Journal of Linguistic Geography (2025), 1–12
doi:10.1017/jlg.2025.10005

CAMBRIDGE UNIVERSITY PRESS

Article

Reassessing the areality of sociative causative markers: A South American feature

Françoise Rose[1] and Marine Vuillermet[2]

[1]Laboratoire Dynamique Du Langage (CNRS/Université Lyon 2), 14 avenue Berthelot, 69007 Lyon, France and [2]Department of Comparative Language Science and Institute for the Study of Language Evolution, University of Zurich

Abstract

An avenue for the progress of areal linguistics in South America is the investigation of the geographical distribution of specific features, such as the expression of sociative causation. Sociative causation is a particular type of causation where the causer not only makes the causee do an action but also participates in it (Shibatani & Pardeshi, 2002). Guillaume & Rose (2010) hypothesize that dedicated sociative causative markers are an areal feature of South America, in particular western South America. The aim of the present paper is to reassess the spatial distribution of these markers based on a large worldwide sample of 325 languages. The results show that dedicated sociative causative markers are significantly more frequent in South America compared to the rest of the world.

Keywords: causation; voice; linguistic area; language contact; Amazonia; valency; morphosyntax

1. Introduction

The South American continent harbors exceptional linguistic diversity, as it shows the highest proportion of language families and isolates[1] per continent (Campbell, 2012; Seifart & Hammarström, 2017:260). This raises questions about the history of the settlement of South America, the most recently populated continent according to current scientific knowledge (e.g. O'Connor & Kolipakam, 2014). At the same time, within South America, some features, such as a five-member vowel system, extensive classifier system, or cross-referencing of only one argument on the verb have been found to be widely shared among languages from different stocks specifically within the area called Amazonia (Dixon & Aikhenvald, 1999:8–10). Such patterns raise the question of whether they are due to contact or are remnants of deep genetic relationships (Epps, 2009), and their close examination is instrumental in the debate on whether the Amazonian basin forms a linguistic area (see a summary in Epps & Michael, 2017).

> As descriptions of Amazonian languages increase in both number and quality, work identifying macro-areas on the basis of relatively fine-grained linguistic phenomena will probably become more common. A promising example of this type is Guillaume & Rose's (2010) suggestion that sociative causatives may be an areal feature of southwest Amazonia, with the distribution of sociative causatives outside this area attributed to the spread of Tupían languages from their southwestern homeland. A systematic examination of the distribution of such morphemes both within South America and beyond is an obvious target for future research. (Epps & Michael 2017:952)

In this work, we are undertaking a re-evaluation of the global distribution of dedicated expression for sociative causation with a worldwide sample of over 300 languages. Our results confirm the main hypotheses of Guillaume & Rose (2010) and establish dedicated sociative causative constructions as a rare phenomenon worldwide with the notable exception of South America.

1.1. Sociative causation

Sociative causation is a particular type of causation where the causer not only makes the causee do an action but also participates in it (Shibatani & Pardeshi, 2002; Zúñiga & Kittilä, 2019).[2] It is further distinguished in three semantic sub-types (Shibatani & Pardeshi, 2002), illustrated by the examples in (1), where sociative causation is expressed by a regular causative marker.

(1) Sociative causation in Japanese (Japonic, Japan; adapted from Shibatani & Pardeshi, 2002:100)

a. Joint action: 'make someone do something by doing it with him/her'

Hahaoya-ga	*kodomo-o*	*asoba-**se**-te*	*i-ru.*
mother-NOM	child-ACC	play-**CAUS**-CONJ	be-PRS[3]

'Mother is making the child play (playing with her).'

b. Assistive: 'help someone do something'

Hahaoya-ga	*kodomo-ni*	*osikko-o*	*sa-**se**-te*	*i-ru.*
mother-NOM	child-DAT	pee-ACC	do-**CAUS**-CONJ	be-PRS

'Mother is making the child pee.'

c. Supervision: 'make someone do something, and supervise this action'

Hahaoya-ga	*kodomo-ni*	*hon-o*	*yoma-**se**-te*	*i-ru.*
mother-NOM	child-DAT	book-ACC	read-**CAUS**-CONJ	be-PRS

'Mother is making the child read a book.'

Corresponding author: Françoise Rose; Email: francoise.rose@cnrs.fr

Cite this article: Rose F and Vuillermet M. Reassessing the areality of sociative causative markers: A South American feature. *Journal of Linguistic Geography* https://doi.org/10.1017/jlg.2025.10005

Shibatani & Pardeshi (2002:148ff.) consider sociative causation as an intermediate category between direct and indirect causation on the causative continuum and show that it is usually expressed either by direct or indirect causative markers, as in (1). Languages can also have markers dedicated to this category (see also Kulikov, 2001:892), as is the case with the sociative causative *ha-* in Alamblak (2b).

(2) Alamblak (Sepik, Papua New Guinea; Bruce, 1984:155–156, cited in Shibatani & Pardeshi, 2002:99)

a. Direct causation
ka*-fkne-më-r-m*
DIR.CAUS-enter-RPST-3SG.M-3PL
'He caused them to enter (something) by physically taking them.'

b. Sociative causation
ha*-fkne-më-r-m*
SOC.CAUS-enter-RPST-3SG.M-3PL
'He caused them to enter (something) by entering with them.'

c. Indirect causation
yima-r ***hay****-noh-më-r-a*
person-3SG.M **give**-unconscious-RPST-3SG.M-1SG
'A man gave me (something) (causing) me (to become) unconscious.'

1.2. Previous research on the distribution of sociative causation

Guillaume & Rose (2010) presented a preliminary worldwide survey of the dedicated expression of sociative causation by grammatical morphemes and listed 17 such cases, the large majority of which were found in South America (Map 1). The authors thus hypothesized that sociative causatives constitute an areal feature of South American languages, more precisely south-western Amazonia. Since more than half of the 15 languages listed in South America belong to the Tupian family, the authors also suggested that the feature could have developed in the Tupian family first, and then diffused to neighboring languages of other stocks. This hypothesis was supported by the likely origin of the Tupian family in the same area (Rodrigues, 1999:108).

Due to its exploratory nature, Guillaume & Rose's (2010) survey has certain limitations, some of which are already acknowledged in the article itself. First, the sample was neither geographically nor genetically balanced, and included only positive examples, i.e. languages where a sociative causative morpheme was attested, making it impossible to quantify the feature's prevalence. Additionally, the results could be biased not only by the authors being Amerindianists, but also by the descriptive tradition of valency-changing mechanisms in general and sociative causation in particular in South America (at least since the sixteenth century with Anchieta, 1595:48–49, on Tupinambá). From our experience in the current work, we have observed that descriptions of South American languages almost systematically account for valency-changing derivations, while such sections are strikingly absent from many grammars of North American languages. As for sociative causation, for Tupian languages (and most notably languages from the Tupi–Guarani branch), the usual template for grammars includes a section on "comitative causative." As a consequence, we think that a dedicated sociative causative marker would be more likely described for South American languages and with a rather transparent label.

Pöllänen (2022, 2024) is a follow-up study to Guillaume & Rose (2010) focusing on the "core" geographical area where most sociative causative markers had been found, and widening the scope to non-morphological means.[4] The survey examines a genealogically balanced sample of 32 languages from a zone covering western and southern Amazonia, the Andes, the dry Chaco Basin area and the Atacama Desert, and finds two more languages with a dedicated sociative causative marker. A detailed account of the discrepancies between Guillaume & Rose (2010), Pöllänen (2022), and the present study is available in Supplementary material 3.

Section 2 describes the aims and methodology of the present study (sample, questions, and coding). Section 3 presents the results and Section 4 discusses them. Section 5 offers a summary of the paper.

