1	SG *mī*	PL *ɲí*	SG *N̄=*	PL *ɲí*	SG *ŋɛ̄*	PL *ɲí*
2	SG *bū*	PL *nɛ̄*	SG *a*	PL *nɛ̄*	SG *(w)ō*	PL *nɛ̄*
3 HUM	SG *gɛ̄*	PL *bō*	SG *ē*	PL *bō*	SG *gɛ̄*	PL *bō*
3 NHUM	SG *gɛ̄*	PL *gɛ̄*	SG *ē*	PL *gɛ̄*	SG *gɛ̄*	PL *gɛ̄*

Tigon Mbembe distinguishes three sets of pronouns: subject, object, and independent pronouns, as shown in (5). All three sets include phonologically free words (except for the 1st singular subject pronoun, a homorganic nasal cliticized on the verb). Both independent pronouns and object pronouns can be used to refer to the P. However, the position of object pronouns is constrained: they must immediately follow the verb. On the contrary, the position of independent pronouns with respect to the verb is less constrained when used to index the P: in the imperative clause in (6a), for instance, an independent pronoun is used for the object, as imperative clauses follow a VSO order, with an element (S) interrupting the VO sequence. The same is true for (6b), where an adverb appears between the verb and the object. The 1st and 2nd person singular forms in the set of object pronouns are segmentally quite different from their independent pronoun counterparts, while the plural forms are identical across all three sets. Furthermore, object pronouns fully coincide segmentally with possessor indexes on nouns, though the latter are affixal, as shown by the contrast between (6c) and (6d):

(6) Tigon Mbembe (Richter 2014: 289, 384, 188, 174)

a. *nāfyě nɛ̄ mī*
help 2PL.SBJ 1SG.INDP
'Help (pl) me!'

b. *ā lē ŋɔ̄ pɔ́-é **mī** mɔ̄ mū*
2SG.SBJ CONS refuse ever-NEG 1SG.INDP DM NEG
'You will never refuse me.'

c. *ē vè **ŋɛ̄***
3SG.SBJ see 1SG.OBJ
'S/he sees me.'

d. *tā-**ŋɛ̄** tē vè-tē tā àfā*
house-1SG.POSSR stand see-STAT house chief
'My house is opposite the chief's house.'

Given the positional constraints on object pronouns, the existence of another set of independent pronouns (which are typically used for emphasis and in contrastive focus), and the full segmental coincidence between object pronouns and possessor indexes, we have included this index set in our dataset as a case of coincidence between argument and possessor indexes.

(b) The index set composed of free forms may meet Creissels' criteria (b) and (c), displaying a syntactic position different from the one of the corresponding noun phrase and a higher degree of interaction with TAM and polarity markers.

This is what happens in Bullom So (cf. Childs 2011: 224–225). The so-called object pronouns in Bullom So given in (7) are a set of free forms that fully coincide segmentally with possessive pronouns:

(7) Bullom So [buy] (Southern Mel; Childs 2011: 58)

Object/possessive pronouns	
1SG	*mì*
2SG	*mɔ̀*
3SG	*wɔ̀*
1PL	*hìn*
2PL	*nyà*
3PL	*ŋà*

Object pronouns, unlike full nominals in object function, must appear pre-verbally, as exemplified in (8a) versus (8b), and show a relatively higher degree of interaction with TAM or polarity markers than nouns in object function. When TAM auxiliaries are present, for instance, object pronouns appear after the auxiliary and before the lexical verb, and are hosts for negation, as in (8c), while nouns appear after the verb in the same syntactic context, and the negative marker is hosted by the auxiliary, as in (8d):

(8) Bullom So (Childs 2011: 67, 213, 209, 247)

a. *yá* ***mɔ̀*** *pɛ̀* *ké*
1SG.SBJ 2SG.OBJ later see
'I will see you later.'

b. *à* *ké* ***nɔ́*** *kò* *tò* *àyì*
1SG.SBJ see.PERF person to forest inside
'I saw someone in the forest.'

c. *à* *cé* ***wɔ̀-én*** *mìntá*
1SG.SBJ IPFV.AUX 3SG.OBJ-NEG fearless
'I am afraid (not fearless) of him.'

d. *ù-pɛ̀ŋkàrɛ́ŋ* *ù-cɛ́* *wɔ̀* *cè-ɛ́ŋ* *kìmɔ̀* *ì-pál* *fúfáfú*
NCM-frog NCM-DEF 3SG.SBJ IPFV.AUX-NEG flee NCM-sun nothing
'The frog does not flee the sun without cause.'

Moreover, much like Romance clitics, Bullom So object pronouns can cooccur with the corresponding noun phrase when the latter is dislocated for topicalization (Childs 2011: 224–225). Similar reasoning led us to include the index set of so-called object pronouns in Ngemba [nge] (Southern Bantoid; Leroy 2007: 207–209) and the

sets of free forms labeled as "possessive" or "genitive" pronouns in two Austronesian languages, Cebuano and Ida'an. In Cebuano [ceb] (Greater Central Philippine), possessive pronouns (free words) can be preposed to verbs in non-actor voice to index the A argument – a feature not shared by nouns in the same syntactic environment (Tanangkingsing 2009: 123). In Ida'an [dbj] (North Borneo), NPs and free pronouns indexing the A argument in the so-called undergoer voice appear in the same syntactic slot, immediately after the verb. Only pronouns, however, can occasionally appear preverbally in this kind of voice, possibly due to the omission of a generic auxiliary in sequences of clauses (Goudswaard 2005: 133, 191).

(c) The index set composed of free forms is actually *mid-way between free forms and clitics*.

In Lundayeh [lnd] (North Borneo; Mortensen 2021: 112–113), only some forms of the set of genitive pronouns, indexing the A argument and the possessor, have the status of free words. Similarly, in Northern Subanen [stb] (Greater Central Philippine; Daguman 2004: 64) the set of genitive forms of personal pronouns, which index the A argument and the possessor, comprises forms that are free words as well as enclitics (some persons have only the free form, while other persons have both forms). Finally, in Cocama-Cocamilla [cod] (Mawetí-Guaraní), short forms of pronouns indexing the A, the S, and the P are defined as free words (Vallejos 2016: 208), though they have an intermediate status within a tripartite system that also comprises long forms and proclitics (Vallejos 2016: 201). All three forms show segmental similarities, suggesting a diachronic cline going from long forms to proclitics, with short forms representing an intermediate stage.

A second terminological concern was the adoption of a convenient term to define the phenomenon. We have opted for the term *coincidence* (and for the verb *to coincide*). When the same index set is used, totally or partially, to refer to the A, the S, and/or the P with verbs and to the possessor with nouns, we will say that indexes for the A, the S, and/or the P and indexes for possessors *coincide*. This choice might sound problematic, as it could lead to the misleading conclusion that there are two index sets that happen to be homophonous, when in fact there is just one set with two functions (the indexation of certain arguments on verbs and of possessors on nouns). While we fully acknowledge this concern, we still believe that this term is the most neutral one to describe the phenomenon in question. Talking of coincidence does not imply any primacy of one function over the other, either synchronically or diachronically.

1.2 Previous studies

As already mentioned, the phenomenon of coincidence between argument and possessor indexes has been discussed in language- or family-specific studies, but the

only typological survey available (Siewierska 1998) presents certain methodological limitations that may affect the robustness of its conclusions. In the next two sections, we will discuss studies on individual languages first (1.2.1), and then we will move on to deal with Siewierska's (1998) typological survey (1.2.2).

1.2.1 Studies on individual languages and language families

The question of coincidence between argument and possessor indexes has been noted several times in the descriptions of single languages or language families (e.g. Gildea 1998; Palmer 2011; Starosta et al. 2009 [1982]). All the studies that have dealt with this phenomenon converge in offering source-oriented explanations of the coincidence, hypothesizing that a nominalized form of the verb is responsible for the shared marking of arguments and possessors. However, these studies diverge in their views on the diachronic processes by which the nominalized form of the verb may have acquired a predicative function. Starosta et al. (2009 [1982]), for instance, propose that in Austronesian voice constructions, the coincidence between argument and possessor indexes results from the conventionalization of copular constructions of the type 'X is **my**-beaten one', originally employed as discursive strategies to foreground different arguments (e.g. the P). Similarly, Gildea (1998: 163–164 and *passim*) explains the possessor-like marking of Ps in various tenses in Cariban languages as emerging from nominalized forms in equative clauses, where the P is marked as a possessor and the A as an oblique. These forms become "the new inflected verb, with the possessor becoming an absolutive argument and the oblique agent phrase becoming an ergative argument" (Gildea 1998: 163).

Another study explicitly addressing the use of possessive or originally possessive morphology to index verbal arguments is Palmer (2011), whose focus is on Northwest Solomonic languages (Oceanic). These languages retain to varying degrees the reconstructed Proto-Oceanic clause structure, consisting of a particle/proclitic that encodes reality status and/or the person and number of the subject, followed by the lexical verb, and a suffix/enclitic encoding the person and number of the object. In addition to this clause structure, many Northwest Solomonic languages have developed an innovative structure involving a postverbal particle indexing the subject of the preceding verb. The subject-indexing morphology in this innovative structure "is also either employed in nominal constructions to index a possessor, or is derived from morphology that had that function at an earlier stage in the language" (Palmer 2011: 687), as illustrated in (1) above from Mono, repeated here as (9):

(9) Mono (Palmer 2011: 687)
  a. *Soipa* *ma-mate* ***e-na***
  S. RED-die IPFV-3SG.SBJ
  'Soipa was dying.'

b. *kai-gu* *sa-**na*** *auau*
brother-1SG.POSSR GENP-3SG.POSSR dog
'my brother's dog'

Semantically, the innovative structures in Northwest Solomonic languages encode a variety of TAM functions, expressing non-past events or imperfective aspect, permission, prohibition, ongoing states/events, negated situations, and more. In other languages of this branch, the innovative structure serves discourse functions: in Kokota [kkk] and Cheke Holo [mrn], for instance, it highlights events having high saliency (Palmer 2011: 710). When used with possessor-indexing function, the suffixes in question attach to two hosts, one for general possession as exemplified in (9b), and another for "consumed" possession. Consumed possession indicates that the possessor "has consumed, is consuming, or will consume the possessum (such as with items eaten or drunk)" (Palmer 2011: 688). In the languages that retain it, the formal distinction between the two hosts may be repurposed in the innovative construction to encode a range of TAM distinctions: for example, the consumed host may be reinterpreted as a perfective marker and the general one as an imperfective marker. The innovative structures in Northwest Solomonic languages are said to have originated from a construction at the level of the protolanguage that was nominal in nature. According to Palmer (2011: 732), "the originating construction […] was a nominalized clause, not a nominalized verb", namely a clause-initial recapitulative nominalized clause, where the A or the S are indexed as possessors, a pattern that can be reconstructed, with some uncertainty, to Proto-Northwest Solomonic (Palmer 2011: 731–734). The function of recapitulative nominalized clauses was to present background information necessary to provide a temporal setting for the action/situation described in the following clause. As argued by Ross (1982), recapitulative clauses typically contain continuing action rather than punctiliar or perfective events (as in *She rolled along the ground*$_{[foreground]}$. ***Rolling along the ground***$_{[background]}$, *her body gets covered with mud*$_{[foreground]}$), which would explain the prevalence of (broadly speaking) imperfective uses of these structures once they evolved into independent clauses.

1.2.2 Siewierska (1998)

Siewierska (1998) is the only typological account of the identity of person marking on nouns and verbs. However, her focus differs from ours: her study aims to identify affinities in the marking of pronominal possessors and subjects/objects, and to correlate these affinities with a language's dominant alignment type. Based on a 157-language sample, she takes into account the cases of correspondence in form between argument and possessor indexes and shows that some form of correspondence is attested in 83 % of her sample with notable areal differences – South America showing the highest incidence of this phenomenon, and Australia the lowest. Her analysis also reveals that

there is no statistically significant difference between the A and the P in terms of their affinity with possessors. Possessors align with the A in 39 % of the cases in which there is correspondence in form between argument and possessor indexes, while the P shows affinities with the possessor in 41 % of the cases. What is significant, however, is the areal distribution of these affinities: possessor/P affinities are more likely to be found in New Guinea, Africa, Southeast Asia and Oceania, whereas possessor/A affinities are more frequent in Meso-America, South America and Eurasia (Siewierska 1998: 21). As far as alignment is concerned, no clear correlations emerge between the dominant alignment type of a language and affinities of the possessor with A versus P (Siewierska 1998: 24), although languages with ergative verbal alignment "do exhibit a preference for possessor affinities with the A (55 %) as compared with just the O (9 %) *[P in our terminology]*" (Siewierska 1998: 36). Similarly, there are no clear tendencies regarding the propensity of inalienable possessor indexes to align with P indexes and of alienable possessor indexes to align with A indexes, although inalienable possession slightly favors the P (Siewierska 1998: 32–34).

All these conclusions are based on a broader definition of similarity/correspondence than the one adopted in this study (see below, Section 3.2), ranging from full/partial segmental coincidence to cases in which there is merely some similarity but no actual segmental coincidence. These cases are problematic because they inflate the figures for the diffusion of the phenomenon. Siewierska's notion of similarity covers as diverse situations as cases in which there is identity between one or two forms in the set, cases in which there is no true identity but most forms show some resemblance, and cases in which resemblance is restricted to just one or two forms. The following index sets from Larike, for instance, are provided as a case of correspondence in form between argument and possessor indexes (Siewierska 1998: 16):

(10) Larike [alo] (Central Malayo-Polynesian; Laidig 1993: 320–321)[3]

	P indexes	Possessor indexes
1SG	*-aʔu*	*aku-*
2SG	*-ne*	*amu-*
3SG.HUM	*-ma*	*mana-*
3SG.NHUM	*-a*	*ir-*
1PL.EXCL	*-ami*	*amir-*
1PL.INCL	*-ite*	*iter-*
2PL	*-imi*	*imir-*
3PL.HUM	*-mati*	*matir-*
3PL.NHUM	*-ri*	*irir-*

3 The Larike data in Siewierska (1998: 16) are displayed in a messy fashion, due to the conflation of the two sets in a single column, which induces the idea that the two sets are segmentally identical.

As is clear from (10), similarity is limited to plural persons, the only difference being a formative *-r-* that can be isolated in the possessive paradigm. The two sets have different morphological status: P indexes are suffixal, whereas possessor indexes are prefixal. Moreover, the P suffixes are segmentally very similar to independent pronouns (cf. (11)), from which they possibly derive through a very common process of grammaticalization (see Givón 1976):

(11) Larike (Laidig 1993: 320–321)

	Independent pronouns
1SG	*aʔu*
2SG	*ane*
3SG.HUM	*mane*
3SG.NHUM	/
1PL.EXCL	*ami*
1PL.INCL	*ite*
2PL	*imi*
3PL.HUM	*mati*
3PL.NHUM	/

If the parameter of similarity with independent pronouns is not taken into account when postulating similarities in argument/possessor indexes, there is a high risk of overgeneralizations, potentially undermining any conclusions drawn from cross-linguistic data. A further problematic case, in which lack of comparison with independent pronouns produces similar distortions, is Tonkawa. The paradigms presented in the first two columns in (12) are considered by Siewierska (1998: 16) to be a case of similarity between patient and possessor indexes. The third column reports the set of independent nominative and accusative pronouns, overlooked by Siewierska (1998) and taken from the primary source (Hoijer 1933):

(12) Tonkawa [tqw] (isolate; Hoijer 1933: 68, 134, 122)

	P indexes	Possessor indexes	Nominative/accusative independent pronouns
1SG	*ge-*	*ca·-*	*ca·-ya/ca·-cik*
2SG	∅	*na·-*	*na·-ya/na·-yak*
3SG	∅	*ʾa·-*	*ʾa·-ye·-le/ʾa·-ye·-lak*
1PL	*geu- (ge-we-)*	*geuca-*	*geu-ca·-ga/geu-ca·-gak*
2PL	*we-*	*wena-*	*we-na·-ga/we-na·-gak*
3PL	*we-*	/	*ʾa·-we·-ga/ʾa·-we·-gak*

In this case as well, the possessive prefixes display rather close segmental similarity with the set of independent pronouns, whereas the P prefixes do not. The only similarity between P and possessive prefixes – the one found in 1st and 2nd person plural forms – is likely too weak to classify Tonkawa as a language displaying similarity between argument and possessor indexes.

Incorporating independent pronouns into the general analysis is essential to discriminate genuine cases of identity between argument and possessor indexes, which may be indicative of a diachronic spread from one type to the other, from cases of parallel development from the same source (the independent pronoun). In the latter case, synchronic similarity is merely a synchronic reflex of two different, and potentially unrelated, diachronic processes.

2 Sample and methodology

The sample includes 225 languages, expanding the 223-language sample discussed in Mattiola (2020). The original 223-language sample follows the Diversity Value technique (Rijkhoff et al. 1993; Rijkhoff and Bakker 1998), which accounts for the internal diversity of families by considering the number of genera within them: families showing more genera are more represented in the sample than families with fewer genera, irrespective of the overall number of languages of the family. In this way, the resulting sample allows the emergence of the maximum degree of variety. A few changes have been made to the internal structure of this sample, based on the availability of reliable grammatical descriptions. When a language needed to be replaced, however, the closest available language with a suitable grammatical description was chosen to preserve the original balancing as much as possible.[4] One

4 The following languages in Mattiola's (2020) 223-language sample were replaced:

Language	Replaced by
Tboli [tbl] (Austronesian, Bilic)	Sula [szn] (Austronesian, West Central Maluku)
Central Palawano [plc] (Austronesian, Greater Central Philippine)	Eastern Tawbuid [bnj] (same genus)
Pampanga [pam] (Austronesian, Central Luzon)	Lundayeh [lud] (Austronesian, North Sarawakan)
Iraya [iry] (Austronesian, North Mangyan)	Kola [kvv] (Austronesian, Aru)
Ibatan [ivb] (Austronesian, Batanic)	Paku [pku] (Austronesian, Basap-Greater Barito)
Tehuelche [teh] (Continental Chonan)	Selk'nam (Insular Chonan)
Sanapaná [spn] (Lengua-Mascoy, Eastern Enlhet-Enenlhet)	Enxet Sur [enx] (Lengua-Mascoy, Lengua)
Chilcotin-Nicola [clc] (Central British Columbia Athabaskan)	Central Carrier [crx] (same genus)
Beria [zag] (Eastern Saharan)	Dazaga [dzg] (Western Saharan)
Makasae-Makalero [mkz] (Timor-Alor-Pantar, East Timor)	Abui [abz] (Timor-Alor-Pantar, Nuclear Alor-Pantar)
Karok [kyh] (isolate)	Taulil [tuh] (the only extant language of the small Taulil-Butam family)
Gaa [ttb] (Atlantic-Congo, Bantoid)	Wawa (same genus)

language was removed without replacement,[5] and three additional languages were included. These additions do not significantly affect the proportional distribution of languages across macro-areas and language families.[6]

For each language in the sample, we have looked for phenomena of coincidence between argument and possessor indexes, whether total or partial (see below, Section 3.2). The analysis of sources was conducted in two stages. An initial review generated the first population of the database. The results were then rigorously tested in a second stage to assess their accuracy and identify any instances that may have been missed in the first step. Each author analyzed half of the languages in the sample, while also reviewing each other's findings to ensure consistency across the analysis. The following discussion will be based only on the languages in which such phenomena are attested.[7]

3 A typology of argument and possessor indexes coincidence

We have found 105 cases of coincidence between possessor and argument index sets in the sample. These cases are found in 87 languages out of 225 (38.67 %), and in 46.52 % of the 187 languages with person indexation in the sample. As discussed above in relation to Siewierska's analysis, where coincidence was found in 83 % of the languages, a more restrictive definition of the phenomenon results in a significantly lower figure for its cross-linguistic incidence. Some languages display more than one instance of coincidence between the two index types. Table 1 displays the areal distribution of these languages.

---

5 The language in question is Adamorobe Sign Language [ads], which was removed due to the theoretical difficulties in identifying the correlates of the phenomenon under investigation in signed languages.

6 The three languages are Domari [rmt] (Indo-Aryan), Tlahuitoltepec Mixe [mxp] (Oaxaca Mixe), and Panare [pbh] (Venezuelan Cariban). The Indo-European languages in the original sample included mostly European languages: the addition of an Indo-Aryan language such as Domari, spoken across the Middle East, was meant to account for more diversity within the family by including a spoken, non-territorial variety in contact with other non-Indo-European languages (Matras 2012: 1–4). The reason for including Tlahuitoltepec Mixe and Panare is that Mixe-Zoque and Cariban languages are known from Wichmann (1995) and Gildea (1998), respectively, to display various cases of coincidence between argument and possessor indexes, with significant differences among different languages of the family that suggest different diachronic trajectories from Proto-Mixe-Zoque and Proto-Cariban to the daughter languages.

7 The dataset is publicly available on Zenodo: https://zenodo.org/records/14187805.

Table 1: The areal distribution of cases of coincidence between possessor and argument indexes.

Area	N. of languages (and %) with cases of coincidence	N. of languages (and %) without cases of coincidence
Africa	8 (20 %)	32 (80 %)
Australia & New Guinea	11 (28.21 %)	28 (71.79 %)
Eurasia	8 (33.33 %)	16 (66.67 %)
North America	19 (50 %)	19 (50 %)
South America	18 (46.15 %)	21 (53.85 %)
Southeast Asia & Oceania	23 (51.11 %)	22 (48.89 %)
Total	87 (38.67 %)	138 (61.33 %)

The phenomenon is unevenly spread: while it is relatively rare in Eurasia, Africa, and Australia/New Guinea, it is quite widespread across the Americas and in Southeast Asia & Oceania. The areal distribution is statistically significant ($\chi^2 = 13.883$, d.f. 5, $p = 0.016$).

In the next sections, we will discuss all the parameters considered in the analysis of the phenomenon. These parameters include the type of verbal argument indexed by the index set that is also used to encode the possessor on nouns (A, S, and/or P), the type of coincidence (total or partial) between the two sets of indexes, the segmental similarity between argument/possessor indexes and independent pronouns (if any), the presence (and nature) of additional morphology in verbs whose indexes coincide with those marking the possessor, the restrictions on coincidence patterns depending on voice and TAM, and the parameter of (in)alienability. Before discussing the individual parameters, however, it is necessary to briefly explain the rationale behind our selection of these analytical parameters.

3.1 Rationale behind the selection of parameters

The selection of the parameters of variation for this cross-linguistic study has been guided by insights from the existing literature on the phenomenon under analysis. For each parameter, we can formulate hypotheses and predictions regarding the range of variation attested across languages. As discussed below, some of these hypotheses are confirmed by the cross-linguistic survey, while others are not, either because the parameter in question turns out not to be very informative or because more data is needed. We have already made explicit in Section 1.2.2 why it is necessary to take into account the similarity between indexes and independent pronouns: without a clear evaluation of this similarity, there is no possibility of discriminating between genuine cases of coincidence between argument and possessor indexes and cases in which both likely derive from the grammaticalization of independent pronouns.

The parameter of partial versus total coincidence that will be discussed at length below enables us to determine whether specific persons in the index sets tend to favor or disfavor coincidence. As is well-known, differences between speech act participants and 3rd persons in terms of morphosyntactic behavior are "ineradicable" (Lyons 1977: 638) and ubiquitous in grammatical systems (Siewierska 2004: 5.8). We may hypothesize that 1st and 2nd person arguments are more likely than 3rd persons to be encoded the way 1st and 2nd person possessors are, because of the "pragmatic delicacy and dangerousness" (Heath 1991: 78) of making overt reference to 1st and 2nd person agents or patients, which gives rise to various avoidance strategies (see DeLancey 2018; Konnerth and Sansò 2021: 15). Specifically, 1st and 2nd person verbal forms are more subject to diachronic renewal through the recruitment of nominalized verb forms as main clause predicates. Nominalizations, in Brown and Levinson's terms, are a well-documented negative politeness strategy (Brown and Levinson 1987: 207–209), i.e. a strategy through which agents and patients are referred only indirectly (or more indirectly than in non-nominalized clauses) as "attributes" of the action (Brown and Levinson 1987: 208).

The reason for including the parameter of additional morphology has to do with the individuation of possible diachronic connections between the coinciding index sets. As will be discussed in Section 3.5, there are languages in which some verbal constructions, besides displaying argument indexes that coincide with possessor indexes, are characterized by the presence of other formatives that are indicative of their nominal source, e.g. formatives that are similar to nominalizing morphemes or relational markers that are used in adnominal possession constructions.

The discussion in Section 1 and the literature survey in 1.2 have already emphasized that the coinciding sets may not be applicable to all verbal constructions, being limited to some TAM or voice constructions. We know from many typological studies that nominalized VPs can give rise to voice constructions (e.g. passive constructions, cf. Bugaeva 2011; Givón 1988; van den Berg and Boerger 2011; Sansò 2016; inverse constructions, Crevels 2011) as well as to various types of TAM constructions (DeLancey 2011; Grunow-Hårsta 2011; Palmer 2011; Rhee 2011; Sansò 2020; Yeh 2011: 419). Accordingly, we expect to find correlations between cases of coincidence of argument/possessor indexes and specific voice or TAM constructions.

Finally, the parameter of (in)alienability has been included because it has been often claimed that alienability splits may affect the coding of arguments in action nominal constructions (Koptjevskaja-Tamm 1993: 210), one of the possible diachronic sources of coincidence between argument and possessor indexes (see Section 4). In particular, Van Lier and Van Rijn's (2018) pilot study on Oceanic languages shows that agentive arguments in action nominal constructions are more likely to be coded by alienable possessor markers, while patientive arguments tend to be coded by inalienable possessor markers.

In the following sections, we will examine each parameter in detail, and we will discuss to what extent the variation attested in our dataset confirms the hypotheses and predictions introduced in this Section.

3.2 Parameter 1: Indexed argument

The first parameter is the argument that is indexed through the same indexes used to mark possessors on nouns. All three arguments A, S and P, as well as all their possible combinations, are attested in the sample, although with different frequencies.

Table 2 shows the indexed arguments in the languages of the sample in which there is coincidence between person indexes on verbs and nouns. In Sections 3.2.1–3.2.6 we will discuss the various patterns in detail.

There are significant differences in the areal distribution of the argument(s) indexed. Fisher's exact test (Monte Carlo simulation with 1,000,000 replicates) was performed on Table 3, yielding a significant p-value of 0.0023 (two-sided test).[8] The analysis of residuals reveals that certain arguments and combination of arguments show statistically significant deviations from what would be expected under the null hypothesis in some macro-areas. These represent key areas of interest for further exploration. In particular, coincidence phenomena involving A + S indexes are much more frequent than coincidence phenomena regarding other arguments or combination of arguments in Eurasia. Coincidence limited to A indexes, on the other hand, is more frequent

Table 2: Distribution of patterns of coincidence between possessor and argument indexes by indexed arguments (S_P = subjects of inactive intransitive verbs; S_A = subjects of active intransitive verbs).

Indexed argument	N. of cases (and %)
A only (including A + S_A)	21 (20 %)
S only (including S_P)	10 (9.52 %)
P only (including S_P + P)	27 (25.71 %)
A + S	36 (34.29 %)
S + P	1 (0.95 %)
A + S + P	8 (7.62 %)
A + P	2 (1.9 %)
Total	105 (100 %)

8 The two cases of coincidence between possessor indexes and A + P indexes have been removed from Table 3 since they are limited to Austronesian languages and have a language-specific diachronic explanation, which makes them genuine outliers.

Table 3: Distribution of patterns of coincidence between possessor and argument indexes by indexed arguments and macro-area.

	A only	P only	S only	A + S	S + P	A + S + P
Australia & New Guinea	2	5	3	2	0	0
North America	2	7	4	11	0	2
South America	3	7	2	5	1	4
Southeast Asia and Oceania	13	2	1	5	0	1
Africa	0	4	0	5	0	0
Eurasia	1	2	0	8	0	1

in Southeast Asia and Oceania, but this is possibly a distortion due to the high number of Austronesian languages in the sample, where coincidence between A and possessor indexes is an inherited feature (see Section 4.1.1). Coincidence limited to S indexes is higher than expected in Australia/New Guinea in comparison with other areas of the world, whereas languages in which possessor indexes are used for all three arguments (A + S + P) are more frequent in South America than elsewhere in the world.

3.2.1 A only (including A + S_A)

Coincidence between possessor and argument indexes may be limited to A arguments. In Highland Popoluca, so-called set A proclitics index the A on the verb and the possessor on the noun, as shown in (13), where the proclitic *ʔan=* is used to refer to the 1st person subject of the transitive verb 'to stir' and to the 1st person possessor of the noun 'dough':

(13) Highland Popoluca [poi] (Gulf Zoquean; de Jong Boudreault 2009: 398)

ʔan *=jeʔy-pa =ʔam*	*jeʔm*	***ʔan*** *=jaʔp-i*
1.ERG+stir-INCOMP+ALR	that	1.POSSR+grind-NMLZ

'I stir my dough.'

In Katukína-Kanamarí [knm] (Katukinan; dos Anjos 2011: 178), there is a set of person prefixes indexing the possessor of inalienable nouns and the A in bivalent verbs. Subjects of monovalent verbs are not encoded by person prefixes (dos Anjos 2011: 151). In various Austronesian languages, the A (and the A only) is indexed by means of possessor indexes in what is variously labelled undergoer, non-actor, or patient voice.[9] Tondano exemplifies this pattern, which is frequent throughout the family and will be discussed in more detail in Section 4.1.1:

9 The argument marked by means of possessor indexes is labelled variously in grammatical descriptions of Austronesian languages depending on the kind of analysis (e.g. ergative, actor, agent). For the sake of uniformity, we call A(gent) the indexed argument in these cases, since it is the most agentive argument of a bivalent predicate.

(14) Tondano [tdn] (Minahasan; Brickell 2014: 337, 341)
a. *sè=tu'a=**ku***
AN.PL-old-1SG.POSSR
'My parents'
b. *padahal pa-èdo-en=**ku** N=pèra'*
although DYN-take-PV=1SG.A INAN=roe
'Although I take the fish eggs.'

In some languages such as Tariana [tae] (Arawakan; Aikhenvald 2003: 122), the same set of person indexes is used to cross-reference the possessor and a set of arguments including the A and active intransitive subjects (S_A). In particular, the class of intransitive verbs whose only argument is marked by possessor indexes includes those transitive verbs that can be used without an overt patient argument (e.g. *hunt*), called ambitransitive in Aikhenvald (2003).