2. Aims and methodology

2.1. Aims

Guillaume & Rose (2010) have hypothesized that the presence of dedicated sociative causative markers is an areal feature of South America, with a cluster in south-western Amazonia, on the basis of their pilot survey of dedicated sociative causative markers in a worldwide convenience sample.

The major aim of the present paper is precisely to reassess the spatial distribution of dedicated sociative causative markers on the basis of a survey on a large worldwide sample of 325 languages. A secondary aim is to widen the scope beyond morphemes by including syntactic constructions that could be dedicated to sociative causation.

2.2. Language sample

The language sample for this study was built within a larger multidisciplinary project (Out of Asia SNSF Sinergia project) which, among other aims, was designed to (re)assess known linguistic areas and discover new ones in the Americas. The 325-language sample designed for this project over-samples American languages. It includes 220 languages of the Americas and 105 languages from other parts of the world, i.e. 25 languages per macro-area.[5] Within each macro-area (as defined in Hammarström & Donohue, 2014), our language sample maximizes phylogenetic diversity and favors isolates over language families, while trying to cover as much geographical space as possible. A consequence of the American bias in the sample and the maximization of phylogenetic diversity is that it is almost only in the Americas that a stock has several representative languages in the sample. Another consequence is that our results are more telling for the Americas than for the other macro-areas, where they may be under- or over-estimated. The geographical and phylogenetic distributions of the languages are illustrated in Map 2 and summarized in Table 1. Language names, genetic and macro-areal classifications, as well as geographical coordinates, follow Glottolog (Hammarström et al. 2022).

2.3. Questions and coding

All languages of the sample have been coded for the three following questions, as part of a larger questionnaire on sociative causation available in Supplementary material 1, also serving as a coding guide.

(i) Does the language have a dedicated construction to express sociative causation?

https://doi.org/10.1017/jlg.2025.10005 Published online by Cambridge University Press

Map 1. Survey of sociative causative markers in the world (Guillaume & Rose, 2010:390).

This question, labelled SocCaus.01 in the questionnaire, targets the **presence of a dedicated construction** for the expression of sociative causation, whatever the means of grammatical expression. We consider a construction to be dedicated when sociative causation is expressed by a grammatical device which exclusively expresses sociative causation. Possible answers are "yes" or "no."

(ii) If yes to the preceding question, what kind of construction is sociative causation expressed with?

This second question (SocCaus.02 in the questionnaire) concerns the **form of the sociative causative construction**, i.e. whether it is a dedicated morpheme or a specialized combination of morphemes. This is in line with Pöllänen (2022, 2024) but contrasts with Guillaume & Rose (2010), who exclusively searched for grammatical morphemes and excluded periphrasis and complex predication (see Supplementary material 3 for details of the discrepancies across studies).

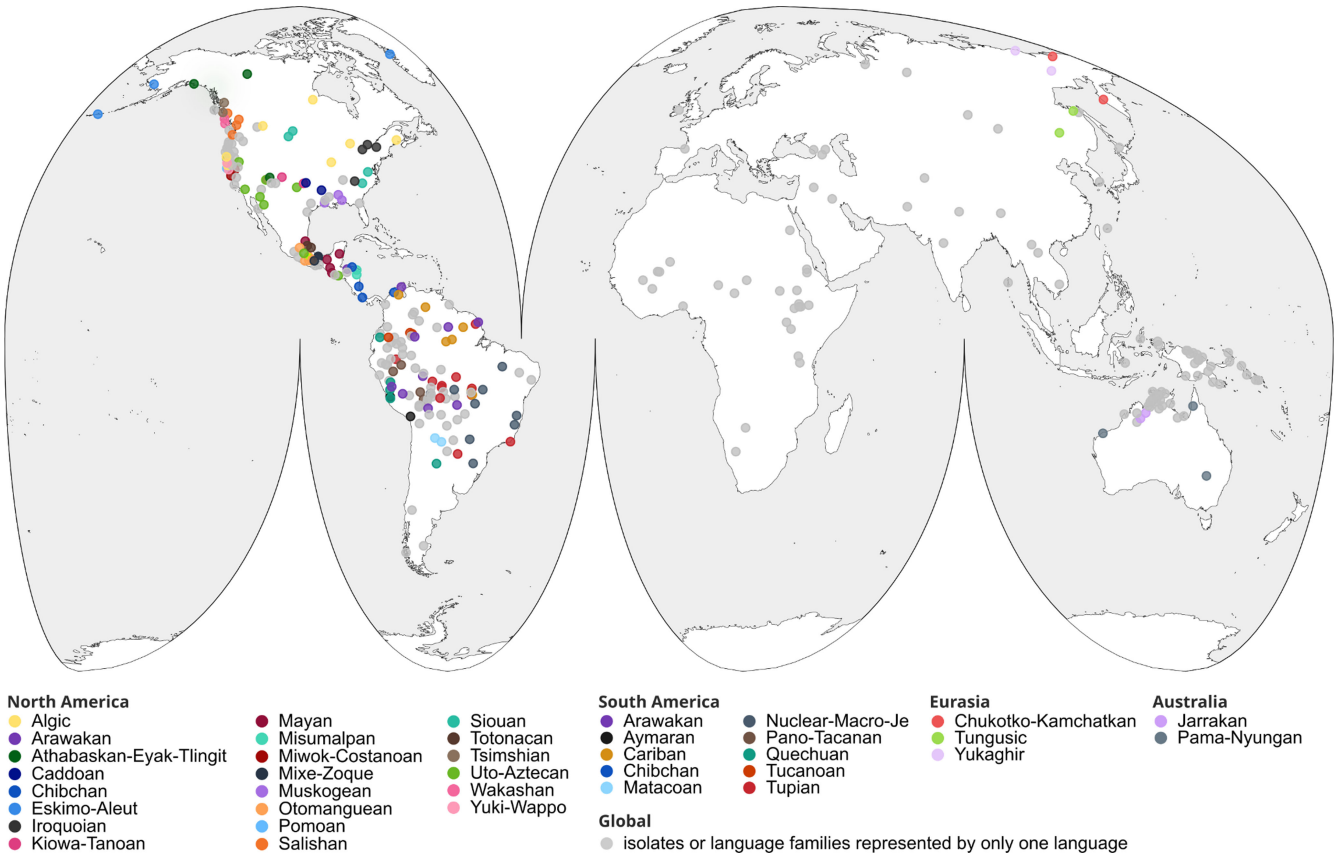
(iii) Does the language use a non-dedicated construction to encode sociative causation?

The last question (SocCaus.04) codes for **non-dedicated expressions of sociative causation**. An example is when a language has a causative marker that *sometimes* entails sociative causation. Beyond its interest for typologists, the aim of this question within the present study is to observe the distribution of the various expressions of causative sociation in different macro-areas, without a restriction to dedicated constructions.

In the remainder of the section, we describe the coding process. Because the concept of sociative causation is still not widespread, it is not always clearly identified as present or absent by grammar authors.[6] Hence the valency-changing sections of grammars were carefully scrutinized, and dozens of keywords like "help," "aid," "together" were systematically searched. Also, the semantics of causation is often not explored.[7] As a result, examples and their contexts were carefully reviewed and possible interpretations were examined and discussed among authors and research assistants (see the case of Mojeño Trinitario in the next paragraph). Consistent sociative causation meaning throughout the examples was judged necessary to consider the construction to be dedicated and yield a "yes" to the first question. In case of inconsistent meaning, we considered the polyfunctional construction not to be a dedicated sociative causation construction ("no" to the first question). Consequently, we sometimes disregarded constructions analyzed as sociative causatives in other papers, i.e. expressing sociative causation *among* other functions, and even as their "primary function." In the case of uncertainty, we contacted the grammar authors when possible (see Acknowledgments). Edge cases were always discussed collectively.