3.2.2 S only (including S_P)

There are languages in which only subjects of intransitive verbs are indexed with the same indexes used for possessors. In Tlahuitoltepec Mixe, subjects of intransitive verbs and possessors are marked the same way in the so-called dependent inflection. Despite its name, this type of inflection is not related to subordination, at least synchronically, and is used when a non-argument (e.g. a locative or an adverbial adjunct) occurs before the verb:

(15) Tlahuitoltepec Mixe [mxp] (Mixe; Romero-Mendez 2009: 178, 442)
a. *apenaxy ja'a n-tsukuj ojts **y**-ook-y*
hardly DEM.DIST 1.POSSR-aunt PST 3SG-die-DEP
'My aunt died with difficulties.'
b. *yë'ë jä'äy **y**-tixytyëjk*
DEM.MED person 3POSSR-woman
'That person's wife.'

The coincidence between intransitive dependent person indexes and possessor indexes in Mixe-Zoque languages can be traced back to Proto-Mixe-Zoque (cf. Wichmann 1995: 100, Table 3.5). In other Mixe-Zoque languages, e.g. in San Miguel Chimalapa Zoque [zoh] (not a sample language), dependent inflection is limited to subordinate clauses (Johnson 2000: 201). Possessor-like marking of arguments in subordinate clauses is commonly attested and is related to the use of nominalized forms of verbs as predicates in these clauses. As will be discussed in Section 4, in Tlahuitoltepec Mixe main clauses with adverbial or other non-argument adjuncts likely result from the reinterpretation of complex clauses (including a main and a subordinate clause) as simple clauses: thus, the

sentence in (15a) may have originated from the reinterpretation of a structure like *[it is with difficulties]$_{MAIN}$ [my aunt **her** dying]$_{SUBORDINATE}$*.

In other languages, only a subset of S arguments is indexed by means of possessor indexes. Lowland Oaxaca Chontal is one such example. This language has a set of indexes that marks the possessor on nouns and the S argument in a set of intransitive non-verbal predicates. In some of these predicates, the root is originally a noun combined with possessor markers. Structures such as *(there is) your illness* or *(there is) their money*, featuring a nominal predicate in a copula-less locational or existential sentence, have been reinterpreted as *you are ill* or *they have money*. That this represents a genuine case of reinterpretation rather than a result of translation bias is supported by the fact that, when used predicatively, as in (16a) and (16b), possessed nouns lack the determiner before the possessor index, which is required when they are used non-predicatively (e.g. *l-ay-ñepo'*, DET-1SG.POSSR-back, 'my back'; O' Connor 2004: 35). This special pattern of argument marking has spread to intransitive subjects in other non-verbal predicates whose nominal origin is synchronically opaque, as in (16c):

(16) Lowland Oaxaca Chontal [clo] (Tequistlatecan; O' Connor 2004: 44–45, 52)
  a. ***ay**-kwana*
     1SG.POSSR-illness
     'I have sickness / I am ill.'
  b. ***ajl**-pufki*
     1PL.POSSR-health
     'We have (good) health, we are healthy.'
  c. ***ajl**-tyuwa'*
     1PL.POSSR-alone
     'We'll stay all alone.'

In Maxakalí, only inactive (i.e. non-agentive) subjects of intransitive verbs are indexed in the same way as possessors on inalienable nouns:

(17) Maxakalí [mbl] (Maxakalian; Campos 2009: 69, 112, 78)
  a. ***'ũg**-pata*
     1SG.POSSR-foot
     'my foot'
  b. ***ũg**-mõyõn putup*
     1SG-sleep IRR
     'I want to sleep.'

Other languages of the sample in which coincidence between possessor and argument indexes is limited to a subset of S arguments include Yuchi, Lavukaleve and North Tukang Besi. In Yuchi [yuc] (isolate), the so-called *dzo*-series of person markers

is used to mark the only argument of intransitive stative verbs (S_P) and possessors in alienable possessor constructions (Linn 2000: 384). In Lavukaleve, possessor indexes mark the only argument of a small set of verbs of motion (*feu* 'go inland'; *hau* 'go shorewards'; *vau* 'go out', and *vo* 'come'; Terrill 2003: 368–372). This indexation strategy is described as a "feature of the casual conversational style characteristic of young people" (Terrill 2003: 371). When combined with possessor indexes, the verbs in question display a suffix *-i/-e* that has strong parallels with the nominalizer suffix. The only difference is that the suffix causes the loss of the final *-u* of the three verbs *feu*, *hau* and *vau*, whereas it does not cause such a change when it is used as a nominalizer, as shown in (18a) and (18b), respectively:

(18) Lavukaleve [lvk] (isolate; Terrill 2003: 369-370)

a. *a-e-se-ge* *urial* *vo-ma-re*
3SG.M.OBJ-SUB-be.full-ANT coconut.crabs 3PL.OBJ-take-NF
nga**-va-**i
1.SG.POSSR-go.seawards-NMLZ(?)
'It [the bag] is full, so taking the coconut crabs I go seawards.'

b. *o-rou* *lo* *foiga* *o-ma* *vau-i* *hi*
3SG.F.OBJ-smoke finish DEM 3SG.F.OBJ-take go.out-NMLZ 3SG.N.FOC
'Having smoked it, okay, you take it out [lit: it's a taken out one].'

Finally, in North Tukang Besi, property-denoting lexemes may index their only argument (S) either by means of person prefixes that are shared with other transitive and intransitive verbs or by means of possessive suffixes, as in (19a) and (19b), respectively. The latter possibility is allegedly stigmatized by speakers, although it is relatively frequent in casual conversation (Donohue 1999a: 156):

(19) North Tukang Besi [khc] (Celebic; Donohue 1999a: 156)

a. *'u-kalu-mo?*
2SG.R-tired-PFV
'Are you tired?'

b. *kalu-**'u**-mo?*
tired-2SG.POSSR-PFV
'Are you tired?'

3.2.3 P only (including S_p + P)

P is the only argument indexed by means of possessor markers in various languages. In Domari [rmt] (Indo-Aryan), P is indexed through a set of indexes that are identical to the set of possessive markers (Matras 2012: 252). In Burushaski [bsk] (isolate), there are three sets of lexically determined inalienable possessor markers that also index

the P in transitive verbs (Yoshioka 2012: 36, 110). Similarly, in Urarina, the same set of proclitics indexes the possessor and the P in transitive verbs, as in (20):

(20) Urarina [ura] (isolate; Olawsky 2006: 215)
  a. ***ka***=*hitɛana*
     1.POSSR=blowgun
     'my blowgun'
  b. ***i***=*tɛuerehe-~~ɯrɯ~~*
     2.POSSR=child-PL
     'your children'
  c. ***ka***=*kuruatahani-a* *kana=kwaaunera=ne* *eseneta-anu*
     1SG.OBJ=help-3SG.A 1PL.EXCL=creator=SUB believe-1SG.A
     'I believe that God (=our Creator) helps me.'
  d. ***i***=*tɛuara-ri-a-ũ=ni=ta*
     2SG.OBJ=see-IRR-NEG-1SG.A=ASS=FRS
     'I will not watch you.'

In To'abaita, only one class of transitive verbs (Class 2) takes P indexes that coincide with possessor indexes (Lichtenberk 2008: 117–122). This class mostly includes compound verbs of the form V + N and verbs with the suffix *-toqo*, whose meaning is to "test/check the referent of the direct object by performing the activity designated by the base verb" (Lichtenberk 2008: 117; cf. (21a)). The suffix *-toqo* has no etymology: as Lichtenberk (2008: 117) notes, "it is conceivable that [it] originated in a noun; however, there is no obvious candidate in the present-day language (or in closely related languages)." Regarding compound verbs, it is possible that the possessor-like index for the P results from the reinterpretation of a possessor suffix on the incorporated noun, which follows the verb in the compound (Lichtenberk 2008: 573–574). Thus, the object index *-ku* in the verb *lafu-lifo* 'extract-tooth' in (21b) may have originally been a possessor index combining with the incorporated noun *lifo*, and it is likely that "this kind of compounding originated as incorporation of the possessum noun […] together with the personal suffix on the possessum noun" (Lichtenberk 2008: 574):

(21) To'abaita [mlu] (Oceanic; Lichtenberk 2008: 573)
  a. *kwai* *lae-toqo-na* *fasi* *tala* *naqi*
     1SG.FUT go-TEST-3OBJ PREC path this
     'I'm going to try walking on this path.'
  b. *kera* *lafu-lifo-**ku***
     3PL.NFUT remove-tooth-1SG.OBJ
     'They extracted my tooth/teeth.'/'They performed tooth-extraction on me.'

In Hidatsa [hid] (Missouri River Siouan; Park 2012: 70–71), the so-called B-set marks possessors, the P in transitive verbs, and the subject of inactive intransitive verbs. In

contrast, subjects of active intransitive verbs are indexed by the so-called A-set, which does not display any coincidence with possessor indexes.

Coincidence between P indexes and possessor indexes may also involve recipient arguments in ditransitive constructions, which are often treated as P across languages. In Yagaria [ygr] (Siane-Yagaria; Haiman 1980: 199, 215), for instance, possessor indexes mark the animate P of transitive verbs and the recipient (inanimate Ps are not marked on the verb in Yagaria).

3.2.4 A + S

In several languages, the arguments indexed by the same indexes used for possessors are A and S, i.e. the most agentive argument of transitive clauses and the only argument of intransitive clauses. In Maco [wpc] (Maco-Piaroa), a specific verb class (whose membership exclusively depends on the phonological shape of the root) has A/S cross-referencing prefixes that coincide with the possessor indexes (Rosés Labrada 2015: 286–287). In Chamorro a bunch of irregular verbs takes possessor suffixes to index S and A: these verbs are *ga'o* 'prefer', *alok* 'say', and *ya* 'like, prefer' (cf. (21a); Topping 1973: 91–92). Moreover, the same indexes are also used with other verbs in interrogatives introduced by the words *hafa* 'what', and *hayi* 'who', as in (22b).

(22) Chamorro [cha] (Chamorro; Topping 1973: 92, 109)
   a. *ga'o-**ku*** *guihan*
      prefer-1SG fish
      'I prefer fish.'
   b. *hafa* *lini'e'-**ña***
      what see-3SG
      'What did he see?'

As argued by Topping (1973: 109), these verbs can be interpreted as precategorial forms that in various constructions (including the interrogative construction) function as nouns, which would justify the marking of their A/S as possessor: a more literal translation of the clauses in (22) would be 'my preference is fish' and 'what was his sight?'.

Itzá has two sets of person markers. The so-called Set A indexes A + S in the incompletive aspect and possessors, as exemplified in (23) with the Set A 1st person index *in(w)-*:

(23) Itzá [itz] (Yucatecan; Hofling 2000: 36, 45)
   a. *k-**inw**-il-ik-∅*
      INCOMP-1SG-see-INCOMP.TR-3OBJ
      'I see it.'

b. *k-**in**-wen-el*
INCOMP-1SG-sleep-INCOMP.INTR
'I sleep.'
c. ***in**-naj*
1SG.POSSR-house
'my house'

The marking of the incompletive status includes the prefix *k-* for both transitive and intransitive verbs, and two suffixes that differ depending on transitivity. The incompletive status suffix for intransitive verbs is *-Vl*, where *V* represents a vowel harmonic with the vowel of the verb stem. The same marker is used to derive verbal nouns from intransitive verbal roots, cf. *jan-al*, eat-NMLZ, 'eating, food', *kon-ol*, sell-NMLZ, 'selling, sale' (Hofling 2000: 105). Similarly, transitive verbal nouns are derived from transitive verbal roots by means of the suffix *-ik*, which is identical in form with the incompletive status suffix used for transitive verbs, cf. *il-ik*, see-NMLZ, 'seeing (it)' (Hofling 2000: 116). The nominal origin of the incompletive status suffix for both transitive and intransitive verbs explains why a set of person indexes also used with nouns is used to mark A or S on these verbs.

3.2.5 S + P (including S_P + P)

In one language of the sample, indexes used to refer to the possessor are also used to refer to S and P arguments on verbs. The language is Apinayé [apn] (Je Setentrional; Cunha de Oliveira 2005: 180), where the same person prefixes refer to the possessor on nouns and to S and P argument on monovalent and bivalent verbs, respectively.

3.2.6 A + S + P

In eight languages in the sample, the person indexes used to refer to the possessor may be used on the verb to refer to all three types of arguments – A, S, and P. In three of these languages, possessor and person indexes appear to be reduced versions of full personal pronouns. In Ute [ute] (Northern Uto-Aztecan), for instance, pronominal clitics, referring to the subject (in transitive and intransitive clauses), the object and the possessor, are "transparently reduced versions of the set of independent pronouns" (Givón 2011: 170). Similarly, person clitics in Cocama-Cocamilla are reduced forms of full pronouns used in colloquial and fast speech (Vallejos 2016: 211) and they can serve as indexes for A, S and P on verbs (in different positions in the sentence), and as possessor indexes on nouns. In Mosetén-Chimané [cas] (isolate), possessive clitics on nouns coincide with person clitics indexing A, S, and P on verbs, and both are clearly reduced forms of independent pronouns (Sakel 2004: 118–119, 122–123).

In other languages, the person indexes referring to A, S, P, and possessors may be diachronically unrelated to full personal pronouns. In Cholón, an extinct language known primarily through an extensive 18th-century grammar (Alexander-Bakkerus 2005: 21–22), the same indexes are used to mark possessors on nouns and A, S, and P on verbs. The order of prefixes on transitive verbs is always AP:

(24) Cholón [cht] (Híbito-Cholón; Alexander-Bakkerus 2005: 208, 209, 130)[10]
  a. ***m-a****-kole-ki*
     2SG.A-1SG.OBJ-love-IMP
     'Love me!'
  b. ***mi****-po-koly-aŋ*
     2SG.M.A-3PL.OBJ-love-INCOMP
     'You love them.'
  c. ***a****-l^{y}iš*
     1SG.POSSR-monkey
     'my monkey'
  d. ***mi****-kt^{y}ok*
     2SG.M.POSSR-box
     'your box'

In Northwest Sahaptin, the same set of person enclitics is used to index A, S, P, and possessors, as exemplified in (25). Even in this case, there is no segmental similarity with independent pronouns:[11]

(25) Northwest Sahaptin [yak] (Sahaptin; Jansen 2010: 79, 141)
  a. *aw=****nash*** *paysh* *wyáłamayk-sha*
     now=1SG.SBJ maybe while.going-lose-IPFV
     'Now I must have gotten lost.'
  b. *wíwnu=****nash*** *i-ni-ya*
     huckleberry-1SG.OBJ 3SG-give-PST
     'She gave me huckleberries.'
  c. *kwnak=****nash*** *wacha* *nishaykt*
     there=1SG COP.PST home
     'My home was there.'

10 3rd person indexes in Cholón are differentiated by argument type. The 3rd person singular index coincides with the corresponding possessor index only when it indexes S or P, whereas the 3rd person plural index coincides with the corresponding possessor index only when it refers to A or S. The other forms of the set are identical for A, S and P (Alexander-Bakkerus 2005: 130).

11 Recipients in ditransitive constructions are marked as objects in Yakama (Jansen 2010: 141).

3.2.7 A + P

In two Austronesian languages in the sample, possessor indexes are used to mark A and P to the exclusion of S. The coincidence, however, is found in different voice types. In West Coast Bajau [bdr] (Sama-Bajaw; Miller 2007: 143), possessor indexes are used to index agents in Undergoer Voice, and patients in Actor Voice, while independent pronouns are used to refer to intransitive subjects. A similar situation occurs in Jambi Malay [jax] (Malayo-Sumbawan; Yanti 2010: 490).

3.2.8 Correlations between arguments indexed and alignment type

As already discussed in Section 1.2, Siewierska (1998: 23–31) considers the possibility that there are significant correlations between the alignment type of a language and the argument(s) indexed by means of possessor indexes. Siewierska explicitly argues that "it could be the case that while affinities with both the A and the O *[P in our terminology]* occur in the cross-linguistically dominant alignment, i.e. accusative, only one or the other type of possessor affinities are displayed in languages with ergative or active alignment" (Siewierska 1998: 23). However, her data do not reveal a clear trend: while there are no evident differences in the affinities between possessor indexes and the A versus the P in languages with accusative or active alignment (Siewierska 1998: 24), in languages with ergative alignment possessor affinities with the P are exceptionally low (Siewierska 1998: 25). Our dataset, based on more restrictive criteria for positing coincidence between possessor and argument indexes, provides a different picture. The distribution of indexed arguments in connection to the alignment type of the language is displayed in Table 4.

Table 4: Distribution of indexed arguments in connection with alignment type.

Ergativity	S	A	A + S	P	S + P	A + S + P
Yes	6	12	0	6	1	2
No	4	9	36	21	0	6

Fisher's exact test performed on the data in Table 4 shows that the distribution is highly significant. The results showed a highly significant association between the variables (p = 1.44e-07), providing strong evidence against the null hypothesis that the argument indexed and the alignment type are not related and suggesting a

significant relationship between the variables. The analysis of Pearson's residuals shows that the most significant deviations from expected values are those regarding:

(a) cases in which the indexed argument is A (and A only), which are much more frequent in languages with ergative alignment;
(b) cases in which the indexed arguments are A + S, which in our dataset are only found in languages with nominative-accusative alignment.

Moreover, cases in which the indexed argument is P are not at all sporadic or exceptional in languages with ergative alignment, contrary to Siewierska's (1998) findings. The tendencies just described allow us to propose at least the following statistical universal implication: if in a language there is coincidence between possessor and A + S indexes (to the exclusion of P indexes), then that language shows nominative-accusative alignment.

3.3 Parameter 2: type of coincidence

The second parameter considered is the type of coincidence. This parameter accounts for the number of possessor/argument indexes within a given set that are identical: we distinguish between instances of total coincidence, in which all possessor/argument indexes in the set coincide, and instances of partial coincidence, which only involves some indexes of the set. We have found total coincidence in 69/105 cases (65.71 %) and partial coincidence in 36/105 cases (34.29 %).

In Sections 3.2.1–3.2.2 we will discuss the two types of coincidence and we will provide some methodological remarks about how the various situations found in the sample have been analyzed and classified.

3.3.1 Total coincidence

Total coincidence is found, for example, in Abui (cf. (26)), where all the person forms of the prefix set indexing the inalienable possessor and the P are identical:

(26) Abui [abz] (Timor-Alor-Pantar; Kratochvíl 2007: 77-78)

	Inalienable Possessor	Object
1SG	*na-*	*na-*
2SG	*a-*	*a-*
3REFL	*da-*	*da-*
3NON-REFL	*ha-*	*ha-*
DISTR	*ta-*	*ta-*
1PL.EXCL	*ni-*	*ni-*
1PL.INCL	*pi-*	*pi-*
2PL	*ri-*	*ri-*

In Lundayeh (Mortensen 2021: 61), like in other Austronesian languages, the whole genitive set of pronouns is used to mark both the possessor on nouns and the A in patient voice, as exemplified for the 2nd person in (27):

(27) Lundayeh (Mortensen 2021: 102; 113)

a. *bala=**muh***
word=2SG.GEN
'your words'

b. *bet-in=**muh*** *kayuh* *inih.*
tie-PV=2SG.GEN log this
'You are tying up this log.'

There are a few languages in the sample in which possessor and argument indexes coincide, but one of the two sets displays an additional segment. Consider the situation in Hiri Motu, exemplified in (28):

(28) Hiri Motu [hmo] (Motu-based pidgin; Dutton and Voorhoeve 1975: 72)

Person	Possessor indexes	Object indexes
1SG	*-egu*	*-gu*
2SG	*-emu*	*-mu*
3SG	*-ena*	*-a*
1PL.INCL	*-eda*	*-da*
1PL.EXCL	*-emai*	*-mai*
2PL	*-emui*	*-mui*
3PL	*-edia*	*-dia*

Possessor indexes in Hiri Motu contain an initial segment *-e* (*-en* in the 3rd person singular, possibly for euphonic reasons) but otherwise coincide with P indexes. Cases like this have been considered as instances of total coincidence, since the difference between the two sets is systematic and the additional segment can be hypothesized to represent a different (albeit opaque) formative.

3.3.2 Partial coincidence

As clarified above, by partial coincidence we mean that only some indexes of the set coincide. For example, in Karajá only first and second persons (with no formal distinction between singular and plural) of the pronominal prefixes used to mark the P on verbs and the possessor on nouns coincide:

(29) Karajá [kpj] (Karajá; Ribeiro 2012: 198)

	Possessive prefixes	P prefixes
1	*wa-*	*wa-*
2	*a-*	*a-*
3	*i-*	∅

In Huehuetla Tepehua, coincidence only involves the 1st person object/possessor prefix *kin-*, as shown in (30). The other persons of the paradigm are segmentally different.

(30) Huehuetla Tepehua [tee] (Totonacan; Kung 2007: 192)

*xa-**kin**-jun-ni-y*	*juu*	***kin**-pay*
PST-1OBJ-tell-DAT-IPFV	ART	1POSSR-father

'My father told me.'

The parameter of total versus partial coincidence encroaches on another sub-parameter, namely the presence of segmental similarities in non-coinciding forms. When a given index set is characterized by partial coincidence, some indexes within the set fully coincide, while others have segmental similarities but are not identical. In Kuot, for instance, 3SG, 1PL.EXCL, 2PL and 3PL forms coincide in both sets, while the other persons are segmentally quite similar:

(31) Kuot [kto] (isolate; Lindström 2002: 213)

Person	Possessor indexes	Object indexes
1SG	*tuo-*	*to-*
2SG	*nuo-*	*no-*
3SG.M	***a-***	***a-***
3SG.F	***o-***	***o-***
1PL.INCL	*buo-*	*bu-*
1PL.EXCL	***pa-***	***pa-***
2PL	***mi-***	***mi-***
3PL	***ma-***	***ma-***

In other languages characterized by partial coincidence, the non-coinciding forms in the two sets are segmentally more divergent. A case in point is Kola:

(32) Kola [kvv] (Central Malayo-Polynesian; de Winne 2013: 31, 39)

Person	Possessor indexes	Object indexes
1SG	***-ng***	***-ng***
2SG	***-m***	***-m**/-ka*
3SG	*-∅*	*-ni*
1PL.INCL	***-sita***	***-sita***
1PL.EXCL	***-ma***	***-ma***
2PL	*-mi*	*-kem*
3PL	*-di*	*-da/-yi*

The distinction between total and partial coincidence only takes into account the number of persons of the set. Nevertheless, it is important to note that in some cases, possessor and argument indexes differ in their morphological status. We have

already discussed the case of Tigon Mbembe (see examples (5) and (6) above), where there is total coincidence between possessor and P indexes, but the former consists of suffixes, while the latter consists of independent words (whose behavior is more constrained than that of independent pronouns, see the discussion in Section 1.1). In North Tukang Besi, there is partial overlap between the possessive index set and the realis subject index set. However, in the former the indexes are suffixal, while in the latter they are prefixal, as shown in (33).

(33) North Tukang Besi (Donohue 1999a: 113)

	Possessive	Realis Subject
1SG	*-su*	*ku-*
2SG	***-'u***	***'u-*** */ nu-*
3SG	***-no***	***no-*** */ o-*
1PA	*-mami*	*ko-*
1PL	*-nto*	*to-*
2PL	*-miu*	*i-*
3PL	***-no***	***no-*** */ o-*

We decided to consider cases like the ones just discussed as instances of total or partial coincidence depending on what proportion of the person-indexing forms coincide, irrespective of the morphological status of the indexes, based on the segmental identity between (indexes belonging to) the two sets.

In Table 5, we have calculated which persons coincide when there is partial coincidence, limiting ourselves to singular and plural persons. A language in which partial coincidence applies solely to dual forms has been excluded, reducing the number of languages with partial coincidence from 36 to 35.[12] First person

Table 5: Persons involved in cases of partial coincidence (percentages are calculated on the total number of cases of partial coincidence in the sample, i.e. 35).

Person	N. of cases (and %)
1SG (including 1st person, number indifferent)	26/35 (74.28 %)
2SG (including 2nd person, number indifferent)	21/35 (60 %)
3SG (including 3rd person, number indifferent)	17/35 (48.57 %)
1PL (including 1PL.INCL, 1PL.EXCL)	24/35 (68.57 %)
2PL	20/35 (57.14 %)
3PL	13/35 (37.14 %)

12 The language in question is Mehri (Afro-Asiatic, gdq; see Rubin 2018: 54–60), where there is coincidence between the set of possessor indexes on singular nouns and a set of person indexes – specifically those used after the 3SG.M and 3PL.F forms of the perfect tense – in all three persons of the dual.

singular indexes, for instance, coincide in 26 cases out of 35 in which there is partial coincidence. The Pearson's Chi-squared test indicates a statistically significant relationship between specific persons and likelihood of coincidence (p = 0.02368): 1st and 2nd person indexes are more frequently involved in coincidence patterns than 3rd person ones.

The hypothesis discussed in 3.1 that 1st and 2nd person argument indexes are more likely to undergo pragmatic renewal as the result of the recruitment of nominalized verb forms as main predicates, though bound to remain speculative in the absence of in-depth diachronic studies, is not contradicted by the figures in Table 5.

3.4 Parameter 3: presence/absence of additional morphology

As already noted in Sections 3.1.2 and 3.2.1 when discussing Lavukaleve and Hiri Motu, respectively, we found 17 cases out of 96 where the index sets exhibit some additional morphology beyond the formative encoding person. An example is found in Ngemba. Here, the set of possessive indexes is hosted by agreement prefixes including a vowel and a consonant whose form depends on the noun indicating the possessed entity. The same set indexes the P when hosted by *ɣ-/ɣo-*:

(34) Ngemba (Leroy 2007: 142; 207)

	Possessive indexes	Object indexes
1SG	|*VC-an*`|	|*ɣ-an*`|
2SG	|*VC-ɔ*`|	|*ɣ-ɔ*`|
3SG	|*VC-i*´|	|*ɣ-i*´|
1DU	|*VC-ɔɣá*`|	|*ɣo-ɔɣá*`|
1PL.EXCL	|*VC-uɣ*´|	|*ɣo-uɣ*´|
1PL.INCL	|*VC-uɣ́nà*|	|*ɣo-uɣ́nà*|
2PL	|*VC-ən*´|	|*ɣo-ən*´|
3PL	|*VC-a*´|	|*ɣo-a*´|

In Ngemba, the person formatives in the index set referring to the possessor and in the index set referring to the P are the same, exemplifying what appears to be an instance of total coincidence as defined in Section 3.2.1. However, in both cases, additional formatives are present that differentiate the two domains (nominal and verbal) in which the coinciding indexes are used.

The presence of additional morphology can manifest in various patterns. First, there may be different formatives in the two sets, like in the case of Ngemba just discussed. Second, an additional formative may occur exclusively in one of the two sets, like in the Hiri Motu case discussed in 3.2.1, where the possessor index set

contains the opaque formative *-e-/-en-*. Finally, the same additional morphology can be used in both sets. An instance of this pattern is found in Cebuano, where the set of indexes referring to the possessor and to the A in non-actor voice both employ the suffixal linker *-nga*, as shown in the following examples:

(35) Cebuano [ceb] (Greater Central Philippine; Tanangkingsing 2009: 38; 83)

a. *na-kawat=siya sa **iya-nga** pagka-pordoy*
AV-steal=3SG.NOM because 3SG.POSSR-LK NMLZ-poor
'He stole because of his poverty.'

b. ***iya-nga** gi-kuha? Ug **iya-nga** gi-kugos*
3SG.POSSR-LK PFV.PV-take and 3SG.POSSR-LK PFV.PV-carry
'He took and carried (the dog).'

From our perspective, the most intriguing cases are those in which the additional morphology used on the index set for a given argument/combination of arguments has other functions in a language that potentially indicate a nominal origin of the possessor-like indexing of arguments. For example, in Lavukaleve, as previously noted (cf. ex. (17)), in a small group of intransitive verbs the S is indexed using prefixes from the possessive prefix paradigm, while the verbs themselves receive a special suffix *-i/-e*. Terrill (2003: 369–370) acknowledges the formal resemblance of the suffix in question with the nominalizer suffix *-i/-e* but interprets the similarity as a case of homophony, arguing that "the Nominaliser suffix does not cause loss of a final *u* of a verb stem". Yet, the similarity is striking and finds parallels in other languages of the sample, suggesting diachronic connections between nominalized forms of the verb and possessor-like indexing of arguments.

Nominalization also appears to play a role also in Lopnor Uighur, where there is partial coincidence between possessive and argument indexes in the past tense, the habitual, the conditional mood and after the copula *idi* (Abdurehim 2014: 154). As discussed by Erdal (2014), the past tense formative to which the possessive suffixes attach in both Uyghur and Old Turkic contains an alveolar consonant that, according to some scholars, "might originally have been that of a verbal noun" (Erdal 2014: 238). Similarly, the linker *-nga* in Cebuano, discussed above, "connects two grammatical entities where one modifies the other" (2009: 144) and its head is mostly nominal. Once again, this supports the hypothesis of a nominal origin of the possessor-like indexing of the A in non-actor voice in Cebuano.

In other cases, the additional morphological material found in the possessor index set originates from a verbal (i.e. non-nominal) pattern of indexation. A notable example is Ainu, where there is total coincidence between the person indexes marking possessors (as in (36a)) and those marking the A (cf. (36b)). In the case of alienable possession, these indexes cliticize on the host *-kor*, forming what some scholars refer to as 'possessive pronouns' (Refsing 1986: 25):

(36) Ainu [ain] (isolate; Refsing 1986: 161, 92)

a. ***ku***=*kor* *menoko*
1SG=HOST woman
'my woman'

b. *kuani* *ka* *cep* *ka* ***ku***=*e*
I too fish too 1SG=eat
'I, too, eat fish as well.'

The host *-kor* is also a full verb meaning 'to have sth/sb'. These "possessive pronouns" can therefore be interpreted as arising from the reanalysis of a (gapped) relative clause expressing ownership in attributive position ('the woman I have' in (36a)). This analysis would account for the presence of verbal person indexes to encode the possessor.

The Ainu case demonstrates that when arguments and possessor indexes coincide, it is not necessarily because argument indexes derive from possessor indexes via the reanalysis of an originally nominalized form of the verb as a predicative element. However, in most cases within the sample, the additional morphology attested in the cases of coincidence between argument and possessor indexes suggests a diachronic pathway in which nominalized forms of the verb acquire a predicative function through time.