An example of a polyfunctional morpheme is the Mojeño Trinitario prefix *im-∼em-*. Its occurrence in (3a) is a good illustration of the expression of sociative causation. However, other occurrences like (3b) express causation only, while some rare occurrences like (3c) seem to express sociative only (or with a very

https://doi.org/10.1017/jlg.2025.10005 Published online by Cambridge University Press



Map 2. Geographical and phylogenetic distributions of the languages of the sample.

Table 1. Geographical and genetic distribution of the languages in the sample

	Languages	Stocks
South America	105	62
North America	115	62
Eurasia	30	27
Africa	25	25
Australia	25	22
Papunesia	25	25
	325	221[a]

[a]Two stocks appear in both North and South America: this is why the total of stocks is reduced by 2.

weak causation). Despite occurrences like (3a), we did not consider *-im* to be a dedicated sociative causative morpheme.

(3) Mojeño Trinitario (Arawak, Bolivia; Rose, 2018)

a. *v-**im**-yon=yore te' to 'p-ochkoy-pa'i*
1PL-**CAUS**-go=FUT PREP.NH ART.NH other-be.close-CLF.ground
'We are going to take him to the other side of the earth.' (19.179)

b. *ene ma tata ty-**im**-yon-nu=po te to*
and ART.M 1SG.father 3-**CAUS**-go-1SG=PFV PREP.NH ART.NH
n-chowo=po te to n-ow-sa
1SG-come_back=PFV PREP.NH ART.NH 1SG-live-NMLZ
'And my father sent me to the . . . , I went back to my village.' (9.011)

c. *mu-**em**-epeno=po*
3M-**CAUS**-die=PFV
'He died with it (could be a debt or that he died with his sins).' (Gill, 1957:154)

Our methodology is thus quantitative in terms of the number of languages surveyed, and qualitative in terms of the care provided in harvesting and coding the data.

3. Results

The raw data are given in Inman et al. (2025)[8] and Supplementary material 2. The feature set 'Categorical genderlect' in Inman et al. (2025) gives the full coding for the three questions and the 325 languages. Supplementary material 2 details each of the dedicated constructions in our sample, with information on the author's label of the construction, its attested semantics and illustrative examples.

This section highlights the major results. Section 3.1 presents the quantitative and geographical distribution of languages with dedicated sociative causative constructions in our sample. Section 3.2 examines the possible forms of the dedicated constructions. Section 3.3 measures the use of a non-dedicated construction to encode sociative causation.

3.1. Dedicated constructions for sociative causation: how many and where

Positive answers about the presence of a dedicated construction for sociative causation amount to 19 out of 325 languages. The detailed list is given in Table 2. Languages with a dedicated sociative

https://doi.org/10.1017/jlg.2025.10005 Published online by Cambridge University Press

Table 2. Languages with dedicated sociative causative constructions

Language	Stock	Macro-area	Dedicated construction type
Nama (Namibia)	Khoe-Kwadi	Africa	dedicated morpheme
Galo	Sino-Tibetan	Eurasia	
Marind	Anim	Papunesia	
Alamblak	Sepik		
Awa-Cuaiquer	Barbacoan	South America	
Shawi	Cahuapanan		
Amarakaeri	Harakmbut		
Guahibo	Guahiboan		
Krenak	Nuclear-Macro-Je		
Hup	Naduhup		
Yauyos Quechua	Quechuan		
Awetí	Tupian		
Karo (Brazil)	Tupian		
Mekens	Tupian		
Mundurukú	Tupian		
Teko	Tupian		
Cavineña	Pano-Tacanan		
Ese Ejja	Pano-Tacanan		compositional morpheme caus + com.appl
Yuracaré	isolate		special causee marking and no valency-changing morpheme

construction amount to 5.8% of our sample, while negative answers amount to 94.2%, with 306 languages. We can conclude that sociative causative constructions are non-marginally present, as one in 20 languages of our sample has them.

However, the spatial distribution of dedicated sociative causative constructions is very skewed towards South America (see Map 3). As detailed in Table 2, of the 19 languages with a dedicated construction for sociative causation in the sample, 15 are spoken in South America. The 15 South American languages with dedicated constructions for sociative causation belong to ten stocks: eight of them individually belong to different stocks (Barbacoan, Guahiboan, Naduhup, Cahuapanan, Harakmbut, Nuclear-Macro-Je, Quechuan, and one isolate, Yuracaré), five are Tupian languages, and two Pano-Tacanan. Note that our sample also includes another six Tupian, two Pano-Tacanan, eight Nuclear-Macro-Je, and four Quechuan languages that do not display a sociative causative construction.[9] The four languages with dedicated sociative causative constructions spoken outside of South America are Nama (Khoe-Kwadi, Namibia, Botswana, and South Africa), Galo (Sino Tibetan, India), Marind (Anim, Papua New Guinea and Indonesia), and Alamblak (Sepik, Papua New Guinea).

This means that among the South American languages of the sample, 14.3%, or one in seven languages, have a dedicated sociative causative construction. On the other hand, if we only consider the languages spoken outside of South America in the sample, the prevalence of dedicated constructions for sociative causation falls to only 1.8%. Even when taking into account the effect of the American bias, the results are suggestive of a certain skewing in the geographical distribution of dedicated sociative causative constructions in the languages of the world.

Our sample is unbalanced for the number of languages and stocks taken into consideration for each macro-area. In order to balance these, we generated 250 random subsamples, with 15 languages of different stocks for each macro-area. The average presence of dedicated sociative causative constructions in each of these macro-areas is very similar to the one in our sample (compare Table 4 with Table 3), confirming that the American focus in our sample is not distorting macro-areal differences. In all cases, the presence of dedicated sociative causative constructions in South America remains noteworthy.

3.2. Forms of the dedicated constructions for sociative causation

Our survey targeted any grammatical construction dedicated to the expression of sociative causation, so as to include languages which express sociative causation with dedicated devices other than a dedicated morpheme, e.g. a combination of a causative morpheme with some other marker on the verb or on the causee.

The results are given in the last column of Table 2. Most languages with a dedicated sociative causative construction encode it through a dedicated morpheme (17/19). Only two languages do not: Ese Ejja combines a causative and a comitative applicative marker, and Yurakaré uses a specific object paradigm associated with the absence of a valency-changing marker. These three types of dedicated constructions are presented next.

Dedicated sociative causative morphemes are illustrated with Marind and Teko in (4).

(4) a. Marind (Anim, Papua New Guinea and Indonesia; Olsson, 2021:333)

*ah-**e**-hwil,* *nok* *da* *mano-ɣol.*

IMP-**SOC.CAUS**-walk 1 sago FUT:1-beat.sago

'You go away with him, I'm going to beat sago.' (The speaker tells her husband to take away their infant son so she can work.)

b. Teko (Tupian, French Guiana; Rose, 2011:258)

*o-**er**-aho* *o-**ero**-ker.*

3SBJ-**SOC.CAUS**-go 3SBJ-**SOC.CAUS**-sleep

'He carries her and makes her sleep with him.'

The two cases of dedicated sociative constructions encoded differently than with a dedicated sociative causative morpheme merit special attention. We first look at Ese Ejja, which expresses sociative causation with a combination of a causative and an applicative marker. The marker *-sawa* is a comitative applicative which can occur on its own to encode co-participation only, as in (5a). It most frequently combines with the causative *-mee* to express sociative causation: in (5b), the subject helps the people to learn (lit. make them know) their language, i.e. participates in their learning.[10]

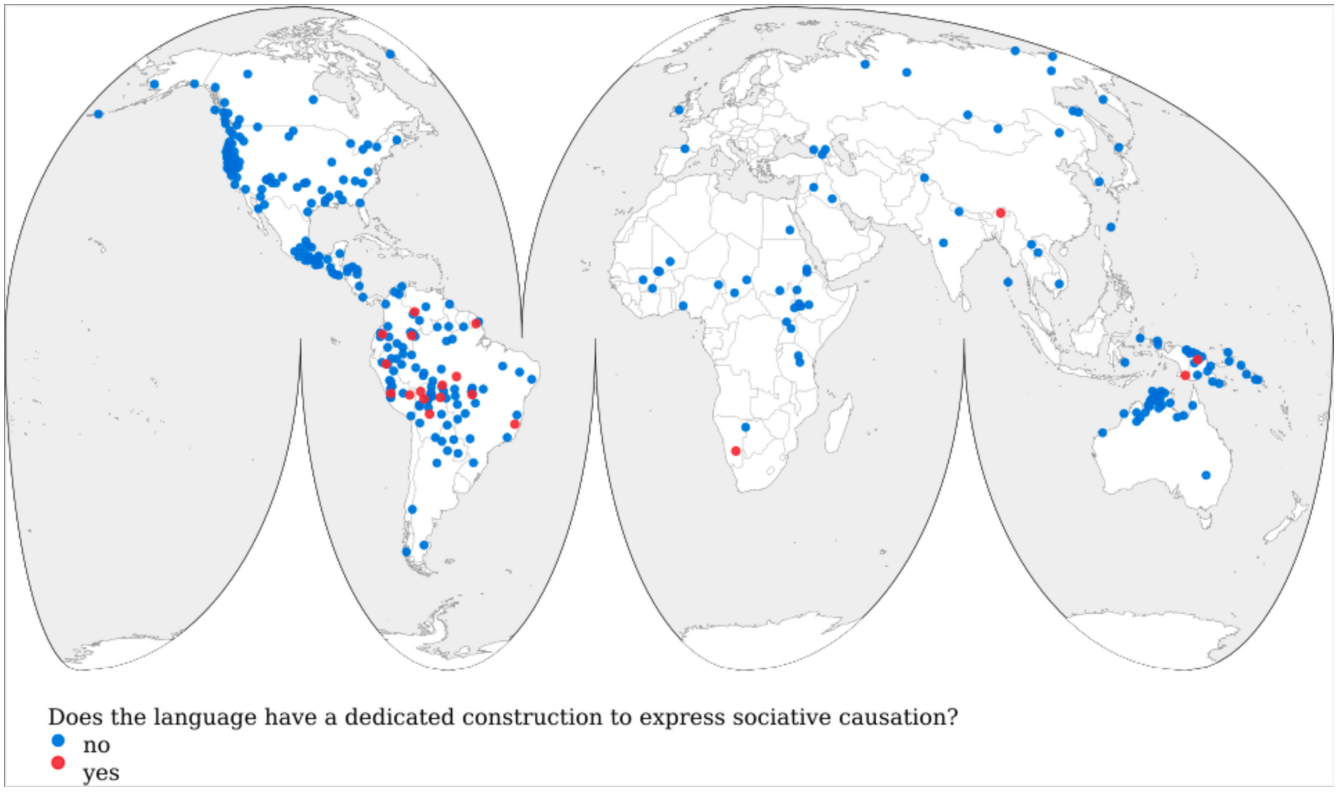
(5) Ese Ejja (Pano-Tacanan, Bolivia and Peru; author's fieldnotes)

a. *Keña=a* *Marina* *poki-**sawa**-ka-naje.*

Keña=ERG Marina.ABS go-**COM.APPL**-3A-PST

'Keña went with Marine.'

https://doi.org/10.1017/jlg.2025.10005 Published online by Cambridge University Press



Map 3. Worldwide distribution of dedicated sociative causative constructions.

Table 3. Dedicated sociative causative constructions across macro-areas

	Counts	Frequency
South America	15/105	14.3%
North America	0/115	0%
Eurasia	1/30	3.3%
Africa	1/25	4.0%
Australia	0/25	0%
Papunesia	2/25	8.0%
	19/325	5.8%

Table 4. Prevalence of dedicated sociative causative constructions across 250 subsamples

	Average	Standard deviation
South America	14	8
North America	0	0
Eurasia	3	0
Africa	4	3
Australia	0	0
Papunesia	8	4
	6	6

b. *Miyaya* *ekwana* *e-sowi*
2SG.ERG 1EXCL.ABS NPF-tongue(ABS)
*'ba-**mee**-**sawa**-añ̃a.*
see-**CAUS-COM.APPL**-PRS.A1/2
'You make us know your words.' (Shoemaker & Shoemaker, 1983)

In Ese Ejja, the combination of the sociative and the causative marker seems to systematically express sociative causation, which is not necessarily the case cross-linguistically. In Yimas, this same combination gives rise to a caused event with an additional causee (in the dative), rather than involving the causer in the caused event.

(6) Yimas (Lower Sepik-Ramu, Papua New Guinea; Foley, 1991:303)
m-n *pu-ŋa-**taŋ-tar**-kwalca-t*
NR.DIST-1SG 3PL.OBJ-1SG.DAT-**COM.APPL-CAUS**-rise-PFV
'He woke them up along with me.'

Finally, Yuracaré expresses sociative causation by prefixing a special paradigm of object indexes onto intransitive verbs. Example (7a) illustrates the intransitive verb root *yupa-* 'go in', which has a zero-marked third person subject. Example (7b) demonstrates that this intransitive root needs some valency-changing process to be used transitively. Here causation is encoded through reduplication of the final syllable, and van Gijn postulates the presence of zero A and P third person affixes. Finally, example (7c) shows this same intransitive verb root with only an overt, third person object prefix *ka-*, which not only transitivizes the verb but also implies the semantics of sociative causation.

https://doi.org/10.1017/jlg.2025.10005 Published online by Cambridge University Press

(7) Yuracaré (isolate, Bolivia; van Gijn, 2006:150, gloss adapted with van Gijn's validation)

a. *yupa-∅* *inel=chi* *choponi*
go.in.SG-3 inside=DIR pigeon
'The pigeon entered [the boat].'

b. *∅-yupa~**pa**-∅* *a-ñu* *a-sip=chi*
3-go.in.SG-**CAUS**-3 3SG.P-baby 3SG.P-house=DIR
'He puts his baby in the house.'

c. ***ka**-yupa-∅* *a-ñu* *a-tupta=y*
3SG.OBJ/SOC.CAUS-go.in.SG-3 3SG.P-baby 3SG.P-mosquito_net=LOC
'He goes into the mosquito net with his child.'

The construction in (7c) is a dedicated sociative causation construction because the use of this object indexing paradigm on intransitive verbs necessarily triggers the sociative causation meaning. Note that this paradigm differs from the regular object indexing paradigm only in the third person (∅- for the 'regular' object indexing paradigm vs. *ka-* for the sociative causative). This means that with first and second person, it is the mere presence of the object index (and the absence of a valency-changing marker) which induces sociative causation. We have consequently considered this grammatical expression to be a construction rather than a single morpheme.

The Yurakaré sociative causative marker partially recalls the "special causee marking" in Punjabi, a type of dedicated construction reported in only one language in Guillaume & Rose (2007), and not attested in our sample. As shown in (8a), the verb form *cukvãiã* is causativized with a causative suffix, which expresses a regular causative if the causee is encoded with an ablative. In (8b), the same causativized verb form *cukvãiã* along with the accusative/dative marker *nũ* marking the causee expresses sociative causation.

(8) Punjabi (Indo-Aryan, Bangladesh, India and Pakistan; Khokhlova, 2003, cited in Guillaume & Rose, 2007)

a. Ablative case: regular causative meaning
us *ne* *[**kulĩ*** *to**ⁿ**]*** *sũṭkes* *cukvãiã*
3OBL ERG **porter** **ABL** suitcase lift+**CAUS**+PP.M.SG
'He made the porter lift (his) luggage.'

b. Accusative/dative case: sociative causative meaning
to *asĩⁿ* *[**us*** *nũ]* *samãn* *cukvãiã*
then we **3OBL** **ACC/DAT** luggage lift+**CAUS**+PP.M.SG
'(The old woman was weak and did not manage to lift her language), so we helped her to lift it.'