3.5 Parameter 4: Similarity with independent pronouns

Similarity with independent pronouns has been included among the parameters in that it is essential to discriminate between cases where independent pronouns independently develop into argument indexes and possessor indexes at the same time, and cases in which other diachronic scenarios are involved. Some methodological remarks, however, are necessary before discussing this parameter in detail.

Firstly, there are cases in which a language has more than one set of unrelated independent pronouns, none of which can be defined primary in terms of distribution or function. A case in point is Chamorro, where there are two sets of independent pronouns:

(37) Chamorro [cha] (Austronesian; Topping 1973: 106–108)

Person	Set I independent pronouns	Set II independent pronouns	Possessor/argument indexes
1SG	***hu***	*yo'*	***-hu***
2SG	*un*	*hao*	*-mu*
3SG	*ha*	*gue'*	*-ña*
1PL.INCL	***ta***	*hit*	***-ta***

1PL.EXCL	*in*	*ham*	*-mami*
2PL	*en*	*hamyo*	*-miyu*
3PL	*ma*	*siha*	*-ñiha*

In cases like this, we have compared the possessor/argument indexes with the set of independent pronouns that exhibit overall greater segmental similarity with them: this is set I in Chamorro, which has been accordingly classified as a case of partial similarity, limited to two persons in the paradigm.

Secondly, when in a language there are two sets of independent pronouns, one of which is used with emphatic/contrastive function, and the two sets have different phonetic and morphological shapes, the non-emphatic/contrastive set is generally considered, unless the latter does not meet the first criterion (greater segmental similarity) introduced above.

Thirdly, there may be cases where possessor/argument indexes display segmental similarity with the genitive forms of the independent pronouns, but not with their nominative/subject forms. In this case, the following scenario can be hypothesized: the genitive forms of independent pronouns have evolved into possessor indexes, which in turn came to function as person indexes in specific contexts. A case in point is Ket, where a class of verbs indicating sound emission takes possessive indexes to index the subject. These indexes show similarity with the genitive forms of independent pronouns but show little or no similarity with their nominative forms:

(38) Ket [ket] (Yeniseian; Georg 2007: 160, 117)

Person	Independent pronouns – Nominative	Independent pronouns – Genitive	Possessor/ argument indexes
1SG	*ād*	*ā**b***	***b-***
2SG	*ū(k)*	*ū**k***	***k-***
3SG.M	*bū*	*bú**da***	***da-***
3SG.F/N	*bū*	*bú**d**i*	***d-***
1PL	*átn*	*átn**na***	***na-***
2PL	*ákŋ*	*ákŋ**na***	***na-***
3PL	*būŋ*	*būŋ**na***	***na-***

In cases like this, we have considered the nominative/subject form as a term of comparison, and we have classified the language accordingly (thus, Ket has been classified as a case in which free forms are different from the coinciding index sets).

Finally, there are cases in the sample in which the independent pronouns appear to be secondary formations, containing a formative to which the bound person

indexes are attached. In Saaroa, for instance, independent pronouns are formed by means of suffixation of person indexes to a formative *ilha-*:

(39) Saaroa [sxr] (Tsouic; Pan 2012: 258)

Person	Independent pronouns	Possessor/argument indexes
1SG	*ilha-ku*	*-ku*
2SG	*ilha-u*	*-u*
3SG	*ilha-isa*	*-isa*
1PL.INCL	*ilha-ta*	*-ta*
1PL.EXCL	*ilha-lhamu*	*-lhamu*
2PL	*ilha-mu*	*-mu*
3PL	*ilha-isa/ilha-lhisa*	*-isa/-lhisa*

In Nisgha [ncg] (Tsimshian), independent pronouns are formed by suffixation of bound person indexes to a base *ṅí-*, which can be traced back to the topicalizer *ṅí*, 'that's' (Tarpent 1989: 333). Given the secondary nature of these independent pronouns from a morphological point of view, we have classified cases like Saaroa and Nisgha as a special category in Table 6 below.

As shown in Table 6, in a significant number of cases (40/105) there is no similarity with free forms, while cases of total identity between person/possessor indexes and free forms are rare (3 cases out of 105). Between these two extremes, several intermediate situations occur. These intermediate cases are more challenging to classify and form a continuum ranging from partial segmental similarity in all or most of the forms of the index set to full segmental similarity limited to a few forms. For the sake of simplification, we have reduced the variety of possible patterns to two intermediate cases, namely (i) cases where more than 50 % of coinciding indexes display identity or similarity with the corresponding independent pronouns (35 cases), and (ii) cases where 50 % or fewer of the coinciding indexes display identity/similarity with their corresponding independent pronouns (9 cases).

Table 6: Similarity/identity with independent pronouns.

Type of identity	n. of languages (and %)
(1) Total identity	3 (2.86 %)
(2) >50 % of the coinciding indexes are identical/similar to independent pronouns	35 (33.33 %)
(3) ≤50 % of the coinciding indexes are identical/similar to independent pronouns	9 (8.57 %)
(4) Independent pronouns are entirely different	40 (38.1 %)
(5) Independent pronouns are secondary formations	18 (17.14 %)
Total	105 (100 %)

Among the languages in the sample in which possessor/argument indexes show >50 % identity with independent pronouns, there are cases in which identity involves all persons except one. In Yimas, for instance, S indexes coincide with possessor indexes, and are segmentally identical to 1st and 2nd person independent pronouns, the only partial exception being the 2nd person singular, where the independent pronoun is *mi* and the bound index is *ma-*:

(40) Yimas [yee] (Lower Sepik; Foley 1991: 200)

Person	Independent pronoun	S/possessor indexes
1SG	*ama*	*ama-*
1PL	*ipa*	*ipa-*
1DU	*kapa*	*kapa-*
2SG	***mi***	***ma-***
2PL	*ipwa*	*ipwa-*
2DU	*kapwa*	*kapwa-*

In Ulwa [yla] (Ulmapo; Barlow 2018: 142–143), identity with independent pronouns involves all forms of the index set referring to the P and the possessor except the 3rd singular: *mï* is the independent pronoun, while *ma=* is the P/possessor enclitic. The same pattern is attested in Abui (Kratochvíl 2007: 77–78), where all the forms of the index set referring to the P and the possessor fully coincide with independent pronouns except the 3rd person singular form (*di* is the independent pronoun, *da-* is the P/possessor bound index). In Warembori [wsa] (South Halmahera-West New Guinea; Donohue 1999b: 28–30), the index set labelled set I, indexing possessors and A + S, displays full segmental identity with independent pronouns only in the plural persons, while the singular persons are possibly reduced forms of the corresponding independent pronouns. In Selk'nam, the whole set of bound indexes indexing the P and the possessor displays segmental similarity with the set of independent pronouns:

(41) Selk'nam [ona] (Insular Chonan; Rojas-Berscia 2014: 77)

Person	Independent pronoun	P/possessor indexes
1SG	*ya*	*y-*
2SG	*ma*	*m-*
3SG	*ta*	*t-*
1PL.INCL	*įkųwa*	*įkų-*

In 9 languages of the sample, there is ≤50 % similarity with independent pronouns. In Maybrat-Karon, for instance, 1st and 2nd singular and 2nd plural forms of indexes referring to A + S and possessors display a phonological segment that is also found in the corresponding independent pronouns, whereas the other forms of the set do not show any segmental similarity with independent pronouns:

(42) Maybrat-Karon [ayz] (isolate; Dol 2007: 62)

Person	Independent pronoun	A+S/Possessor index
1SG	***t***uo	***t-***
2SG	***n***uo	***n-***
3SG.M	ait	y-
3SG (unmarked)	au	m-
1PL	amu	p-
2PL	a***n***u	***n-***
3PL	ana	m-

Cases of free forms entirely or almost entirely different from the person/possessor indexes are, among others, Highland Popoluca and Halkomelem. The Set A proclitics (indexing A in transitive clauses and possessors) and the independent pronouns of Highland Popoluca are given in (43), while Halkomelem independent pronouns and person/possessor indexes are given in (44). In this latter case, there is only some segmental similarity between the 2PL independent pronoun and its corresponding index:

(43) Highland Popoluca (de Jong Boudreault 2009: 155, 162)

Person	Independent pronoun	Set A proclitic
1	*ʔɨch*	*ʔan=*
2	*mich*	*ʔin=*
3	*jeʔ*	*ʔi=*

(44) Halkomelem [hur] (Central Salish; Suttles 2004: 321)

Person	Independent pronouns	Person/possessor indexes
1SG	*ʾə́nθə ~ ʾə́ nθə*	*nə-*
2SG	*nə́w̓*	*ʾən- ~ ʾəT-*
3SG/PL	*ƛ̓á / ƛ̓áləm*	*-s*
1PL	*łníməł*	*-ct*
2PL	*łwə́ləp*	*-ələp*

In summary, the prevalence of cases where there is little or no similarity between the coinciding indexes and independent pronouns or where independent pronouns are secondary formations (67 cases out of 105, types (3)–(5) in Table 6) makes the parallel grammaticalization scenario of independent pronouns into bound indexes (for arguments and possessors) discussed in Section 1 unlikely, or at least questionable, for about two thirds of the cases in which argument and possessor indexes coincide. This leaves open the possibility that other diachronic scenarios may better explain the observed patterns of coincidence. Some of these scenarios will be explored in Section 4.

3.6 Parameter 5: TAM and voice

As already discussed, coincidence may be limited to some verbal constructions, defined in terms of TAM and voice. In the two following sections, we will describe the cases of limitation of coincidence between argument and possessor indexes to specific verbal constructions that we have found in the sample.

3.6.1 Coincidence between argument/possessor indexes dependent on TAM

There are 17 cases in the sample in which the coincidence between possessor and argument indexes depends on specific TAM constructions. In Lopnor Uyghur, as already discussed, coincidence between argument and possessor indexes is limited to the past tense (Abdurehim 2014: 154), a limitation that is attested already in Old Turkic (Erdal 2004: 237). Modality/reality status seems to be relevant only in a couple of cases: one of them is North Tukang Besi, where the coincidence involves the realis subject indexes (cf. (33)). As for aspect, we have already mentioned the case of Itzá (cf. example (23)), where coincidence between argument and possessor indexes is limited to the incompletive aspect (Hofling 2000: 35–36). Another case is Panare, where Set II prefixes index the possessor and the P in the non-pastperfective aspect, as exemplified in (45):

(45) Panare [pbh] (Cariban; Payne and Payne 2013: 213, 77, 105)

a. ***y-ánkë-ñe*** *tunko* *Pekoro* *në-tú-jpë* *inken*
3-take-NONSPEC.TR bird Pedro DERG-give-PART.PST.INFR child
'The bird carries what Pedro gave her kids.'

b. ***y-eña-n***
3POSSR-hand-POSS
'his/her hand'

c. *y-ïkïtï-**ñe*** *uya*
INF-cut-NMLZ.A DAT
'towards the picker/towards the one that is picking'

In non-pastperfective aspect, transitive verbs express their P argument with a set of prefixes (Set II) that is also used in NPs to encode the possessor, as shown in (45a) and (45b) for a third person possessor/P (marked by *y-*). Moreover, the verb is also marked by a set of suffixes encoding specific aspectual sub-types: (45a) exemplifies the so-called 'non-specific aspect' suffix, used, among other things, to encode ongoing situations in the present and situations in the immediate future (Payne and Payne 2013: 212). As will be discussed in Section 4, many aspectual suffixes in Panare are synchronically isomorphic with nominalizers (Payne and Payne 2013: 159), including the non-specific aspect suffix *-ñe*, which is also used as an agent nominalizer (cf. (45c)).

Limitations to specific TAM constructions may therefore depend on the semantics of the source nominalization, whose reinterpretation is favoured in specific TAM contexts. As demonstrated by the Panare case, for instance, the restriction to the non-pastperfective aspect may result from the reinterpretation of a copula-less equative sentence containing an agent nominalization (e.g. *the bird [is] the carrier etc.* in example (45a)) as a construction referring to ongoing, incomplete actions (see also the discussion in Section 4).

3.6.2 Coincidence between argument/possessor indexes dependent on voice

As already noted in Section 3.1.1, voice is a domain in which there is coincidence between argument ad possessor indexes in several Austronesian languages, where A is indexed by means of possessor indexes in what is variously labelled undergoer, non-actor, or patient voice. In our sample, 11 Austronesian languages exhibit this pattern, exemplified by Tondano in example (12) above. In two of these languages, the pattern also extends to actor voice, but in this case the possessor indexes are used to refer to the P. In West Coast Bajau, the so-called Set I comprises pronominal indexes that encode "the possessive function, and also the non-subject actor" in undergoer voice (Miller 2007: 87):

(46) West Coast Bajau (Miller 2007: 99, 140)

a. *emma'=**ku*** *ai* *nebong* *poon* *saging* *e*
father=1[SET I] PERF AV.cut.down tree banana DEM
'My father cut down the banana tree.'

b. *boi* *∅-boo=**ku*** *iyo* *pitu.*
COMPL UV-bring=1[SET I] 3[SET II] to.here
'I brought him/her here.'

c. *aku* *boi* *moo* *iyo* *pitu.*
1[SET II] COMPL AV-bring 3[SET II] to.here
'I brought him/her here.'

The same first-person pronominal index *=ku* is used in (46a) to mark possession and in (46b) to mark the A in undergoer-voice clauses. In actor-voice clauses, the A is encoded by a different set of indexes (called 'Set II'), as shown in (46c), and the Set I encodes inanimate pronominal Ps (while animate Ps are expressed by means of Set II, Miller 2007: 142–143), as shown in (47) where the Set I clitic *=ni* refers to the inanimate undergoer 'it (the food)'.

(47) West Coast Bajau (Miller 2007: 142)
∅-sele'-sele' *gai* *nanduk=**ni*** *engko'* *sawan.*
UV-replace-RED 3PL AV.ladle=3[SET I] PREP glass
'They took turns ladling it (the food) out with a glass.'

A similar situation is found in Jambi Malay, where there is partial coincidence for the first and third person between the indexes used to refer to the possessor and those expressing the P in the active voice and the A in the passive voice (Yanti 2010: 508; 509).

Austronesian voice is a case of symmetrical voice, in which neither of the alternating voice types is morphologically unmarked and both arguments in actor and undergoer voice remain direct core arguments. Coincidence between argument and possessor indexes depending on asymmetrical voice constructions is only sporadically observed in our sample. If we consider inversion phenomena as a type of voice,[13] perhaps the clearest example in our dataset is found in Trio. In this language, the indexes referring to the possessor and the P are used exclusively in inverse configurations, i.e. when the P outranks the A in the person hierarchy. The forms in (48a), for instance, are used when a 3rd person A acts on a 1st or 2nd person P. The same prefixes index possessors on nouns, as illustrated in (48b):

(48) Trio [tri] (Cariban; Meira 1999: 283, 202)
  a. ***j**-eta* 's/he/it has heard me'
     ***ë**-eta* 's/he/it has heard you'
     ***k**-eta* 's/he/it has heard us (1+2)'
  b. ***j**-ewa* 'my rope'
     ***ë**-ewa* 'your rope'
     ***k**-ewa* 'our (1+2) rope'

3.7 Parameter 6: alienability

Languages may have two sets of possessor indexes, one that is used with inalienable items such as kin terms, body parts, and other items that cannot be easily separated or given away, and one that is used with alienable items (food, house, etc.). When such a situation occurs, coincidence with argument indexes may involve only one of the sets. In most cases, it is the inalienable set that coincides with person indexes on verbs (27 languages out of 30 in which this parameter is relevant). In three languages, coincidence involves the alienable set of possessor indexes. The three languages in

13 Inverse constructions can be subdivided into two types: semantic and pragmatic inverses. The former are obligatory when the P outranks the A on a given hierarchy (e.g. animacy), as in Trio, whereas the latter depend on the relative topicality or discourse-prominence of the arguments (Givón 1994: 22). According to Gildea (1994: 187), semantic inverses do not qualify as genuine voice constructions, because voice constructions are typically optional; instead, they are better understood as alignment systems. However, the two types may be diachronically connected: semantic inverses can emerge through a gradual obligatorification of pragmatic inverses.

question are Ainu, Yuchi, and Onondaga. As already discussed, in Ainu (cf. example (36)) person and possessor indexes coincide in the alienable possessive construction, but not in the inalienable possessive construction, as the result of the reinterpretation of a gapped relative clause. In Yuchi and Onondaga, both the inalienable and the alienable set are used to index different syntactic functions.