The choice of including formal means other than dedicated morphemes does not affect the results much: the great majority of dedicated expressions of sociative causation are morphemes, found in 17 of the 19 languages. Pöllänen (2024:23) rightly notes that this is expected given the tendency towards rich and agglutinative verbal morphology in western South American languages.

3.3. Non-dedicated constructions to encode sociative causation: where?

This last result targets the non-dedicated constructions encoding the sociative causation meaning. By non-dedicated construction, we mean that the language makes use of a construction which can but does not systematically express sociative causation. Example (9) illustrates the use of the comitative marker *mu-* in Totontepec Mixe, which adds an instrument to the argument structure of the verb in (9a), as a marker of sociative causation with motion verbs involving inanimate objects, as in (9b).

Table 5. Non-dedicated sociative causative constructions across macro-areas

	Counts	Frequency
South America	33/105	31.5%
North America	26/115	22.6%
Eurasia	11/30	36.7%
Africa	12/25	48%
Australia	7/25	28%
Papunesia	10/25	40%
	99/325	30.5%

(9) Totontepec Mixe (Mixe-Zoquean, Mexico; Guzmán, 2012:231, 240)

a. *'ëts* *tsöjx* *nmutuump*
'ëts *tsöx* *n=**mu**-tön-py*
1SG machete 1A=**COM**-work-INCOMPL.TR
'I work with a machete.'

b. *mu-mats*
COM-come
'bring it'

We will not go into the details of the particular constructions because our goal in surveying these non-dedicated constructions within this article is to contrast their geographical distribution with that of dedicated constructions (Section 3.1), rather than to refine the existing typology of the expression of the sociative causation meaning.[11]

Table 5 presents the distribution of non-dedicated constructions for sociative causation in the different macro-areas. It is not geographically skewed, and in particular, there is no strong bias towards South America: on average, 31.5% of the South American languages of the sample present a non-dedicated construction, while the average for the total sample is 30.5%.

4. Discussion

Section 4.1 discusses the worldwide presence and distribution of dedicated sociative causative constructions, while Sections 4.2 and 4.3 discuss their genetic and geographical distribution, respectively.

4.1. Worldwide presence of dedicated sociative causative constructions

Our large sample allows us to evaluate the overall frequency of dedicated sociative causative constructions. They are found in 19 languages out of 325, about 5.8% of the languages of the sample. As such, they cannot be considered extremely rare. However, because their presence is denser in South America, their prevalence outside of South America is particularly low, and even negligible, with four cases out of 220 languages spoken outside of South America (< 2%).

The present paper is not meant to be a comprehensive, worldwide report of dedicated sociative causative markers in the literature, but a survey of its distribution in a large, balanced and worldwide sample. For a fuller inventory, one should add to the 19 languages of our study the nine additional ones listed in Guillaume

https://doi.org/10.1017/jlg.2025.10005 Published online by Cambridge University Press

& Rose (2010), and the one additional language in Pöllänen (2022, 2024). However, this would require an assessment of all reported cases, as we do not necessarily endorse the analyses by these authors. Supplementary material 3 documents the case of disagreement between the three studies, but does not assess the analysis of languages which are not part of our sample.

The absence of sociative causative constructions from North American languages is very telling: they are completely absent from a total of 115 languages distributed in 62 language stocks. By contrast, we cannot exclude that sociative causative constructions might be more present in Africa, Papunesia, and Eurasia because the number of languages investigated in these macro-areas is much lower (25 to 30). Sociative causatives are plausibly present in more than one or two languages in these areas, especially given the biases explained in the methodology. This represents an exciting topic to investigate in the future.

4.2. Genetic distribution

Languages with dedicated sociative causative constructions in our sample belong to 14 different stocks (13 language families and an isolate, Yuracaré). Of these 13 families, only four, all from South America, are represented by more than one language in our sample. Two of the four families display a dedicated sociative causative construction in several languages: two out of four Pano-Tacanan languages, and five out of 11 Tupian languages.

The third and fourth families, the Nuclear-Macro-Je and Quechuan families, albeit represented by nine and five languages respectively, have only one language each with a dedicated marker (Krenak and Yauyos Quechua).

Within the Pano-Tacanan stock, Cavineña and Ese Ejja, both of the Tacanan branch, show a dedicated construction. However, the two constructions are not directly related diachronically: the construction in Cavineña is based on a dedicated sociative causative morpheme *-kere* and the construction in Ese Ejja is bimorphemic with the causative *-mee* and the sociative *-sawa*.[12] Note that the two languages from the Panoan branch present in the sample have non-dedicated constructions to express sociative causation, so that the semantic domain of sociative causation might be particularly salient in the Pano-Tacanan stock, whether expressed by a dedicated construction or not.

As for the Tupian stock, a dedicated sociative causative morpheme is found in five languages of the survey, belonging to four different branches of the family: Teko and Aweti belong to the large Maweti–Guarani branch, Karo to the Purubora-Ramarama branch, Mekens to the Tuparic branch and Mundurukú to the Mundurukuic branch. As mentioned above, these dedicated markers are traditionally called "comitative causative" and are considered cognates and reflexes of the Proto-Tupi verbal prefix **er^{j}e- ~er^{j}o-* reconstructed by Rodrigues & Cabral (2012:509, 531–533).[13] As for the six Tupian languages within our sample that do not show a dedicated sociative causative marker, they belong either to branches with no reflex of **er^{j}e- ~er^{j}o-* (Gavião do Jiparaná, Jurúna, Karitiâna), or to the Maweti–Guarani branch (Tupinambá, Paraguayan Guarani, and Cocama-Cocamilla). In Tupinambá, the functions of the reflex of **er^{j}e- ~er^{j}o-* cover non-causative meanings, so that we do not treat it as a dedicated sociative causative marker (see Section 2.3 on the exclusion of polyfunctional morphemes). Paraguayan Guarani and Cocama-Cocamilla have lost the dedicated sociative causative marker. In Paraguayan Guarani, the reflex of Maweti–Guarani **er^{j}e-* (Corrêa de Silva, 2010:218) is not fully productive and its combination with some verb roots acquired a conventionalized meaning (Estigarribia, 2020:218–219). In Cocama-Cocamilla, it has fossilized; we uncovered it in a few verb roots, but it is not recognized as an independent morpheme in grammar descriptions. The Tupian family thus constitutes a nice showcase for dedicated sociative causative markers, with a suggested reconstructed form, a hypothetical lexical source for it,[14] inheritance with formal differentiation through a number of branches,[15] and some examples of loss of the dedicated marker.

4.3. Geographic distribution

Section 3.1 states that the distribution of dedicated sociative causative constructions is skewed towards South America. In contrast, Section 3.3 highlights that there was no such bias for non-dedicated constructions. This shows that the general concept of sociative causation is expressed to a similar degree around the world, but tends to grammaticalize almost exclusively in South America, confirming the very areal status of dedicated sociative causative constructions.