In Yuchi, there are three index sets: one, called the *di-* actor set and used to index A and S, is also used to index possessors in inalienable possession relations, as in (49a-b) and (49c), respectively. Another set, the so-called *dzo*-set, indexes the patient in transitive verbs and the possessor in alienable possession relations, as in (49d) and (49e). Finally, a third set, the so-called *dze*-set, is used to refer to the only argument of stative one-place predicates and to the possessor in an idiosyncratic set of possession relations, including items that are close enough to the possessor but can still be alienated (*strength*, *clan*, *name*, etc.), as demonstrated in examples (49f) and (49g), respectively:

(49) Yuchi (Linn 2000: 132, 143, 384, 151, 136, 389)

a. ***di****-wawa*
1SG-chew
'I chew.'

b. *se-****di'****-nẽ*
3SG.PAT-1SG.A-see
'I see her.'

c. ***di****-thoci*
1SG.POSSR-head
'my head'

d. *s'e-****dzo****-shi*
3SG.A-1SG.PAT-pity
'He pities me.'

e. ***dzo****-yõthl'i'e*
1SG.POSSR-knife
'my knife'

f. ***dze****-sh'o*
1SG-be.tired
'I'm tired.'

g. ***dze'****-yada*
1SG.POSSR-clan
'my tribe'

In Onondaga, the possessor in alienable possession relations is indexed by the set that also indexes the only argument of stative intransitive clauses (cf. (50a) and

(50b)), while the possessor in inalienable possession relations is indexed by the set that also indexes the only argument of active intransitive clauses (cf. (50c) and (50d)):

(50) Onondaga [ono] (Northern Iroquoian; Woodbury 2018: 286, 148, 287, 212)

a. *hohú·wa?*
ho-hu̯w-a?
3SG.M-canoe-NSF
'his canoe'

b. *hoya?da·yę́sdih*
ho-ya?d-a-yęsd-ih
3M.SG-body-JN-look.nice-STAT
'He is nice looking.'

c. *hgihi?ná?ge*
k-gihi?n-a?=ge
1SG.POSSR-knee-NSF=LOC
'on my knee'

d. *hya de?khé·tha?*
hya de?-**k**-he·t-ha?
NEG NEG-1SG-smoke-HAB
'I don't smoke.'

Thus, the dataset does not support the prediction, already discussed in Section 3.1 and based on Van Lier and Van Rijn's (2018) study on Oceanic languages, that agentive arguments (A, S_A) tend to correlate with alienable possessor markers, while patientive arguments (P, S_P) tend to be coded by inalienable possessor markers. Instead, in both Yuchi and Onondaga the pattern is precisely the opposite of what was expected.

3.8 Other dimensions of variation

Other dimensions of variation appear to influence the phenomenon of coincidence between argument and possessor indexes, albeit more sporadically. We have found one case in which coincidence between argument and possessor indexes depends on event semantics. In Pilagá, there is partial coincidence between the prefixes used to mark possessors and those marking the S:

(51) Pilagá (Vidal 2001: 83, 141)

	Possession	Subject
1SG	***ñi-ƀiaq***	***ñ-oƀid'***
	1.POSSR-forest	l-come
	'my forest'	'I arrive.'

2SG	***an-****ƀiaq*	***an-****oƀi'*
	2.POSSR-forest	2-come
	'your forest'	'You arrive.'
3SG	***n-****biaq*	***n-****oƀi'*
	3.POSSR-forest	3-come
	'his forest'	'He arrives.'

The indexes exemplified in (51) do not always index the S in intransitive verbs. In addition to the set of prefixes exemplified in (51) (so-called Set B), there is another set of S indexes, called set A: while some verbs can only take one of these two sets, most verbs can take both and use one or the other depending on the semantic properties of the event. In particular, "when set B forms are used, the subject is generally understood as a semantic Undergoer or as an Experiencer" (Vidal 2001: 177) rather than a semantic agent. For example, set B is used when the subject is not deliberately performing an action (cf. (52a)), while set A is used when the subject is performing an action purposefully or is more aware of the action denoted by the verb (cf. (52b)).

(52) Pilagá (Vidal 2001: 200)

a. ***ña****-nom-ta*
1[SET.B]-know-ASP.RSLT
'I know.'

b. ***sa****-noma*
1[SET.A]-know
'I have knowledge (i.e., I am capable of/I know what I am capable of).'

There is also a small group of intransitive verb roots that can only take set B prefixes to mark the S. These verbs are "intransitive predicates where the subject is an Undergoer" and "there is no sense of 'control' on the part of the grammatical subject" (Vidal 2001: 184):

(53) Pilagá (Vidal 2001: 184)

a. ***ñi****-tonaʃak*
1[SET.B]-be_happy
'I am happy.'

b. ***ñi****-doč-i-ñi*
1[SET.B]-be_sincere-EP-ASP
'I am sincere.'

Another dimension of variation, though sporadically attested, concerns the distribution of the phenomenon of coincidence across the verbal lexicon. While most cases of coincidence involve all verbs in a language, there are cases where

coincidence is limited to a specific verbal class or to a small number of verbs. Some of these restrictions have been already briefly mentioned: in Lowland Oaxaca Chontal (cf. ex. (16)), coincidence between possessor indexes and S indexes only occurs in a few intransitive non-verbal predicates; similarly, in North Tukang Besi possessor-like marking of the S is limited to some property-denoting lexemes (cf. (19)); coincidence in To'abaita only involves a specific verb class that includes V + N compound verbs; in Maco, coincidence is only found in one verb class whose membership is determined by the shape of the root, whereas in Chamorro (cf. ex. (22)) the phenomenon in question only targets a few irregular intransitive verbs.

Finally, although a sociolinguistic characterization of the phenomenon is beyond the scope of this article, it is worth noting that in some cases coincidence patterns are described as innovative (and sometimes stigmatized by the speakers): in Lavukaleve, possessor-like marking of arguments with a few intransitive motion verbs is a "feature of the casual conversational style characteristic of young people" (Terrill 2003: 371), and a similar trend is observed in North Tukang Besi (cf. examples (18) and (19), respectively).

4 Diachronic scenarios

As illustrated by some of the cases discussed above, it is sometimes possible to reconstruct the diachronic trajectories that have led to the patterns of coincidence between argument and possessor indexes. The aim of this section is to review the available evidence concerning the pathways of development of possessor-like indexation of arguments, with a focus on determining whether these pathways are cross-linguistically recurrent and share common characteristics, among which the (combination of) arguments targeted and the diachronic steps of the reinterpretation process. In the absence of historical data for many of the languages discussed in this article, the diachronic scenarios discussed below can only be considered robust if they are both recurrent and similar across languages.

Some of these diachronic trajectories originate from a nominalized form of the verb, while others do not. In the following sections, we will discuss these trajectories in detail, aiming to identify (i) the scenarios in which a nominalized form of the verb might have been reinterpreted as a finite form of the verb that retains the possessive indexing of arguments, and (ii) the scenarios in which the possessive indexing of arguments does not derive from a nominalized verb form.

4.1 Scenario 1: grammaticalization from independent pronouns

As previously discussed, one of the most common sources of person indexes on verbs and nouns is independent pronouns (Givón 1976). Some cases of coincidence between possessor and argument indexes result from a parallel grammaticalization process where independent pronouns develop into argument indexes on the one hand, and into possessor indexes on the other.

In Abui, the index set given above in (26) refers to possessors in inalienable possessive constructions and to the P when prefixed to verbs. This set is clearly derived from the paradigm of independent pronouns, as discussed in Section 3.4. The coincidence between possessor and P indexes can be reconstructed to Proto-Alor-Pantar (Klamer 2017: 16), and the bound forms clearly derive from the reconstructed free forms, presented in (54):

(54) Proto-Alor-Pantar reconstructed free forms (from Klamer and Kratochvil 2018: 74)

1SG	**naN*
2SG	**aN*
3	**gaN*
1PL.INCL	**piN*
1PL.EXCL	**niN*
2PL	**iN*

Patient prefixation in the daughter Alor-Pantar languages is not obligatory with all patients. One of the factors determining the use of a patient prefix instead of a free form when a pronominal patient is involved is animacy (e.g. in Teiwa, cf. Klamer 2017: 17–18). The convergent behavior of patient prefixes in the daughter languages allows to reconstruct a Proto-Alor-Pantar embryonic system of DOM, where some divalent verbs (i.e. those taking a human/animate P) take patient prefixes and others express P with a free form (Klamer and Kratochvil 2018: 77). Given this reconstructed picture, it is possible to hypothesize the following scenario to explain the coincidence between patient and possessor indexes in Alor-Pantar:

(a) the Proto-Alor-Pantar free forms gave rise to possessor prefixes on nouns;
(b) independently from (a), the same free forms evolved into bound patient indexes when the patient was discursively prominent (e.g. human/animate patients).

We can hypothesize that the grammaticalization-from-independent-pronouns scenario may be at work in those cases in which there is clear synchronic

similarity between the coinciding indexes and independent pronouns (38 cases out of 105 in our dataset, see the discussion in Section 3.4 above). In particular, in three cases where the coinciding indexes refer to all three arguments (A, S, and P) the bound indexes appear to be the result of a (rather recent) process of cliticization of independent pronouns, following the diachronic path discussed by Givón (1976), cf. Section 3.2.6. However, even when clear correspondences in form exist between the two coinciding sets and independent pronouns, the grammaticalization-from-independent-pronouns scenario should be approached with caution, as alternative scenarios always need to be considered. In the two Northern Gyalrong languages, Japhug (no ISO code available; Glottocode japh1234) and Tshobdun (no ISO code available; Glottocode tsho1240), for example, there is a set of person suffixes that almost fully coincides with the set of possessor prefixes. The only differences are the absence of 2nd and 3rd person singular forms and a slight variation in the 1st plural suffix in the former set (cf. Jacques 2021: 612). Furthermore, both sets exhibit similarities with independent pronouns. Internal reconstruction and comparison with other Gyalrong languages suggest that the two sets originally only partially coincided. Their near-complete segmental coincidence today is the result of paradigm levelling between the two sets, driven by their shared function of referring to the same persons in different syntactic domains (Jacques 2021: 614–615).

In the majority of the cases of coincidence attested in our sample, lack of similarity with independent pronouns suggests different diachronic scenarios. It is to the discussion of these scenarios that we now turn.

4.2 Nominalization scenarios

The diachronic processes most commonly invoked to explain the coincidence between possessor and argument indexes involve a nominalized form of the verb or a nominalized clause as the source and a predicative structure that retains nominal-like marking of arguments as the target. But in what contexts is a nominalized verb or clause reinterpreted as a predicative form? In this section, we examine the mechanisms driving these diachronic processes. To do so, we distinguish different scenarios based on the verbal argument(s) that are indexed by possessor indexes and on the TAM restrictions on the target structures. We will demonstrate that some cross-linguistic diachronic regularities emerge and that some diachronic pathways recur across languages and language families.

4.2.1 Scenario 2: reinterpretation of patient-oriented nominalizations in copular constructions

Starosta et al. (2009 [1982]) proposed a scenario involving a nominalized form of the verb to explain the possessor-like marking of A in non-actor voice in Austronesian languages. Simplifying somewhat, and leaving aside the diachronic details, their arguments can be summarized as follows:

(a) agents in non-actor voice constructions across Austronesian subgroups are marked similarly to possessors of underived nouns;
(b) reflexes of Proto-Austronesian non-actor voice morphemes largely consist of nominal derivational affixes, which retain their nominal function in various daughter languages;
(c) the reinterpretation of originally nominal elements as verbal ones likely occurred in equative clauses.

In this reconstruction, non-actor voice constructions derive from object or place nominalizations with meanings such as "the one beaten *by the enemies*", "the thing to be burned up *by the vandals*", or "*my* sweeping place": all these nominalizations share a possessor-like marking of A (indicated in italics), which is preserved in the resulting voice construction. The onset of the process is to be found in equative constructions such as "I [am] the one beaten by the enemies" or "this [is] my sweeping place". At some stage in Proto-Austronesian such constructions represented alternative ways of describing an event through "recentralization" (Starosta et al. 2009: 395), to be intended as a way of placing "a new situational role in the perceptual center of the stage" (*ibidem*).

Starosta's et al. (2009 [1982]) hypothesis has sparked significant debate and has faced various criticisms in the Austronesianist literature. Critics argue that there is insufficient historical or comparative evidence to support the claim that voice markers originated as nominalizers, and some scholars suggest that the voice system could have evolved through alternative grammaticalization processes that do not involve nominalization (see, among many others, Ross 2002). While a full review of this debate is far beyond the scope of this article, the survey presented in this article highlights the existence of similar diachronic pathways in different language families, which seems to corroborate, at least typologically, the plausibility of Starosta's et al. proposal. Take, for instance, the case of Katukina-Kanamari, a language in which A (and only A) is indexed by means of indexes that are also used for possessors, much like in various Austronesian languages. Consider (55a) and (55b):

(55) Katukina-Kanamari (dos Anjos 2011: 179)

a. ***ha****-wadi:k* *Hu*
3SG-name Hu
'Hu was her name.'

b. ***ha****-hak* *tabi*
3SG-arrow(v.) Penelope
'He arrowed the Penelope (a kind of bird).'

According to Queixalós (2010), the use of possessor indexes to index the A in Katukina-Kanamari is the result of the reinterpretation process schematized in (56), similar to the one hypothesized by Starosta et al. (2009) to account for the possessive indexing of A in non-actor voice in Austronesian:

(56) a. *grandfather (is) [**your** killed one]* > b. ***you**$_{poss}$ killed grandfather*

In (56a), the patient-oriented nominalization serves to obliterate/reduce the salience of the A. As a result of morphosyntactic crystallization, (56a) eventually evolves into the primary means to describe an event, as in (56b). In this innovative structure, the A argument retains its possessor-like marking. Queixalós' hypothesis account for various facts of Katukina morphosyntax: firstly, in some divalent clauses in Katukina the pronominal agent is not encoded by possessor indexes preceding the verb but instead by an independent pronoun appearing after the verb, as in (57):

(57) Katukina-Kanamari (Queixalós 2010: 275)

koya	*o*	***adu***
pap	drink	1SG

'I drink pap.'

The clause in (57) likely represents a remnant of the original divalent clause formation strategy before the emergence of the structure in (56b): divalent clauses with generic objects such as (57) did not transition to the ergative pattern represented by clauses in which A receives possessor-like marking, thereby giving rise to a split ergativity pattern (dos Anjos 2011: 267–268; Queixalós 2010: 254). This can be explained by the fact that generic objects are typically non-salient in discourse, and thus they do not require the lowering of the agent's salience.

To sum up, the two diachronic explanations proposed for Austronesian and Katukina exhibit similar motivations and dynamics: for primarily discourse-related reasons, a non-agent argument is encoded as the pivot of a sentence containing a patient-oriented nominalization where the agent is encoded as a possessor. Over time, this way of encoding the event becomes the unmarked strategy to do so, and A maintains what synchronically looks like a possessor-like marking.

4.2.2 Scenario 3: reinterpretation of agent or action nominalizations in copular constructions

A nominalized source has been invoked to explain the coincidence between S and P indexes and possessor indexes in Panare (see above, Section 3.5.1; Payne and Payne 2013: 158–159). Drawing on the synchronic isomorphism between certain non-pastperfective suffixes and nominalizers, Payne and Payne (2013: 159; see also Gildea 1998) argue that non-pastperfective aspect formatives represent earlier nominalizations (including agent and action nominalizations). Given the synchronic coincidence between the non-specific aspect marker *-ñe* and the agent nominalizer *-ñe*, we might venture the hypothesis that in example (45a), repeated below as (58), a former "The bird is a carrier of what Pedro etc." has been reinterpreted as "The bird carries what Pedro etc.", with the P still being marked as a possessor:

(58) Panare (Payne and Payne 2013: 213)

y*-ánkë-*ñe	*tunko*	*Pekoro*	*në-tú-jpë*	*inken*
3-take-NONSPEC.TR	bird	Pedro	DERG-give-PART.PST.INFR	Child

'The bird carries what Pedro gave her kids.'

Another non-pastperfective suffix isomorphic with an action nominalizer is *-n* (cf. (59a)), used just like *-ñe* as a marker of non-specific aspect but limited to intransitive verbs, as in (59b). In this case, the reinterpretation proceeds from an original "there is my going/my going is" to a resulting "I'm going":

(59) Panare (Payne and Payne 2013: 99, 214)

a. *yu-tach ima-**n***
INF-dance(v.)-NMLZ
'dancing, dance (n.)'

b. *wu-të-**n*** *yu* *oromaepun-to-yaka*
1SG-go-NONSPEC.INTR 1SG talk-NMLZ.PURP-toward
'I'm going where they are talking.'

A nominalization-based diachronic scenario, however, does not necessarily account for all cases in which the argument indexed as a possessor is the P, and it is not clear how robust and typologically generalizable the scenario proposed by Payne and Payne (2013) for Panare is. In Creek, for instance, P indexes are also used with inalienably possessed nouns:

(60) Creek [mus] (Muskogean, Martin 2011: 134, 169, 453)

a. ***ca**-cokhátpi* 'my lip'
***po**-hapó:* 'our camp'

b. ***ca**-na:fk-íck-a'*
1SG.PAT-hit.LGR-2SG.A-Q
'Are you hitting me?'

c. ***po****-na:fk-ís*
1PL.PAT-hit.LGR-IND
'He/she is hitting us.'

According to the reconstruction proposed by Martin (1993: 450), the pattern of coincidence attested in Creek (and in other Muskogean languages) has a verbal origin: inalienable nouns taking the person prefixes exemplified in (60) are a restricted group, whose core mostly includes former nominalizations. These nominalizations originally carried verbal prefixes corresponding to the P argument: for example, *ca-poca:si* 'my lord' can be traced back to a nominalization of the type 'the one who masters me'.

4.2.3 Scenario 4: insubordination of nominalized clauses/stand-alone nominalizations developing predicative functions

The fourth diachronic scenario includes processes by which nominalizations, whether embedded or not, evolve into predicative forms through various mechanisms. One such mechanism is insubordination, in which formally subordinate clauses acquire conventionalized uses as main clauses (Evans 2007: 367). Among other features, e.g. reduction in TAM inflection, one of the possible formal characteristics of subordinate clauses is the presence of a nominalized form of the verb. In some languages of the sample, the coincidence between possessor and argument indexes makes it possible to hypothesize that the possessor-like marking of arguments is the result of a process of insubordination from a source subordinate clause containing a nominalization. (This might explain the so-called dependent marking in Tlahuitoltepec Mixe, which, at least synchronically, is unrelated to subordination.), if we hypothesize that other Mixe-Zoque languages, in which dependent marking is limited to subordinate clauses (Johnson 2000: 201), reflect the original Proto-Mixe-Zoque distribution.

Insubordination typically involves nominalizations that are embedded as arguments of a matrix clause. However, beyond insubordination proper, non-embedded (or "stand-alone") nominalizations can function as complete and independent utterances, with staging effects in narratives and/or stance functions, among others (see Yap et al. 2011: 26–27). As discussed in Section 1.2.1, Palmer (2011) hypothesizes that the nominal marking of arguments in certain TAM constructions in Northwest Solomonic languages is a persistent feature of the source construction, a nominalized clause used with recapitulative function and providing background information. Other languages in which this scenario could be tentatively invoked to explain the coincidence between person and possessor indexes are Itzá and Halkomelem. In Itzá, coincidence is limited to the incompletive aspect in both transitive and intransitive clauses. The markers of

incompletive aspect are similar in shape to nominalizing morphemes, as discussed in 3.2.4. Incompletive status in Itzá is predominantly used in discourse in clauses that provide background information to what follows (Hofling 2000: 357–360). The set of possessive indexes in Halkomelem, given in (61), can be used to refer to A + S in nominalized clauses, formed by prefixing *s-* NMLZ to the first word of a predicate.

(61) Halkomelem [hur] (Central Salish; Suttles 2004: 63)

1SG	*nə-*
2SG	*ʔən- ~ ʔəT-*
3SG/PL	*-s*
1PL	*-ct*
2PL	*-ələp*

Nominalized clauses can appear as arguments of complex clauses, as is usual for this kind of clauses, but they can also appear in independent clauses after adverbial elements such as *yéł* 'just now' or *wənáy* 'only', as in (62a), or standing alone as complete sentences in narratives, signaling a sudden event as in (62b):

(62) Halkomelem (Suttles 2004: 99, 105)

a. *yéł* *ṁə* ***nə**-s-técəl*
just.now CERT 1SG.POSSR-NMLZ-arrive.here
'I just got here.'

b. ***ʔəT**-s-wə-łx̌íləx*
2SG.POSSR-NMLZ-EST-stand
'(Then) you stand up.'

4.3 Scenarios not involving nominalizations

In other cases, a source of the possessor-like marking of verbal arguments other than nominalized forms of the verb can be hypothesized. In this subsection, we deal with these cases.

4.3.1 Scenario 5: reinterpretation of existential sentences with possessed nouns as stative predicates

As discussed above, in Lowland Oaxaca Chontal the possessor-like marking of arguments is limited to intransitive subjects of stative predications such as 'I'm ill' or 'I have money'. As argued by O'Connor (2004), the origin of this pattern is to be found in copular structures with possessed nominals such as 'my illness (is)' (=I'm ill) or 'my money (is)' (= I have money). The extension of this argument-marking pattern to

cases in which a nominal is no longer synchronically available (as in example (16c) above) suggests that the possessive marking of verbal arguments may have originated from copular structures with an existential or locative function. A similar pathway, from a possessor index in a possessed noun to a verbal index in existential sentences, can be hypothesized in the case of Chamorro (cf. (22) above), where structures like "my preference is fish" and "what was his sight?" were reinterpreted as "I_{POSSR} prefer fish" and "what did he_{POSSR} see?".

4.3.2 Scenario 6: reinterpretation of possessor marking of incorporated nouns

In To'abaita, verbs originating from the incorporation of a noun with the structure V + N have object suffixes that are identical to possessor indexes. As argued by Lichtenberk (2000: 376), this nominal-like marking of objects must be traced back to possessor indexing originally belonging to the incorporated noun following the process schematized in (63):

(63)

they	extract	teeth	my	>	they	extract-teeth	me_{POSS}	
	V	N	POSS			V+N	OBJ	
'They extracted my teeth.'					'They performed tooth-extraction on me.'			

The To'abaita case shows that possessor marking on incorporated nouns may represent the gateway through which possessive indexes extend to verbal arguments.

5 Conclusions

The survey conducted in this article highlights that the phenomenon of coincidence between possessor indexes and argument indexes is relatively widespread across the world's languages, as it appears in 87 out of 225 languages sampled. However, it is not as prevalent as suggested by the only typological survey currently available, that of Siewierska (1998), which is based on a much broader definition of "affinity" than our concept of coincidence.

The survey also reveals that the phenomenon is unevenly distributed across different regions of the world and that there are correlations between alignment type and the types of argument indexes that coincide with possessor indexes. The reasons for this correlation, possibly diachronic in nature, have yet to be investigated in detail.

In about one-third of the investigated cases of coincidence, the segmental similarity between the coinciding indexes and independent pronouns suggests that both coinciding indexes may derive, through independent grammaticalization

processes, from independent pronouns, as already hypothesized by Givón (1976). However, in the majority of identified cases, alternative diachronic scenarios can be proposed. In some of these scenarios, a nominalized form of the verb is reinterpreted as a predicative form. It is possible to identify recurring pathways across languages that account for certain regularities: for example, the use of possessor indexes to refer to the A seems to presuppose, in various languages, an initially innovative and subsequently conventionalized use of a nominalized form (e.g., "your killed one") in equative sentences ("I'm your killed one") to reduce the salience of the agent (cf. Section 4.2.1). Similarly, in at least some cases where the coincidence involves imperfective forms, the diachronic pathway may resemble that proposed by Palmer (2011) to explain these phenomena in the Northwest Solomonic languages: a nominalization used in discourse as a recapitulative strategy becomes autonomous, evolving into a standalone predicative form, retaining from its original usage the characteristic of being employed in imperfective sentences that depict background events in the narrative (cf. Section 4.2.3). Likewise, there are recurring pathways in which nominal forms with a possessor are reinterpreted as stative sentences, where the sole argument (S) retains the marking as a possessor (cf. Section 4.3.1). All in all, diachrony appears as an important factor to understand the cross-linguistic variation of the phenomenon, allowing us to propose some explanations of the coincidence patterns found across languages.

Finally, some predictions based on the literature, discussed in Section 3.1, were confirmed by the survey, while others were not. For example, the hypothesis that the indexes most frequently coinciding are those of speech act participants is supported by the data (recall the statistical significance of the distribution in Table 5, Section 3.3). On the contrary, the hypothesis that the indexes of alienable possessors tend to correlate more with indexes for agentive argument, while those of inalienable possessors tend to correlate more with indexes for non-agentive argument, is not corroborated by the survey data, which showed that in the vast majority of cases, the possessor indexes involved in the identified coincidence phenomena are those of inalienable possessors.

Abbreviations

1, 2, 3	1st, 2nd, 3rd person
A	agent
ADV	adverbial formative
ALR	already (aspectual marker)
AN	animate
ANT	anterior
ART	article

ASP	aspect
ASS	assertive
AUX	auxiliary
AV	actor voice
CERT	certainty particle
COMPL	completive
CONS	consecutive
COP	copula
DAT	dative
DEF	definite
DEM	demonstrative
DEP	dependent aspect/mood
DERG	de-ergative
DET	determiner
DIST	distal
DISTR	distributive
DM	discourse marker
DU	dual
DYN	dynamic
EP	epenthetic vowel
ERG	ergative
EXCL	exclusive
EST	'established' aspectual prefix
F	feminine
FOC	focus marker
FRS	frustrative
FUT	future
GEN	genitive
GENP	generic possession marker
HAB	habitual
HOST	host for person indexes
HUM	human
IMP	imperative
INAN	inanimate
INCL	inclusive
INCOMP	incompletive
IND	indicative
INDP	independent
INF	infinitive
INFR	inferential
INT	intentional
INTR	intransitive
IPFV	imperfective
IRR	irrealis
JN	stem joiner
LGR	lengthened grade
LK	linker

LOC	locative
M	masculine
MED	medial
N	neuter
NCM	noun class marker
NEG	negative/negation
NF	non-finite
NFUT	non-future
NHUM	non-human
NMLZ	nominalizer
NOM	nominative
NON-REFL	non-reflexive
NONSPEC	non-specific
NSF	noun suffix
OBJ	object
PART	participle
PAT	patient
PERF	perfect
PFV	perfective
PL	plural
POSS	possessed
POSSR	possessor
PREC	precedentive
PREP	preposition
PST	past
PURP	purpose
PV	patient voice
Q	question marker
R	realis
RED	reduplication
REL	relative
RSLT	resultative (aspect)
SBJ	subject
SG	singular
STAT	stative
SUB	subordinator
TEST	test
TR	transitive
UV	undergoer voice.

Research ethics: No relevant issues.
Informed consent: No relevant issues.
Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.
Use of Large Language Models, AI and Machine Learning Tools: None declared.

Conflict of interest: The authors state no conflict of interest.
Research funding: Università dell'Insubria – Assegni di ricerca Junior, bando 2021.
Data availability: The dataset is sent as a separate file to the journal's editor and is freely accessible at the following URL: https://zenodo.org/records/14187805.

References

Abdurehim, Esmael. 2014. *The Lopnor dialect of Uyghur: A descriptive analysis*. Helsinki: University of Helsinki Ph.D. dissertation.
Aikhenvald, Alexandra. 2003. *A grammar of Tariana, from northwest Amazonia*. Cambridge: Cambridge University Press.
Alexander-Bakkerus, Astrid. 2005. *Eighteenth-century Cholón*. Utrecht: Landelijke Onderzoekschool Taalwetenschap.
Applegate, Richard Brian. 1972. *Ineseño Chumash grammar*. Berkeley: University of California Berkeley Ph.D. dissertation.
Barlow, Russell. 2018. *A grammar of Ulwa*. Mānoa: University of Hawai'i at Mānoa Ph.D. Dissertation.
Brickell, Timothy C. 2014. *A grammatical description of the Tondano (Toundano) language*. Melbourne: LaTrobe University Ph.D. dissertation.
Brown, Penelope & Stephen C. Levinson. 1987. *Politeness. Some universals in language usage*, 2nd edn. Cambridge: Cambridge University Press.
Bugaeva, Anna. 2011. A diachronic study of the impersonal passive in Ainu. In Andrej Malchukov & Anna Siewierska (eds.), *Impersonal constructions. A cross-linguistic perspective*, 517–546. Amsterdam & Philadelphia: John Benjamins.
Campos, Carlo Sandro de Oliveira. 2009. *Morfofonêmica e morfossintaxe da Língua Maxakalí*. Belo Horizonte: Universidade Federal de Minas Gerais Ph.D. dissertation.
Childs, George Tucker. 2011. *A grammar of Mani*. Berlin & New York: Mouton de Gruyter.
Creissels, Denis. 2005. A typology of subject and object markers in African languages. In F. K. Ehrard Voeltz (ed.), *Studies in African linguistic typology*, 43–70. Amsterdam & Philadelphia: John Benjamins.
Crevels, Mily. 2011. Who did what to whom in Magdalena (Itonama). *International Journal of American Linguistics* 77(4). 577–594.
Cunha de Oliveira, Christiane. 2005. *The language of the Apinajé people of Central Brazil*. Eugene: University of Oregon Ph.D. dissertation.
Daguman, Josephine. 2004. *A grammar of Northern Subanen*. Melbourne: LaTrobe University Ph.D. dissertation.
de Jong Boudreault, Lynda J. 2009. *A grammar of Sierra Popoluca (Soteapanec, a Mixe-Zoquean language)*. Austin: University of Texas at Austin Ph.D. dissertation.
DeLancey, Scott. 2011. Finite structures from clausal nominalizations in Tibeto-Burman. In Foong-Ha Yap, Karen Grunow-Hårsta & Janick Wrona (eds.), *Nominalization in Asian languages. Diachronic and typological perspectives*, 343–360. Amsterdam & Philadelphia: John Benjamins.
DeLancey, Scott. 2018. Deictic and sociopragmatic effects in Tibeto-Burman SAP indexation. In Sonia Cristofaro & Fernando Zuñiga (eds.), *Typological hierarchies in synchrony and diachrony*, 345–375. Amsterdam & Philadelphia: John Benjamins.
de Winne, David. 2013. *A morphosyntax sketch of Kola, an Austronesian language of Aru, eastern Indonesia*. Leiden: Rijksuniversiteit te Leiden MA thesis.