The prevalence of dedicated sociative causative constructions in South America cannot be explained by chance, as they are almost absent from the rest of the world, nor only by genetic inheritance, since they are found in many different stocks in South America. Consequently, their particular geographical distribution seems to be explainable as a result of diffusion across languages. A further contribution of this paper is to point out that dedicated sociative causative constructions may in the future serve as an excellent feature to observe diffusion of linguistic features in an area, and to argue for a linguistic area. As already presented in the introduction, shared linguistic features have already been argued to account for linguistic areas in South America: sometimes for Amazonia as a whole, recently more often for a western/eastern divide of South America, and for more reduced areas such as the Guaporé–Mamoré region. Aikhenvald (2012) posits a divide between three major areas: Andean, Amazonian, and Southern Cone. Nevertheless, the possibility of the whole Amazonian basin forming a single linguistic area has been a matter of debate (see a summary in Epps & Michael, 2017), with some quantitative studies supporting instead a western/eastern division of South America (see e.g. Birchall, 2014).[16]

Map 4 zooms in on the distribution of languages with dedicated sociative causative constructions within our South American sample. We observe that most cases are found within the area generally defined as "Greater Amazonia,"[17] with the two exceptions Awa-Cuaiquer and Yauyos Quechua, which are spoken on the western slopes of the Andes. Within Greater Amazonia itself, the highest density of cases is found in south-western Amazonia. However, we are not yet in a position to comment further on the areal distribution of the sociative causative within South America, because a denser sample would be necessary to allow for a finer delimitation of areas comprising the languages with a dedicated construction. In fact, the main goal of the major research project Out of Asia, of which the present study is part, is to analyze the distribution of a set of linguistic features, including the sociative causative, with a Bayesian algorithm (sBayes, by Ranacher et al., 2021) to detect areal signal, controlling for universal preference and genetic inheritance.

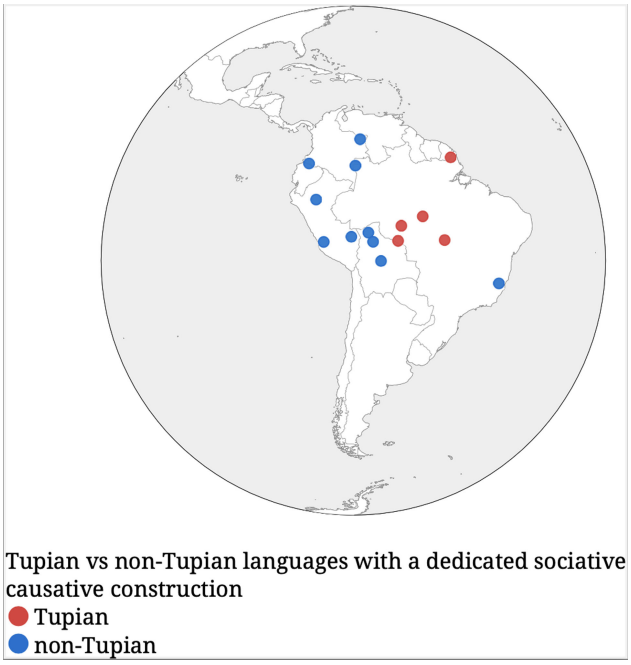
Our survey has not uncovered a formal resemblance between dedicated sociative causative morphemes or constructions across stocks, so we can only suggest that a potential diffusion of this feature would not have taken place through borrowing of a

https://doi.org/10.1017/jlg.2025.10005 Published online by Cambridge University Press

Map 4. Geographical distribution of dedicated sociative constructions in South America.

construction or morpheme, but through replication of a pattern (the plain fact of having a morpheme or a special construction for the specific function of sociative causation).[18] Each innovative language would have used inherited material to create a dedicated construction on the basis of an external model, and the diffusion would have resulted from a series of individual replications from one language to the other. It is an important result for contact linguistics, because "little attention has been granted in the literature to the borrowing of features belonging to the domain of verbs" (Matras, 2007:44). Our results conform to the generalization that "contact phenomena in the area of voice and valency are almost exclusively pattern-oriented" (Matras, 2007:47). Pattern replication is often characteristic of linguistic convergence that can lead to the building of linguistic areas (Matras & Sakel, 2007). Finding evidence of some concrete examples from contact-induced transfer is far beyond the reach of the present study, but we definitively call for such an investigation.

A final contribution our survey offers to further research on the diffusion of a dedicated sociative causative construction is to highlight the importance of Tupian languages within the cases of dedicated sociative constructions, with five cases out of 19. Moreover, Map 5 shows that these Tupian languages are central to the area where languages with dedicated sociative causative constructions are found. This is suggestive of a potentially central role of Tupian languages in that diffusion process (across stocks, but maybe also within the Tupian stock; see note 15). Very briefly, the large and dense geographical diffusion of Tupian languages from a core in south-western Amazonia (Noelli, 2008; Rodrigues & Cabral, 2012; dos Santos et al., 2015; O'Hagan et al., 2019), their dedicated sociative causative morpheme that seems to have been rather stable through time (Rodrigues & Cabral, 2012:509, 531–533), and their many contact situations (Cabral, 1995; Rodrigues, 1996; Muysken, 2012) could have made them instrumental in the



Map 5. Tupian vs. non-Tupian languages with a dedicated sociative causative construction.

diffusion of dedicated sociative causative constructions. The specificities of this hypothetical scenario are left for future research.

5. Summary

The goal of this paper was to reassess, on the basis of a worldwide sample of over 300 languages with a focus on the Americas, the

https://doi.org/10.1017/jlg.2025.10005 Published online by Cambridge University Press

areal relevance of sociative causative markers hypothesized by Guillaume & Rose (2010). The scope of our investigation was wider than that of Guillaume & Rose (2010), by including expressions dedicated to sociative causation other than morphemes. The result is that out of 325 worldwide languages, 19 show a dedicated construction for sociative causation, and in 17 cases out of 19, it is expressed by a plain dedicated sociative causative marker. Importantly, 15 of these 19 languages are spoken in South America. The present study has confirmed that dedicated sociative causative constructions is a rara outside of South America, but not in South America. The prevalence of dedicated constructions in this macro-area is all the more telling that non-dedicated constructions for sociative causation are evenly distributed across macro-areas, reaching an average of one in four languages: South America is the only macro-area where sociative causation is frequently grammaticalized.

Anchored within a project aiming at uncovering historical contact among languages in the Americas, our survey provides data that may, in conjunction with comparable data on other linguistic features, be instrumental in informing future research on areal patterns in South America.

Supplementary material. To view supplementary material for this article, please visit https://doi.org/10.1017/jlg.2025.10005

Acknowledgments. This work is based on the ATLAs database (Inman et al., 2025) which was funded primarily by the Out of Asia SNSF Sinergia project CRSII5_183578 and the Swiss National Centre of Competence in Research Evolving Language. We would like to warmly thank David Inman and Natalia Chousou-Polydouri for invaluable discussions on the questionnaire, and revision of our work, as well as Kellen Parker van Dam for generating the maps. The coding work by Oscar Cocaud-Degrève and Raphaël Luffroy was also fundamental to this study. It was in very large part based on grammars generously shared by Harald Hammarström. Finally, we would also like to thank the linguists who provided additional information about particular individual languages, namely Alexandra Aikhenvald on Tariana, Les Bruce on Alamblak, Fernando de Carvalho and Emerson José Silveira da Costa on Tupinambá, Eva Lindström on Kuot, Florian Lionnet on Laal, Andrey Nikulin Guzmán on Borum (Krenak), Bruno Olsson on Marind, Mark Post on Galo, Yvonne Treis on Kambaata and An Van linden on Harakmbut.

Competing interests. The authors declare none.

Notes

1 In this paper, we define a "language family" as languages with a shared ancestry and "isolates" as language families with only one member language. The term "stock" covers both language families and isolates (see Nichols, 1992, or Whalen & Simons, 2012).

2 Alternative terms that have been used for the phenomenon include causation of involvement (Dixon, 2000), comitative causation in the descriptive tradition of the Amazonian language family Tupi (Adam, 1896; Rodrigues, 1953; Jensen, 1998; Rodrigues & Cabral, 2012), and assistive (Kulikov, 1993:131, 2001:892).