Dol, Philomena. 2007. *A grammar of Maybrat: A language of the Bird's Head Peninsula, Papua Province, Indonesia*. Canberra: Pacific Linguistics.
Donohue, Mark. 1999a. *A grammar of Tukang Besi*. Berlin & New York: Mouton de Gruyter.
Donohue, Mark. 1999b. *Warembori*. Munich: Lincom Europa.
dos Anjos, Zoraide. 2011. *Fonologia e Gramática Katukina-Kanamari*. Amsterdam: Vrije Universiteit Amsterdam Ph.D. dissertation.
Dutton, Tom E. & Clemens L. Voorhoeve. 1975. *Beginning Hiri Motu*. Canberra: Australian National University.
Erdal, Marcel. 2014. *A grammar of Old Turkic*. Leiden: Brill.
Evans, Nicholas. 2007. Insubordination and its uses. In Irina Nikolaeva (ed.), *Finiteness. Theoretical and empirical foundations*, 366–431. Oxford: Oxford University Press.
Foley, William. 1991. *The Yimas language of New Guinea*. Stanford: Stanford University Press.
Georg, Stefan. 2007. *A descriptive grammar of Ket. Part I: Introduction, phonology and morphology*. Kent, UK: Global Oriental.
Gildea, Spike. 1994. Semantic and pragmatic inverse: 'Inverse alignment' and 'inverse voice' in Carib of Surinam. In Talmy Givón (ed.), *Voice and inversion*, 187–230. Amsterdam & Philadelphia: John Benjamins.
Gildea, Spike. 1998. *On reconstructing grammar: Comparative Cariban morphosyntax*. Oxford: Oxford University Press.
Givón, Talmy. 1976. Topic, pronoun, and grammatical agreement. In Charles N. Li (ed.), *Subject and topic*, 149–188. New York: Academic Press.
Givón, Talmy. 1988. Tale of two passives in Ute. In Masayoshi Shibatani (ed.), *Passive and voice*, 417–440. Amsterdam & Philadelphia: John Benjamins.
Givón, Talmy. 1994. The pragmatics of de-transitive voice: Functional and typological aspects of inversion. In Talmy Givón (ed.), *Voice and inversion*, 3–44. Amsterdam & Philadelphia: John Benjamins.
Givón, Talmy. 2011. *Ute reference grammar*. Amsterdam & Philadelphia: John Benjamins.
Goudswaard, Nelleke. 2005. *The Begak (Ida'an) language of Sabah*. Amsterdam: Vrije Universiteit Amsterdam Ph.D. dissertation.
Gravelle, Gloria. 2010. *A grammar of Moskona: An East Bird's Head language of West Papua, Indonesia*. Amsterdam: Vrije Universiteit Amsterdam Ph.D. dissertation.
Grunow-Hårsta, Karen. 2011. Innovation in nominalization in Magar. A Tibeto-Burman language of Nepal. In Foong Ha Yap, Karen Grunow-Hårsta & Janick Wrona (eds.), *Nominalization in Asian languages. Diachronic and typological perspectives*, 215–254. Amsterdam & Philadelphia: John Benjamins.
Haiman, John. 1980. *Hua: A Papuan language of the Eastern Highlands of New Guinea*. Amsterdam: John Benjamins.
Haspelmath, Martin. 2013. Argument indexing: A conceptual framework for the syntactic status of bound person forms. In Dik Bakker (ed.), *Language across boundaries*, 209–238. Berlin & New York: Mouton de Gruyter.
Heath, Jeffrey. 1991. Pragmatic disguise in pronominal-affix paradigms. In Frans Plank (ed.), *Paradigms: The economy of inflection*, 75–89. Berlin & New York: Mouton de Gruyter.
Hofling, Charles A. 1990. Possession and ergativity in Itzá Maya. *International Journal of American Linguistics* 56(4). 542–560.
Hofling, Charles A. 2000. *Itzaj Maya grammar*. Salt Lake City: University of Utah Press.
Hoijer, Harry. 1933. Tonkawa, an Indian language of Texas. In Franz Boas (ed.), *Handbook of American Indian languages*, Part 3, 1–148. Glückstadt-Hamburg-New York: J.J. Augustin & Columbia University Press.
Jacques, Guillaume. 2021. *A grammar of Japhug*. Berlin: Language Science Press.

Jansen, Joana Worth. 2010. *A grammar of Yakima Ichishkíin/Sahaptin*. Eugene: University of Oregon Ph.D. dissertation.
Johnson, Heidi Anna. 2000. *A grammar of San Miguel Chimalapa Zoque*. Austin: University of Texas at Austin Ph.D. dissertation.
Klamer, Marian. 2017. The Alor-Pantar languages: Linguistic context, history and typology. In Marian Klamer (ed.), *The Alor-Pantar languages: History and typology*, 2nd edn., 1–48. Berlin: Language Science Press.
Klamer, Marian & František Kratochvíl. 2018. The evolution of differential object marking in Alor-Pantar languages. In Ilja A. Seržant & Alena Witzlack-Makarevich (eds.), *Diachrony of differential argument marking*, 61–85. Berlin: Language Science Press.
Konnerth, Linda & Andrea Sansò. 2021. Towards a diachronic typology of individual person markers. *Folia Linguistica Historica* 42(1). 1–24.
Koptjevskaja-Tamm, Maria. 1993. *Nominalizations*. London: Routledge.
Kratochvíl, František. 2007. *A grammar of Abui. A Papuan language of Alor*. Leiden: Rijksuniversiteit te Leiden Ph.D. dissertation.
Kung, Susan Smythe. 2007. *A descriptive grammar of Huehuetla Tepehua*. Austin: University of Texas at Austin Ph.D. dissertation.
Laidig, Wyn D. 1993. Insights from Larike possessive constructions. *Oceanic Linguistics* 32(2). 311–351.
Leroy, Jacqueline. 2007. *Le mankon*. Paris: Peeters.
Lichtenberk, Frantisek. 2008. *A grammar of Toqabaqita*. Berlin & New York: Mouton de Gruyter.
Lindström, Eva. 2002. *Topics in the grammar of Kuot*. Stockholm: Stockholm University Ph.D. dissertation.
Linn, Mary Sarah. 2000. *A grammar of Euchee (Yuchi)*. Lawrence: University of Kansas Ph.D. dissertation.
Lyons, John. 1977. *Semantics*. Voll. I and II. Cambridge: Cambridge University Press.
Martin, Jack B. 1993. "Inalienable possession" in Creek (and its possible origin). *International Journal of American Linguistics* 59(4). 442–452.
Martin, Jack B. 2011. *A grammar of Creek (Muskogee)*. Lincoln: University of Nebraska Press.
Matras, Yaron. 2012. *A grammar of Domari*. Berlin: Mouton de Gruyter.
Mattiola, Simone. 2020. *Two language samples for maximizing linguistic variety*. Bologna: University of Bologna. Available at: http://amsacta.unibo.it/id/eprint/6504.
Meira, Sérgio. 1999. *A grammar of Tiriyó*. Doctoral dissertation, Houston, Rice University.
Miller, Mark T. 2007. *A grammar of West Coast Bajau*. Arlington: University of Texas at Arlington Ph.D. dissertation.
Mortensen, Christian J. 2021. *The Kemaloh Lun Bawang language of Borneo*. Mānoa: University of Hawai'i at Mānoa Ph.D. dissertation.
O'Connor, Loretta. 2004. *Motion, transfer, and transformation: The grammar of change in Lowland Chontal*. Santa Barbara: University of California at Santa Barbara Ph.D. dissertation.
Olawsky, Knut. 2006. *A grammar of Urarina*. Berlin & New York: Mouton de Gruyter.
Palmer, Bill. 2011. Subject-indexing and possessive morphology in Northwest Solomonic. *Linguistics* 49(4). 685–747.
Pan, Chia-jung. 2012. *A grammar of Lha'alua, an Austronesian language of Taiwan*. Townsville: James Cook University Ph.D. dissertation.
Park, Indrek. 2012. *A grammar of Hidatsa*. Bloomington: Indiana University Ph.D. dissertation.
Payne, Thomas E. & Doris L. Payne. 2013. *A typological grammar of Panare. A Cariban language of Venezuela*. Leiden: Brill.
Queixalós, Francesc. 2010. Grammatical relations in Katukina-Kanamari. In Spike Gildea & Francesc Queixalós (eds.), *Ergativity in Amazonia*, 235–283. Amsterdam & Philadelphia: John Benjamins.

Refsing, Kirsten. 1986. *The Ainu language. The morphology and syntax of the Shizunai dialect*. Aarhus: Aarhus University Press.
Rhee, Seongha. 2011. Nominalization and stance marking in Korean. In Foong-Ha Yap, Karen Grunow-Hårsta & Janick Wrona (eds.), *Nominalization in Asian languages. Diachronic and typological perspectives*, 393–422. Amsterdam & Philadelphia: John Benjamins.
Ribeiro, Eduardo Rivail. 2012. *A grammar of Karajá*. Chicago: University of Chicago Ph.D. dissertation.
Richter, Doris. 2014. *A grammar of Mbembe*. Leiden: Brill.
Rijkhoff, Jan, Dik Bakker, Kees Hengeveld & Peter Kahrel. 1993. A method of language sampling. *Studies in Language* 17(1). 169–203.
Rijkhoff, Jan & Dik Bakker. 1998. Language sampling. *Linguistic Typology* 2(3). 263–314.
Rojas-Berscia, Luis Miguel. 2014. *A heritage reference grammar of Selk'nam*. Nijmegen: Radboud Universiteit Nijmegen MA thesis.
Romero-Mendez, Rodrigo. 2009. *A reference grammar of Ayutla Mixe (Tukyo'm ayuujk)*. Buffalo: State University of New York at Buffalo Ph.D. dissertation.
Rosés Labrada, Jorge E. 2015. *The Mako language: Vitality, grammar and classification*. London, Canada: University of Western Ontario Ph.D. dissertation.
Ross, Malcolm D. 1982. The development of the verb phrase in the Oceanic languages of the Bougainville region. In Amran Halim, Lois Carrington & Stephen A. Wurm (eds.), *Papers from the Third International Conference on Austronesian Linguistics*, vol. 1: *Currents in Oceanic*, 1–57. Canberra: Pacific Linguistics.
Ross, Malcolm. 2002. The history and transitivity of western Austronesian voice and voice-marking. In Fay Wouk & Malcolm Ross (eds.), *The history and typology of Western Austronesian voice systems*, 17–62. Canberra: Pacific Linguistics.
Rubin, Aaron D. 2018. *Omani Mehri: A new grammar with texts*. Leiden: Brill.
Sakel, Jeanette. 2004. *A grammar of Mosetén*. Berlin & New York: Mouton de Gruyter.
Sansò, Andrea. 2016. Agent defocusing constructions from nominalized VPs. A cross-linguistic type? *Studies in Language* 40(4). 894–954.
Sansò, Andrea. 2020. Routes towards the irrealis. *Transactions of the Philological Society* 118(3). 401–446.
Seiler, Hansjakob. 1983. Possessivity, subject and object. *Studies in Language* 7(1). 89–117.
Seržant, Ilja A. 2021. Cyclic changes in verbal person-number indexes are unlikely. *Folia Linguistica Historica* 42(1). 49–86.
Siewierska, Anna. 1998. On nominal and verbal person marking. *Linguistic Typology* 2. 1–55.
Siewierska, Anna. 2004. *Person*. Cambridge: Cambridge University Press.
Starosta, Stanley, Andrew Pawley & Lawrence A. Reid. 2009 [1982]. The evolution of focus in Austronesian. In Elizabeth Zeitoun (ed.), *Formosan linguistics: Stanley Starosta's contributions*, vol. 2, 329–481. Taipei: Academia Sinica.
Suttles, Wayne. 2004. *Musqueam reference grammar*. Vancouver: University of British Columbia Press.
Tanangkingsing, Michael. 2009. *A functional reference grammar of Cebuano*. Taipei: National Taiwan University Ph.D. dissertation.
Tarpent, Marie-Lucie. 1989. *A grammar of the Nisgha language*. Wellington: University of Victoria, Wellington Ph.D. dissertation.
Terrill, Angela. 2003. *A grammar of Lavukaleve*. Berlin & New York: Mouton de Gruyter.
Topping, Donald M. 1973. *Chamorro reference grammar*. Honolulu: University of Hawai'i Press.
Vallejos, Rosa. 2016. *A grammar of Kukama-Kukamiria: A Language from the Amazon*. Leiden: Brill.
Van den Berg, Ren & Brenda H. Boerger. 2011. A Proto-Oceanic passive? Evidence from Bola and Natügu. *Oceanic Linguistics* 50(1). 221–246.
Van Lier, Eva & Marlou van Rijn. 2018. Alienability splits in action nominal constructions. *Sprachtypologie und Universalienforschung* 71(4). 631–677.

Vidal, Alejandra. 2001. *Pilagá Grammar (Guaykuruan Family, Argentina)*. Eugene: University of Oregon Ph.D. dissertation.

Wichmann, Søren. 1995. *The relationship among the Mixe-Zoquean languages of Mexico*. Salt Lake City: University of Utah Press.

Woodbury, Hanni. 2018. *A reference grammar of the Onondaga language*. Toronto: University of Toronto Press.

Yanti, {no name}. 2010. *A reference grammar of Jambi Malay*. Newark: University of Delaware Ph.D. dissertation.

Yap, Foong-Ha, Karen Grunow-Hårsta & Janick Wrona. 2011. Introduction: Nominalization strategies in Asian languages. In Foong-Ha Yap, Karen Grunow-Hårsta & Janick Wrona (eds.), *Nominalization in Asian languages. Diachronic and typological perspectives*, 1–57. Amsterdam & Philadelphia: John Benjamins.

Yeh, Maria Mei-Li. 2011. Nominalization in Saisiyat. In Foong-Ha Yap, Karen Grunow-Hårsta & Janick Wrona (eds.), *Nominalization in Asian languages. Diachronic and typological perspectives*, 561–588. Amsterdam & Philadelphia: John Benjamins.

Yoshioka, Noboru. 2012. *A reference grammar of Eastern Burushaski*. Tokyo: Tokyo University of Foreign Studies Ph.D. dissertation.