3 Examples are from a variety of sources. We have translated the glosses in the examples when they were not in English in the original work. We have homogenized the glosses using the Leipzig Glossing Rules with the addition of the gloss SOC.CAUS for the markers (re)analyzed as sociative causative. This is especially relevant for the sociative causative markers that had not been analyzed as such in their original work. Abbreviations absent from the Leipzig Glossing Rules are as follows: CONJ = conjunction; DIR = direction; DIR.CAUS = direct causative; INCOMPL = incompletive; NH = non-human; NPF = nominal prefix; NR.DIST = near-distal; PP = perfective participle; PREP = preposition; RPST = remote past; SOC.CAUS = sociative causative.

4 Guillaume & Rose (2007) had highlighted that sociative causation could be encoded through other valency-changing mechanisms, such as applicative, reciprocal, or sociative/cooperative markers.

5 Eurasia has an additional five languages from Siberia because of their crucial interest to the Out of Asia hypothesis.

6 In most grammars, the sociative causative marker was not glossed as such. For instance, it was glossed ACPN for 'accompaniment' in Marind (Olsson, 2021:333) and DE.CAUS for direct event causative in Alamblak (Bruce, 1984:155). For the sake of clarity, as already indicated in note 3, we have modified the heterogeneous glosses to SOC.CAUS.

7 In her paper on a semantic map of causation, Levshina (2022:181, our bold) similarly observes that descriptive works focus on "the most typical factitive implicative causation (i.e. making something happen or someone do something)" and neglect other types like the "permissive, accidental curative, **assistive**, directive, non-implicative, etc."

8 The data are available at https://zenodo.org/records/15227808 and are repeated on the ATLAs website (http://atlas.evolvinglanguage.ch).

9 These languages are namely Gavião do Jiparaná, Jurúna, Tupinambá, Cocama-Cocamilla, Karitiâna, and Paraguayan Guaraní for the Tupian languages, Matsés and Shipibo-Conibo for the Pano-Tacanan languages, Djeoromitxí, Xavánte, Kaingang, Canela-Krahô, Karajá, Maxakalí, Ofayé, and Rikbaktsa for the Nuclear-Macro-Je languages, and Huallaga Huánuco Quechua, Imbabura Highland Quichua, North Junín Quechua, and Santiago del Estero Quichua for the Quechuan languages.

10 Note that we could also have analyzed Ese Ejja as having a dedicated morpheme *-meesawa* synchronically, historically made up of a combination of two morphemes. This analysis was dispreferred because of the productivity of *-mee* as a regular causative marker, and transparency of its causative contribution in combination with *-sawa*, and the possibility of inserting a morpheme between the two.

11 In a nutshell, the non-dedicated constructions attested in our sample are mostly valency-changing markers, with the most widely attested being regular causative markers, and the second most common applicative markers.

12 The three other languages from the Tacanan branch all have a dedicated sociative causative construction which is also often translated as "help" (Guillaume, 2012:209, to appear; Pitman, 1980:57). It is marked with *-tsawa*, cognate with the Ese Ejja sociative marker *-sawa*.

13 In fact, the comitative causative prefix of Tupari languages is taken as a reflex of *erje-/erjo-* by Rodrigues & Cabral (2012) in one part of their paper (p. 509), but considered to have a different origin later in the paper (pp. 532–533).

14 Rodrigues & Cabral (2012:532) draw a historical link between the Tupi "comitative causative" verbal prefix and a comitative postposition.

15 One should not discard the possibility that the reconstructed form did not have a sociative causative meaning to start with, but that when this function arose later, it developed in parallel in related languages and/or diffused among related languages through contact.

16 Regional linguistic areas at smaller scale are widely accepted and result from contact zones with frequent interaction between speakers of multiple languages (Campbell, 2012; Epps & Michael, 2017), for example the Vaupés (Aikhenvald, 2002; Epps, 2007), the Chaco (Comrie, Golluscio, Gonzáles & Vidal, 2010), and the Guaporé–Mamoré area (Crevels & van der Voort, 2008).

17 Greater Amazonia is the area covered by the hydrographic basins of the Amazon and its affluents and pushing northward and eastward to the littoral, thus also including the Guianas and Orinoco basin, as well as the eastern plains (Denevan, 1976).

18 See Matras & Sakel (2007) for pattern replication in linguistic convergence

References

Adam, Lucien. 1896. *Matériaux pour servir à l'établissement d'une grammaire comparée des dialectes de la famille Tupi (Bibliothèque Linguistique Américaine XVIII)*. Paris: Librairie-Éditeur J. Maisonneuve.

Aikhenvald, Alexandra Y. 2002. *Language contact in Amazonia*. Oxford: Oxford University Press.

Aikhenvald, Alexandra Y. 2012. *Languages of the Amazon*. Oxford & New York: Oxford University Press.

Anchieta, Joseph de. 1595. *Arte de Gramatica da Lingua mais Usada na Costa do Brasil*. Coimbra: Antonio Mariz.

Birchall, Joshua. 2014. *Argument marking patterns in South American languages*. Radboud Universiteit Nijmegen doctoral dissertation.

Bruce, Les. 1984. *The Alamblak language of Papua New Guinea (East Sepik) (Pacific Linguistics, Series C, 81)*. Canberra: Research School of Pacific and Asian Studies, Australian National University.
Cabral, Ana Suelly. 1995. *Contact-induced language change in the western Amazon: The non-genetic origin of the Kokama language*. Pittsburgh: University of Pittsburgh Ph.D. dissertation.
Campbell, Lyle. 2012. Classification of the indigenous languages of South America. In Lyle Campbell & Verónica Grondona (eds.), *The indigenous languages of South America: A comprehensive guide*, 59–166. Berlin & Boston: De Gruyter Mouton.
Comrie, Bernard, Lucía A. Golluscio, Hebe Gonzáles & Alejandra Vidal. 2010. El Chaco como área lingüística. *Estudios de lenguas amerindias 2*. 85–130.
Corrêa da Silva, Beatriz Carretta. 2010. *Mawé/Awetí/Tupí-Guaraní: Relações linguísticas e implicações históricas*. Universidade de Brasília Ph.D. dissertation.
Crevels, Mily & Hein van der Voort. 2008. The Guaporé–Mamoré region as a linguistic area. In Pieter Muysken (ed.), *From linguistic areas to areal linguistics (Studies in Language Companion Series 90)*, 151–179. Amsterdam & Philadelphia: John Benjamins.
Denevan, William M. 1976. *The native population of the Americas in 1492*. Madison, WI: University of Wisconsin Press.
Dixon, R. M. W. 2000. A typology of causatives: Form, syntax and meaning. In R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *Changing valency: Case studies in transitivity*, 30–83. Cambridge: Cambridge University Press.
Dixon, R. M. W. & Alexandra Y. Aikhenvald. 1999. Introduction. In R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *The Amazonian languages*, 1–21. Cambridge: Cambridge University Press.
Epps, Patience. 2007. Grammatical borrowing in Hup. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective (Empirical Approaches to Language Typology 38)*, 551–565. New York: Mouton de Gruyter.
Epps, Patience. 2009. Language classification, language contact, and Amazonian prehistory. *Language and Linguistics Compass 3*(2). 581–606.
Epps, Patience & Lev Michael. 2017. The areal linguistics of Amazonia. In Raymond Hickey (ed.), *The Cambridge handbook of areal linguistics (Cambridge Handbooks in Language and Linguistics)*, 934–963. Cambridge: Cambridge University Press.
Estigarribia, Bruno. 2020. *A grammar of Paraguayan Guarani*. London: UCL Press.
Foley, William A. 1991. *The Yimas language of New Guinea*. Stanford: Stanford University Press.
Gijn, Rik van. 2006. *A grammar of Yurakaré*. Radboud Universiteit Nijmegen Ph.D. dissertation.
Gill, Wayne. 1957. *Trinitario grammar*. San Lorenzo de Mojos: Misión Nuevas Tribus.
Guillaume, Antoine. 2012. Maropa (Reyesano). In Mily Crevels & Peter Muysken (eds.), *Las lenguas de Bolivia*, 191–229. La Paz: Plural Editores.
Guillaume, Antoine. to appear. Takanan languages. In Patience Epps & Lev Michael (eds.), *Amazonian Languages. An International Handbook* (Handbooks of Linguistics and Communication Science), vol. 4. De Guyter Mouton.
Guillaume, Antoine & Françoise Rose. 2007. A typology of sociative causative: Between causatives and applicatives. Presented at the Association for Linguistic Typology VII, Paris, September 25–28, 2007.
Guillaume, Antoine & Françoise Rose. 2010. Sociative causative markers in South-American languages: A possible areal feature. In F. Floricic (ed.), *Essais de typologie et de linguistique générale: Mélanges offerts à Denis Creissels* (Langages), 383–402. Lyon: Presses de l'École Normale Supérieure.
Guzmán, Verónica. 2012. *Las construcciones aplicativas en el mixe de Totontepec*. México D.F.: Centro de Investigaciones y Estudios Superiores en Antropología Social Master thesis.
Hammarström, Harald & Mark Donohue. 2014. Some principles on the use of macro-areas in typological comparison. *Language Dynamics and Change 4*(1). 167–187.
Hammarström, Harald, Robert Forkel, Martin Haspelmath & Sebastian Bank. 2022. *Glottolog database 4.6*. https://glottolog.org/
Inman, David, Natalia Chousou-Polydouri, Marine Vuillermet, Kellen Parker van Dam, Shelece Easterday, Françoise Rose, Alena Witzlack-Makarevich, et al. 2025. The Areal Typology of Languages of the Americas (ATLAs) database. *Scientific Data 12*(933). http://atlas.evolvinglanguage.ch
Jensen, Cheryl. 1998. Comparative Tupí–Guaraní morpho-syntax. In Desmond Derbyshire & Geoffrey Pullum (eds.), *Handbook of Amazonian languages*, vol. *IV*, 490–603. Berlin: Mouton de Gruyter.
Khokhlova, L. 2003. Infringement of morphological and syntactic operations' pairing in "second causative" formation. *Indian Linguistics 64*(1). 1–17.
Kulikov, Leonid. 1993. The "second causative": A typological sketch. In Bernard Comrie & Maria Polinsky (eds.), *Causatives and transitivity*, 121–154. Amsterdam, Philadelphia: John Benjamins.
Kulikov, Leonid. 2001. Causatives. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language typology and language universals*, vol. *2*, 886–898. Berlin, New York: Walter de Gruyter.
Levshina, Natalia. 2022. Semantic maps of causation: New hybrid approaches based on corpora and grammar descriptions. *Zeitschrift für Sprachwissenschaft 41*(1). 179–205.
Matras, Yaron. 2007. The borrowability of structural categories. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective (Empirical Approaches to Language Typology 38)*, 31–73. Berlin: Mouton de Gruyter.
Matras, Yaron & Jeanette Sakel. 2007. Introduction. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective (Empirical Approaches to Language Typology 38)*, 1–13. Berlin: Mouton de Gruyter.
Muysken, Pieter. 2012. Contacts between indigenous languages in South America. In Lyle Campbell & Verónica Grondona (eds.), *The indigenous languages of South America: A comprehensive guide*, 235–258. Berlin & Boston: Mouton de Gruyter.
Nichols, Johanna. 1992. *Linguistic diversity in space and time*. The University of Chicago Press.
Noelli, Francisco Silva. 2008. The Tupi expansion. In Helaine Silverman & William H. Isbell (eds.), *Handbook of South American archaeology*, 659–670. New York: Springer.
O'Connor, Loretta & Vishnupriya Kolipakam. 2014. Human migrations, dispersals, and contacts in South America. In Loretta O'Connor & Pieter Muysken (eds.), *The native languages of South America*, 29–55. Cambridge: Cambridge University Press.
O'Hagan, Zachary, Natalia Chousou-Polydouri & Lev Michael. 2019. Phylogenetic classification supports a northeastern Amazonian Proto-Tupí-Guaraní homeland. *LIAMES: Línguas Indígenas Americanas 19*. 1–29.
Olsson, Bruno. 2021. *A grammar of Coastal Marind (Mouton Grammar Library 87)*. Berlin: Mouton de Gruyter.
Pitman, Donald. 1980. *Bosquejo de la Gramática Araona* (Notas Lingüísticas de Bolivia 9). Riberalta, Bolivia: Instituto Lingüístico de Verano en colaboración con el Ministerio de Educación y Cultura, Dirección Nacional de Antropología.
Pöllänen, Roosa. 2022. *Sociative causative in South American languages: Formal–functional analysis*. Helsinki: University of Helsinki Master's thesis.
Pöllänen, Roosa. 2024. Sociative causative in South American languages: Formal–functional analysis. *LIAMES: Línguas Indígenas Americanas 24*. 1–33.
Ranacher, Peter, Nico Neureiter, Rik van Gijn, Barbara Sonnenhauser, Anastasia Escher, Robert Weibel, Pieter Muysken & Balthasar Bickel. 2021. Contact-tracing in cultural evolution: A Bayesian mixture model to detect geographical areas of language contact. *Journal of the Royal Society Interface 18*(181).
Rodrigues, Aryon Dall'Igna. 1953. Morfologia do verbo Tupí. *Letras 1*. 121–152.
Rodrigues, Aryon Dall'Igna. 1996. As línguas gerais sul-americanas. *Papia 4*(2). 6–18.
Rodrigues, Aryon Dall'Igna. 1999. Tupí. In R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *The Amazonian languages*, 107–124. Cambridge: Cambridge University Press.
Rodrigues, Aryon Dall'Igna & Ana Suelly Arruda Câmara Cabral. 2012. Tupían. In Lyle Campbell & Verónica Grondona (eds.), *The Indigenous languages of South America: A comprehensive guide*, 495–574. Berlin & Boston: Mouton de Gruyter.
Rose, Françoise. 2011. *Grammaire de l'émérillon teko: Une langue tupi-guarani de Guyane française (Langues et sociétés d'Amérique traditionnelle 10)*. Leuven & Walpole, MA: Peeters.

https://doi.org/10.1017/jlg.2025.10005 Published online by Cambridge University Press

Rose, Françoise (collector). 2018. Corpus mojeño trinitario. ORTOLANG (Open Resources and TOols for LANGuage). https://hdl.handle.net/11403/corpus-mojeno-trinitario/v2

Santos, Eduardo José Melo dos, Andréa Luciana Soares da Silva, Paloma Daguer Ewerton, Louise Yukari Takeshita & Maria Helena Thomaz Maia. 2015. Origins and genetic dynamics of Tupí expansion: A genetic tale. *Boletim do Museu Paraense Emílio Goeldi: Ciências humanas 10*(2). 217–228.

Seifart, Frank & Harald Hammarström. 2017. Language isolates. In Lyle Campbell (ed.), *Language isolates in South America (Routledge Language Family Series)*, 260–286. London & New York: Routledge, Taylor & Francis.

Shibatani, Masayoshi & Prashant Pardeshi. 2002. The causative continuum. In Masayoshi Shibatani (ed.), *The grammar of causation and interpersonal manipulation (Typological Studies in Language 48)*, 85–126. Amsterdam: John Benjamins.

Shoemaker, Jack & Nola Shoemaker. 1983. *El discurso en el idioma ese ejja: Relaciones comunicacionales en la gramática ese ejja*. La Paz: Instituto Nacional de Estudios Lingüísticos, ILV.

Whalen, D. H. & Gary F. Simons. 2012. Endangered language families. *Language 88*(1). 155–173.

Zúñiga, Fernando & Seppo Kittilä. 2019. *Grammatical voice (Cambridge Textbooks in Linguistics)*. Cambridge: Cambridge University Press.

https://doi.org/10.1017/jlg.2025.10005 Published online by Cambridge University Press